Vol. 13 2020 No. 01

GEOGRAPHY ENVIRONMENT SUSTAINABILITY

Special Issue «Border and Coastal Areas of Greater Eurasia: Environmental and Socio-economic Challenges»
Special Issue «Spatiotemporal Analysis in Medical and Veterinary Geography»

EDITORIAL BOARD

CONTENTS

SPECIAL ISSUE
«BORDER AND COASTAL AREAS OF GREATER EURASIA: ENVIRONMENTAL AND SOCIO-ECONOMIC CHALLENGES»

Guest editors:

•**Prof. Vladimir Kolosov**, Deputy Director and Head of Laboratory at the Institute of Geography of Russian Academy of Sciences;

•**Prof. Alexander Druzhinin**, Director of the North Caucasian Institute of Socio-Economic Studies of the Southern Federal University, Rostov-on-the Don;

•**Dr. Paul Fryer**, Adjunct Professor and Academic Head of «Border Crossing's Master's Degree Programme, University of Eastern Finland;

Göran Roos, Ksenia Y. Voloshenko, Tatiana E. Drok, Yury M. Zverev
EUROPEAN COUNTRIES' TYPOLOGY BY THE INTENSITY OF TRANSBOUNDARY COOPERATION AND ITS IMPACT ON THE ECONOMIC COMPLEXITY LEVEL 6

Andrei G. Manakov, Jaak Kliimask
RUSSIAN-ESTONIAN BORDER IN THE CONTEXT OF POST-SOVIET ETHNIC TRANSFORMATIONS 16

Tadeusz Palmowski and Gennady M. Fedorov
THE POTENTIAL FOR DEVELOPMENT OF RUSSIAN-POLISH CROSS-BORDER REGION........ 21

Denis A. Dirin and Paul Fryer
THE SAYAN BORDERLANDS: TUVA'S ETHNOCULTURAL LANDSCAPES IN CHANGING NATURAL AND SOCIOCULTURAL ENVIRONMENTS 29

Alexander G. Druzhinin, Tatyana Yu. Kuznetsova, Andrey S. Mikhaylov
COASTAL ZONES OF MODERN RUSSIA: DELIMITATION, PARAMETRIZATION, IDENTIFICATION OF DETERMINANTS AND VECTORS OF EURASIAN DYNAMICS........ 37

Alexander N. Pilyasov
ARCTIC ENTREPRENEURSHIP DEVELOPMENT FACTORS........ 46

P.Ya. Baklanov and M.T. Romanov
TRANSBOUNDARY REGIONS OF EAST ASIA: GEOGRAPHICAL AND POLITICAL PRECONDITIONS AND LIMITATIONS TO LONG-TERM DEVELOPMENT........ 57

Maria V. Zotova and Anton A. Gritsenko
NEIGHBOURHOOD AND PERCEPTIONS IN SMALL CITIES ON DIFFERENT RUSSIAN BORDERS........ 64

Alexander B. Sebentsov
CROSS-BORDER COOPERATION ON THE EU-RUSSIAN BORDERS: RESULTS OF THE PROGRAM APPROACH........ 74

Vladimir A. Kolosov and Maria V. Zotova
MULTIPLE BORDERS OF NAGORNO-KARABAKH........ 84

SPECIAL ISSUE

«SPATIOTEMPORAL ANALYSIS IN MEDICAL AND VETERINARY GEOGRAPHY»

Guest editors:

- •Prof. Paula Santana, Department of Geography and Tourism, Faculty of Arts and Humanities of the University of Coimbra;
- •Prof. Svetlana Malkhazova, Faculty of Geography, Lomonosov Moscow State University;
- •Prof. Vladimir Tikunov, Faculty of Geography, Lomonosov Moscow State University;
- •Dr. Fedor Korennoy, Federal Center for Animal Health.

Elena A. Grigorieva and Anna B. Sukhoveeva

QUALITY OF LIFE, REPRODUCTIVE HEALTH AND SOCIAL SECURITY: MEDICAL AND SOCIAL ENVIRONMENT AT THE RUSSIAN FAR EAST 92

Yelena A. Ogneva-Himmelberger and Madeline Haynes

SPATIAL PATTERNS OF ADVERSE BIRTH OUTCOMES AMONG BLACK AND WHITE WOMEN IN MASSACHUSETTS – THE ROLE OF POPULATION-LEVEL AND INDIVIDUAL-LEVEL FACTORS 99

Jaime Bosch, Irene Iglesias, Marta Martínez, Ana de la Torre

CLIMATIC AND TOPOGRAPHIC TOLERANCE LIMITS OF WILD BOAR IN EURASIA: IMPLICATIONS FOR THEIR EXPANSION 107

Svetlana M. Malkhazova, Polina V. Pestina, Vladimir S. Tikunov

EMERGING AND RE-EMERGING NATURAL FOCAL DISEASES OF EUROPEAN RUSSIA (TYPOLOGICAL CLASSIFICATION OF NOSOLOGICAL PROFILES AND DYNAMICS OF INCIDENCE) 115

Ricardo Almendra and Paula Santana

MORTALITY FROM CARDIOVASCULAR DISEASES IN THE MUNICIPALITIES OF MAINLAND PORTUGAL: SPATIOTEMPORAL EVOLUTION BETWEEN 1991 AND 2017 128

REGULAR ISSUE

Dmitry V. Zhitin, Stanislav S. Lachininskii, Anna A. Mikhaylova, Alexander V. Shendrik

URBAN TRANSFORMATION OF A POST-SOVIET COASTAL CITY: THE CASE OF SAINT PETERSBURG 145

Diana Dushkova and Maria Ignatieva

NEW TRENDS IN URBAN ENVIRONMENTAL HEALTH RESEARCH: FROM GEOGRAPHY OF DISEASES TO THERAPEUTIC LANDSCAPES AND HEALING GARDENS 159

Yurij Vasil'chuk, Julia Chizhova, Natalia Frolova, Nadine Budantseva, Maria Kireeva, Alexander Oleynikov, Igor Tokarev, Ekaterina Rets, Alla Vasil'chuk

A VARIATION OF STABLE ISOTOPE COMPOSITION OF SNOW WITH ALTITUDE ON THE ELBRUS MOUNTAIN, CENTRAL CAUCASUS 172

Alexey Yu. Sidorchuk and Tatiana A. Matveeva

PERIGLACIAL GULLY EROSION ON THE EAST EUROPEAN PLAIN AND ITS RECENT ANALOG AT THE YAMAL PENINSULA 183

Vyacheslav V. Dolotov, Yuri N. Goryachkin, Andrey V. Dolotov
ANALYSIS OF THE CURRENT STATE OF UNSTABLE GEOMORPHOLOGICAL STRUCTURES WITH MODERN METHODS........195

Vladislav V. Sysuev
GEOPHYSICAL ANALYSIS OF LANDSCAPE POLYSTRUCTURES........200

Natalia N. Matinian, Anastasia L. Gusareva, Kseniia A. Bakhmatova, Anastasia A. Sheshukova
MICROBIOLOGICAL INDICATORS AND HEAVY METAL CONCENTRATION IN ECOLOGICAL ASSESSMENT OF URBAN SOILS OF SAINT PETERSBURG, RUSSIA........214

Nataliia V. Terekhina and Margarita D. Ufimtseva
LEAVES OF TREES AND SHRUBS AS BIOINDICATORS OF AIR POLLUTION BY PARTICULATE MATTER IN SAINT PETERSBURG........224

Olga B. Popovicheva, Elena Volpert, Nikolay M. Sitnikov, Marina A. Chichaeva, Sara Padoan
BLACK CARBON IN SPRING AEROSOLS OF MOSCOW URBAN BACKGROUND........233

Tatiana Gladenkova
BEAUTY AND PERSONAL CARE TRANSNATIONALIZATION: MAIN CHANGES IN ITS SPATIAL STRUCTURE........244

Oxana A. Klimanova and Olga I. Illarionova
GREEN INFRASTRUCTURE INDICATORS FOR URBAN PLANNING: APPLYING THE INTEGRATED APPROACH FOR RUSSIAN LARGEST CITIES........251

Olga A. Samonova and Elena N. Aseyeva
PARTICLE SIZE PARTITIONING OF METALS IN HUMUS HORIZONS OF TWO SMALL EROSIONAL LANDFORMS IN THE MIDDLE PROTVA BASIN – A COMPARATIVE STUDY........260

Göran Roos[1], Ksenia Y. Voloshenko[2], Tatiana E. Drok[2], Yury M. Zverev[2*]
[1] Australian Industrial Transformation Institute, Flinders University, Adelaide 5042, Australia.
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia.
***Corresponding author:** yzverev@kantiana.ru

EUROPEAN COUNTRIES' TYPOLOGY BY THE INTENSITY OF TRANSBOUNDARY COOPERATION AND ITS IMPACT ON THE ECONOMIC COMPLEXITY LEVEL

ABSTRACT. Over recent years, it has become increasingly obvious that the countries, regions and individual systems are now developing within the framework of the emerging technological paradigm. The key elements for their development are knowledge and capabilities, being transformed into the products exported by a given country, these constitute the core of the economic complexity theory. In this article, the authors attempt to assess the long-term correlations between economic complexity and transboundary intensity drawing on the example of European countries. The authors developed a European Countries' Typology according to their transboundary cooperation intensity. The paper establishes that the influence of the transboundary factor weakens as the economic complexity increases, and under certain conditions, it has a negative impact. It substantiates that the revealed relationships are due to the increasing role of global processes rather than transboundary ones as the economy becomes more complex and oriented towards the global market.

KEY WORDS: typology, European countries, transboundary cooperation, economic complexity, heterogeneous panel analysis

CITATION: Göran Roos, Ksenia Y. Voloshenko, Tatiana E. Drok, Yury M. Zverev (2020) European Countries' Typology By The Intensity Of Transboundary Cooperation And Its Impact On The Economic Complexity Level. Geography, Environment, Sustainability, Vol.13, No 1, p. 6-15
DOI-10.24057/2071-9388-2019-66

INTRODUCTION

The European countries have created an open market providing free movement of goods, services, capital and ensuring the absence of confrontation between neighbouring countries. The Euroregions practice has proven to be effective in the field of economic development and production of joint solutions to emerging economic problems. Transboundary cooperation in the EU was aimed at supporting remote and peripheral border areas. National prosperity is directly related to economic complexity level, which is determined by the diversity and complexity of the country's export portfolio, which depend on the level of its technological development.

Thus, future national prosperity is a function of current and future economic complexity: the higher the level of economic complexity, the higher the potential for the creation of prosperity in a nation. In order for high economic complexity to develop into high national prosperity, the following conditions must be met (Hausmann et al., 2011; Roos 2017):

– a broad portfolio of highly-exported products and services;

– a high level of uniqueness for all these products, that is, few other countries (preferably none) can produce and export these products and services;

– the high complexity of products, meaning that their production requires a set of capabilities, a significant proportion of which in terms of value added exists within the national economy, since they are unique or have an extremely high cost, thus national suppliers are more preferred to international ones.

The development and production of advanced products take place in the context of interdependence, which requires cooperation between the various participants. In this regard, the question arises as to the extent to which the economic complexity influences not only internal but also external interactions of the participants, which is most clearly manifested at the level of transboundary cooperation. However, in recent years, some European countries have demonstrated a slight decrease in the intensity of transboundary interactions and fluctuations in economic complexity indicators. Therefore, there is a need for an empirical test of the hypothesis about the possible impact of transboundary interactions (taking into account the countries' typology by its intensity) on economic complexity. The paper does not consider Eastern Europe as defined by German Standing Committee on Geographical Names (StAGN).

The study aims to assess the impact of the Proposed Transboundary Intensity Index (TII), i.e. the neighbouring countries' share in the foreign trade, on the Economic Complexity Index (ECI) drawing on the example of European countries. The authors of the work are trying to answer the following questions:

1) What are the criteria for measuring the Transboundary Cooperation Intensity?

2) What methods and models should be used for cluster analysis and identification of typological groups of European countries?

3) Is there a relationship between economic complexity level and transboundary cooperation intensity level?

Following the logic of the study, the article considers the basic concepts of the economic complexity theory and transboundary cooperation, it provides the European

Countries' Typology based on clustering using Gaussian mixture distributions by the transboundary cooperation intensity criterion, it features the analysis of long-term relationship between economic complexity and transboundary cooperation intensity using the panel data models.

LITERATURE REVIEW

The economic complexity theory was formulated and developed by Hausmann & Hidalgo (Hidalgo et al. 2009; Hidalgo et al. 2007; Hausmann et al. 2013; Hidalgo et al. 2016; Hidalgo 2018; Hidalgo et al. 2018). Economic complexity is a measure of the knowledge in a society that gets translated into the products it makes (Hidalgo et al. 2016). According to the Center for International Development at Harvard University (Hausmann et al. 2011), «the economic complexity of a country is calculated based on the diversity of exports a country produces and their ubiquity, or the number of the countries able to produce them (and those countries' complexity). Countries that are able to sustain a diverse range of productive know-how, including sophisticated, unique know-how, are found to be able to produce a wide diversity of goods, including complex products that few other countries can make».

The economic complexity is now being studied in the context of various social and economic issues, including to what extent it explains and influences the economic growth and prosperity of the countries (Hidalgo et al., 2007; Hausmann et al. 2011; Zhu et al. 2017; Stojkoski et al. 2017), income inequality (Felipe 2012; Fortunato et al. 2014; Hartmann et al. 2017), labour market (Roos 2017), structural shifts (Roos et al. 2014). There has been a link identified between it and productivity (Sweet & Eterovic 2019), foreign direct investment (FDI) (Javorcik et al. 2017), it has been discovered that more populated cities export proportionately more skill-intensive and complex goods than less populated citiesz-Lanchas, Llano, (Minondo et al. 2018). Recent years saw also works connecting it to environmental issues (see, for example, Neagu et al. 2019). The authors examine the long-term relationship between economic complexity, energy consumption structure, and greenhouse gas emission. Recent economic complexity studies concentrate on the principle of relatedness in the economic diversification, the dominance of technological innovation and optimization of knowledge diffusion (Boschma 2017; Hidalgo 2018; Hidalgo et al. 2018; Alshamsi, Pinheiro et al. 2018; Utkovski et al. 2018; Hartmann et al. 2019; Roos 2019). At the same time, it is necessary to recognize the technical issues in measuring the economic complexity level. Some of them are associated with the use of the product space method (Radosevic 2017). Evaluation of the modern practice of application and development of the economic complexity today reveals a disadvantage of limited attention to its economic and geographical context. In particular, economic complexity does not consider the transboundary cooperation impact. In broad terms, it refers to the neighbouring countries' interaction (primarily economic one) at different territorial levels. Given the close cooperation between the countries, the processes of regionalization and globalization, this factor can be crucial for economic complexity level. The study uses the concept of «transboundary intensity», meaning the country's active support for transboundary cooperation, including trade relations, which can directly or indirectly affect the level of economic complexity.

There are three basic types of transboundary interactions identified in the literature (Practical Guide to Cross-border Cooperation, 2000: 40; Scott 2017):

– Cross-border cooperation between neighbouring authorities is intended to develop cross-border economic and social centres through joint strategies for sustainable territorial development. The Euroregion is an example of a spatial form of international cooperation created within its framework (Scott 2000; Perkmann 2003). Interreg A program coordinates this cooperation of the EU territorial units at the NUTS III level is coordinated.

– Transnational cooperation is a cooperation between national, regional and local authorities aiming to promote a higher degree of territorial integration across large groupings of European regions, with a view to achieving sustainable, harmonious and balanced development in the Community and better territorial integration with a candidate and other neighbouring countries. This cooperation leads to the formation of large territories that include several regions of different EU member states (for example, Alpine Space, Danube, North Sea, etc.). In the EU, cooperation at this level is coordinated through the Interreg B program.

– Interregional cooperation is intended to improve the effectiveness of policies and instruments for regional development and cohesion through networking, particularly for regions whose development is lagging behind and those undergoing conversion. This involves projects that cover all EU member states. They are coordinated through the Interreg C program.

The first type is primarily the interaction of local authorities on both sides of the border, the second – the interaction mainly between regions of different countries, and the third – transboundary interaction at the country level. The latter, that is the economic interaction at the level of neighbouring countries, seems to primarily influence the economic complexity and, in turn, is influenced by it. However, for the purposes of this work this level of interaction should be considered not as interregional, but as transnational since it is a question of the interaction of neighbouring nation-states across state borders. In this sense, the authors share the position of V.S. Korneevets (Korneevets 2010: 19), who suggested to give the spatial combinations formed in the course of cooperation between states the name of transnational regions. However, transnational interactions between neighbouring states are transboundary ones and are largely determined by cross-border interactions at lower hierarchical levels (in-country regions and local authorities' levels).

The need to take into account the influence the transboundary factor has on the economic complexity level, in our opinion, mainly arises from the specific nature of the formation and development of cross-border ties (Korneevets 2010). Firstly, it is the ongoing increase in the interactions between border areas of different countries and their contact function (connected with the transit potential and the combination of various competitive factors characteristic of regions of different countries, and their specific resources) that inevitably contributes to the development of transboundary relations. This suggests that the countries with a higher intensity of transboundary cooperation have higher diversification of exports. Secondly, the border regions of neighbouring countries can often act as competitors, since similar resources and development conditions determine the production of uniform goods and services to be sold on the external market. This, on the contrary, can adversely affect the level of economic complexity. Accordingly, for certain types of countries, there may be no interdependence between transboundary cooperation and economic complexity. Thirdly, it is the current geopolitical turbulence (Druzhinin et al. 2017) and the change in the global geopolitical situation, observed over the past 20 years. It is reflected in the change in the intensity of transboundary relations and, in turn, in different periods it can have different effects on the economic complexity.

METHODOLOGY AND DATA

The research identifies and examines the effect of transboundary cooperation on the economic complexity on the example of European countries in 1997–2017. The considered countries are those having maritime and land borders. The study does not cover the microstates (Andorra, the Vatican, Liechtenstein, Monaco, San Marino) and the island states (Great Britain, Ireland, Iceland, Malta). The preliminary sample included 32 countries: Austria, Albania, Andorra, Belgium, Bosnia and Herzegovina, Bulgaria, Czech Republic, Croatia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxembourg, Montenegro, the Netherlands, Norway, Poland, Portugal, Romania , Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, TFYR of Macedonia (Republic of North Macedonia since February 12, 2019).

Searching for the possible metrics for transboundary cooperation intensity, the authors identified the following criteria: 1) the number of border crossings; 2) the number of checkpoints per kilometre of the border; 3) the share of international trade with neighbouring countries in the total international trade turnover; 4) the country's participation in the free trade zone (customs union) with neighbouring countries; 5) visa or visa-free regime with neighbouring countries; 6) the number and age of Euroregions on the common border. However, due to the lack of available statistical data or their recording by the above indicators, as well as the data for their calculation, there was only one criterion chosen, that is the neighbouring countries' share in the foreign trade. This indicator is widely known as the coefficient of cross-border specialization of foreign trade. Empirical and theoretical justifications for its use in the analysis of transboundary (cross-border) links are given in numerous papers (see, for example, Mezhevich & Zhuk 2013). Recognizing the scientific problem of determining the intensity of transboundary relations (Korneevets 2010), this paper makes an assumption of foreign trade specialization sufficiency for the characteristics of the interaction between neighbouring countries. The coefficient of cross-border specialization of foreign trade was calculated for each country as the ratio of foreign trade with adjacent territories to the country's total foreign trade turnover. In this study it is referred to as the Transboundary Intensity Index (TII). As a source of information, the authors used the UN Comtrade Database.

The dynamics of the economic complexity of selected countries was analysed using the – Economic Complexity Index (ECI). It is calculated annually by the Center for International Development at Harvard University. ECI is a rank of countries based on how diversified and complex their export portfolio is. The raw trade data on goods are derived from countries' reporting to the United Nations Statistical Division (COMTRADE). A description of the ECI calculation methodology is available in the works of Hausmann & Hidalgo (Hidalgo et al. 2007; Hidalgo & Hausmann 2009; Hausmann et al. 2011). The annual data are available at two official Internet resources: The Atlas of Economic Complexity and The Observatory of Economic Complexity.

Based on the results of the collection and analysis of available statistical data, 2 countries were excluded from the sample. The first one is Montenegro, since there is no ECI data for it on the above mentioned resources, and the paper does not include the index calculation. Second is Luxembourg , as the ECI is calculated for Luxembourg and Belgium together. The final sample consisted of 29 countries.

Identification and evaluation of the relationship between transboundary intensity and economic complexity were carried out in the following sequence. All calculations were performed using StatSoft Statistica v10.0, analysis of panel data – EViews v9.0.

Step 1. At the initial stage, the TII and ECI values for each country for were analysed for each year in the interval 1997–2017. It has been established that the relationship between the variables is weak (R2 on average varied within 0.2–0.3). Further analysis included the study of the variable dynamics for each country in the studied time period. The results obtained varied in significance and closeness of the relationships between variables. This led to the conclusion that there is a need for separate assessment for particular typological groups.

Step 2. The experimental clustering of countries was carried out according to different criteria: R2 values, correlation coefficient, TII and ECI values (separately and jointly). Calculations were carried out both with and without standardization of variable values. The best results on the statistics of typological groups were obtained by TII. The countries were later classified into clusters using the EM-algorithm basing on Gaussian mixture distributions. There were 3 clusters of countries (subpanels) identified, including 2 subpanels with a positive linear relationship: i) countries with low ECI and TII; ii) countries with medium ECI and TII and 1 subpanel with a negative linear relationship (or its absence); iii) countries with a high ECI. The transition from the first to the third group (cluster) involves the increase in the value of ECI and TII.

Step 3. Panel data analysis was carried out. There are several reasons for using panel data analysis in the study. Firstly, the results of the regression analysis conducted at the previous stage produced unsatisfactory results. Secondly, the study established a dataset with temporal and spatial dimensions. Thirdly, panel data analysis allows controlling heterogeneity and serial correlation (see, for example, Wooldridge 2007). Moreover, it is known that due to their specific structure panel data allow building flexible and informative models. They also provide the opportunity to take into account and analyse individual differences between economic units, which cannot be done within standard regression models (Baltagi 2005). As a result, the empirical analysis and the obtained panel models for typological groups made it possible to investigate the long-term correlation between the changes in economic complexity and changes in transboundary intensity.

We will provide a brief description of panel data analysis.

An array of variable values for the countries in 1997–2017 was used to form a balanced panel. The data exist for all countries for all periods of time. The panel data analysis was carried out in a known sequence (see, for example, Wooldridge 2007) and included the following main stages:

1) Cross-section Dependence in Panel Data. First of all, the presence of dependencies between the variables inside each cross-section of the data was checked. In the study, this suggests the presence of a correlation between countries in terms of TII and ECI, cross-section dependence was tested by means of Pesaran scaled LM and CD test. The test was carried out using other tests in EViews v9.0.

2) Description of Panel Data Structure. This included panel cointegration conditions testing. The main hypothesis of stationarity has been preliminary examined taking into account possible differences in the rate of convergence and the assumption of cross-national correlations of residuals. EViews provides convenient tools for computing panel unit root tests: Levin, Lin and Chu (2002), Breitung (2000), Im, Pesaran and Shin (2003), Fisher-type tests using ADF and PP tests – Maddala and

Wu (1999), Choi (2001), and Hadri (2000), as well as all three modifications (with constant, trend, constant and trend). Hadri Z-stat was analysed using Newey-West estimator taking into account possible heteroscedasticity. The cointegration between panel variables was tested using the Pedroni test (Pedroni 2004).

3) Evaluation of Panel Data Models. The key stage was the identification of panel analytic models type: the pooled model, the fixed effect model and the random effect model. Additionally, there were fully modified ordinary least squares (FMOLS) and the panel dynamic ordinary least squares (DOLS) models built (Baltagi 2005). The problem of model selection is solved using the standard hypothesis testing technique (Brillet 2011). Additionally, the constructed panel models were monitored. This included their quality assessment (Wooldridge 2007). The authors used the coefficient of determination (R2), the Fisher criterion, the panel analogue of the Durbin-Watson test, the Jarque-Bera test, etc.

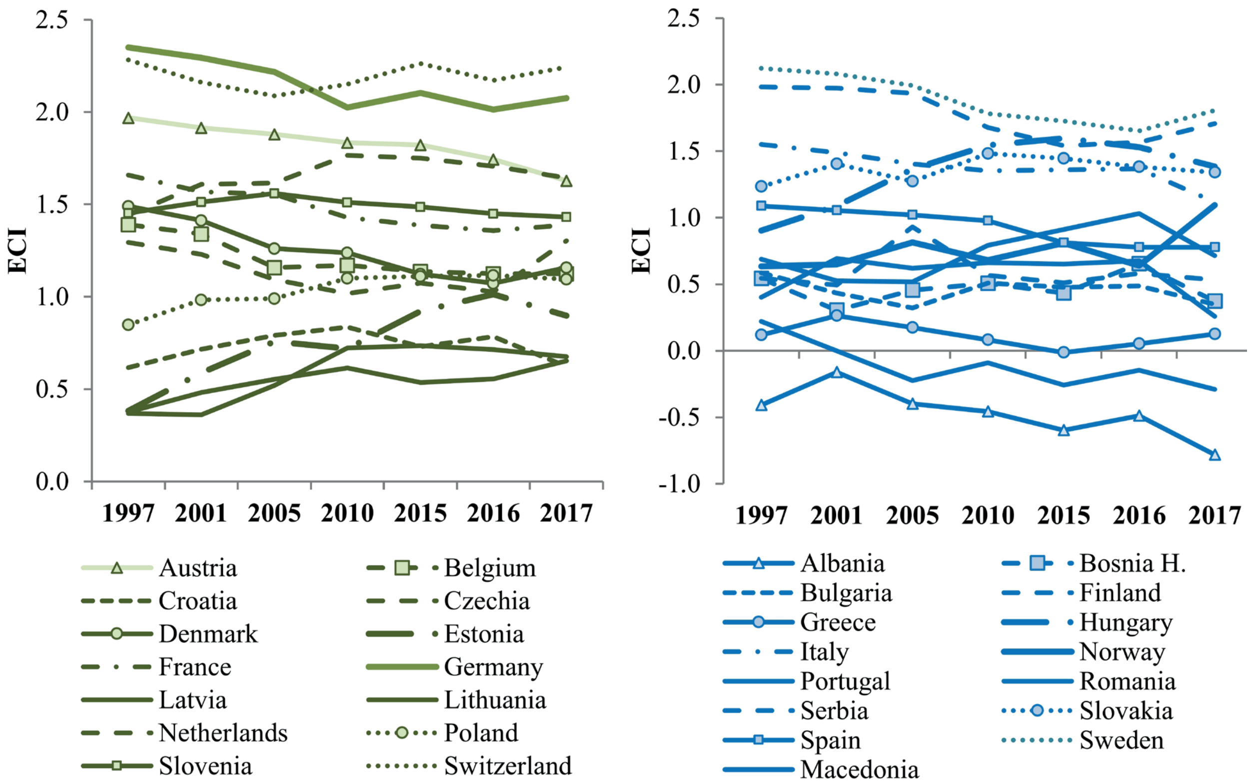
RESULTS AND DISCUSSION

Analysis of changes in the transboundary intensity and economic complexity in European countries in 1997–2017

In recent years, a decrease in the intensity of transboundary interaction processes has been observed in almost all analysed groups of European countries. This manifested itself in the dynamics of cross-border specialization of international trade turnover. For ease of assessment of the transboundary intensity changes, all countries in the sample were divided into two groups: i) countries with a high – TII (more than 30%); (ii) Countries with a low TII – (less than 30%).

All the countries with the exception of Latvia and Croatia demonstrated a decrease in TII in 1997–2017. While the two countries show an increase in TII by 16.5 and 6.7% respectively. It should be noted that in the same period European countries showed similar ECI dynamics (Fig. 1 a and b).

Given that the ECI is measured by the ubiquity and diversity of manufactured and exported products, we can put forward the following explanation: the rapid growth of economic complexity in the studied interval 1997–2017 was observed in those countries that initially had much smaller values for the indicators of both economic complexity, well-being and the pace of economic development (traditionally measured in terms of GDP per capita). These are for example Croatia, Czechia, Estonia, Greece, Hungary, Latvia, Lithuania, Netherlands, Poland, Romania and Slovakia. Therefore, the development of these countries in 1997–2017, associated with changes in their production structure and production capabilities due to the expansion of their knowledge and capability base, led to the development of more technologically sophisticated products, which accordingly affected the growth of ECI. At the same time, the so-called effect of a low (or zero) base was not typical for countries with a higher level of development and ECI. Given that the number of countries producing complex products grew, and the emergence of new technologically complex types of products occurs at a slightly lower rate, countries with initially high ECI in 1997 somewhat reduced their positions and their index values went down. These are the representative examples of the following European countries with the largest total (cumulative) ECI decrease: Italy (-0,434), Austria (-0,341), Denmark (-0,333), Sweden (-0,315), Finland (-0,276), Germany (-0,275), France (-0,269), Belgium (-0,265). For example, Fig. 2 shows a comparison of the Product Space within the framework of Economic Complexity for Estonia and Italy in 1997 and 2016. These countries demonstrated (respectively) the largest increases and decreases in ECI among the European countries. It is indicative that during the study period, Estonia showed the increase in ICT from 5.19 to 13.04%, and in electrical line telephonic and telegraphic apparatus from 2.84 to 7.65%, and the increase in the production structure complexity (Fig. 2.1). At the same time, export structure in Italy has remained relatively stable, with the shares being redistributed among



a) the first group – TII is more than 30%

b) the second group – TII is less than 30%

Fig. 1. European countries' ECI in 1997–2017

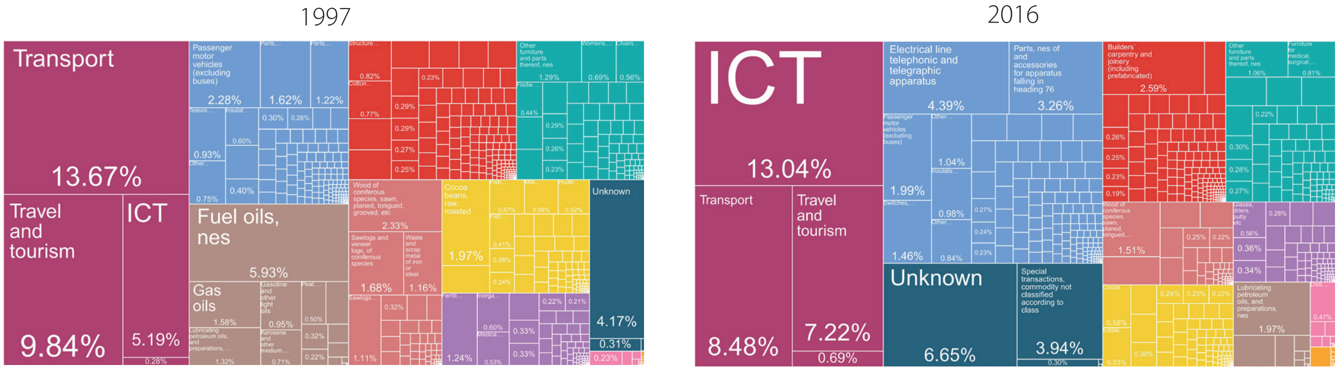
Source: The Atlas of Economic Complexity, The Observatory of Economic Complexity

ICT, Transport, Travel and Tourism within 2–3% (Fig. 2.2), the ratio between complex products types has changed but not fundamentally. Thus, it can be assumed that the main reason for the reduction in ECI in the group of more developed countries was the decrease in the prevalence of complex products accompanied by insufficient pace of economic diversification.

The analysis of the ECI and TII dynamics reveals that not all countries in the sample have a direct correlation between them, the growth of cross-border trade is not always accompanied by the increase in economic complexity and vice versa. This is particularly true in relation to such countries as Czech Republic, Estonia, Lithuania, Netherlands, Poland, Norway, Slovakia, where in 1997—2017 the level of economic complexity increased while the transboundary intensity went down. Bulgaria, Finland, Portugal and Sweden showed an opposite trend – the transboundary intensity was increasing while the economic complexity was decreasing (Fig. 3).

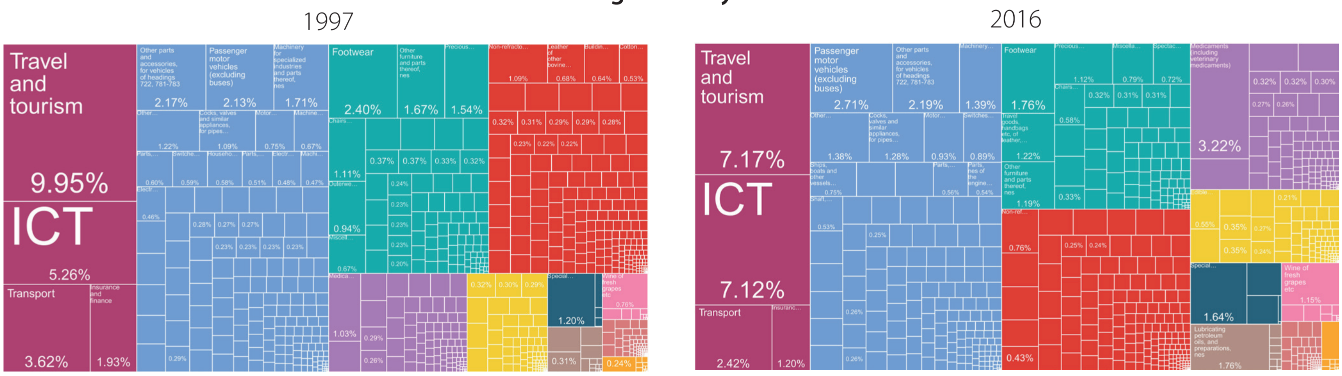
The direct and inverse correlation between the economic complexity and transboundary cooperation can be fixed at the level of different typological groups of countries by the degree of intensity of the latter. In order to assess the dependence of the variables, the countries were classified by the transboundary intensity level, this was followed by a panel data analysis at the level of selected groups of countries (subpanels).

Fig. 2.1. Estonia



Export, %

Fig. 2.2. Italy



Export, %

Fig. 2. Comparing the Product Space within the framework of Economic Complexity

Note: X-axis –distance, Y-axis – complexity

Source: The Atlas of Economic Complexity. www.atlas.cid.harvard.edu

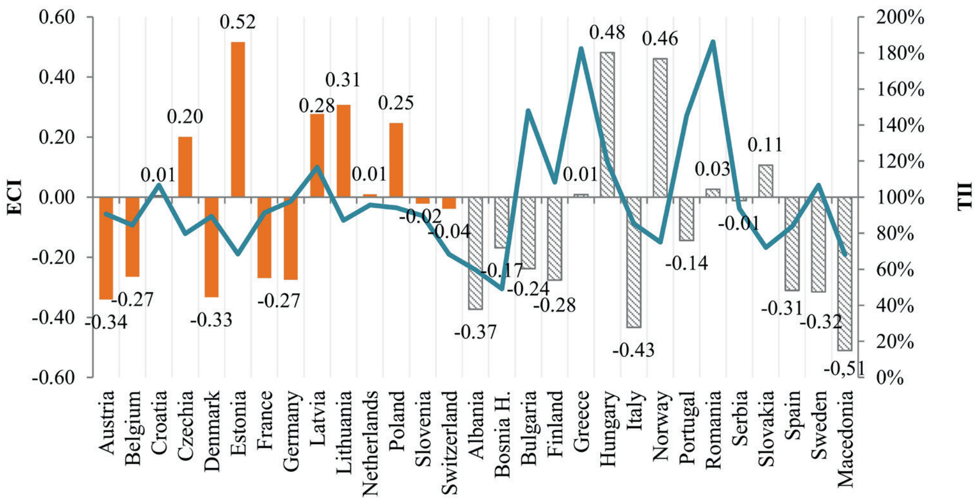


Fig. 3. Comparison of cumulative changes in ECI and TII of European countries in 1997—2017 (1997 – 100%)

Source: calculated and compiled by the authors

Countries' typology by transboundary intensity

To classify the European countries by the level of transboundary intensity, a regression analysis of the data on the initial base for the entire population and each country was carried out separately for the 1997–2017.

According to the results of the preliminary regression analysis for the studied countries, an ambiguous situation has developed. Some countries are characterised by a strong linear relationship, while in the others the relationship between transboundary intensity and economic complexity is either absent or non-linear. For the entire sample, the Pearson correlation coefficient (r) was 0.4386, and the coefficient of determination (R^2) was respectively 0.1924 ($p = 0.000$). Kolmogorov-Smirnova and Shapiro-Wilka test reject the hypothesis about the normal distribution for the entire sample of countries (for TII: $p_{w-s}<0.01$ $p_w=0.00$; for ECI: $p_{w-s}<0.1$ $p_w=0.00$). It was revealed that the linear moderate and strong ECI and TII connections are not found in all the countries. There are considerable differences at the level of correlation significance (p – values). It should be noted that the estimated values of R^2 and Pearson's coefficients have the disadvantages of low resistance to outliers in the original sample, and also impose limitations on linearity and monotony of data changes. In this regard, non-parametric statistics were evaluated using the Spearman coefficient. For the entire sample of countries, the values turned out to be slightly lower than the Pearson coefficient: 0.4136 ($p = 0.000$). Due to unsatisfactory results, further analysis by country is not given.

In this regard, in the next step, using a Gaussian mixture of distributions (GMM), countries were clustered in order to identify subpanels for the subsequent panel data analysis. In recent years, finite Gaussian mixtures in real processes and phenomena modelling has been used in growing number of fields of science and practice: in pattern and speech recognition, biology, medicine, geography, physics and chemistry, neuroinformatics, sociology and economics, etc. (see for example, Kozhevnikova et al. 2012; Seethalakshmi et al. 2014; Compiani et al. 2016). GMM-based sample objects grouping has a number of advantages comparing to other known approaches, for example, k-medians clustering, such as smoothness (infinite differentiability), identifiability, completeness, resolution (Aprausheva et al. 2015).

The obtained country grouping is validated by Gaussian mixture distribution itself. Even with a strong mutual penetration or intersection of classes (types, groups), the algorithm shows a adequate result on the maximum likelihood of the original sample (Aprausheva et al. 2015). In the process of clustering, normally distributed groups within the total sample are analysed (Reynolds 2009).

GMM and data research is performed in this work using the sci-kit-learn package of the Python programming language. The EM-algorithm for sequential addition of components was used to determine the optimal clustering for the subsequent distribution of countries between the optimal number of groups according to the level of transboundary activity. Estimation of the distribution made it possible to establish that the initial sample of countries is best described when divided into three clusters.

The first cluster (type): Albania, Bosnia Herzegovina, Bulgaria, Croatia, Netherlands, Norway, Portugal, Serbia, Slovenia, Switzerland, TFYR of Macedonia. The second cluster (type) group: Austria, Belgium, Denmark, France, Hungary, Italy, Latvia, Lithuania, Romania, Spain. The third cluster (type): Czech Republic, Estonia, Finland, Germany, Greece, Poland, Slovakia, Sweden.

For the first group (subpanel), the smallest values of the variables are characteristic; for the second and third, the level of transboundary intensity is almost the same, Fig.4).

At the same time, the third cluster (subpanel) includes the countries with the negative correlation between economic complexity and transboundary intensity identified using Spearman-Pearson coefficient (7 out of 10 countries have negative ECI and TII correlation).

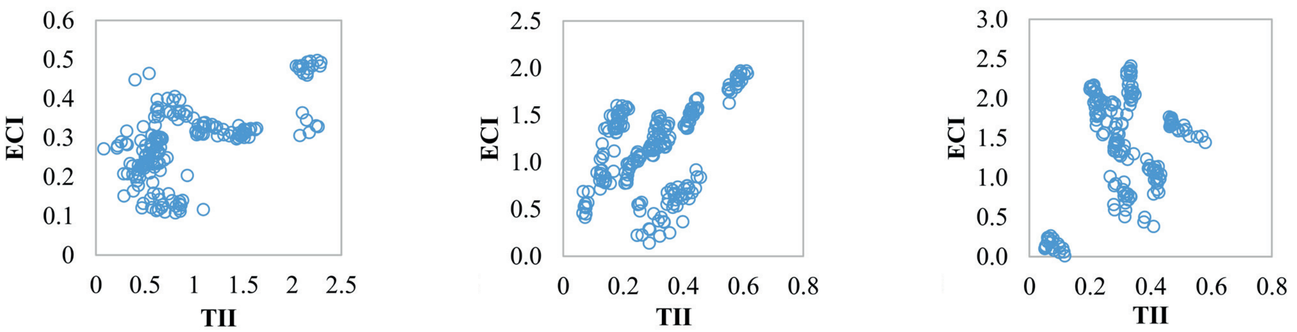
It has been noted that the closeness of the connection between the transboundary intensity and economic complexity decreases with the growth of the latter. The greater is the level of economic complexity, the less the country's transboundary activities affect its change. To test this assumption and assess the direction of the transboundary intensity effect on the economic complexity (direct or inverse correlation), panel data were analysed and econometric models were constructed in accordance with the algorithm described in Section 3.

Panel data analysis: the impact of transboundary intensity on economic complexity by country type

It is known that, in contrast to simple regression models, the panel data analysis makes it possible to take into account individual differences between objects (Verbeek 2003), or between countries in this case. In addition, it allows to reduce the dependence between explanatory variables, and, consequently, standard estimation errors, and to prevent the aggregation displacement, which occurs during the analysis of time series and spatial data (Ratnikova 2006).

The brief results obtained in accordance with the above mentioned approach can be found below.

1. Source data quality evaluation. The main problems that may arise when using panel data, causing offsets and bias in regression, include the following: cross-sector dependence, non-stationarity, heteroscedasticity, lack of



a) First subpanel (r = 0.5987) *ECI* = 0.720, *TII* = 0.277
b) Second subpanel (r = 0.4239) *ECI* = 1.138, *TII* = 0.304
c) Third subpanel (r = 0.3542) *ECI* = 1.358, *TII* = 0.303

Fig. 4. Scatter plots, European countries groups, 1997–2017

Source: author's calculations

cointegration. Therefore, a series of tests were carried out to detect these problems. Thus, according to the cross-sector dependence test, the first panel at 10% significance level can be considered spatially correlated; for the second and third panels it is possible to accept the null hypothesis of the absence of spatial correlation. Basing on the obtained values of test statistics, it can be concluded that the indicators of economic complexity and transboundary intensity are the implementation of a non-stationary random process with the order of integration 1 (I (1)). The identified relationships are not apparent and are generated by unit roots. In this study, it is impossible to accept the hypothesis of the absence of heteroscedasticity for all panels at a significance level of 5%, therefore, the country data are heterogeneous. The tests confirmed the panel cointegration of baseline data by country. This means that, despite the random (poorly predictable) nature of changes in objects, there is a long-term correlation between them, which leads to interlinked changes (Ratnikova 2006; Wooldridge 2007).

The initial data testing leads to the following conclusion: there is a problem of spatial correlation and heteroscedasticity, however, the series is stationary in terms of first difference and there is joint integration indicating the possibility of estimating the long-term correlation between the variables. In long-term variables can be regressed in the same model without taking the difference.

2. Panel models construction. Next, to analyse the long-term correlation between the variables, the authors estimated the parameters of the equations of the pooled model, fixed effect model and random effect model. The analysis was performed using the method of least squares (LS – Least Squares (LS and AR)). The tests indicate the best values for the fixed effect models. This shows that a change in economic complexity has individual time-invariant country-level effects depending on transboundary activities intensity. In general, the results confirm the intuitive idea of the inconsistency of the combined or random models for the description of panel data. The testing for a unit root in the regression residuals confirms the cointegration between the variables under study. In all the subpanels, Durbin-Watson test and Jarque-Bera test (JB-test) for normality as well as the evaluation of the correlogram of random variations reveal the presence of autocorrelation in the models. This can be connected both with a long time range, and the effect of other random factors that are not taken into account in the model. Therefore, the cointegrating coefficients are estimated using the between-dimension fully modified ordinary least squares (FMOLS) and dynamic ordinary least squares (DOLS) techniques as proposed by (Pedroni 2000, 2004). This allows eliminating endogeneity in the regressors and serial correlation in the errors. This method has already been applied to the economic complexity assessment in previous works (see, for example, Neagu et al. 2019; Dogan et al. 2017).

To create FMOLS and DMOLS models, their various modifications were used for the trend specification (constant (level), trend), for the panel method (pooled, pooled (weighted), grouped), lag method (for example, fixed or Schwarz). The next step was the selection of models with the best R2 values, the regression coefficient of at least 5% significance, positive results of stationary residuals testing and Jarque-Bera statistics for normal distribution (for df =2and significance level 5%, not more than 5.991) (Green 2005). As a result, the quality and reliability of DMOLS were higher compared to the FMOLS and OLS models, and these models were used for further analysis.

3. Analysis Results. The following results were obtained on the basis of the constructed DMOLS model parameters.

The first sub-panel of countries is characterized by a direct dependence of economic complexity on transboundary intensity. According to the model specification and equation, a 1% ECI change occurs when transboundary intensity increases by only 0.45%.

For the second subpanel countries, economic complexity changes when transboundary intensity grows by 1.21%. Given that the average ECI value in the second subpanel countries is higher, it can be concluded that with increasing economic complexity, the influence of the transboundary factor becomes weaker.

For the third group, given that it includes countries with a negative correlation between economic complexity and transboundary cooperation, there was a model describing a particular correlation constructed. In the third sub-panel, the growth of economic complexity is associated with a 2.43% decrease in transboundary intensity. For the same reason, the conclusions cannot be extended to all the countries surveyed, and there is a need for a more in-depth analysis. For example, it is necessary to study whether the negative correlation of variables is a consequence of the characteristics of transboundary processes in countries, their economic structure or the product space, or whether the correlation with the transboundary factor becomes negative when a certain level of economic complexity and level of economic development is reached. Of course, finding answers to these questions requires consideration of a much larger number of factors in the model, as well as an increase in the spatial (geographical) dimension of the sample of the countries.

At the same time, it should be recognized that the assessment and analysis of changes in the transboundary intensity and economic complexity of European countries made it possible to establish not only the existence of the connection between them. The calculations confirm the speculative conclusions of the changing role of transboundary relations as not only the complexity of the economy increases, but also the economy in general grows. For example, for emerging economies, the reduction of trade barriers leads to the intensification of transboundary interactions with neighbouring countries (Foster et al. 2011). This becomes a powerful driver and an important source of diversification and economic growth for a country. However, the subsequent growth of production capacities in a country in the context of the globalization inevitably leads to its active integration and inclusion into global production chains. The geography of international relations has significantly changed. The priority has shifted to the integration into the global economy rather than further strengthening and supporting transboundary interactions. For example, the above features of transboundary development can be traced quite well in such European countries as the Czech Republic, Estonia, Lithuania, Poland, etc. Corresponding dependencies can be also observed in the countries' economic complexity levels. In the early stages of the increase of economic complexity, transboundary links are indeed important, both in terms of export capacity increase and active scientific and technological exchange, for the growth of the countries' own production capabilities. This is confirmed by the calculations at the level of European countries of the first type (cluster). They are characterized by high elasticity of changes in economic complexity with increasing transboundary intensity (e.g., Bosnia and Herzegovina, Bulgaria, Croatia, Slovenia, Switzerland, North Macedonia). The gradual increase in the economic complexity reduces the economy's dependence on transboundary interactions, since other factors, in particular, those contributing to the expansion of the knowledge and capability base, have a greater influence on the growth of economic complexity. These processes are characteristic of the second type of European countries, as the impact the transboundary intensity on economic complexity decreases, i.e. a greater increase in transboundary interactions is required to increase economic complexity. The indicators of economic complexity for developed countries and exporters of complex products (machinery, equipment and tools, metals and materials, chemistry and pharmaceuticals) are no longer dependent on the intensity of transboundary interactions. For the countries of the third group including Germany, Sweden, Finland, Estonia, Poland, the Czech Republic, and Greece, the correlation between economic complexity and transboundary interactions intensity becomes inverse. Thus, transboundary interactions, initially being a factor for the growth of economic complexity, are being replaced by the global economic processes.

CONCLUSIONS

The concept of economic complexity uses the idea of interconnections in the global market, presenting international trade data as a two-way network in which countries are connected by exported products, with the diversification of exported products being a major factor. The analysis base is empirical observations showing that the most competitive countries have diversified exports, while developing countries export only a few products, as a rule, already exported by many other countries. According to economic complexity theory, the level of export diversification reflects the level of the country's industrial development.

This study focuses on the economic and geographical aspect of economic complexity, which is expressed in its dependence on transboundary intensity. The prerequisite for this hypothesis is a number of empirical and theoretical studies (A more united and stronger central Europe… 2018; Hörnström et al. 2015; Land-based spatial planning… 2015; Socio-economic challenges… 2018; Wassenberg et al. 2015; etc.) confirming the importance of transboundary cooperation development. At the same time, the technological cooperation between countries can have the key role in economic complexity.

To study the relationship between transboundary cooperation and economic complexity, the authors apply the concept of transboundary intensity. It is assumed that the more actively the neighbouring countries are involved in various types of mutual international trade operations, the higher the transboundary intensity is. The search for a transboundary intensity metric included the consideration of various criteria. However, taking into account the data availability for all the countries in the sample for the period of 1997–2017, the authors have selected the coefficient of cross-border specialization of foreign trade as the main indicator, referred to in this paper as a Transboundary Intensity Index (TII). The Economic Complexity Index (ECI) was used to measure economic complexity. The data sources used were the UN Comtrade Database and The Atlas of Economic Complexity.

The study used various approaches to distinguish between the types of countries. The TII basis provided the best statistical results for the typological group. The countries were classified into clusters using Gaussian mixture distributions. The method was chosen due to such advantages as smoothness (infinite differentiability), identifiability, completeness, resolution. In addition, the method is the best one for the identification of clusters for the purposes of stabilization of the maximum likelihood of the initial sample. The method was implemented and the study of the data was carried out using the scikit-learn, a machine learning package in the Python programming language. There were 3 clusters identified using the criterion of transboundary intensity. The results of a preliminary assessment of the types of countries (subpanels) showed the following dependence: the higher the economic complexity, the less its change is influenced by the country's transboundary intensity, and the correlation between variables becomes inverse. This hypothesis was tested by the analysis of panel data and the construction of models.

In general, the construction and evaluation of panel models verify the authors' hypothesis about the dependence (and its change) between the transboundary intensity and economic complexity. This is confirmed by the differences in the results obtained at the level of identified country types. For example, the increase in the initially low transboundary intensity is associated with significant growth in the economic complexity level. This is typical for the first type of countries. As the transboundary intensity increases, its impact on economic complexity weakens. This situation is typical for the second type of countries. While a further increase in economic complexity has an inverse connection with transboundary intensity. These are the results for the third type of countries.

The results of the study suggest that as the economy becomes more complex, its dependence on transboundary interaction decreases, since other factors, in particular, those contributing to expanding their knowledge and capability base, have a greater influence on the growth of economic complexity. Thus, the countries with low and high transboundary intensity are characterized by different connections with and effects on economic complexity. In the first case, the strengthening of transboundary relations is a driver for the growth of economic complexity. On the contrary, for the second type, the orientation to cross-border markets will significantly hamper the potential for growth of the number of complex products in the export basket. In terms of their role and importance, international relations and globalization processes begin to prevail over transboundary ties and regionalization processes, despite the fact that in earlier studies special importance was attached to the contractual functions of the territories (Korneevets 2010).

However, it should be noted that the study and the above approach has certain limitations. These are the insufficient number of indicators used to assess the transboundary intensity and the need to verify the findings using a larger number of geographic areas of various types of countries. In addition, the cluster analysis and identification of typological groups should be expanded to include other indicators characterizing the level of socio-economic development of countries (GDP, investment, employment, income, etc.). All these issues are certainly of methodological and practical interest. However, the main objective of this study was to identify the fact and characteristics of the connections and the interdependence of transboundary intensity and economic complexity.

The results of the study are significant for understanding the relationship between transboundary links and economic complexity. However, they make an even greater contribution to the development of new approaches and to the study of factors and conditions, the potential and constraints of the development of territories that differ in transboundary relations closeness.

Notes

1. Atlas learning resources. The Atlas of Economic Complexity. [online]. Available at: www.atlas.cid.harvard.edu/learn/glossary
2. Summaries of the Operational Programmes adopted by the European Commission and financed under the European Regional Development Fund or the Cohesion Fund. [online]. Available at: www.ec.europa.eu/regional_policy/index.cfm/en/atlas/programmes?search=1&keywords=&periodId=2&countryCode=ALL&regionId=ALL&objectiveId=11&themeId=ALL
3. UN Comtrade Database. [online]. Available at: www.comtrade.un.org
4. The Atlas of Economic Complexity. [online]. Available at: www.atlas.cid.harvard.edu
5. The Observatory of Economic Complexity. [online]. Available at: www.atlas.media.mit.edu
6. EViews 9.5 New Econometrics and Statistics: Testing and Diagnostics. [online]. Available at: https://www.eviews.com/EViews9/ev9ecdiag_n.html
7. EViews – User's Guide: Panel Unit Root Testing. [online]. Available at: www.eviews.com/help/helpintro.html#page/content/advtimeser-Panel_Unit_Root_Testing.html#ww185476
8. В скобках указано абсолютное кумулятивное изменение ECI за 1997–2017гг.
9. StatSoft. Texbook. [online]. Available at: www.statsoft.com/Textbook/Basic-Statistics#Correlationsb

ACKNOWLEDGEMENTS

The reported study was funded by RFBR and the government of the Kaliningrad region according to the research project № 19-410-390002 «Economic Complexity and Sector-Specific Strategy Selection by the regions of Russia in the context of the emerging value creation paradigm, using the case of the Kaliningrad region».

REFERENCES

A more united and stronger central Europe needs transnational cooperation (2018). Input paper of the Interreg CENTRAL EUROPE Programme. [online]. Available at: www.interreg-central.eu/Content.Node/discover/Input-Paper.pdf [Accessed 11 May 2019].

Alshamsi A., Pinheiro F.L. & Hidalgo C.A. (2018). Optimal diversification strategies in the networks of related products and of related research areas. Nature Communications, 9 (1). DOI: 10.1038/s41467-018-03740-9.

Aprausheva N.N. & Sorokin S.V. (2015). (Notes on gauss mixtures), VC RAN (in Russian.). DOI: 10.13140/RG.2.2.33609.34404.

Baltagi B.H. (2005). *Econometric Analysis of Panel Data*. John Wiley & Sons.

Boschma R. (2017). Relatedness as driver of regional diversification: A research agenda. Regional Studies, 51 (3), 351–364. DOI: 10.1080/00343404.2016.1254767.

Brillet J.L. (2011). Structural Econometric Modelling: Methodology and Tools with Applications under EViews 2016. [online]. Available at: www.eviews.com/StructModel/structmodel.pdf [Accessed 14 Apr. 2019].

Compiani G. & Kitamura Y. (2016). Using mixtures in econometric models: a brief review and some new results. *The Econometrics Journal*, 19, 95–127. DOI: 10.1111/ectj.12068.

Díaz-Lanchas J., Llano C., Minondo A. & Requena F. (2018). Cities export specialization. *Applied Economics Letters, Taylor & Francis Journals*, 25 (1), 38–42. DOI: 10.1080/13504851.2017.1290784.

Dogan, E. & Aslan A. (2017). Exploring the relationship among CO2 emissions, real GDP, energy consumption and tourism in the EU and candidate countries: Evidence from panel models robust to heterogeneity and cross-sectional dependence. *Renewable and Sustainable Energy Reviews*, 77, 239–245. DOI: 10.1016/j.rser.2017.03.111.

Druzhinin A.G., Gorochnya V.V., Gontar N.V., et al. (2017). Transboundary Clusters in the Coastal Cones of the European Part of Russia: Inventory, Typology, Factors, and Prospects, *Baltic Region*, 9 (4), 29–44. DOI: 10.5922/2074-9848-2017-4-2.

Felipe J. (2012). Tracking the Middle-Income Trap: What is It, Who is in It, and Why? Part 2, No 307, *ADB Economics Working Paper Series*, Asian Development Bank.

Fortunato P. & Razo C. (2014). *Export sophistication, growth and the middle-income trap*. In Salazar-Xirinachs J. M., Nübler I. and KozulWright R. (eds.), Transforming Economies: Making Industrial Policy Work for Growth, Jobs and Development, International Labour Office, Geneva: ILO.

Foster N., Hunya G., Pindyuk O., Richter S. (2011). *Revival of the Visegrad Countries' Mutual Trade after their EU Accession: a Search for Explanation*. Wiener Institut für Internationale Wirtschaftsvergleiche (wiiw). Research Reports 372. [online]. Available at: www.wiiw.ac.at/revival-of-the-visegrad-countries-mutual-trade-after-their-eu-accession-a-search-for-explanation-dlp-2449.pdf [Accessed 11 May 2019].

Green W.H. (2005). Econometric Analysis. Pearson Education International.

Hartmann D., Bezerra M., Lodolo B. & Pinheiro F.L. (2019). *International trade, development traps, and the core-periphery structure of income inequality*. [online]. Available at: SSRN: www.ssrn.com/abstract=3312097 or DOI: 10.2139/ssrn.3312097.

Hartmann D., Guevara M.R., Jara-Figueroa C., Aristara´N, M. & Hidalgo C.A. (2017). Linking Economic Complexity, Institutions, and Income Inequality. *World Development*, 93, 75–93. DOI: 10.1016/j.worlddev.2016.12.020.

Hausmann R., Hidalgo C.A., Bustos S., Coscia M., Simoes A. & Yildirim M.A. (2011). *The Atlas of Economic Complexity:* Mapping Paths to Prosperity. MIT Press.

Hidalgo C.A. (2018). From useless to keystone. *Nature Physics*, 14, 9–10. DOI: 10.1038/nphys4337.

Hidalgo C.A. et al. (2018). *The Principle of Relatedness*. In: Morales A., Gershenson C., Braha D., Minai A., Bar-Yam Y. (eds) Unifying Themes in Complex Systems IX. ICCS 2018. Springer Proceedings in Complexity. Springer, Cham. DOI: 10.1007/978-3-319-96661-8_46.

Hidalgo C.A. & Hartmann D. (2016). *Economic complexity, institutions, and income inequality*, OECD Insights. [online]. Available at: www.oecd-ilibrary.org/economics/debate-the-issues-complexity-and-policy-making_9789264271531-e [Accessed 31 Mar. 2019].

Hidalgo C.A. & Hausmann R. (2009). The building blocks of economic complexity. *PNAS*, 106 (26), 10570–10575. DOI: 10.1073/pnas.0900943106.

Hidalgo C.A., Klinger B., Barabási A.-L. & Hausmann R. (2007). The Product Space Conditions the Development of Nations. *Science*, 317 (5837), 482–487. DOI: 10.1126/science.1144581.

Hörnström L., Diş A.T., Berlina A. (2015). Added value of cross-border co-operation. Cross-border co-operation: still an undeveloped potential? *Nordregio News*, 1, 3–7.

Javorcik B., Lo Turco A. & Maggioni D. (2017). New and improved: Does FDI boost production complexity in host countries? *Economic Journal*, 128 (614), 2507–2537. DOI: 10.1111/ecoj.12530.

Korneevets V.S. (2010). *(International regionalization in the Baltic Sea*), Saint Petersburg (in Russian.).

Kozhevnikova I., Shveikina V. & Domínguez E. (2012). Modelling Caspian Sea fluctuations. *J. Flood Risk Manage*, 5, 3–13. DOI: 10.1111/j.1753-318X.2011.01116.x.

Land-based spatial planning and the added value of cross-border cooperation (2015). Background paper. Baltic SCOPE. 42 p. [online]. Available at: www.balticscope.eu/content/uploads/2016/11/Land-based-spatial-planning-and-the-added-value-of-cross-border-cooperation.Final_.Nov09.pdf [Accessed 11 May 2019].

Mezhevich N.M., Zhuk N.P. (2013). Cross-border specialization of interregional interaction: applying new assessment methods. *Baltic Region*, 1 (15), 38-52. DOI: 10.5922/2074-9848-2013-1-3.

Neagu O. & Teodoru M.C. (2019). The Relationship between Economic Complexity, Energy Consumption Structure and Greenhouse Gas Emission: Heterogeneous Panel Evidence from the EU Countries. *Sustainability*, 11 (2), 497. DOI: 10.3390/su11020497.

Neagu O. & Teodoru, M.C., (2019). The Relationship between Economic Complexity, Energy Consumption Structure and Greenhouse Gas Emission: Heterogeneous Panel Evidence from the EU Countries. Sustainability, *MDPI, Open Access Journal,* 11 (2), 1–29. DOI: 10.3390/su11020497.

Pedroni P. (2000). Fully Modified OLS for heterogeneous cointegrated panels. In: Baltagi B.H., Nonstationary panels, panel cointegration and dynamic panels, Amsterdam: Elsevier, 93–130.

Pedroni P. (2004). Panel cointegration: asymptotic and finite sample properties of pooled time series with an application to the PPP hypothesis. *Economic Theory*, 20 (3), 597–625. DOI: 10.1017/S0266466604203073.

Perkmann, M. (2003). *The rise of the Euroregion. A bird's eye perspective on European cross-border co-operation*. Lancaster University. Department of Sociology. 19 p. [online]. Available at: www.pdfs.semanticscholar.org/f420/d587e6e9801a95e450634970f33005ec1479.pdf [Accessed 11 May 2019].

Practical Guide to Cross-border Cooperation. Third Edition (2000). Association of European Border Regions (AEBR). European Commission. [online]. Available at: www.aebr.eu/files/publications/lace_guide.en.pdf [Accessed 11 May 2019].

Radosevic S. (2017). *Assessing EU Smart Specialization Policy in a Comparative Perspective*. In Advances in the Theory and Practice of Smart Specialization, Academic Press, 1–36. DOI: 10.1016/B978-0-12-804137-6.00001-2.

Ratnikova T.A. (2006). The introduction to the econometric analysis of panel data (in Russian), The Economic magazine of Higher school of economics, 3, 492–519.

Reynolds, D. (2009). Gaussian Mixture Models. In: Li S.Z., Jain A. (eds) Encyclopedia of Biometrics. Springer, Boston, MA.

Roos G. (2017). *Technology-Driven Productivity Improvements and the Future of Work: Emerging Research and Opportunities*. Hershey, PA: IGI Global, 1-255. DOI: 10.4018/978-1-5225-2179-2.

Roos G. & Kennedy N. (2014). *Global Perspectives on Achieving Success in High and Low Cost Operating Environments*. IGI Global, Hershey, PA, USA. DOI: 10.4018/978-1-4666-5828-8.

Roos G., Cheshire A., Nayar S., Clarke S. M. & Zhang W. (Eds.) (2019). *Harnessing Marine Macroalgae for Industrial Purposes in an Australian Context: Emerging Research and Opportunities*. Hershey, PA: IGI Global.

Scott J.W. (2000). Euroregions, Governance and Transborder Co-operation within the EU. Borders, regions and people. *European research in regional science,* 10, 104–115.

Scott J.W. (2017). *Cross-Border, Transnational, and Interregional Cooperation*. Richardson D. (Ed.) The International Encyclopedia of Geography. DOI: 10.1002/9781118786352.wbieg0616.

Seethalakshmi R., Saavithri V. & Vijayabanu C. (2014). Gaussian Scale Mixture Model for Estimating Volatility as a Function of Economic Factor, *World Applied Sciences Journal,* 32 (6), 1035–1038. DOI: 10.5829/idosi.wasj.2014.32.06.783.

Socio-economic challenges, potentials and impacts of transnational cooperation in central Europe (2018). Final Report. The Vienna Institute for International Economic Studies (wiiw) and Interreg CENTRAL EUROPE. 266 p. [online]. Available at: www.interreg-central.eu/Content.Node/events/Final-Report-181018-full.pdf [Accessed 11 May 2019].

Stojkoski V. & Kocarev L. (2017). The Relationship Between Growth and Economic Complexity: Evidence from Southeastern and Central Europe. *MPRA Paper,* University Library of Munich, Germany.

Sweet C. & Eterovic D. (2019). Do patent rights matter? 40 years of innovation, complexity and productivity. *World Development*, 115 (C), 78–93. DOI: 10.1016/j.worlddev.2018.10.009.

Utkovski Z., Pradier M.F., Stojkoski V., Perez-Cruz F. & Kocarev L. (2018). Economic complexity unfolded: Interpretable model for the productive structure of economies. *PloS one*, 13 (8), e0200822. DOI:10.1371/journal.pone.0200822.

Verbeek M. (2003). A Guide to Modern Econometrics. John Wiley & Sons.

Wassenberg B., Reitel B., Peyrony J., Rubió J. (2015). *Territorial Cooperation in Europe - A Historical Perspective.* Luxembourg: Publications Office of the European Union.

Wooldridge J.M. (2007). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

1. Zhu S. & Li R. (2017). Economic complexity, human capital and economic growth: empirical research based on cross-country panel data. *Applied Economics*, 49 (38), 3815–3828. DOI: 10.1080/00036846.2016.1270413.

Received on June 17th, 2019 Accepted on January 21th, 2020

Andrei G. Manakov[1,2*], Jaak Kliimask[3]
[1] Pskov State University, Pskov, Russia.
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia.
[3] Estonian University of Life Sciences, Tartu, Estonia.
***Corresponding author:** region-psk@yandex.ru

RUSSIAN-ESTONIAN BORDER IN THE CONTEXT OF POST-SOVIET ETHNIC TRANSFORMATIONS

ABSTRACT. The authors explore the main trends and regional peculiarities of ethnic transformations during the post-Soviet period in the Russian and Estonian borderlands. Special emphasis is placed on the dynamics of the share of two dominant ethnic groups – Russians and Estonians. It is argued that the main trend of ethnic transformations is an increase of the share of the dominant ethnic groups in the structure of population. The almost Russian-speaking Ida-Viru county in Estonia is an exception, where together with a small growth in the share of Estonians, there was growth in the number of Russians. The authors analyse the dynamics of smaller Baltic-Finnic nations – Ingrians and Setus (Setos), living in the Russian border area with Estonia.

KEY WORDS: border territories, Estonia, Russia, population, ethnic structure, post-Soviet period

CITATION: Andrei G. Manakov, Jaak Kliimask (2020) Russian-Estonian Border In The Context Of Post-Soviet Ethnic Transformations. Geography, Environment, Sustainability, Vol.13, No 1, p. 16-20
DOI-10.24057/2071-9388-2019-68

INTRODUCTION

The demise of the Soviet Union gave an impetus to massive ethnic transformations in the post-Soviet area. Demographic processes, especially migration, started to have a considerable impact on the pace of ethnic transformation (Khrushchev 2010). The formation of new state borders in the post-Soviet area resulted in changes in the ethnic structure of the population in border territories. The growth of the dominant ethnic groups has already become a general trend in border territories, though it has not been equally strong in all parts of the Russian-Estonian border area. The objective of this research is to determine the regional peculiarities of the ethnic and demographic processes taking place in the Russian and Estonian borderlands during the post-Soviet period.

MATERIALS AND METHODS

There is an extensive body of research literature on the post-Soviet transformation of the ethnic structure of the population of Russia as a whole and of that in individual regions (Manakov 2016; Orlov 2013; Safronov 2014, 2015; Streletskiy 2011, 2017, 2018 etc.). Many research works explore ethnic and demographic processes in Estonia during the post-Soviet period (Sushchy 2018; Tammaru and Kulu 2003; Zhitin and Ivanova 2017 etc.) and study ethnic and social problems related to them (Berg 2001; Hallik 2011; Włodarska-Frykowska 2016 etc.).

The analysis of ethnic statistics has its own specificity since the data on the national composition of the population in the Soviet Union and in post-Soviet Russia was obtained only during censuses. The last Soviet population census was in 1989. In post-Soviet Russia, there have been only two population censuses – in 2002 and in 2010. In Estonia, the population censuses have been conducted in 2000 and in 2011. Due to the fact that the national composition of the population in Estonia is calculated every year, researchers can simultaneously analyse the dynamics of the ethnic structure of population in the border administrative units and towns in Russia and Estonia during the period 1989–2010. Additionally, we have included ethnic statistics on Estonia for 2011 and 2016 in the research.

To sum up, we analysed the results of the USSR census in 1989 and in the Russian Federation in 2002 and 2010, which are available on Demoscope Weekly[1], as well as the ethnic statistics from Estonia for 1989, 2002, 2010, 2011 and 2016, published on Population statistics of Eastern Europe & former USSR[2].

RESULTS AND DISCUSSION

Post-Soviet ethnic and demographic transformations in Russia can be divided into two types, which contributed to the growth of monoethnicity in the border territories and at the same time to its decline.

The general pace of demographic processes in border areas is defined by the polarization of population in Russia – its concentration in largest cities. That means there has been a massive migration outflow of Russian population from border areas mainly to St. Petersburg and Moscow as well as their agglomerations.

The growth of monoethnicity of border territories in Russia was caused by:

– migration outflow of Ukrainian and Belarusian population, especially in the 1990s; as well as fast assimilation of the remaining representatives of these nations in Russia;

– inflow of Russian population from the former Soviet republics, mainly in the 1990s, including Estonia and Latvia;

[1] Demoscope Weekly. Available at: http://www.demoscope.ru/weekly/ssp/census.php?cy [Accessed 19 Apr. 2019].
[2] Population statistics of Eastern Europe & former USSR. Available at: http://pop-stat.mashke.org/ [Accessed 19 Apr. 2019].

– outflow of Non-Russians from the border areas due to the concentration of dominant ethnic groups in their national autonomies, basically, in the republics of the Russian Federation.

The only process that influenced the reduction of monoethnicity of Russian borderlands was an inflow of non-ethnic population from the former Soviet republics, especially from Moldova, the Transcaucasus and Central Asia. The inflow could not compensate the outflow and assimilation of the Ukrainian and Belarusian population. It resulted in the growth of the Russian population in Russian border areas in the post-Soviet period.

Post-Soviet ethnic and demographic transformations in Estonia mainly dealt with the growth of the Estonian share of the population and decreasing number of Russians and other nationalities. But in border counties of Estonia there was an exact opposite process, which we talk about below. Generally, there was an intensive outflow of population from Estonia in the 1990s – mainly Russian speaking (Russians, Ukrainians and Belarusians). At the beginning of the 21st century that outflow, including the Russian population, decreased considerably. At the same time since 2004, after the accession of Estonia to the EU, there has been a strong migration outflow of mainly Estonians from the country to EU states. Eventually the growth of the Estonian share in the country slowed down, and in the second decade of the 21st century we can see a decrease of that share in certain years.

Generally speaking, in the second decade of the 21st century, the migration outflow from Estonia was decreasing, unlike that from Latvia and Lithuania. It was a natural consequence of a more favourable economic situation in Estonia compared to the other Baltic states. However, there was a considerable decrease of migratory outflow of Russians that led to a relative balance of Russians and Estonians in the ethnic structure of the population in Estonia.

Both in Russia and in Estonia the polarization of population increased in the post-Soviet period, as people moved to Tallinn and surrounding Harju county. Estonian counties bordering on Russia suffered the most. Ida-Viru county despite its natural resources and industrial traditions, became the most problematic area in Estonia in terms of low average salary and poor employment conditions. That pushed out the young generation (mostly non-Estonians) and made an inflow from other parts of Estonia highly unlikely, resulting in changes of the national structure of the population[1].

Demographic processes in the Russian-Estonian borderlands in the post-Soviet period.

According to the results of the census of 1989, the population of the whole Russian-Estonian borderlands, now covering Pskov and five administrative areas of Russia, as well as five counties of Estonia, was 918.5 thousand inhabitants, including 506.8 thousand people (55.2%) in Estonia, and 411.7 thousand people (44.8%) in Russia. By 2010, the total of the population of the whole Russian-Estonian border region decreased to 788.3 thousand people, i.e. by 14.2%. The fastest rate of population decrease was registered in the border counties of Estonia – by 16.3% - whereas in the Russian border area the population decrease was 11.5%. As a result, in 2010 the population of the Estonian borderlands was 424 thousand (53.8%) and in the Russian one – 364.3 thousand people (46.2%).

The most intensive decrease in the population number was in Estonian Ida-Viru county in the period 1989-2010 (by 23.7%). It was caused by industrial decline in a number of sectors, which were established in the Soviet period. Consequently, there was a massive outflow of population. Jõgeva, Põlva and Võru counties were characterized by an average rate of population loss. In Tartu county the population decreased only by 7.3%, because the city of Tartu is the second largest and important population centre in Estonia.

In the Russian part of the borderlands, the most considerable decrease in population during the same period was registered in peripheral districts located far from St. Petersburg and Pskov: Gdov (by 36.7%) and Slantsy (by 29.2%). The population of Kingisepp district decreased by 15%. The minimal loss of population was observed in Pskov city and neighbouring Pskov and Pechory districts (about 5%).

Post-Soviet ethnic transformations in the Russian-Estonian borderlands.

In the Soviet period, Estonia experienced a considerable inflow of Russian-speaking population (mainly Russians, Ukrainians and Belarusians), who moved to the capital of the republic and to the territory of modern Ida-Viru county, where many industries were developing rapidly, for instance, energy production, mining, manufacturing, shale oil excavation, the production of construction materials, chemicals, consumer goods, etc. Ida-Viru county became the most Russian-speaking county of Estonia. In 1989, Russians accounted for almost 70% of its population. The high share of Russians was registered in all large cities of the region: in Narva the share of Russians was 86%, in Sillamäe – 86.5%, in Kohtla-Järve – 64.7%, and in Kiviõli – 51.7%. Moreover, almost all the non-Estonian population in these cities was Russian-speaking.

In addition to Ida-Viru county, an increased share of Russians was observed in Tartu (Fig. 1), which is part of the Soviet ethnic legacy. But in Tartu county, as well as in neighbouring Jõgeva county there are two Russian-speaking towns (Kallaste and Mustvee) and a number of smaller settlements located on the coast of Lake Peipsi, which were founded by Russian Old Believers at the end of the 17th century. But due to the small population (less than 2,000 inhabitants), these towns and settlements of Old Believers have not had a large impact on the ethnic population structure in the Estonian counties along the western coast of Lake Peipsi. The lowest share of Russians was registered in Põlva and Võru counties.

In 1989–2010, the most considerable decrease in the share of the Russian population occurred in the city of Tartu. In those Estonian counties where the share of Russians was initially low (Jõgeva, Põlva и Võru), the decrease in the number of Russians was lower. In this context, it was unexpected to see a growth of the share of Russian population in Ida-Viru county, which occurred despite an intensive outflow of the population from that county. A similar demographic process was typical of all Russian-speaking towns of the region: Narva, Kohtla-Järve, Sillamäe and Kiviõli, in spite of the loss of a quarter to a half of their population. The main cause of this process was a decrease in the share of Ukrainians and Belarusians in Ida-Viru county due to their assimilation to the Russian population (Sushchy 2018). As a result, the Russian population experienced less considerable losses compared to other ethnic groups, excluding Estonians. It is worth noting that the largest outflow of population from Ida-Viru county was observed in the 1990s. Then it slowed down in the first decade of the 21st century and then increased

[1] *Regional development and policy. (2017). Regional development strategy for 2014-2020. Available at: https://www.rahandusministeerium.ee/en/regional-development-and-policy [Accessed on 19 Apr. 2019].*

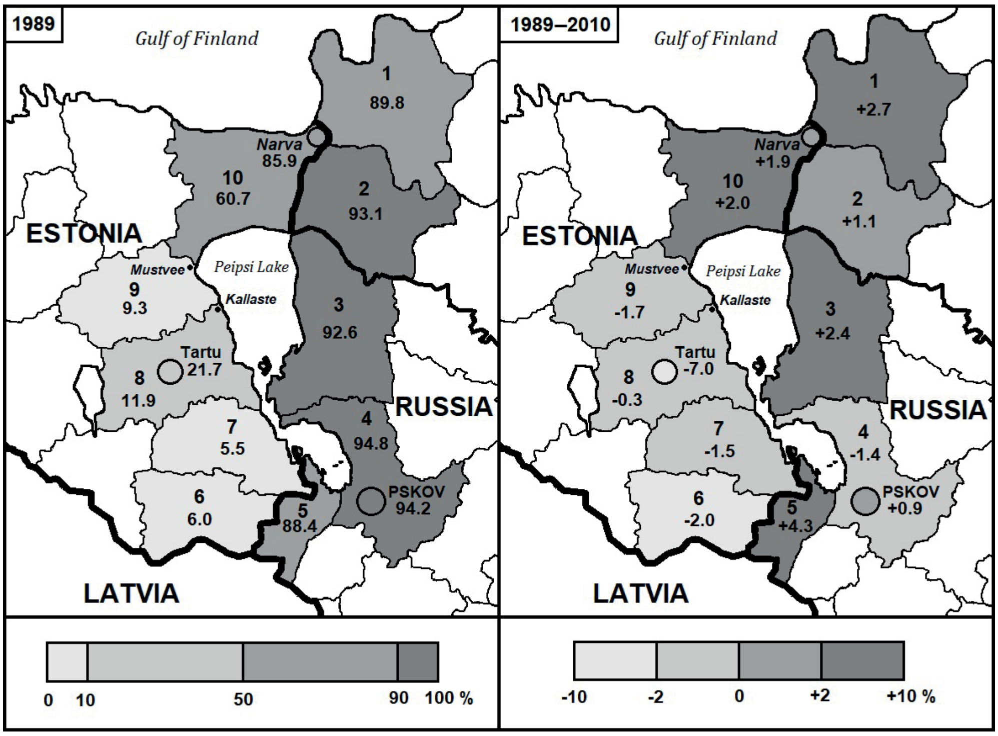


Fig. 1. The share of Russian population in border administrative units (modern borders) of Russia and Estonia based on the results of the census in 1989 and dynamics of the share from 1989 to 2010, in %

Districts of Leningrad region of the Russian Federation: 1 – Kingisepp, 2 – Slantsy; districts of Pskov region of the Russian Federation: 3 – Gdov, 4 – Pskov, 5 – Pechory; counties of Estonia: 6 – Võru, 7 – Põlva, 8 – Tartu, 9 – Jõgeva, 10 – Ida-Viru.

again after 2010. In 2010–2016, the share of Russians in the county increased by 1.6%, reaching 73%. During the same period there was a decrease in the share of Estonians to 19.1% (-0.5%) in Ida-Viru county, which occurred for the first time in the post-Soviet period.

At the end of the Soviet period in most border areas of Russia the share of Russians exceeded 90% of the population. The Pechory district in Pskov region and the Kingisepp district in Leningrad region were the only exceptions. The border areas cannot be called completely monoethnic due to the important share of two ethnic groups: Ukrainians (2.5% of the population of the Russian borderlands) and Belarusians (1.7%). In the post-Soviet period there was a notable decrease in the shares of these ethnic groups (by 0.8 and 0.6% respectively). So even the rather massive inflow of non-Russian population from the former Soviet republics of the Transcaucasus and Central Asia did not lead to a decrease in the ethnic Russian population share.

The inflow of non-Russian migrants in the post-Soviet period resulted in the lower share of Russians in Pskov district, but in the city of Pskov and other border districts the share of Russians increased. On average, the proportion of the Russian population increased by 1.4%, with the Pechory district standing apart mainly due to an outflow of the local Baltic-Finnic population, which is discussed below.

In 1989, the share of Estonians in the territory of modern border counties of Estonia was more than half of the population (54.5%), notably less than in the republic as a whole (61.5%). At the end of the Soviet period among border counties of Estonia, only Põlva county and Võru county could be called monoethnic since the share of Estonians exceeded 90% (Fig. 2). Jõgeva county was also close to being a monoethnic one, but in Tartu county the share of Estonians was only 76.1%. In the territory of these two counties, the share of Estonians was particularly low in the towns of Mustvee (43,3%) and Kallaste (19.7%). But the lowest share of Estonians at the end of the Soviet period was in modern Ida-Viru county – 18.5%, reaching its lowest share in the Russian-speaking towns of the region – Narva (4.0%) and Sillamäe (3.2%).

During 1989–2010, the growth in the share of Estonians in the border counties (+5.8%) was lower than in Estonia as a whole (+7.3%). The largest growth in the number of Estonians was registered in Tartu, mainly due to the outflow of Russians. Jõgeva county, Võru county and Põlva county demonstrated a considerable increase in the proportion of Estonians. Since in Kallaste and Mustvee, towns with a notable population loss, the migration outflow was mostly composed of the local Estonian population, the share of Estonians became much smaller. A minimal growth in the Estonian population occurred in Ida-Viru county (+1.1%). The same process was also typical for the following towns of the region: Narva, Kiviõli (+0.6%), Sillamäe (+1.6%), and in Kohtla-Järve, where the share of Estonians even decreased (-4.8%). Generally speaking, the growth in the number of Estonians in Ida-Viru county was caused by a decrease in the share of other minor ethnic groups and a notable migration outflow of Estonians from the region. Consequently, the demographic ratio of Russians and Estonians in Ida-Viru county changed slightly in the post-Soviet period.

In 1989 in the Russian border area, the share of the Baltic-Finnic nations increased only in Kingisepp of Leningrad region (mostly Ingrians), Gdov district of Pskov region (Estonians) and in Pechory district (Estonians and Setus). In Kingisepp district, there were about 700 Ingrians at the end of the Soviet period. In the post-Soviet period, the process of their assimilation and migratory outflow to Estonia and Finland increased and by 2010, their numbers almost halved. In Gdov district of Pskov region, the number of Estonians during the same period decreased 4-fold – from 200 to 50 people. In other border areas of Russia, the Baltic-Finnic population also assimilatedand migratory outflow. The only exception was in Slantsy district with a slight inflow of Finno-Ugrian population.

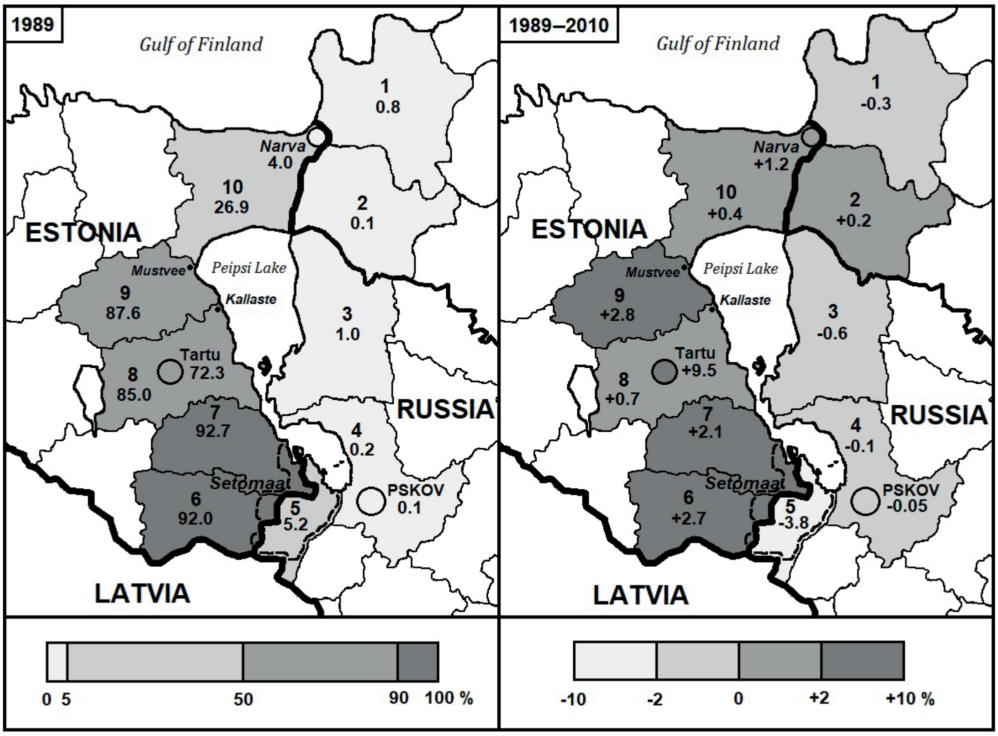


Fig. 2. Population structure of the border administrative units (modern borders) of The Baltic-Finnic nations in Russia and Estonians in Estonia based on the results of the census of 1989 and the dynamics of their share in 1989-2010, in %
Districts of Leningrad region of the Russian Federation: 1 – Kingisepp, 2 – Slantsy; districts of Pskov region of the Russian Federation: 3 – Gdov, 4 – Pskov, 5 – Pechory; counties of Estonia: 6 – Võru, 7 – Põlva, 8 – Tartu, 9 – Jõgeva, 10 – Ida-Viru. The dashed line indicates the historical territory of Setumaa.

We should also mention that Pechory area in Pskov region is the homeland of the Baltic-Finnic people called Setus (or Setos). This nation started to form in the 13th century due to the establishment of a political border between the Pskov lands with the Livonian Order to the west of the modern Russian-Estonian border. This part of the Baltic-Finnic population accepted Orthodox Christianity in the 16th century. As a result, Setus developed a different material and spiritual culture compared to Estonians, who first accepted Catholicism and later Lutheranism. In 1920, the whole territory of Setus became a part of the Republic of Estonia, in which Petseri county (or Setumaa) was formed. In 1944, the territory of Setumaa was divided between the Estonian and Russian Soviet Federative Socialist Republics (Manakov and Mikhaylova 2015). In Estonia some territories of Setumaa became part of Põlva county and Võru county, and in Russia – part of Pechory district.

According to the census of 1989, there were 1,140 Estonians in the Pechory district of Pskov region, of which about 950 of them are Setus. After the formation of the border between Russia and Estonia a migration outflow of Setus to Estonia began and continued until 2005. Our research shows that in 1999 there were about 500 Setus in Pechory district. In 2005, there were about 250 people (Manakov and Potapova 2013). Further on the migration outflow of Setus to Estonia almost stopped and the current number of Setus in Pechory district is a bit more than 200 people (Suvorkov 2017). In 2010, by the Decree of the Government of the Russian Federation Setus were included in the List of small-numbered peoples of Russia.

In Estonia, the number of Setus can be examined using the results of the census of 2011, when the population speaking different dialects of the Estonian language was counted. In Estonia, the census registered 12.5 thousand people speaking the Setu subdialect of the Võru dialect of the Estonian language. But within the Estonian part of Setumaa there are only about 1.8 thousand representatives of the Setu subdialect. In Põlva county in Setrmaa there are two rural municipalities (Mikitamäe and Värska) with 1,240 Setus (4.5% of the county population). In Võru county in Setumaa there is Meremäe rural municipality and a part of Misso rural municipality, where there are 550 representatives of the Setu subdialect living (1.6% of the county population). These rural municipalities were included in the single municipality of Setumaa making part of Võru county. Therefore, the total number of Setus in the Russian and Estonian part of Setumaa can be estimated at 2,000 people.

CONCLUSION

The results of our research confirmed the general trend of the transformation of the ethnic structure of the population in the border areas of Russia and Estonia in the post-Soviet period. This trend manifested in the increase of the share of the dominant ethnic group (Estonians in Estonia and Russians in the border areas of Russia). However, in the post-Soviet period also the processes that could not be considered as part of the described common trend of ethnic transformation of Russia's and Estonia's population were observed. In Estonia, the growth of the share of ethnic Estonians in the population of border districts (+5.8%) was lower than the average in the country. The most Russian-speaking county of Estonia, Ida-Viru experienced a catastrophic decrease of population (almost by one quarter) as a result of a significant downscaling of the mining and manufacturing industry. However, the intensity of the migration outflow of Russians and a certain

[1] *RL0445: Population with Estonian as their mother tongue by ability to speak a dialect, age group and place of residence, 31 December 2011. Available at: http://andmebaas.stat.ee/Index.aspx?lang=en&SubSessionId=db4eee2c-dc1c-4d51-a603-eefe3dba90dd&themetreeid=7 [Accessed 19 Apr. 2019].*

part of Ukrainians and Belarusians calling themselves 'Russians', was twice as small as the loss in the county population as a whole. It resulted in a small growth in the Russian population, especially in the towns of Ida-Viru county. A similar phenomenon was observed in two Russian-speaking towns located in the western coast of Peipsi (Chudskoe) lake: Kallaste (Tartu county) and Mustvee (Jõgeva county). There was also a significant decrease in the population, but the local Estonian population was more involved in the migration outflow, which led to a decline in the share of Estonians in them. In the overwhelming majority of the regions of Russian Federation bordering Estonia the share of Russians increased while the share of Baltic-Finnish ethnic groups diminished. The most visible these processes were in Pechora district of Pskov region. The establishment of the state border between Russia and Estonia had a negative impact upon the demography of Setus. The territory of their residence is now divided between a Russian part (Pechory district of Pskov region) and an Estonian part (Võru county). In the post-Soviet period, there was a considerable outflow of Setus to Estonia from Russia, which resulted in the fourfold decrease in their number.

ACKNOWLEDGEMENTS

This research was carried out with the support of the Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University.

REFERENCES

Berg E. (2001). Ethnic Mobilisation in Flux: Revisiting Peripherality and Minority Discontent in Estonia. Space & Polity, 5(1), 5-26.

Hallik K.S. (2011). Russians in Estonia. In: Russians: ethno-sociological research. Moscow: Nauka, 90-119 (in Russian with English summary).

Khrushchev S.A. (2010). The ethnic aspects of demographic processes in the Baltic region states. Baltic Region, 4, 81-91.

Manakov A.G. and Potapova K.N. (2013). An ethnic-social portrait of the Setus of Pechorskii District, Pskov Region (through Case Studies of 1999-2011). Etnograficeskoe Obozrenie, 2, 177-187 (in Russian with English summary).

Manakov A.G. and Mikhaylova A.A. (2015). Changes in the territorial and administrative division of Northwest Russia over the Soviet period. International Journal of Economics and Financial Issues, 5, 37-40.

Manakov A.G. (2016). Dynamics of the national composition of the population of the North-West of Russia in 1939-2010. Pskov Journal of Regional Studies, 3, 74-96 (in Russian with English summary).

Orlov A.Yu. (2013). Historical and geographical aspects of the transformation of the ethnic structure of the population of the Russian Federation. Regional Studies, 2, 120-124 (in Russian with English summary).

Safronov S.G. (2014). Transformations in ethnic population composition in Russia in 1989-2010. Regional Research of Russia, 4(2), 38-46.

Safronov S.G. (2015). Russian population ethnic structure: trends and transformations. Baltic Region, 3, 106-120.

Streletsky V.N. (2011). Shifts in the ethnic settlement in Russia at the end of the 20 – beginning of the 21 centuries and their cultural and geographical aspects. South Russian Forum: economics, sociology, political science, social and economic geography, 1(2), 51-72 (in Russian)

Streletsky V.N. (2017). Ethnic, confessional and cultural patterns of regionalism in the post-Soviet Russia. Hungarian Geographical Bulletin, 66(3), 219-233.

Streletskiy V. (2018). Ethnic settlement shifts in Russia in 1989-2017: Main trends and cultural-geographical consequences. Practical Geography and 21 Century Challenges International Geographical Union Thematic Conference dedicated to the Centennial of the Institute of Geography of the Russian Academy of Sciences, 532.

Suvorkov P.E. (2017). Ethno-social portrait of the Setus of the Pechory district of the Pskov region (based on the materials of the expedition 2016). Pskov Journal of Regional Studies. 2017, 3, 55-79 (in Russian with English summary).

Sushchy S.Ya. (2018). Russians in the Baltic States – geodemographic trends of the post-Soviet period and prospects for the first half of the 21st century. Narodonaseleniye, 21(3), 21-36 (in Russian with English summary).

Tammaru T. and Kulu H. (2003). The Ethnic Minorities of Estonia: Changing Size, Location, and Composition. Eurasian Geography and Economics, 44(2), 105-120.

Włodarska-Frykowska A. (2016). Ethnic Russian Minority in Estonia. International Studies Interdisciplinary Political and Cultural Journal, 18(2), 153-164.

Zhitin D.V. and Ivanova A.A. (2017). Demographic development of the Baltic countries. Moscow: Russian Book Publishers Association (in Russian with English summary).

Received on June 21th, 2019 Accepted on January 21th, 2020

Tadeusz Palmowski[1] and Gennady M. Fedorov[2]*
[1] University of Gdansk, Poland
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia
***Corresponding author:** GFedorov@kantiana.ru

THE POTENTIAL FOR DEVELOPMENT OF RUSSIAN-POLISH CROSS-BORDER REGION

ABSTRACT. The problems of the formation of international regions on the borders of Russian Federation and EU countries after the deterioration of relations between them in 2014 became more complicated due to the reduction of mutual economic, social, political and other cross-border ties. However, such links remain, especially at the local level, as both sides benefit from them. Polish and Russian authors are trying to find common approaches in assessing the situation and explaining the need in the development of relations between cities, territories and businesses located on both sides of the border, which contributes to the formation of cross-border regions. The authors use literature, materials of cross-border cooperation programs and their own research experience, identifying factors and features of cross-border interactions at the Russian-Polish border. The article presents a SWOT analysis of the formation of the Russian-Polish cross-border region – a comparison, on the one hand, of strengths and weaknesses, and on the other, opportunities and threats to its development. It is shown that in 2014–2019 political factors prevailed over socio-economic ones, which negatively affected the development of the regions along the border. Nevertheless, in 2018 the implementation of joint projects within the framework of the Russia-Poland cross-border cooperation program co-financed by the EU and both countries continued. Although the number of mutual crossings of the border has decreased, it remains quite important. In Kaliningrad, there is a Polish visa center that promptly issues Schengen visas, free of charge for scientists and teachers, students and some other categories of the population. In the summer 2019, free electronic visas were established in Kaliningrad region, which increased the influx of tourists, including Polish. The authors hope that the objective laws of the world market will lead to the intensification of mutual relations and the formation of the Russian-Polish cross-border region, which would contribute to increasing the international competitiveness of its parts on both sides of the border.

KEY WORDS: cross-border region, Russian-Polish relations, South-Eastern Baltic, Kaliningrad region, Warmian-Masurian voivodeship, Pomeranian voivodeship

CITATION: Tadeusz Palmowski and Gennady M. Fedorov (2020) The Potential For Development Of Russian-Polish Cross-Border Region. Geography, Environment, Sustainability, Vol.13, No 1, p. 21-28
DOI-10.24057/2071-9388-2019-70

INTRODUCTION

It is widely accepted that the increasing regionalisation of the late 20th century, including the formation of international regions of various hierarchical levels, was caused by growing globalisation. In Eastern Europe, globalisation and the fall of the Iron Curtain between capitalist and socialist countries in the late 1980s/early 1990s weakened the barrier functions of borders and strengthened the contact ones. However, there were significant differences between countries (Herrschel 2011). Fundamental changes in favour of weakening the barrier and enhancing the contact function of the border between states have occurred within the European Union. The specifics of the situation in different countries are revealed by Kolosov and Więckowski (2018), who have identified a number of areas in cross-border research. In our article, we add another aspect that generalizes many others: the formation of a cross-border Russian-Polish region. New transnational and cross-border regions were emerging. However, the euphoria surrounding this process, which was felt by many researchers, was replaced by equivocal expert statements. A barometer of these changes was the transparency of the Russian–Polish border. Since the 1990s, it has been alternating between stronger contact function and predominant barrier function. Kolosov et al. (2018a, 2018b) give a clear picture of the dynamics of change, which was caused by shifts in political relations between the parties, at the Russia–EU and accordingly the Russian–Polish border. These change affected the conditions in which the Russian–Polish cross-border region comprising Russia's Kaliningrad region and Poland's Warmian-Masurian and partly Pomeranian voivodeships was developing.

In this article, we consider the factors that determine the dynamics of Russian–Polish cross-border relations. We pay special attention to the balance between strengths and weaknesses and opportunities and threats to the development of cooperation and the formation of a cross-border Russian–Polish region. To give a practical perspective to our theoretical framework, we analyse cases of successful, failed, and promising cooperation projects.

METHODS

Methodologically, our work relies on a systemic approach to studying socio-economic processes taking place in border regions of cooperating countries. We consider cross-border regions and other spatially localised entities (euroregions, clusters, and others) as territorial socio-economic systems that have strong internal ties and respond to external stimuli as a single whole.

Socio-economic infrastructure, people, and authorities on either side of a border are engaged in more or less active collaborations. Researchers have identified various spatial forms of cross-border cooperation: euroregions, associations of local authorities and regions, large regions, growth triangles, arches, development corridors, mega-corridors, cross-border corridors and bridges, cross-border districts and clusters, bipolar and tripolar cross-border systems, and cross-border cities (Association 2004; Druzhinin 2017; Kaledin at al. 2008; Kivikari 2001; Klemeshev at al. 2006; Lechevalier at al. 2013; Mikhaylov 2014; Palmowski 2010; Sohn at al. 2009). The most general term to refer to a territory brought together by mutual ties is cross-border region (Fedorov at al.2009; Ganster at al. 1994; Ganster at al. 1994; Groß at al. 1994; Perkmann 2003; Schmitt-Egner 1996; Scott 1999; Van der Velde at al. 1997). Sometimes, cross-border regions are viewed as part of a single geosystem, which includes both socioeconomic and environmental components (Baklanov at al. 2008). The theory of cross-border region formation, which has introduced these terms, is the methodological framework of this study.

National regions theory distinguishes between homogeneous and coherent regions. Cross-border regions are usually defined as a range of areas belonging to neighbouring countries and brought together either by territorial homogeneity (homogeneous regions) or by strong ties between administrative units of bordering states (coherent regions). Comprising homogeneous border territories, cross-border regions are very similar to homogeneous national regions. A region with homogeneous physiographical characteristics that creates a physiographical continuum is the Baltic / Vistula Spit divided by the Russian–Polish border.

Common natural features are shared by the sections of the South-Eastern Baltic on either side of the Russian–Polish border. Stretching across the northwest of the Baltic Upland towards the Baltic Sea through coastal lowlands, this area can be considered a homogeneous cross-border region.

Sustained by internal connections, coherent cross-border regions (most of which belong to the socio-economic type) differ dramatically from coherent national regions. The primary distinction is that their agents (companies, institutions, organisations) have closer ties with national rather than international partners (Fedorov at al. 2009; Klemeshev at al. 2015). At the same time, relations between the territories of the neighbouring states are developing quite successfully in the Russian–Polish region, whereas the contact function of the border between the two countries, according to the authors of this article, is stronger than the barrier one.

The formation of cross-border regions is most intensive in the countries of the EU, where Union bodies encourage closer cross-border ties between territorial units and municipalities and facilitate the development of cross-border territorial communities – euroregions. The very first one, called EUREGIO, appeared at the German-Dutch border as early as 1958 (EUREGIO 2019).

Increasing globalisation intensifies regionalisation. The EU has used this process to expand and strengthen cross-border ties. The first half of the 1990s saw a surge of publications portraying the EU as a 'Europe of regions'. They stressed that cross-border cooperation at a regional and municipal level would contribute to a stronger integration of the EU.

After the dissolution of the Soviet bloc, as the barrier function of national borders weakened and the contact function strengthened, cross-border ties started to develop along the borders between all European countries, particularly, between EU member states and their neighbours, including Russia. Economic and social relations between Russian regions and neighbouring countries were established along other Russian borders. Similar processes were taking place across the world. Cross-border ties are the key to the development of many border regions. Studies into cross-border ties gained momentum in the mid-1990s, facilitating the emergence of a theory of cross-border regional formation (Gabbe 1997; Kolosov at al. 1997; Perkmann 1997, 2003; Raich 1995; Rees 1997; Van der Velde at al. 1997).

In the Baltic, cross-border regions are rapidly emerging at the borders of Sweden and Denmark, Germany and Poland. A favourable situation for their formation is within the Russia–Finland–Estonia and Russia–Poland–Lithuania border area triangles. Earlier, they had a good chance to develop between Russia and Finland, Russia and Estonia, Russia and Poland, and Russia and Lithuania. Cross-border cooperation is increasing between the neighbouring regions of the EU and Russia. In geographical terms, such cooperation compensates for the shortcomings of the peripheral position occupied by border regions in their countries. Border territories of neighbouring countries may evolve into international development corridors (Fedorov 2018a; Fedorov at al. 2015; Klemeshev at al. 2004), which follow Friedmann's model of national development corridors (Friedmann 1966) connecting core regions within one country.

Cross-border regions have an objective foundation: benefits for production through cooperation, which increase the competitiveness of economic entities on either side of the border, benefits for the social sphere, and exchange of experience in governance. At the same time, their formation is spurred by subjective factors: actions taken by the authorities, NGOs, and non-profits to develop international cooperation.

Cross-border regions emerge at meso- and micro-territorial levels (Fedorov and Korneyevets 2009; Korneyevets 2010; Kropinova 2016; Palmowski 2006; Studzieniecki at al. 2016). A mesoregion is developing along the Russian-Polish border, where almost all spatial forms of cross-border relations are either present or likely to appear. This region is identified based on an assessment of the density of mutual connections between the Russian (Kaliningrad region) and the Polish (Warmian-Masurian and Pomeranian voivodeships) agents of cooperation. In this article, we describe the factors that determine the rate of cooperation development and discuss emerging cross-border forms of economic organization, as well as the most productive joint projects. Based on this, we carry out a SWOT analysis of the conditions and factors behind the development of a Russian–Polish region and assess the prospects for its development.

RESULTS AND DISCUSSION

Before 1945, the territories of the Kaliningrad region and the Warmian-Masurian and most of the Pomeranian voivodeships were part of German East Prussia. On either side of the border, there are remnants of the past era – elements of the architectural environment, the settlement system, and the transport networks. Thus, the South-Eastern Baltic, which is a homogenous region in this respect, has distinctive cultural and historical commonalities. However, when identifying this region, we focus primarily on the socio-economic ties between the border parts of Russia and Poland. That is, we consider it as a coherent region.

To evaluate the factors behind the formation of a Russian–Polish cross-border region, we carried out SWOT analysis. Initially, a technique used in strategic management (Andrews 1971), it is employed today in strategic planning (Table 1).[1] A two-by-two matrix was chosen as its most proper modification (Chermack at al. 2007; Lowy at al. (2019).

[1] SWOT (strengths, weaknesses, opportunities, threats) analysis is a technique in which the factors of development of the object studied are analysed.

Table 1. SWOT analysis of the conditions and factors behind the formation of a Russian–Polish cross-border region

SWOT analysis (by K.Andrews)	*Strengths (internal)* S1. Benefits through cross-border industrial cooperation S2. The development of a cross-border regional market S3. The parties are interested in cooperation in the social sphere (education, research, healthcare, culture, sports) S4.The parties cooperate in solving common environmental problems.	*Weaknesses (internal)* W1. The barrier function is stronger than the contact one W2. The legacy of the command economy W3. Relatively poor development of the 10–15 km border area
Opportunities (external) O1.Integration of transport infrastructure O2.Proximity between economic agents ready to cooperate O3.A wider market for sales O4.Bilateral and multilateral documents and cooperation development programmes	*Opportunities for employing strengths* O1-S1. Joint participation in the maintenance of the North–South and East–West traffic O2-S2. Business clusters specialising in shipbuilding, furniture production, agriculture, innovative enterprises and tourism O3-S2. More goods and services produced O4-S3. Development of cross-border socio-cultural relations. O4-S4. Creation of cross-border conservation areas	*Opportunities for overcoming weaknesses* O1-W1. Increasing throughput and building new border crossings. The effective operation of the Polish Visa Application Center in Kaliningrad; introduction of free electronic tourist visas to the Kaliningrad region by the Russian authorities. O2-W2. Exchange of experience between regional and municipal authorities, international conferences, organization of foreign student practices, international summer schools O4-W3. Joint development of projects for the development of border areas, objects.
Threats (external) T1. Instability in Russian-Polish political relations	T1-S1.Employing strengths to eliminate threats T1-S2.Joint efforts in the arena of Baltic international organisations (Council of Baltic Sea States, HELCOM, etc.) T1-S3.Reciprocal visits by representatives of regional, municipal authorities and representatives of socio-cultural organizations.	Eliminating weaknesses to reduce threats T1-W1. A visa-free regime (starting with the resumption of local border traffic) T1-W2. Broader cultural exchange T3-W3. Joint programmes for the development of cross-border territories (including as part of the Cross-Border Cooperation Programme Poland-Russia)

Using the «two by two» matrix, we compared, on the one hand, the strengths and weaknesses of the territory, and on the other hand, the external opportunities and threats that were identified as a result of studies conducted by the authors. Then we compared external capabilities with internal forces (O-S) and weaknesses (O-W), external threats with internal forces (T-S) and weaknesses (T-W). We determined how the emerging Russian-Polish cross-border region can benefit from its strengths and overcome disadvanages of the weaknesses as well as benefit from the strengths and eliminate the weaknesses to reduce threats.

In fact, there is only one major threat: instability in Russian-Polish political relations. It negates the effects of positive factors. Only the development of mutually beneficial economic and cultural ties, the expansion of contacts between authorities, business, social institutions, public organizations of neighboring territories of the two countries can counteract political differences and ensure the formation of a cross-border Russian-Polish region.

Below, we will analyse two cases, one of them demonstrating how the potential for cooperation can be exploited amid political tensions and the other showing how an earlier successful project was terminated for political reasons.

Case 1. Cross-border cooperation programme

Projects launched within cross-border cooperation programmes help to identify promising areas for collaboration in solving problems of mutual interest, as well as to develop joint actions. In Europe, these projects are initiated by the European Regional Development Fund, which has been running the Interreg programme since 1989. Scheduled for 2014–2020, Interreg V brings together the twenty-eight counties of the EU and twenty-seven non-EU partners, including Russia. All the non-EU members (both the states and their organisations involved in the project, although the latter to a much lesser extent) take part in co-financing the programmes (Interreg 2019).

The Kaliningrad region is covered by the Baltic Sea region sub-programme and Russia–Poland cross-border cooperation programme. As of the beginning of 2019, all the projects of the Baltic Sea Region programme were approved and underway (Russian 2019). Particularly, several projects involve Russian and Polish regions, as well as those of other Baltic Sea countries.

In the first half of 2019, a call for projects for the Poland–Russia 2014–2020 cross-border cooperation programme was concluded. The programme priorities include (Russia-EU 2019):

1. cooperation to promote historical, natural, and cultural heritage and cross-border development;

2. protection of the environment in the cross-border region;

3. accessibility of the regions and reliable cross-border traffic and communications;

4. joint action to ensure the efficiency and security of borders.

Fig. 1 shows the territorial scope of the programme.

Hatching indicates the areas of bordering regions supporting the cooperation mechanisms, which were developed within earlier programmes (Poland ... 2019).

Case 2. Local border traffic

For many years, the Polish–Russian border served as a major physical and intellectual barrier. The situation changed in the 1990s. In 1991–2003 (until October 2003), a visa-free regime existed there (Agreement 2003). A new attempt at a visa-free regime (this time, for the Kaliningrad region and the neighbouring Polish territories [fig. 2]) was local border traffic, which was in effect from July 27, 2012, to July 3, 2016 (Ministry 2019a, 2019b).

The visa-free regime was a chance for the residents of the border area to improve their material prosperity, since the novelty granted access to cheaper goods available across

Fig. 1. The territorial scope of the Poland–Russia 2014–2020 cross-border cooperation programme

Fig. 2. The area of local border traffic between Russia and Poland

Source: (Local 2019).

the border. Moreover, it gave Kaliningraders an opportunity to earn extra cash. Growing cross-border travel contributed to brisk economic activity and the development of small enterprises. The border had a major effect on the functioning of the territories on its either side. The opening of the border boosted the socio-economic development of the borderland towns and villages (Studzińska 2014).

The Polish–Russian borderlands are a special territory from a historical, social, and economic perspective, as well as in terms of their geopolitical position. On the one hand, the Polish–Russian border is local, since it is crossed primarily by the residents of the border areas. On the other hand, it serves as a mirror of Russian-Polish bilateral relations and the EU policy towards its neighbours. The four years of the local border traffic regime were a success. Moreover, changes to the EU rules extended the local border traffic area to the whole territory of the Kaliningrad region and the major urban and academic (Tricity) and tourism centres (the Masurian Lakes in Poland and the seacoast of the Kaliningrad region). The benefits from expanding the local border traffic exceeded all expectations and made a significant contribution to cross-border integration (Kolosov at al. (2018a). In 2014–2015, local border traffic ID cards were used for half of the crossings of the Russian-Polish border. There were fewer than 2.5 million crossings of the border in 2011 and more than 6.5 million in 2014 (Anisiewicz at al. 2016). In 2017, after the termination of the local border traffic, only 3.9 million people crossed the border, including 2.5 from the Russian side (Biuletyn 2017; Gumenyuk at al. 2018).

The local border traffic regime has received a positive response on either side of the border. Nevertheless, on July 3, 2016, the Polish authorities decided to suspend the regime, and the Russian authorities responded accordingly. In this case, the political factor had a negative effect on the development of cross-border ties. There is still hope that the local border traffic regime will be restored over time.

The contribution of the Kaliningrad region and the authorities of the Pomeranian and Warmian-Masurian voivodeships to the launch of local border traffic and the role they played in its functioning are a good example of how the efforts of all levels of authorities can work together. The local border traffic area could have become the touchstone of cooperation between state and local authorities in borderlands. However, the decision of the Polish authorities to suspend the local border traffic regime prevented this. Thus, socio-economic initiatives in border areas depend on the decisions of central authorities. However, in April 2019, Poland started to discuss the possibility of resuming the local border traffic regime with Russia: the Civic Platform party declared that it would resume local border traffic with the Kaliningrad region as soon as it came to power (In Poland 2019).

Case 3. New spatial forms of organization of the economy.

The cross-border region is a new spatial form of economic organisation. Its most common types are euroregions, growth triangles, cross-border clusters, and bi-, tri-, and multipolar systems of international cities. All of them are developing in the Baltic region, many with Russian and Polish participation. These forms bring together regions, municipalities, economic entities, businesses, social welfare institutions, and non-profit organisations.

Euroregions coordinate the joint activities of their constituents, primarily so in social welfare and environmental protection. They rely on the European Outline Conventionon Trans-border Co-operation between Territorial Communities or Authorities. Joint efforts are coordinated by special bodies making non-binding decisions. Russia and Poland together participate in four euroregions: Baltic, Łyna-Ława, Šešupė, and Neman (On the activity 2019).

An active player is the euroregion Baltic established in 1998. It brings tougher the Kaliningrad region of Russia, the Pomeranian and Warmian-Masurian voivodeships of Poland, and administrative units of Lithuania, Sweden, and Denmark. Their collaborations cover environmental protection, youth projects, small entrepreneurship, living standards improvement, and sharing experience in support for disadvantaged social groups.

The Neman euroregion, which was established in 1997, comprises the Podlaskie voivodeship of Poland, the eastern municipalities of the Kaliningrad region (since 2002), and some regions and municipalities of Lithuania and Belarus. The euroregion Łyna-Ława (2003) includes Russian and Polish border territories. A remarkable event is the annual canoeing regatta on the River Łyna-Ława, which gave its name to the euroregion.

As a combination of proximate and horizontally linked economic entities of two or more countries, trans-border clustering has occurred in tourism only. However, clusters may emerge in shipbuilding, furniture production, and agriculture (Druzhinin 2017; Mikhaylov 2014).

Under certain conditions, a bipolar city/agglomeration system may connect Tricity (Gdansk – Gdynia – Sopot) and Kaliningrad (Palmowski 2006). Collaborations are possible in manufacturing (shipbuilding and the food industry), transport, tourism, education and science (particularly, ocean studies), and healthcare. This system may incorporate Klaipeda, thus becoming a tripolar structure (Fedorov 2010).

Growth triangles are the joint efforts of three partners that have different kinds of resources (natural, human, or investment ones) and create together somewhat of a manufacturing cluster. Whereas Russia, represented by the Kaliningrad region, has the necessary natural resources, Poland has the human resources. Thus, the structure is lacking a partner with investment resources. This may be Germany, Sweden, or Denmark. In this case, the idea of a growth triangle in the South-Eastern Baltic will become viable (Kivikari 2001).

CONCLUSION

Amid increasing competition in the world market and growing inter-civilisation tensions, Cross-border cooperation is, firstly, an important factor enhancing the competitiveness of border regions and, secondly, a means to learn about the culture and everyday life of neighbours and thus to ensure mutual understanding. Although Poles and Russians are associated with different civilisations (the Western and Orthodox ones), they speak similar Slavic languages, have similar tastes in food and a similar mindset. All this contributes to international contacts (although some pages of the common history complicate them).

The South-Eastern Baltic, where a Russian–Polish cross-border region is developing, has a very beneficial economic and geographical position (fig. 3). Transport routes running along the southern and eastern coast of the Baltic Sea meet there. This territory may once carry the traffic of the New Silk Road (Druzhinin at al. 2018; Fedorov 2018a; Kolosov at al. 2017). However, this will require the modernisation of roads, railways, port facilities, and checkpoints at the Russian–Polish border.

Enhancing the transport component of the South-East Baltic and the area taking part in transcontinental traffic may boost the development of manufacturing companies processing cargoes and contribute to the formation of industrial clusters, which will be more effective than isolated businesses. Clusters are likely to emerge in agriculture, the fishing industry, and shipbuilding.

Fig. 3. The geographical location of the South-East Baltic: Baltic Crossroad

Joint projects within cross-border cooperation projects and bilateral Russian–Polish agreements will facilitate the development of industrial and social infrastructure, the creation of international tourist routes, and growing expertise of social workers and managers.

Objective market patterns and subjective efforts made by authorities, economic entities, and non-profits should ultimately result in constructive political relations and the formation of a globally competitive cross-border Russian–Polish region in the South-East Baltic.

In any case, although mutual investment in the economy does not increase remaining very poor, in 2017 – 2018trade between the two countries increased. In the years 2000 – 2015 Poland's share in the volume of Russian foreign trade turnover declined from 8% to 2.6%, and then in 2016 and 2017 it increased to 2.8%, and in 2018 to 3.2%.Russia is on the third place in Poland's foreign trade, preceded by Germany and China. Despite the fact that the local border traffic between the Kaliningrad region and neighboring Polish regions, which operated in 2012 – 2016, has not been restored, there are large mutual tourist flows. Visas are issued promptly by the Polish Visa Application Center in Kaliningrad. Since July 1, 2019 free electronic visas have been operating in the Kaliningrad region. This fact contributes to a significant increase in the number of foreign tourists, including Polish, arriving in the region.

ACKNOWLEDGEMENTS

The study was supported by a grant from the Russian Science Foundation, 18-17-00112, «Ensuring the Economic Safety of Russia's Western Borderlands amid geopolitical Turbulence».

REFERENCES

Agreement between the Government of the Russian Federation and the Government of the Republic of Poland on the conditions of mutual trips of citizens of the Russian Federation and citizens of the Republic of Poland (concluded in Warsaw, 18.09.2003) (2003). [online] Available at: www.consultant.ru/document/cons_doc_LAW_124416/ [Accessed 23.04.2019) (In Russian).

Andrews K.R. (1971). The concept of corporate strategy. Homewood. Illinois: Dow Jones – Irwin.

Anisiewicz R. and Palmowski T. (2016). Współpraca polski z obwodemKaliningradzkim Federecii Rosyjskiej jako istotny element integracii bałtyckiej. Prace i Studia Geograficzne, 61(1), 13-28. [online] Available at: www.wgsr.uw.edu.pl/wgsr/wp-content/uploads/2018/10/61.1art.2.pdf [Accessed 23.04.2019].

Association of European Border Regions (2004). Synthesis report: towards a new community legal instrument facilitating public law based transeuropean co-operation among territorial authorities in the European Union. Gronau: AEBR.

Baklanov P.Y. and Ganzey S.S. (2004). Border and transboundary territories as an object of geographical research. Proceedings of the RAS. Ser. geogr., 4, 27-34.

Biuletyn Statystyczny Straży Granicznej za 2017 r. (2017). Warszawa, 2018. Available at: BIULETYN_ STATYSTYCZNY_STRAZY_ GRANICZNEJ_ZA_2017_ROK.pdf.

Chermack T.J. and Kasshnna B.K. (2007). The use and misuse of SWOT Analysis and implications for HRD professionals. Human Resource Development, 10 (4), 383-399.

Druzhinin A.G. (2017). Transboundary clustering in the coastal zones of European Russia: factors, models, and economic and ecistic effects. Rostov-on-Don: Southern Federal University Press (in Russian).

Druzhinin A.G. and Dong Y. (2018). One Belt – One Road Initiative: A Window of Opportunity for Russia's Western Border Regions. Baltic region, 39-55.

EUREGIO. [online] Available at: www.euregio.eu/en [Accessed 19.04.2019].

Federal Service State Statistics. [online] Available at: www.gks.ru/ [Accessed 08.11.2019]. (In Russian).

Fedorov G.M. (2010). The Kaliningrad dilemma: a «development corridor» or a «double periphery? The geopolitical factor of the Development of the Russian exclave on the Baltic Sea. – Baltic region, 2, 4-12.

Fedorov G.M. (2018a). Kaliningrad region in the implementation of the project Greater Eurasia. Pskov regional journal. №. 4 (36), 33-43. (In Russian).

Fedorov G.M. (2018b). On the «international development corridors» and their formation in the border regions of the Russian Federation. – Regional studies, 3 (61), 76-83 (In Russian).

Fedorov G. Belova A., Osmolovskaya L. (2015). On the future role of Kaliningrad Oblast of Russia as an «international development corridor». – Eurolimes, 19, 57-68.

Fedorov G.M. and Korneevets V.S. (2009). Trans-Border Regions in the System of the Regional Hierarchy: the Systemic Approach. Baltic region, 2, 26-33. DOI: 10.5922/2079-8555-2009-2-3.

Friedmann J. (1966). Planning as Innovation: The Chilean Case. Journal of the American Institute of Planners, 32(4), 194-204. DOI: 10.1080/01944366608978495.

Gabbe J. (1997). Europäische Modelle interregionaler und grenzüberschreitender Kooperation. INTERREGIONES, 6, 7-12.

Ganster P., Sweedler A., Scott J. and Eberwein W.-D. (eds.). (1994). Borders and border regions in Europe and North America. San Diego: San Diego State University Press.

Groß B. and Schmitt-Egner P. (1994). Europas kooperierende Regionen. Rahmen bedingungen und Praxis transnationaler Zusammenarbeit deutscher Grenzregionen in Europa. Baden-Baden: Nomos.

Gumenyuk I.S. and Studzieniecki T. (2018) Current and Prospective Transport Connections between Poland's Border Voivodeships and Russia's Kaliningrad Region: Baltic Region, 10(2), 114-132. DOI: 10.5922/2079-8555-2018-2-8.

Herrschel T. (2011). Borders in post-socialist Europe: Territory, Scale, Society. Farnham: Ashgate.

In Poland, the opposition party supported the return of WFP with Kaliningrad (2019). New Kaliningrad. [online] Available at: www.newkaliningrad.ru/news/briefs/politics/22003272-v-polshe-oppozitsionnaya-partiya-vystupila-za-vozvrashchenie-mpp-s-kaliningradom.html [Accessed 14.04.2019]. (In Russian).

Interreg Europe (2019). [online] Available at: URL: www.interregeurope.eu/ [Available at: 14.04.2019].

Kaledin N.V., Korneevets V.S., Chekalina T.N. (2008). Network cooperation as a fundamental factor in the formation of trans-border regions. Bulletin of St. Petersburg University. Ser. Geology, geography, 4, 130-39 (In Russian).

Kivikari U. (2001). A Growth Triangle as an Application of the Northern Dimension Policy in the Baltic Sea Region. Russian-Europe Centre for Economic Policy. Policy Paper. May 2001. [online] Available at: http://www.etela-suomi.fi/english/pdf/kivikari.pdf [Accessed 03.04.2019].

Klemeshev A.P. and Fedorov G.M. (2004). From an isolated exclave – to a «development corridor». Alternative development strategies of the Russian exclave on the Baltic Sea. Kaliningrad: Kaliningrad University.

Klemeshev A.P. and Fedorov G.M. (Eds.) (2015). New forms of international economic cooperation and their role in expanding Russia's participation in international economic cooperation and international integration in the Baltic. Kaliningrad: IKBFU (In Russian).

Klemeshev A.P. and Gutnik A.P. (2006). The Baltic region as a pole of economic integration of the North-West of the Russian Federation and the European Union. Kaliningrad: IKBFU (In Russian).

Kolosov V.A., Suocheng D., Portyakov V.Ya., et al. (2017). The Chinese initiative «The Belt and Road»: A geographical perspective. Geography, environment, sustainability, 10(1), 5-12.

Kolosov V.A., Turovsky R.F. (1997). Modern state borders: new functions in the context of integration and cross-border cooperation. Proceedings of the RAS. Ser. geogr., 5, 106-113 (In Russian).

Kolosov V., Sagan I., Studzińska D., Zotova M. et. al. (2018a). The local border traffic zone experiment as an instrument of cross-border integration: The case of polish-russian borderland – Geographia Polonica, 91(1), 79-86.

Kolosov V., Więckowski M. (2018). Border changes in Central and Eastern Europe: An introduction. Geographia Polonica, 91(1), 5-16. DOI: 10.7163/GPol.0106.

Kolosov V.A., Zotova M.V., Popov F.A., Gritsenko A.A., Sebentsov A.B. (2018a). The post-Soviet borderlands of Russian between the East and the West (an analysis of political discourse). Part 1. Looking West. Polis. Studies, 5, 42-59.

Kolosov V.A., Zotova M.V., Popov F.A., Gritsenko A.A., Sebentsov A.B. (2018b). The post-Soviet borderlands of Russian between the East and the West (an analysis of political discourse). Part 2. Looking East. Polis. Political studies, 5, 57-69.

Korneevets V.S. (2010). International regionalization in the Baltic. St. Petersburg: St. Petersburg University (in Russian).

Kropinova E.G. (2016). Cross-border tourist and recreational regions in the Baltic. Kaliningrad: IKBFU (in Russian).

Lechevalier A. and Wielgohs J. (eds). (2013). Borders and Border Regions in Europe: Changes, Challenges and Chances. Bielefeld: Transcript Verlag.

Local transborder movement Poland Russia (2019). [online] Available at: www.google.com/search?tbm=isch&sa=1&ei=XwQIXYOulq-rmwWClbPAAw&q=local+transborder+movement+poland+russia&oq=local+transborder+movement+poland+russia&gs_l=img.12...212842.240422..244683...0.0..0.344.3391.34j5j0j1......0....1..gws-wiz-img.......0j0i67j0i30j0i19.NmX8Za72u1U#imgrc=0TBhrKbzPKQryM: [Accessed 23.05.2019].

Lowy A., Hood Ph. (2019). The Power of the 2 x 2 Matrix. [online] Available at: www.2x2matrix.com/downloads/swot.pdf [Accessed 23.04.2019].

Mikhaylov A.S. (2014). Geography of international clusters in the Baltic region. Baltic Region, 1, 113-123. DOI: 10.5922/2079-8555-2014-1-10.

Ministry of Foreign Affairs of the Russian Federation (2019). [online] Available at: www.mid.ru/ru/maps/pl/-/asset_publisher/D2CPYayAgyuG/content/id/559654 [Accessed 24.04.2019]. (In Russian).

Ministry of foreign affairs Republik of Poland (2019). [online] Available at: www.msz.gov.pl/en/ [Accessed 24.04.2019].

Palmowski T. (2010). Problems of cross-border cooperation between Poland and the Kaliningrad Oblast of the Russian Federation. Quaestiones Geographicae 29(4). Poznań: Bogucki Wydawnictwo Naukowe, 75-82.

Palmowski T. (2006). Trojmiasto – Kaliningrad: bipolarny Europol? Pomorski przeglad gospodarczy, 5. [online]. Available at: www.ppg.ibngr.pl/autorzy/palmowski-tadeusz?hilite=%27Palmowski%27 [Accessed 03.05.2019].

Perkmann M. (1997). Cross-border cooperation, euroregions', and the governance of crossborder economies. Lancaster Sociology. Working Paper 1, Lancaster University.

Perkmann M. (2003). Cross-Border Regions in Europe: Significance and Drivers of Regional Cross-Border Co-Operation. – European Urban and Regional Studies, 10(2), 153-171. [online] Available at: DOI: 10.1177/0969776403010002004 [Accessed 15.04. 2019].

Poland – Russia PS 2014–2020. Program document (2019). Available at: id.gov39.ru/projects/Программный%20документ%20-%20копия.docx [Accessed 14.04. 2019]. (In Russian).

Raich S. (1995). Grenzüberschreitende und interregionale Zusammenarbeit in einem «Europa der Regionen». Baden-Baden.

Rees N. Interregional Co-operation in the EU and Beyond. European Planning Studies, 5, 385-406.

Russia-EU Cross-Border Cooperation Programs (2019). Ministry of Economic Development of the Russian Federation. [online] Available at: www.ved.gov.ru/interreg_cooperation/cooperation_program/ [Accessed 14.04. 2019]. (In Russian).

Russian National Sub-Committee of the Baltic Sea Interreg Program 2014–2020 (2019). [online] Available at: www.rnsc.leontief-centre.ru/ [Accessed 14.04.2019]. (In Russian).

Schmitt-Egner P., Die «Europäische Kompetenz» von Regionen – ein Paradigma des Transnationalen Regionalismus? – Interregiones 5/1996, 7-56. (in German)

Scott J.W. (1999) European and North American Contexts for Cross-border Regionalism. Regional Studies, 33(7), 605-617.

Sohn C. and Walther O. (2009). The construction of cross-border metropolitan regions in Europe. A comparative perspective. SSRN Electronic Journal. [online]. Available at: www.researchgate.net/publication/272226523_Cross-Border_Metropolitan_Integration_in_Europe_Luxembourg_Basel_and_Geneva (Accessed 21.05.2019).

Studzińska D. (2014). Ruch bezwizowy a rozwój turystyki na pograniczu polsko-rosyjskim. Przegląd Geograficzny, 86(4), 525-540.

Studzieniecki T., Palmowski T. and Korneevets V. (2016). The System of Cross-border Tourism in the Polish-Russian Borderland. Procedia Economics and Finance, 39, 545-552.

The Poland – Russia 2014–2010 Cross-Border Cooperation Program Poland–Russia (2019). [online] Available at: www.plru.eu/ru/pages/11 [Accessed 05.04.2019]. (In Russian).

Van der Velde B. and Martin R. (1997). So many regions, so many borders. A behavioural approach in the analysis of border effects. Paper prepared for the 37th European Congress of the European Regional Science Association. Rome, Italy, Aug. 26-29, Rome.

Received on June 24th, 2019 Accepted on January 21th, 2020

Denis A. Dirin[1,2*], Paul Fryer[3]
[1] Tyumen State University, Tyumen, Russia, Russia
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia
[3] University of Eastern Finland, Joensuu, Finland
***Corresponding author:** d.a.dirin@utmn.ru

THE SAYAN BORDERLANDS: TUVA'S ETHNOCULTURAL LANDSCAPES IN CHANGING NATURAL AND SOCIOCULTURAL ENVIRONMENTS

ABSTRACT. The paper is devoted to ethno-cultural landscapes of the Republic of Tuva. Ethnocultural landscapes (ECLs) are specific socio-environmental systems that developed as a result of the interaction of ethnic groups with their natural and social environments and are in a constant process of transformation. An attempt is made to identify the mechanisms of the formation, functioning and dynamics of ethnocultural landscapes in the specific conditions of the intracontinental cross-border mountain region, as well as to establish the main factors-catalysts of their modern changes. For the first time an attempt is made to delimit and map the ethnocultural landscapes of Tuva. For this, literary sources, statistical data and thematic maps of different times are analyzed using geoinformation methods. The results of 2014-2018 field studies are also used, during which interviews with representatives of different ethno-territorial, gender, age and social groups were taken. It is revealed that the key factors of Tuva's ethnocultural landscape genesis are the natural isolation of its territory; the features of its landscape structure; the role of government; population migrations from other regions and the cultural diffusion provoked by them. 13 ethnocultural landscapes are identified at the regional level. Their modern transformation is determined by the shift of climatic cycles, aridisation, globalisation of sociocultural processes, changes in economic specialisation and ethno-psychological stereotypes.

KEY WORDS: ethnocultural landscapes; borderlands; Tuva; climate change; socio-cultural processes; spatial organisation

CITATION: Denis A. Dirin, Paul Fryer (2020) The Sayan Borderlands: Tuva's Ethnocultural Landscapes In Changing Natural And Sociocultural Environments. Geography, Environment, Sustainability, Vol.13, No 1, p. 29-36
DOI-10.24057/2071-9388-2019-76

INTRODUCTION

Tuva is a border region not only in geopolitical terms, neighbouring Mongolia and near China's restless Xinjiang, but also in natural and cultural senses at the global scale. In natural terms, the Sayan Mountains of Tuva are part of the Eurasian climatic zone and watershed. Their rivers feed the Arctic water basin and form the boundary between the Siberian taiga and steppes and semi-deserts of Inner Asia. Moreover, Tuva is situated at the nexus of three cultures (civilisations): the Christian-Slavonic industrial-agricultural world (European Russia), the Turkic-Finno-Ugrian animistic forest world (Siberia) and the Turkic-Mongol Buddhist-Islamic nomadic world (Inner Asia). At the same time, in Tuva different ethnocultural traditions have been maintained, which has continued to produce specific ethnocultural landscapes.

The term 'ethnocultural landscape' (ECL) is understood as a natural-cultural territorial complex, formed because of the evolutionary interaction of nature with the local ethnic community, practically, semantically and symbolically developing and transforming geographical space according to its spiritual and material needs. An important feature of the ethnocultural landscape is the interpenetration of the ethnic culture into the surrounding space and this space into the ethnic culture. The study of these processes of ethnic groups' interaction with their geographic environments has been very important in geography, ethnology, anthropology, and other fields.

The ethnic history and culture of Tuva has long attracted the attention of researchers, but the aim of this study is to identify patterns in the formation, functioning and development of ethnocultural landscapes in the Republic of Tuva, which has not been undertaken before. This research on the ethnocultural landscapes of Tuva involves analysing the key factors in their formation, establishing the patterns of spatial and functional organisation, and identifying the important causes of and trends in the current situation.

MATERIAL AND METHODS

At the beginning of the 20th century, at the meeting point of geography, cultural studies and ethnography (ethnology), almost simultaneously in Russia (Berg 1915), Germany (Schlüter 1920) and the USA (Sauer 1925) the concept of cultural landscape began to develop. The research focus was the 'cultural landscape' as a natural-social system in which all components (human community, its economy, the natural environment, elements of material and spiritual culture, etc.) inextricably are linked and interdependent; the resulting cultural landscapes themselves are due to this interaction of society and nature. Thus, the main properties of cultural landscapes are their consistency and spatial character (territorial localisation). At the same time, besides a community's purely practical material relations with the natural environment, spiritual relations with cultural landscapes are considered.

Within the framework of this concept, the idea of the ethno-historical essence of the cultural landscape appeared quite quickly (e.g., Lowenthal et al. 1965; Häyrynen 1994; Kalutskov 1998; Yamskov 2003, etc.). Developing out of this idea, an 'ethnocultural approach' to the study of cultural landscapes emerged, which within a short time became one of the most recognised and developed in Russian geographical study. The focus of this approach is the spatial expression of the culture of an ethnic group – an ethnocultural landscape. Viewing the development of an ethnic group through the prism of the ethnocultural landscape is, in fact, borrowing the commonly used landscape indication method from natural science in cultural geography. The most famous proponent of the ethnocultural landscape concept in cultural geography is V.N. Kalutskov. Importantly, his efforts have formed the theoretical and methodological basis of modern ethnocultural landscape research (e.g., Kalutskov, Ivanova and Davydova 1998; Kalutskov 2008; 2011). Additionally, theoretical and applied research in the field of ethnocultural landscapes has been conducted by A.N. Yamskov, T.M. Krasovskaya, V.V. Kuklina, A.V. Lysenko, Zh.F. Degteva, and D.A. Dirin.

Ethnocultural landscape research is profoundly orientated towards field research, resulting in the use of methodologies of related disciplines, such as ethnography, ethno-linguistics, toponymy, folklore studies, etc. Most often, the ethnocultural landscape approach is used in areas with compact populations of ethnic or sub-ethnic groups – especially if they are characterised by a traditional way of life – as well as in multi-ethnic territories. It is not by chance that ethnocultural landscapes are actively examined in the Russian North (Kalutskov 1998b; 2005), the Russian Far North (Klokov, Krasovskaya and Yamskov 2002; Krasovskaya 2012), Sakha-Yakutia (Degteva 2017), south Siberia (Kuklina 2006; Dirin 2008; 2011; 2014) and the Caucasus (Lysenko 2009; Salpagarova 2003).

Outside of Russia, this approach has not yet become widespread. The concept of 'ethnic landscape' as adopted by Anglo-American cultural geography (e.g., Lehr 1990; Noble 1992; Cross 2017) is not identical to the Russian geographical concept of 'ethnocultural landscape'. The ethnic landscape appears as a purely humanitarian-geographical concept – localised ethnic groups in their space with an emphasis on the social characteristics of the community itself, but with largely a complete disregard for nature in the cultural landscape.

Several methods and sources of empirical materials are utilised in this study. To identify the key factors in the ethnocultural development of Tuva, an extensive literature review was conducted on the natural features of the territory and their historical dynamics, settlement history, natural resource management, cultural development, the interaction of indigenous Tuvans with other peoples, and political transformations, especially those influencing the republic's economy and culture. Additionally, an analysis of historical maps was made to follow the dynamics of territorial organisation.

A study of the territorial and functional organisation of Tuva results in the delimitation of boundaries between various ethnocultural landscapes, but also identifies common and specific features in their functioning and connections between them. Maps (cartographic methods) with various themes have also been analysed to delimit the ethnocultural landscapes of Tuva, often used in conjunction with one another: physical maps indicating physiographic boundaries, including watersheds, landscape types, altitudinal belts, climatic zones, etc.; ethnic maps highlighting areas of settlement, sub-ethnic groups, ethnic contact zones; economic maps, including natural resource management systems of various types; religious maps of main confessions, important centres and peripheries; political maps with administrative boundaries and centers); cultural-historical maps showing historical boundaries and places of historical memory; toponymic-linguistic maps identifying language distribution, dialects and marking the geocultural space through geographical names that suggest a relationship to a particular ethnocultural community; mental maps reflecting local perceptions of a territory, as well as the location of ethnocultural boundaries. Geoinformatics (GIS) are used to represent the ethnocultural landscapes cartographically and to model the spatial processes. An ArcGIS research database was created and information recorded in GIS attribute tables is displayed in the form of map layers.

A comprehensive study of modern ethnocultural landscapes cannot be achieved without the use of statistical data. The analysis of regional statistical information was conducted, reflecting the spatial organisation of the population and economy, the sectoral structure of the economy, as well as their dynamics. The toolkit of the Statistica software package was used to perform calculations. To determine historical patterns and trends, a statistical analysis of significant indicators from the perspective of cultural and geographical processes was made to identify changes. The key sources of statistical data are the regional statistical body (http://krasstat.gks.ru/wps/wcm/connect/rosstat_ts/krasstat/ru/statistics/tuvStat/) and the Unified Interdepartmental Information and Statistical System 'State Statistics' (https://fedstat.ru/).

Fieldwork, which involved, in addition to observation, the use of sociological methods (questionnaires, interviews), also was employed. During expeditionary fieldwork in Tuva from 2014 to 2018, modern ways of adaptation by different ethnoterritorial populations to different types of landscapes was observed; the regime of daily and periodic household activities related to the functioning of ethnocultural landscapes. Questionnaires were administered to local communities to identify the main features of the spatial and functional organisation of today's ethnocultural landscapes, as well as to isolate important trends. A total of 382 people took part in the survey, living in 23 settlements (including temporary ones) in the republic. Moreover, 32 interviews with representatives from different ethno-territorial, gender, age and social groups provided additional details on ethnocultural landscape development that were not evident from the results of the survey.

RESULTS AND DISCUSSION

The factors that determined the modern mosaic of ethnocultural landscapes of Tuva are numerous and varied. They include the natural features of the territory and processes of social development. For each historical period, different processes and phenomena had greater influence in ethnocultural developments. However, the key factors that ensured the formation of the characteristics of the varying ethnocultural landscapes that currently exist in Tuva can be identified.

Key factors of formation of ethno-cultural landscapes of Tuva

Geocultural permeability: A system of mountain ranges covers the territory of Tuva, with uplands and large intermontane basins. At the same time, the general regularity of the relief structure is such that from the west and north Tuva is limited by the high mountains of the Altai and the Western and Eastern Sayans, while in the south and the east there are the much lower ranges of the Western and Eastern Tannu-Ola mountains, the Sangheli Highlands, and the Ubsunur Hollow. Thus, the territory of Tuva historically is rather strongly isolated from cultural influences from the north and north-west – Christian-Slavonic industrial-agricultural world (European Russia) and Turkic-Finno-Ugrian animistic forest world (Siberia). At the same time, Tuva is open to influence from the south and the Turkic-Mongol Buddhist-Islamic nomadic cultures of Inner Asia. Due to this natural factor, Tuva is historically and culturally the peripheral part of Inner Asia, not Siberia. The centuries-old close relationship of the Tuvan population with other nations of the Great Steppe predetermined the commonality of their material and spiritual culture, especially in terms of activities, everyday life, folklore, prevailing religious views, common behaviours, etc. The isolation of Tuva as a part of Inner Asia from other cultural worlds, especially the Russian and Chinese, impeded cultural diffusion, while maintaining originality in local cultures; there was a socio-economic lag compared to more open territories. Although important, this isolation was not absolute. The development of transport infrastructure, scientific and technical progress, changing socio-economic conditions, and historical and political processes made the mountain barriers more permeable to cultural diffusion, as well as allowing for migrating ethnocultural communities.

Landscape economy: The economy serves as the foundation for cultural formation and development. It is the economy that provides an economic basis for human existence, as well as the conditions for social stratification, traditions, folklore and much more. Economic activity is also a key factor in the spatial organisation of ethnocultural landscapes, dividing them into functional zones, etc. (Salpagarova, Chomaeva and Uzdenova 2014). In the pre-industrial era, the economy was formed almost exclusively by environmental management. Accordingly, the economy depends entirely on the natural resource potential of the territory. Thus, one of the key factors in the formation and development of the ethnocultural landscapes of Tuva, indirectly through natural resource management, was natural landscapes. The landscape structure of Tuva is dominated by steppe and semi-arid steppe landscapes (about 40% of the total area of the republic) and mountainous taiga forests of different subtypes (about 50% of the territory). The upper relief is occupied by subalpine meadows and woodlands, alpine meadow-tundra landscapes, and mountainous uplands.

The presence of large tracts of steppe landscapes contributed to the spread of nomadic and semi-nomadic animal husbandry in Tuva. Sheep, goats and horses traditionally dominate, though yaks and camels also are bred in some parts of Tuva. Since the nineteenth century cattle breeding has spread. Hunting has always been of utmost importance. In the mountainous taiga landscapes of the north-east of Tuva, perhaps influenced by the Tofalars of southern Siberia, the environmental management system of Tozhu Tuvans, who are hunter-gatherers and reindeer herders, has been formed. The gentle topography of large intermontane basins along with and chernozem and chestnut soils drew Russian settlers at the end of the nineteenth – turn of the twentieth centuries. Agriculture was introduced to Tuva, albeit at a small scale. In general, the high altitudes of the landscape structure determine the complexity and differentiation of the economy (Traditional knowledge 2009).

Migration and innovation diffusion: Migrants, when occupying a new territory, create a new ethnocultural landscape for themselves, which reflects the specifics of the spiritual and material culture of the immigrant community, formed in another territory, and new features that result from adaptation to the new natural conditions and borrowings from the indigenous population. At the same time, the immigrants themselves are carriers of innovations that may be adopted by the local population, influencing the established ethnocultural landscapes.

From the nineteenth century Russians began to settle Tuva, the first being Old Believers seeking isolation and sanctuary from religious persecution by the Russian authorities. According to some sources, isolated settlements of Old Believers appeared already in the eighteenth century (Storozhenko 2004). Old Believers were able to adapt farming to local conditions, cultivating rye, oats, potatoes and some other crops, though playing a minor role in their traditional economy. Their economic basis was cattle breeding and forestry, and later included red deer (Siberian Maral) breeding. The majority of Old Believers settled in the taiga, along the Ka-Khem (Little Yenisei) river. From the late 1830s, Russian goldminers began to move to Tuva, founding prospecting settlements on the Sastyg-Khem, Serlig, and Seskier rivers, while in the 1870s the first Russian trading posts were established in Shagonar and Chaa-Khol. These formed very specific ethnocultural mining and industrial landscapes, new to this territory. The intensive migration of landless Russian peasants to Tuva began after 1885. The first Russian agricultural settlements were founded in Turan, Uyuk, and Tarlyk. However, the peak of migration was reached in the years of the Stolypin reform, which coincided with the establishment of a Russian protectorate over Tuva (Bumbazhay 1999). The first Russian villages emerged in Tandinsky kozhuun ('district', Upper Nikolikoye, Nizhneenikolskoye, Sosnovka, Atamanovka) and in Ka-Khem kozhuun (Fedorovka, Boyarovka, Znamenka, Gryaznukha, etc.). By 1917, about 9 thousand Russians lived in Tuva, which accounted for 15% of the population (Pavlova 2013).

The local population borrowed from the Russians, in addition to certain activities (for example, hay harvesting) and tools (agricultural and hunting equipment), entire branches of natural resource management (for example, fishing and beekeeping). Cultural diffusion is not always associated with migration and innovations can be introduced into local culture as a result of trade or military contacts with other nations, missionary activity, etc. This is perhaps how Buddhism arrived in Tuva, occuping an important place in the spiritual and material culture of Tuvans. Buddhist religious sites (stupas, temples), along with the surviving attributes of shamanism (obo cairns, sacred trees, shamanistic groves) and later Orthodoxy (churches, graveyards, roadside crosses) are not only elements in the functional structuring of space – representing a kind of saced zone – but also important markers of 'belonging' in ethnocultural landscapes.

State administration of territory: The influence of state power on the transformation of existing and the emerging ethnocultural landscapes is very strong. After the incorporation of Tuva into the USSR in 1944, its socialist reorganisation quickly unfolded. The abolition of private ownership of land and the formation of large enterprises (collective and state farms), the sedentarisation of Tuvans, the industrial development of mineral and forest resources, the growth of urban settlements, the persecution of religion and the introduction of the ideology of dialectical materialism all became features of the Soviet era that radically changed the region's ethnocultural landscapes and created new ones – industrial-urban, mining, forestry. The symbolism of space changed. In place of 'archaic' symbols of a religious nature, came the symbols of socialism — red flags, steles, monuments and memorials.

Modern ethnocultural landscapes of Tuva

Ethnocultural landscapes are hierarchical systems. Therefore, they can be considered at several levels: neighbourhood, local, subregional, regional, interregional. To fully characterise the components of a territory's ethnocultural landscapes, above all else it is necessary to delimit them; that is, to determine their position in space and fix their boundaries. In this study, the delimitation and characterisation of Tuva's ethnocultural landscapes at the regional level was carried out. As outlined above, key factors of geocultural differentiation that determine the isolation of ethnocultural landscapes and can be reflected on maps include natural, ethnic, economic, religious, political-administrative, cultural and historical, toponymic-linguistic, and mental factors.

When identifying individual ethnocultural landscapes at the regional level and delineating their boundaries, all the above factors of geocultural differentiation were analysed. Taking into consideration the various mutually overlapping frontiers, as well as centre-peripheral links, 13 ethnocultural landscapes were identified on the territory of Tuva at the regional level (Fig. 1).

1. *The Bai-Taiginsky ECL* is located in the west of Tuva, occupying the eastern macro-slope of the Shapshal ridge and the western part of the Alash plateau within the Bai-Taiga mountain range. Dry stony steppes, larch taiga, and mountain tundra predominate the natural landscape. The population is ethnic Tuvan who specialise in sheep and goat breeding. Hunting and fishing is important. Shamanistic traditions are significant. Most of the settlements are in the Khemchik and Alash river valleys. Lake Kara-Khol is a significant recreational destination, but also is sacred to the local population.

2. *The Mongun-Taiginsky ECL* occupies the southwestern part of the republic, comprising the mountain range of the same name. The most important feature of this territory is the wide distribution of glaciers and ancient glacier landscapes. Within this territory, tributaries of the Mogen-Buren river originate. The majority of the population is Tuvan who embrace a mixture of Buddhist and shamanist beliefs and rituals. A significant part of the ECL belongs to the cluster section of the biosphere reserve «Ubsunur Hollow – Mongun-Taiga». As the territory does not contain a nature reserve, pasturing livestock, especially sheep and sarlyk yaks, is widespread on high mountains of 2000–2500m with valuable fodder grasses, water sources, and the absence of midges in summer (Sat 2016). In recent years, the development of recreational activities has been significant with recreational facilities at Lake Khindiktig-Khol and the mountainous glacial valleys of the Mongun-Taiga massif.

3. *Alash-Khemchiksky ECL*. In the basins of the Alash and Khemchik rivers, the Alash-Khemchiksky ECL occupies a significant part of the Alash Plateau and the Tuva Basin. In the Tuva Basin, dry steppes prevail, while stony steppes and larch forests cover the slopes of individual upland peaks. At present, Buddhist Tuvans dominate the population. Cattle and sheep make use of local resources, while the Khemchik floodplain and floodplain terraces cultivate grain for fodder. In some places, irrigation is practiced. There is a mining industry (LLC Tuvaasbest in Ak-Dovurak) and an industrial-urban type of spatial organisation comprising a local ECL. Some recreational resources exist, the most famous sites are Lake Sut-Khol and the, the water spring («arzhan») Ulug-Dorgun.

4. *The Central Tuva ECL* comprises a large part of the Tuva Basin and the slopes of the surrounding ridges connected through economic ties. Steppe and dry steppe landscapes prevail in the basin and on the mountain slopes – larch forests. Despite the arid climatic conditions, the territory is rather densely populated and developed. The floodplain and floodplain terraces are used for plowing, settlements

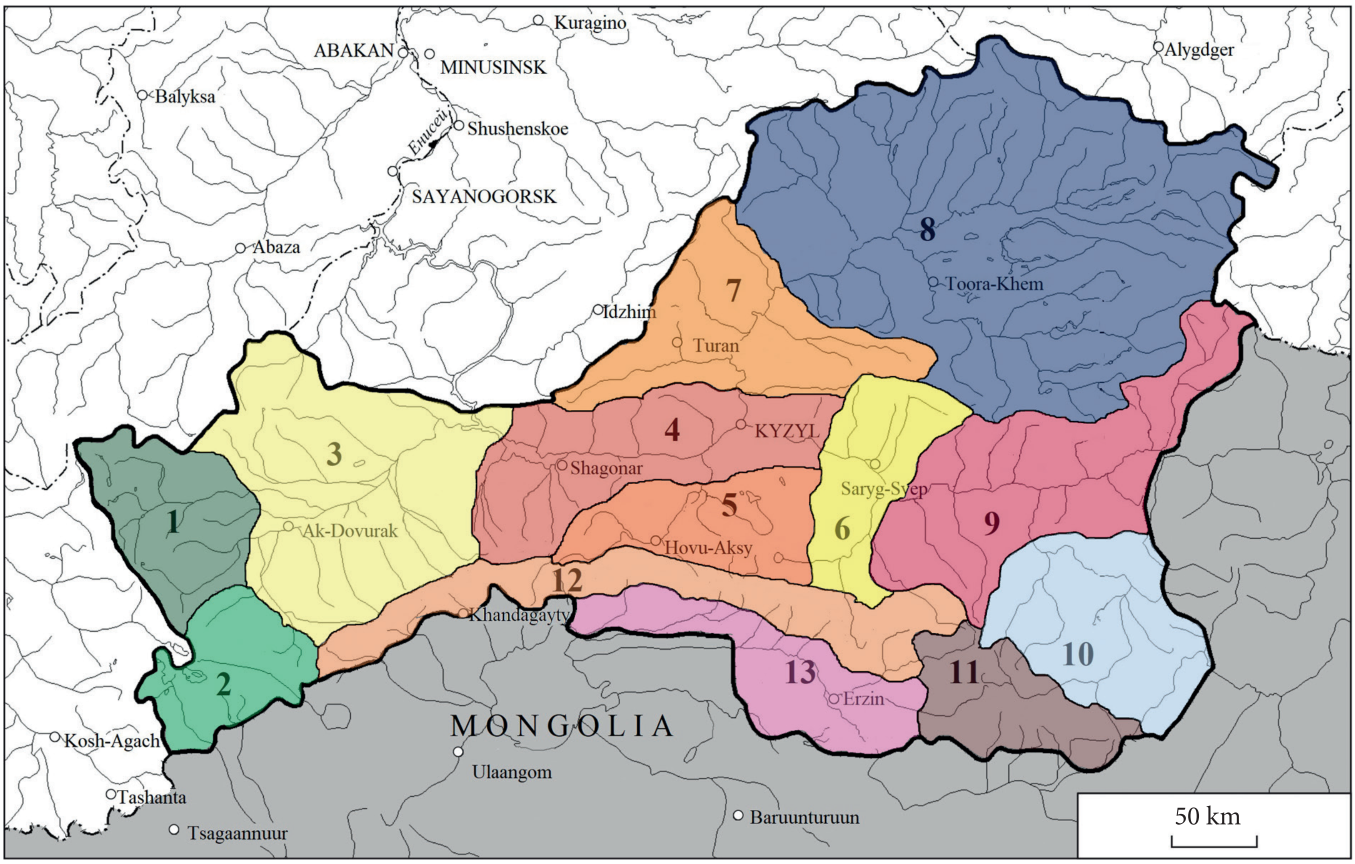


Fig. 1. Ethnocultural landscapes of Tuva (regional level)

**Numbers refer to corresponding individual ethnocultural landscapes in the text.*

and grazing for most of the year. Agriculture is the basis of natural resource management: transhumance animal (sheep and goat) husbandry and farming. On the Sayan-Shushensky Reservoir there is commercial fishing. The population is mostly Tuvan, although there is a small percentage of Russians found mainly in the cities. Religiously, this ECL is clearly associated with Buddhism.

5. *Elegest ECL*. Located in the Elegest river basin, occupying the northern taiga-forested macro slope of the Eastern Tannu-Ola range and the arid steppe southern part of the Tuva Basin is the Elegest ECL. Buddhist Tuvan sheep and goat herders predominate in the population. There are traces of former mining production, for example the derelict Tyvakobalt mine and accompanying infrastructure in Khovu-Aksy. The territory of this ECL quite actively is used for recreation by residents of Kyzyl and other surrounding settlements, most significantly the Chagytai and Khadyn lakes.

6. *The Soi-Burensky ECL* contains the middle course of the Little Yenisei River and its tributaries, the largest of which are the Soi and Buren rivers. Most of this territory is occupied by forested medium-sized mountains. The western part reaches the steppe and forest-steppe landscapes of the Tuva Basin. The peculiarity of this ECL is that most of its settlements were founded by Russian immigrants in the nineteenth century. Until the mid-twentieth century, this territory was ethnoculturally dominated by Russians. However, today Tuvans are the majority population with Russians comprising about 25%. Nevertheless, alongside Tuvan culture and Buddhist religious sites, elements of the Russian presence, such as Orthodox churches, Russian architecture, and some toponyms, are preserved in this ethnic contact zone. Natural resource use is focussed around cattle breeding, horticulture, and forestry.

7. *Turan-Uyuksky ECL*. Located in the north of Tuva on the border with Krasnoyarsk Region is the Turan-Uyuksky ECL. It occupies the steppe of the Turan-Uyuk Hollow, as well as the surrounding larch and mixed cedar-larch forested slopes of the Kurtushibinsky and Uyuksky ranges and the western spurs of the Academician Obruchev Ridge. This ECL is a relic of the Russian development of Tuva, as most of its settlements were founded by Russian immigrants who practiced atypical natural resource management. Today the Turan and Uyuk steppes are almost completely plowed up; farming is an important economic sector, along with cattle breeding, though there are several small gold mines locally. Tuvans currently outnumber Russians, but the latter's share is significant. Many 'Russian' elements (toponyms, religious sites, architectural styles, etc.) are preserved in this ECL. The mountain slopes are used for logging, collecting pine nuts and berries, as well as for local recreation. In particular, winter recreational centres and ski slopes are found here (Kuular and Mongush 2010). Also in this ECL are the Erbek Nature Reserve and Taiga Nature Park.

8. *The Todzhinsky ECL* is the largest regional ethnocultural landscape of Tuva, occupying vast spaces of the Todzha Basin and the surrounding medium-size mountain ridges – from the north and the spurs of the Western Sayan, from the north-east of the Great Sayan Ridge, and from the south of the Academician Obruchev Ridge. In this territory, taiga landscapes prevail and the total forest cover is 67%. The upper relief mountain meadows and tundra, while in the basin there are steppes and floodplain meadows. The population is majority Tozhu Tuvan, which are close economically to the Evenki and Tofalars of the nearby mountains of southern Siberia. This ECL is distinguished by a high level of preservation of the traditional way of life. The established ethnocultural system is based on ancestral forms of environmental management, social organisation and culture in general, as well as the maintenance of customs and beliefs. Economically, locals specialise in mountain taiga reindeer breeding, hunting, gathering wild plants and mushrooms, and fishing. Other animal husbandry (dairy and meat production, horse breeding) is of secondary importance. Central functions are carried out in the villages of Adyr-Kezhig, Ii, and Toora-Khem. Together with their peripheries, which act as industrial and grazing lands (including rather remote ones), they represent the ECL at the local level. Institutionally Tozhu Tuvans are organised into the clan communities 'Ulug-Dag', 'Odugen', and 'Khimsara', which guide traditional nature management. There are forestry enterprises engaged in logging and primary wood-processing. The ECL is popular with tourists and contains the strict nature reserve Azas.

9. *Little Yenisei ECL*. Occupying an area in the middle of the river basin of the same name, the Little Yenisei ECL is bounded by the Academician Obruchev Ridge to the north, and to the south by the Khorumnug-Taiga Ridge. Most of the territory is mountainous taiga covered medium-size mountains. Tundra highlands occupy the high peaks. Development of this territory followed the Little Yenisei River. Here, on the terraces and alluvial deposits of its tributaries, there are 'Chasovennye' Old Believer settlements, who continue to dominate the local population. The most significant centers of settlement within the ECL are Erzhey, Ust-Nashp, Unzhey, Shivey, and Sizim, where Old Believers have not changed their way of life significantly. Upstream are more closed small settlements or hermitages. The main occupation of the population in the winter is hunting, while in the summer – mixed agriculture of cattle breeding, beekeeping, horticulture and fishing. Tourism bases have appeared.

10. *Tere-Kholsky ECL*. In the southeastern part of Tuva, in the basins of the Kargy and Balyktykh-Khem rivers, is the Tere-Kholsky ECL. Lake Tere-Khol is located in the central area, while mountainous taiga landscapes predominate. In the upper reaches of the ridges there are upland landscapes. The population – Tozhu Tuvans – through the 'Emi' clan community maintains the traditional way of life of hunters and reindeer herders of the taiga and shamanism. Logging for local needs is conducted and there is a gold mine co-operative 'Oina'. This ECL has extremely poor transport accessibility.

11. *Sangilensky ECL* is located within the highlands of southeast Tuva. On the northern slopes grow cedar-larch forests; on the southern slopes - steppes, reaching an altitude of 1800-2000m with mountain meadows and tundra. Tuvans live in this extremely sparsely populated ECL. The main branch of the economy is traditional nomadic sheep herding. On the territory of the highlands there is there is the isolated «Ular» protected cluster in the strict nature reserve «Ubsunurskaya Hollow».

12. *The Tannu-Olinsky ECL* occupies the southern steppe macro slope of the Eastern and Western Tannu-Ola ranges. The western part of the territory opens up to Mongolia and visually does not differ from it. The population is Tuvan Buddhist engaging in sheep rearing and hunting. The territory is known for its mineral springs – the radon waters of Ulaatai and saline waters of Torgalyg and Khuregecha. Within the ECL there is the protected area cluster «Aryskannyg» of the strict nature reserve «Ubsunurskaya Hollow».

13. *Ubsunursky ECL*. In the Russian part of the Ubsunur Basin on the border with Mongolia, the Ubsunursky ECL is very sparsely populated. Traditionally, it was used by Tuvans to pasture their sheep. However, at present, a significant part of it is included in the Ubsunurskaya Hollow strict nature reserve, which has UNESCO World Heritage Site status. Within the ECL there are significant reserves of rock salt.

Modern Dynamics of Tuvinian Ethnocultural Landscapes

Each ethnocultural landscape has developed uniquely, however today it is possible to discuss some common factors and development patterns in Tuva beginning from the post-Soviet period – from 1991 to the present.

Natural factors: Nature's dynamics change the conditions and effectiveness of economic activity. It requires adaptation, which can be expressed in new methods and technologies of environmental management, or changes in the specialisation or mode of activity.

The most dynamic component of the natural environment is climate. In ethnocultural landscapes, climatic changes are manifested indirectly, through environmental-economic grounds and their characteristics such as biological productivity, resistance, species composition of biota, etc. The warming of our climate is the main change in the natural environment in recent decades. Since 1990, the average January temperature in different parts of the Altai-Sayan highlands has increased by 1.5-4.0°C, and in July by 0.5-2.2°C. At the same time, especially meaningful changes are typical in the intermontane basins – Todzhinskaya, Tuvinskaya, Ubsunurskaya – of Tuva with a severe continental climate (Izmenenie klimata 2011). The area of glaciation and ice mass in the mountain glaciers of Tuva has decreased sharply (Chistyakov, Ganyushkin and Kurochkin 2015). Due to the more intensive melting of glaciers, some rivers drastically have increased runoff, while others have reduced in those basins where there are almost no glaciers. In these processes, the potential for conflict between land users is set, since the availability of water sources is the most important condition for the traditional management of livestock. The degradation of permafrost is activating geomorphological processes, primarily erosion, which affects the territorial infrastructure and the everyday activities of the population, changing the appearance of ethnocultural landscapes. Numerous thermokarst lakes are being formed, which in some areas is the basis for the development of a new branch of economic specialisation – fish farming. Freezing processes are activated in winter, limiting the functioning of the transport system, etc. Weather events have become very unstable: thaws are often replaced by frosts and droughts recur with greater frequency. The transformation of the landscape structure is manifested in the gradual transformation of its species composition, replacing some ecosystems with others and changing the relationships between types (Izmenenie klimata 2011). Many respondents in forested areas of Tuva noted one of the serious problems is the regular scorching of hayfields in the summer due to the absence of rains, which decreases winter fodder, forcing farmers to buy hay that makes livestock farming (especially cattle breeding) unprofitable and decreases herd size.

In internal regions of the territory under consideration (especially in the intermontane basins of Tuva), aridisation of landscapes is even more pronounced. Desertification is occurring, leading to a decrease in the biological productivity of environmental complexes and a decrease in their resistance to anthropogenic impact, pasture degradation, and a reduction in livestock. At the same time, the reduction in the thickness of snow cover even in the forest-steppe medium-size mountains is permitting year-round grazing. Accordingly, over a significant part of the territory under consideration, grasslands gradually are being converted into pastures. Frequent summer droughts and dry winds lead to a reduction in the cultivated area and a degradation of agriculture, reduced mainly to fodder production. In general, the unpredictability of climate processes is creating great risks in economic activity, reflected in the everyday culture, behaviour, and perceptions of the population. According to the results of the survey, the rapidly changing natural environment is one of the most important reasons for people to move out of mountain villages and/ or change traditional activities to others that are less dependent on nature.

Socio-economic factors: The collapse of the Soviet Union at the end of the twentieth century provoked global geopolitical, socio-economic and geocultural transformations around the world. As Tuva represents a periphery in the borderlands, these processes had particularly catastrophic consequences. Most large enterprises have ceased to exist, while the abolition of collective and state farms was promoted by post-Soviet legislative changes (in particular, the return of the right of private enterprise) and the land reform, which permitted private land rental or ownership. In addition, the distribution of land and property shares among former employees of collective farms was of great importance. As a result, many landowners appeared who did not have the material means to conduct a profitable business. Gradually, a return to extensive, but low-cost and inherently traditional forms of enterprise began. The owners of land shares and small herds of cattle began to unite into public (often family-tribal) brigades for joint economic activities. Also, under market conditions and the breakdown of the USSR's internal economy, numerous industrial enterprises went bankrupt, including such giants as the Tuva-Cobalt industrial complex in Khovu-Aksy village. To a large extent, the bankruptcy of basic enterprises and the degradation of the social infrastructure contributed to the outflow of the ethnic Russian population from the republic. Accordingly, many ethnocultural landscapes transformed that 30-40 years ago were 'Russian'.

Additionally, in Tuva new forms of environmental management, which were previously not present, are appearing. For example, a tourism and recreational sector has begun to develop. Numerous fishing, ecological and ethnographic tours are being promoted on the tourist portal of the republic. A part of the Great Sayan Ring branded route passes through the territory of Tuva. While the flow of tourists is insignificant now, its projected increase will make it one of the leading factors in the transformation of traditional ethnocultural landscapes. The development of tourism, in addition to improvements in the socio-economic conditions of the region, is greatly changing the outlook of the local population. Material and mental cultural diffusions gradually are transforming the traditional consciousness of the highlanders. This is manifested in the growth of materialistic and consumer attitudes towards the natural environment and the penetration of elements of popular culture into traditional ethnocultural landscapes (Dirin and Golyadkina 2016).

Geopolitical and geoeconomic shifts in the global and macro-regional order also have a significant transforming effect on the ethnocultural landscapes. After the collapse of the USSR, China became remarkably strong both economically and geopolitically. It should not be forgotten that until 1914 Tuva was a part of China, which until 1944 did not recognise its separation. Therefore, the growing power of China again can be used to incorporate Tuva into its orbit of influence. Vectors of interstate co-operation between Russia, China and Mongolia (the latter performing a buffer function between Russian Tuva and Chinese Xinjiang) are fully capable of defining the role of this region as a territory of active cross-border co-operation or as a border outpost on the edge of the country. Tuva falls within the scope of globally-important socio-economic development projects. Of note is the proposal to construct the Kuragino – Kyzyl – Elegest railway with a further extension to Urumqi via western Mongolia. The implementation of this project would allow linking 'two mainland transport bridges' – the Trans-Siberian Railway and the so-called Eurasian Railway (Shanghai – Beijing – Urumqi – Almaty – Osh and further through Uzbekistan, Turkmenistan, Iran, Turkey, Oman and Qatar, and via another branch through Moscow to Europe). This meets China's development strategy for its western territories, as well as supports the Great Silk Road project. Improving the transport accessibility of the territory of Tuva and its inclusion in the system of trade flows on a global scale can significantly alter the geocultural image of the republic.

CONCLUSIONS

Modern ethnocultural landscapes of Tuva are embodied in space as a result of the ongoing process of ethnocultural development, due to the many interrelated factors of a natural, socio-economic and socio-psychological nature, the combination of which is unique for each specific ethnic area. Their study faces many difficulties, as it is necessary to analyse the effect of many natural and social factors that are of mutual influence. Yet ethnocultural landscape research is necessary not only to preserve ethnocultural diversity and the heritage of the material and spiritual culture of ethnic groups, but also to provide an opportunity to optimise the management of a territory's development, minimising the potential for conflict between land users who are representatives of different ethnocultural groups.

The findings of this study suggest several conclusions. The modern ethnocultural landscapes of Tuva were formed as a result of a long process of ethnocultural development. The key factors determining their specificity were the natural isolation of the territory, peculiarities of the landscape structure of the territory, which predetermined the environmental management specialisations of the main ethnic groups, the state administration of the territory's development, and the contributions by immigrants and the accompanying cultural diffusion. Currently, most of Tuva's ethnocultural landscapes are characterised by traditional cultural resistance and the low susceptibility of local communities to cultural diffusion.

The spatial differentiation of Tuva's ethnocultural landscapes is determined mainly by the ethnic composition of the population, the structure of the natural landscapes, and the specificities of environmental management. However, the main elements of the socio-economic infrastructure of the territory, which structure the geocultural space, are becoming more and more important. At present, 13 regional ethnocultural landscapes function in Tuva. Among them are that of the Tozhu Tuvan traditional environmental management involving hunting and taiga reindeer herding, the Russian Old Believer ECL that is dominated by forestry and cattle breeding, an agrarian-industrial ECL of Tuvans with a significant share of the Russian population and an ECL encompassing the traditional environmental management of Tuvans (sheep, horse, camel, and yak breeding).

Currently, the ethnocultural landscapes of Tuva are being transformed through climate change (aridisation) and socio-cultural change from globalisation. In all likelihood, in future Tuva's transformation will include space-time displacements of economic activity, change in the specialisation of economic activity, and change in the region's ethno-psychological makeup that determines the transformation of spatial behaviour models.

ACKNOWLEDGEMENTS

This research was supported by the Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University; by RFBR grant (project No.17-33-01159 «Delimitation of the geo-cultural space and cultural-geographical zoning of the Altai-Sayan mountain region»); by Academy of Finland mobility grant No.325773 «Comparative ethno-cultural landscapes in post-Soviet borderlands».

REFERENCES

Berg L.S. (1915). Subject and tasks of geography. Proceedings of the Russian Geographical Society, 51(9), 463-475. (in Russian).

Bumbazhai A. K.-O. (1915). Russia and Tuva: the problem of formation of relations in the late XIX – early XX centuries. Moscow. (in Russian)

Chistiakov K.V., Ganiushkin D.A. and Kurochkin Iu.N. (2015). Current state and dynamics of Nival-glacial systems of the Mongun-Taiga and Tavan-Bogdo-Ola massifs. Ice and Snow, 1(129), 49-60. (in Russian).

Cross J.A. (2017). Ethnic Landscapes of America. Springer International Publishing. DOI: 10.1007/978-3-319-54009-2.

Degteva Zh.F. (2017). Spatial organisation of the ethnocultural landscapes of Yakutia. Irkutsk: Institut geografii im. V.B. Sochavy SO RAN. (in Russian).

Dirin D.A. (2008). Ethnoecological problems of Mountainous Altai. Geography and nature use of Siberia, 10, 43-71. (in Russian).

Dirin D.A. (2011). Factors in the cultural genesis and formation of Altai's cultural landscapes. Bulletin of Altai State University, 3/2(71), 114-119. (in Russian).

Dirin D.A. (2014). Ethnocultural landscapes of Altai: peculiarities of the formation, spatial organisation and development. Questions of Geography, 138, 327-345. (in Russian).

Dirin D.A. and Goliadkina E.I. (2016). The Role of tourism in the transformation of the socio-economic space of the Republic of Altai. Sociology in the modern world: science, education, creation, 8(1), 64-68. (in Russian).

Häyrynen M. (1994). National landscapes and their making in Finland. Topos, 3, 6-15.

Izmenenie klimata (2011). Climate change and its impact on ecosystems, population and economy of the Russian part of the Altai-Sayan Ecoregion: assessment report. Moscow: Vsemirnyi fond dikoi prirody (WWF Rossii). (in Russian).

Kalutskov V.N. (1998). Ethnocultural landscape studies and the concept of cultural landscape. In: Cultural landscape: questions of theory and methodological research. Moscow–Smolensk: Izdatel'stvo SGU, 6-13. (in Russian).

Kalutskov V.N., Ivanova A.A. and Davydova Iu.A. (1998). The cultural landscape of the Russian North: Pinezhye, Pomorye. Moscow: Izdatel'stvo FBMK. (in Russian).

Kalutskov V.N. (2005). Russian North: regional cultural landscape and its age-old cycles. Proceedings of the Russian Geographical Society, 137(5), 43-56. (in Russian).

Kalutskov V.N. (2008). Landscape in cultural geography. Moscow: Novyi khronograf. (in Russian).

Kalutskov V.N. (2011). Ethnocultural landscape studies. Moscow: Geograficheskii Fakul'tet MGU. (in Russian).

Klokov K.B., Krasovskaia T.M. and Yamskov A.N. (2002). Ethno-cultural aspects of nature management in the Arctic region of Russia. Geography and Natural Resources, 4, 23-29. (in Russian).

Krasovskaia T.M. (2012). Ethnocultural conflicts of nature use management in the Russian Arctic. Bulletin of the State Polar Academy, 2(15), 51-54. (in Russian).

Kuklina V.V. (2006). Local communities of southern Siberia in a multi-ethnic environment: cultural and geographical cross-section. Novosibirsk: Izdatel'stvo Sibirskogo otdeleniia RAN. (in Russian).

Kuular Kh.B. and Mongush G.R. (2010). Perturbing factors affecting the forests of the Uyuk Ridge. Contemporary high technologies, 9, 130-131. (in Russian).

Lehr, J.C. (1990). Preservation of the Ethnic Landscape in Western Canada. Prairie Forum, 15(2), 263-276.

Lowenthal, D. and Prince, H.C. (1965) English landscape tastes. Geographical Review, Vol. 55, 186-222.

Lysenko A.V. (2009). Cultural landscapes of the North Caucasus: structure, features of formation and development trends. Stavropol': Stavropol'skii gosudarstvennyi universitet. (in Russian).

Noble A.G. (1992). To Build in a New Land: Ethnic Landscapes in North America. Baltimore: Johns Hopkins University Press.

Pavlova A.V. (2013). Russian immigrants as a cause of socio-economic and migration changes in Tuva (Uryankhai region) in the early twentieth century. Bulletin of Irkutsk State University. Series 'Political Science and Religion Studies', 1(10), 147-151. (in Russian).

Salpagarova S.I. (2003). Formation of the ethnic and cultural landscape of Karachai (XIX – early XX century)]: Stavropol': Stavropol'skii gosudarstvennyi universitet. (in Russian).

Salpagarova S.I., Chomaeva M.N. and Uzdenova Kh.I. (2014). Structure and factors in the formation of ethnocultural landscapes of mountain territories. Natural and technical sciences, 9/10(77), 130-132. (in Russian).

Sat A.M. (2016). Traditional environmental use management of the south-western part of Tuva (with the example of the Mongun-taiga region). Bulletin of TuvSU, 2, 66-74. (in Russian)

Sauer C. (1925). The Morphology of Landscape. University of California Publications in Geography, 22, 19-53.

Shlüter O. (1920). Geography in relation to the natural sciences and humanities. Geographical indicators, 21(7/8), 145-152. (in German).

Storozhenko A.A. (2004). Old Believers of Tuva in the second half of the XIX - first quarter of the XX centuries. Kyzyl. (in Russian).

Traditsionnye znaniia (2009). Traditional knowledge of indigenous peoples of the Altai-Sayan Ecoregion in the field of environmental management: information and methodological guide. Barnaul: Izdatel'stvo Azbuka. (in Russian).

Yamskov A.N. (2003). Ethnoecological studies of culture and the concept of cultural landscape. in: Cultural landscape: theoretical and regional research. Moscow: Izdatel'stvo MGU, 62-77. (in Russian).

Received on July 1st, 2019 Accepted on January 21st, 2020

Alexander G. Druzhinin[1,2*], Tatyana Yu. Kuznetsova[2], Andrey S. Mikhaylov[2,3]
[1] Southern Federal University, Pushkinskaya 160 Str., Rostov-on-Don, 344006, Russian Federation
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia
[3] Saint Petersburg Electrotechnical University «LETI», St. Petersburg, Russia
***Corresponding author:** alexdru9@mail.ru

COASTAL ZONES OF MODERN RUSSIA: DELIMITATION, PARAMETRIZATION, IDENTIFICATION OF DETERMINANTS AND VECTORS OF EURASIAN DYNAMICS

ABSTRACT. The resource potential of the oceans has historically had a fundamental impact on the development and spatial organization of mankind. The role of the «marine factor» in economic activity and the formation of settlement systems has increased even more in the modern period, including in Russia, which has long sea coasts that fulfill the country's most important transport, communication, economic and resource, residential and military infrastructure functions. Since the mid-2000s, in Russia there has been a steady increase in foreign trade activity and the marine economy. The author summarizes and develops theoretical concepts created in Russian science about the functions, boundaries, structure of coastal zones as special geographical areas. Based on GIS analysis and the study of a vast array of demographic and economic statistics, the coastal zone of post-Soviet Russia was delimited. Particular attention is paid to the innovative potential of coastal zones, the features of its localization and formation. It is shown that coastal zones and large cities act as a significant environment for building cross-border interactions in the scientific and innovative sphere. The author argues that a further «shift to the sea» of economic activity and the population of Russia is inevitable. It is provoked by geo-economic and geopolitical changes in modern Eurasia which include the processes of integration and disintegration in the Baltic, Black Sea basin, and the Caspian region, the intensification of geopolitical rivalry in the Arctic, and implementation of the Chinese initiative «One Belt - one Road». However, the development of the country's coastal zones will be unstable, not universal and will be accompanied by a further concentration of socio-economic potential in the few leading coastal centers - St. Petersburg, Rostov-on-Don, Sochi, Vladivostok, Kaliningrad, Makhachkala, etc.

KEY WORDS: coastal zone, marine economy, socio-economic development, innovation, geoinformation analysis, Russia, Eurasia

CITATION: Alexander G. Druzhinin, Tatyana Yu. Kuznetsova, Andrey S. Mikhaylov (2020) Coastal Zones Of Modern Russia: Delimitation, Parametrization, Identification Of Determinants And Vectors Of Eurasian Dynamics. Geography, Environment, Sustainability, Vol.13, No 1, p. 37-45
DOI-10.24057/2071-9388-2019-81

INTRODUCTION

Our Planet, primarily, is the vast, prevailing oceanic and sea areas – 71% of the entire earth's surface (Slevich 1988), encircling dispersedly localized disseminations of land massifs, which, in fact, are archipelagoes of «islands» of various configurations and sizes. An extremely long «junction» of land and sea (the total coastline of top-ten countries by this indicator is more than 460 thousand kilometers[1]) appears in this context as the most important border, contact space on a planetary scale of a with a complex of specific factors and characteristics (including for men, their settlement, economic activity) and, at the same time, a key component of the territorial organization of society.

Coastal location has a universal impact on the spatial organization of economic activity, the localization of infrastructure and the distribution of settlement systems. The marine factor and the tendencies to concentration of the economy and the population on the sea and ocean coasts determine coastal zones as an acute issue for research. Scholars estimate that the worldwide population share living within the boundaries of a coastal zone is around 50-70% (Amos et al. 2013; Cetin et al. 2008; Cracknell 1999; El-Sabh et al. 1998; Kurt 2016; Pak and Majd 2011; Pernetta and Elder 1992; Turner et al. 1996). With that, great variations exist with regard to particular countries under study. In Australia, 83-85% of the population reside within 50 km of the coast (Jacobson et al. 2014; Lyth et al. 2005; Wang et al. 2011; Wescott 2009). Up to 75% of the total population of Mozambique (Ngoile et al. 1993) and 70% of Thailand (Tookwinas 1999) concentrate in the coastal areas. The eastern part of coastal China (Hindrichsen 1998; Wang et al. 2011) as well as the coastal zone of Indonesia (Siry 2007) account for 60% share of population. Approximately 55% of the population of Lebanon (Antipolis 2001) and 53% of population of the United States (Bulleri and Chapman 2010; Crowel et al., 2007; Lam et al. 2009) live in the coastal plains. Nearly 40% share of the total national population is found in Croatia and other Mediterranean region states (Bowen et al. 2006). A study held in the scope of Europe suggests that an average population share of coastal regions is 42% (Mikhaylov et al. 2018).

Despite of the coastal zone covering under 10% of the total land surface, it generates as much as 25% of global primary production and 43% of the total value of global ecosystem services (Costanza et al. 1997; Turner et al. 1996). In the United States, coastal counties comprise of 42% of the total employment and are among the nation's fastest growing (Beatley et al. 2002; Lam et al. 2009). In Europe, 43% of gross

[1] The top-20 countries of the world with the longest coastline URL: https://geographyofrussia.com/20-stran-mira-s-samoj-protyazhennoj-beregovoj-liniej

regional product (GRP) in purchasing power parity (PPP) is concentrated around coastal regions with the highest GRP (PPP) per sq.km – 2.3 million euro, adding up to 18 trillion Euro (Mikhaylov et al. 2018).

The specificity of coastal areas is fully manifested in Russia. The whole spatial dynamics of this country (originally «intracontinental»; Savitskiy 1997, it is perceived primarily as an «ocean of land»; Ilyin 1934, and for the last three centuries serving as one of the leading maritime powers; Druzhinin 2016a) is historically very closely associated with an c «incrementation» and development of national coasts (Druzhinin 2017). Their geopolitical significance was fully appreciated already a century ago (Semenov Tyan Shanskiy 1915), and the natural-ecological and socio-economic specifics began to be steadily comprehend in the Russian scientific discourse since the early 1970s (in the wake of the intensive development of the geography of the World Ocean at that perio; Salnikov 1984), being conceptualized in such invariant categories as «waterside zone» (Anikeev 2012; Aybulatov 1989), «littoral zone»(Fadeev 1998), «seaside territory» (Gogoberidze 2008; Makhnovsky 2014), «marine coast» (Arzamastsev 2009), «sea-land contact zone» (Dergachev 1980) and, finally, «coastal zone» (Bondarenko 1981). Nowadays, in the context of geo-economically and geopolitically motivated significant increase in the role of the «sea factor» for the Russian Federation and its spatial development (Druzhinin 2016a; Druzhinin 2019) the corresponding problem (finding an increasingly pronounced socio-geographical emphasis) again logically goes to the research forefront (Druzhinin 2016c; Fedorov et al. 2017). The purpose of the proposed article is the delimitation of Russian coastal zones based on GIS technologies, assessment of their positioning and «weight» in the socio-geographical structure of the country, as well as identification of trends in the socio-economic and demographic dynamics of coastal zones (in the macro-and meso-levels) in the geo-economic and geopolitical context of Eurasia.

MATERIAL AND METHODS

Contemplating over the «coastal zone of Russia» as a socio-geographical category and its delimitation are associated with a number of conceptual points. The first is the inherent in the national scientific tradition considering «zone» as «one of the types of geographic taxon regions» (Gorkin 2013) and, accordingly, a specific territory «characterized by the presence and intensity of the phenomenon» (Alaev 1977). In the case of the actual seaside zone, the specificity of the territorial structures related to it (in the dual unity of their functionality and spatial localization) is predetermined by the boundary («joint») of the sea and the land, its natural-ecological, resource and other capabilities and properties.

The differences between estimated values within and between countries occur both due to objective factors – the established national settlement system, and subjective ones – the date of data collection, the source of data, the delimitation principles of the coastal zone. The first attempts to delimit coastal zones were made back in the 1950s by the Polish geographer I. Staszewski. His approaches were further developed by a Russian scholar V.V. Poksishevsky and receiving a complex assessment in the cycle of publications by L.A. Bezrukov. The latter gave a quantitative assessment of the population distribution of Russia by the degree of remoteness from the sea and following the corresponding spatio-temporal trend. For example, a research design on the Pacific islands set the 1.5 km limit of width for defining the coastal zone (Mimura et al. 2007). A study of Pak and Farajzadeh (2007) on coastal management of Iran has considered 10 km bandwidth coastal zone of the Caspian Sea. Indian scholars focus on a distance of 50 km from the coast (Qasim et al. 1988; Susanta 2013). The exponential increase in population share within a limited scope of the coastal area places a particular importance to the research period. The «coastal sprawl» (Beach 2002) that happens over the past few decades suggests a rapid change of the coastal environment and the community lifestyle. OECD (2003) recommendations state that boundaries of a coastal zone are mobile and should be defined with respect to economic and social characteristics associated with objectives of coastal management. In solidarity with this approach, we note that in Russian science, ideas about the «depth» (width) of the coastal zones (remoteness from the sea of their «intracontinental» border) vary over a very wide range (from 50 to 200 km; Salnikov 1984; Arzamastsev and Sorokin 2008) and are generally equivalent to the existing spheres «intracontinental» socio-economic dominance of coastal cities (or the boundaries of coastal urban agglomerations).

Being of predominantly a «strip-like» contour, the maritime zones appear, at the same time, as special *limological (boundary) structures* («small spaces that separate large spaces»; Rodoman 1999), and their combination is a borderland invariant, result of one of the many types of boundaries, being a «fundamental social phenomenon», as noted by the leading Russian limologologist V.A. Kolosov (2018).

The coastal zones (and due to the continually discrete geographical existence in practice, it is the actual *multiplicity of the coastal zones* that takes place) are confined to the land-sea contact strip. Nevertheless, not every sea (oceanic) coast (especially in the Russian Arctic and the Far East) can be identified precisely as a «coastal zone», which is primarily a socio-geographical phenomenon, particularly being settled and economically developed as a result of involvement in specific (associated with the use of the «sea factor» in the economy, geopolitics, development of the residential environment) communications, relationships, processes. It is the latter that predetermine universal (but not ubiquitous) manifestations of the economy and the population talasaatatractiveness (sustainably emphasized by representatives of both Russian and foreign science; Pokshishevskiy 1982; Druzhinin 2017; Green 2009; Small and Nicholls 2003; Pak and Majd 2011), as well as the various institutional, economic, socio-cultural, spatial-planning and other effects associated with it.

Among them, the structural and functional specificity of the coastal zones (concentrating the main activity on the development of the resource potential of the World Ocean; Slevich 1988; Zalogin 1984; Pokshishevskiy 1982) is of paramount importance: the «biased» sectoral structure of the economy in favor of those directly connected with the sea, its resource capabilities and the transport and communicative properties of the types and spheres of management (sea ports, shipbuilding, extraction of marine bioresources, «seaside» recreation, etc.). This «bias» manifests itself both on the scale of the country. This enables us to identify «coastal regions» with the allocation of their two typological varieties: territories characterized by developed coastal zones (maritime) and other coastal areas (coastal) with the minimum influence of the sea factor on economic and residential patterns (Fig. 1).

It is at the municipal (local) level that the position-localization and economic characteristics of the coastal zones are organically combined with the organization of the space, with the culture of the marine economic activity, with a special «maritime» identity. This makes it possible, in particular, by identifying a very numerous grouping of coastal cities, coastal municipalities and coastal regions (prevailing in southern Russia and, even more in the north-west of the country, in the Baltic region – figure 2) in Russian conditions, to isolate in their aggregate the most important typological invariant – the «thalasso-centered» (Druzhinin 2016d) or «sea-oriented» (Druzhinin 2019) territories.

Fig. 1. «Maritime component» of the regional structure of the Russian Federation

Fig. 2. Coastal municipalities in the Russian sector of the Baltic Sea

The functionality of the coastal zones corresponds with their *localization*, confined directly to the sea coast (generating properties of borderland, interworlds, contrasting environment and requiring special geoadaptive approaches when solving economic, infrastructure, residential, environmental, defense and other tasks), as well as with *configuration features* predetermined physiographic (winding of the coastline, orography, etc.) and social conditions (economic potential and profile of coastal centers, transportation network scheme, delivery technology of passengers and cargo, etc.). The «sea orientation» is also projected on *the structural and spatial uniqueness* of the coastal zones: being a basic component of the territorial organization of society, they, in turn, represent their own spatial framework – the nodal centers of marine economic activity localized on the coast, as well as the intermodal communications ensuring their functioning (including the marine routes). The development of this framework (as well as in general), the positive dynamics of the coastal zones at the same time is directly predetermined by the marine (economic and geopolitical) interests of the country, the extent and scale of their implementation.

RESULTS

For modern Russia (located at a complex geo-economic and geopolitical «crossroads»), the role of the World Ocean, its resource potential (and, above all, communication, resource and military-strategic capabilities) is increasingly acute. Against this background, the significance of the coastal zone of the country rises, its polycentricity deepens, an intensive «stratification» of coastal territories is observed in terms of conditions, vectors and rates of socio-economic development.

In their overwhelming majority, the coastal zone of Russia is a narrow, winding and intermittent tape: almost 90% (36.8 of 41 thousand km) of the maritime borders of the Russian Federation extend along its Pacific and Arctic coasts, remote from the main centers of socio-economic activity, are weak developed (with the «island» nature of urbanization) and generally being unfavorable to human life in the natural-climatic relation (Fig. 3).

The situation is significantly different in the west and south-west (in the Black Sea and in the Baltic), where the «sea factor» is fully felt not only for coastal municipalities (in the Russian Federation as a whole, they account for 22.4% of the territory and 11.5% demographic potential; Druzhinin 2017), but also at the macro level: in the Southern Federal District, 67% of the population lives within 200 km of accessibility from the coast, and 96% live in the North-Western District. Due to natural and historical circumstances, the pronounced asymmetry of the Russian coastal zones along the south-west-northeast axis (in economic potential, demographic dynamics, and therefore in the width of the strike area) has steadily increased throughout the post-Soviet period. It is symptomatic that over the past three decades, out of 34 cities located on the Arctic and Pacific coasts of the country, only 4 showed positive demographic dynamics (Yuzhno-Sakhalinsk, Naryan-Mar, Salekhard and Artyom).

The fragmented, non-universal «shift to the sea» of the economy and population (with the corresponding «expansion» and «compaction» of the coastal zones) almost completely corresponds to the geo-economic realities (including the successful arrangement of the most important transport and logistics corridors; Radvanyi 2017) and is accompanied by a concentration of population and socio-economic activity in the largest multifunctional coastal cities. So, in particular, in St. Petersburg for the post-Soviet period, the population «increased» by 530 thousand (or 11.3%), in Makhachkala – by 393 thousand (2.2 times), in Sochi – by 132 thousand (by 39%), in Rostov-on-Don – by 96 thousand (by 9.6%), Kaliningrad – by 52 thousand (13%). In general, currently 14 existing urban group settlement systems are functioning on the coasts of the Russian Federation, which differ in their demographic and economic «mass», spatial structure and functionality. The size of their population (including rural settlements in agglomeration) reaches a total of 15.5– 15.8 million people, which is equivalent to 92–93% of the total «coastal» population of Russia. Inherent in the coastal zones of the country, the hyper-urbanization and dominance of urban agglomerations, generating additional prerequisites for advancing socio-economic dynamics, simultaneously create a «dependency track» of increasing marine economic

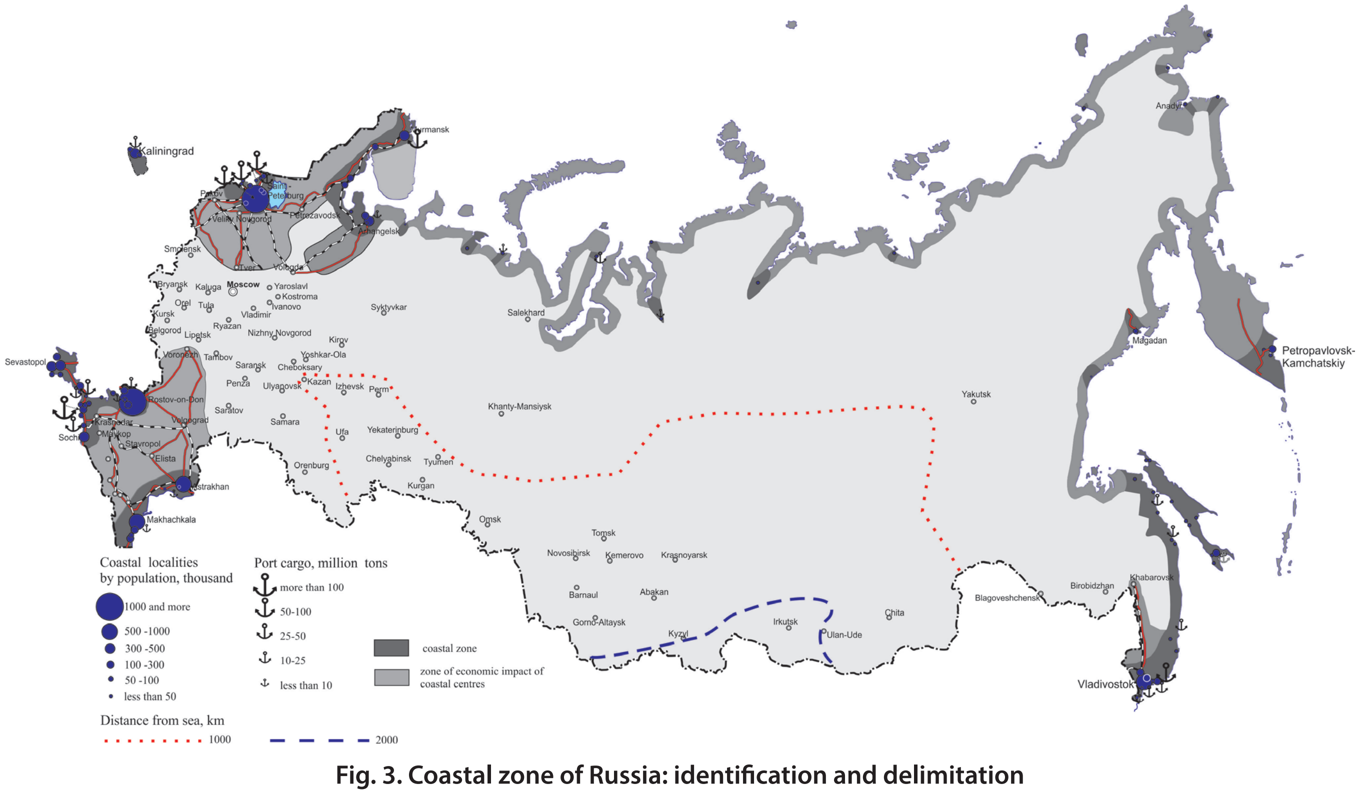


Fig. 3. Coastal zone of Russia: identification and delimitation

activity (and, consequently, the further development of the entire coastal zone) on the infrastructure, investment and human resources of a few leading coastal centers, on their ability to take on the mission of borderland outposts, seaside «facades», to become the nodal elements of the «development corridors», as well as centers for generating and transferring various kinds of innovations.

Innovation activity has a clear polarization towards core regions and major urban agglomerations. These are the capital cities, leading financial and industrial centers, key logistic corridors where most activities of high added value concentrate. As is it noted by Turner et al. (1996), two-thirds of the largest cities of the world with population above 2.5 million inhabitants are coastal. Eight of the top ten largest cities are located at the coast (Reed 2010). The fundamental question is what are the determining factors for the ability of coastal zones to act as priority «platforms» of innovation activity, fostering generation and transfer of technological, industrial, institutional and other types of innovations.

Coastal cities and agglomerations are the natural gates for foreign direct investments (FDI) and knowledge exchange. In China, firms of the eastern provinces take advantage of active trans-aquatic trade and industrial plants of transnational corporations by using reverse engineering for products and technologies, knowledge spillover via labor turnovers, intensification of local R&D facilities (Cheung at al. 2004). An example of Quebec's coastal maritime industry suggests that innovation within a maritime sector spans across a broad range of marine-related activities, such as transportation, shipbuilding, equipment manufacturers, marine products, food industry, fishery and fish farming, etc. (Doloreux at al. 2008). With that, while marine activities are not solely based in coastal areas, the economies of the coastal regions are not limited to maritime sector (Morrissey at al. 2012). The regional innovation system of a coastal region should be considered from the lens of an enabling environment for cross-border, transnational and trans-aquatic cooperation. It has an important role of knowledge and innovation hub for multilateral transfer of resources (financial, intellectual, human, information, etc.) and their absorption for the benefit of the local and national communities.

In the context of Russia, coastal regions are the crucial networking points with the global market. The western coasts are the cooperation bridges with the European Union member states and other European countries, while the eastern regions are naturally focused on Asia Pacific. The role of a «development corridor» (Fedorov 2010) played by the coastal regions is difficult to evaluate with respect to innovation dynamics, since they are not being necessarily utilized on spot (Mikhaylova 2019). One of the possible indicators is the level of research collaboration. Figure 4 presents data on the share of scholarly output implemented in international and inter-regional cooperation over the period of 2013-2017.

International collaboration serves as an indication of knowledge inflow and the degree of openness of the region to foreign counterparts. Inter-regional research networking shows the degree of knowledge transfer within the national innovation system, both affected by the institutional readiness and the relative value of the knowledge possessed. Of the top-10 regions by the share of publications implemented in international collaboration, eight are coastal: Leningrad region – 78.8%, Magadan region – 43.7%, Chukotka Autonomous District – 40.7%, St. Petersburg – 35.3%, Kaliningrad region – 34.8%, Republic of Crimea – 33.6%, Nenets Autonomous District – 33.3%, and Sakhalin region – 30.9%. Out of 22 coastal regions only six fall below the national average value of 19.1%: two at the Black Sea: Rostov region – 19.0% and Krasnodar region – 18.5%, one at the Okhotsk Sea: Khabarovsk Krai – 14.5%, and three at the Caspian Sea: Republic of Dagestan – 13.9%, Republic of Kalmykia – 13.6%, and Astrakhan region – 10.9%.

Data in inter-regional collaboration suggests that seven of the 22 regions are highly integrated in R&D with other Russian regions (with the regional average value of 23.0% publications). These are the Republic of Karelia – 40.5%, the Republic of Kalmykia – 34.8%, Primorsky Krai – 33.7%, Kaliningrad region – 33.4%, Arkhangelsk region – 25.1%,

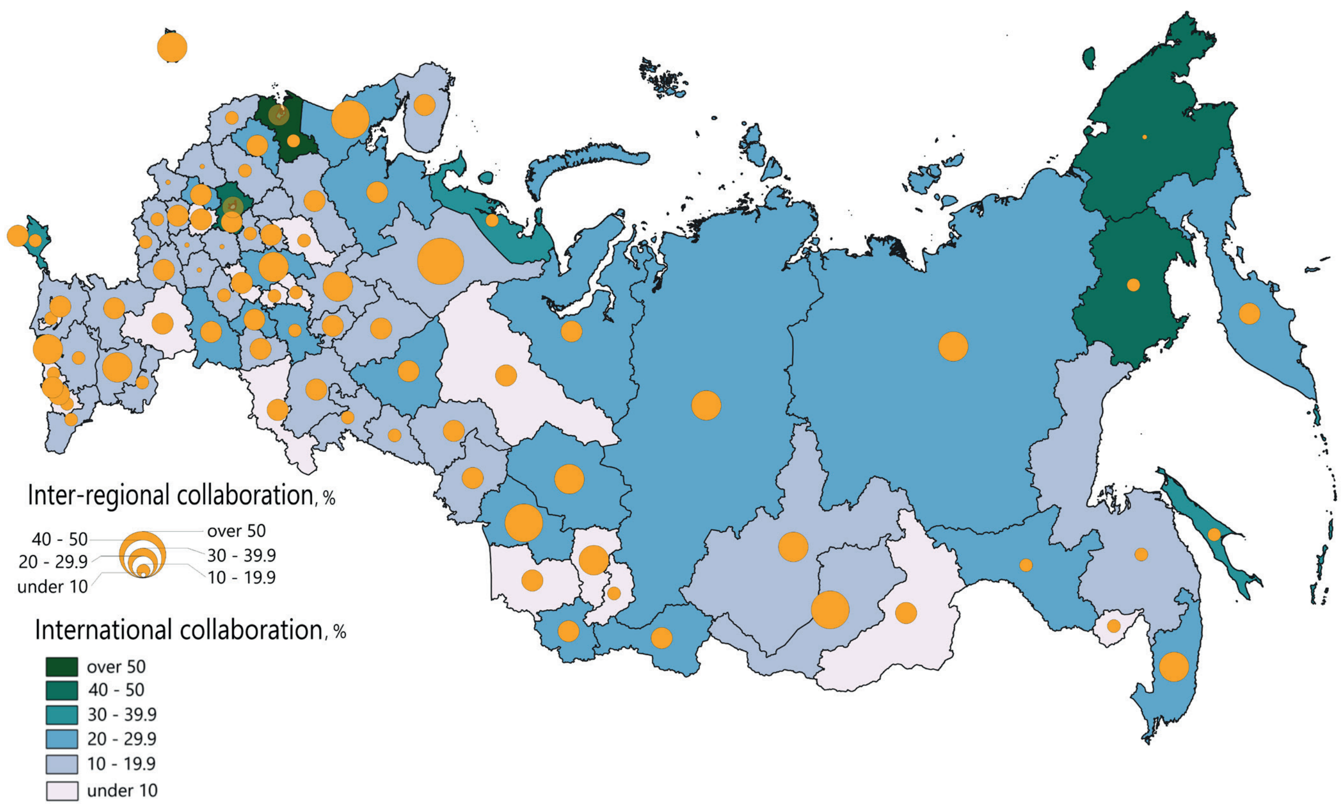


Fig. 4. International and inter-regional cooperation in research

Krasnodar Krai – 24.4%, and Rostov region – 23.7%. It is notable that some of these coastal regions underperform in international collaboration but are in the top-charts of inter-regional research collaboration. For instance, the republics of Karelia and Kalmykia, as well as Primorsky Krai are in the top-10 performers. In this regard, it is important to correlate the publication activity data with information on patents and advanced manufacturing technologies being generated and used.

Figures 5-6 illustrate the localization and weight of regions as divided by their core functions – knowledge-generating or knowledge-commercializing. The data on defragmentation of the national innovation system to generation and commercialization of knowledge presented are the average values for 2013-2017 – identical to those acquired for publication activity window. The patents considered includes data on inventions, utility models, and industrial designs (Fig. 5).

Data suggests that coastal regions account for 17.8% of the total patents generated, with St. Petersburg exceeding the national average value. However, coastal territories show moderate performance with regard to commercialization dynamics. This pattern correlates with the one on research collaboration indicating strong capacity of knowledge generation, including the international spillover effect, and a knowledge transfer function within the national innovation system.

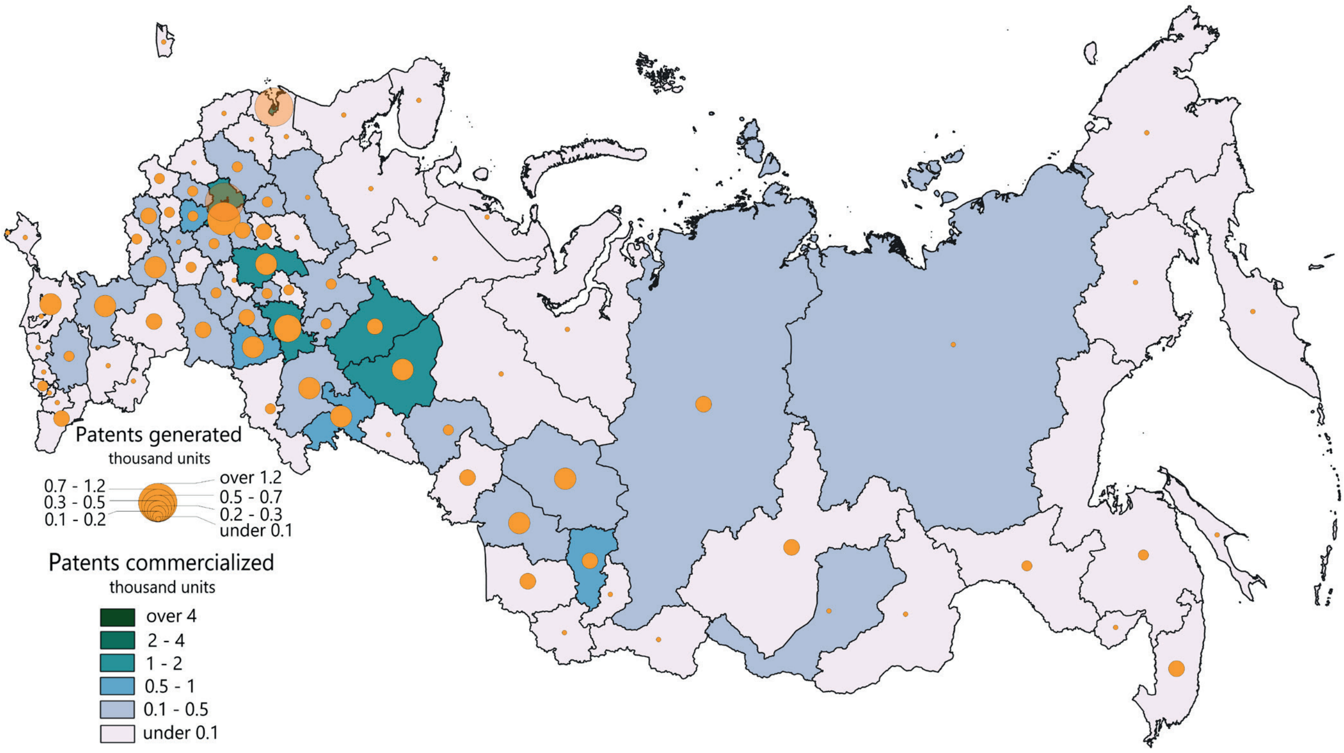


Fig. 5. The volume of patents generated and commercialized by regional innovation systems

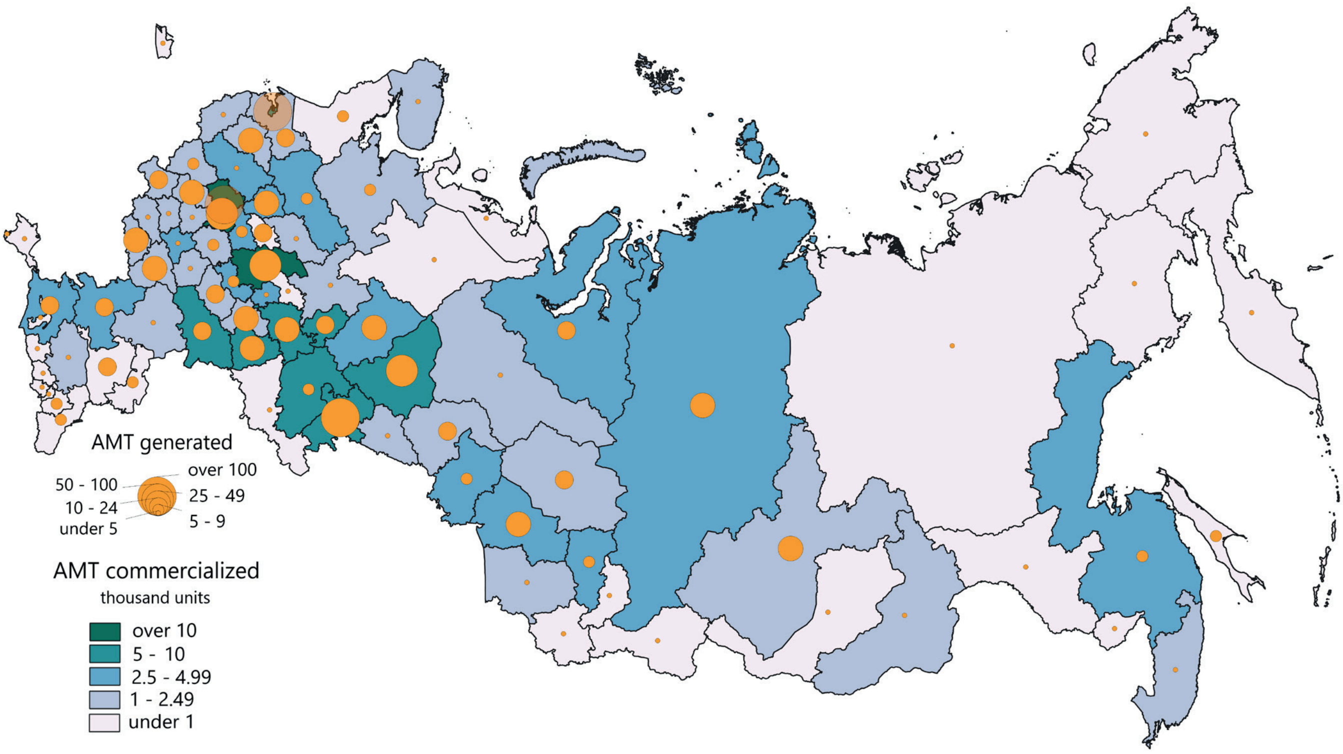


Fig. 6. The volume of advanced manufacturing technologies generated and commercialized by regional innovation systems

Data on the volume of advanced manufacturing technologies suggest that coastal regions implement higher volume of innovative technologies as compared to those generated. Moreover, apart from the industrial centers of the central Russia, coastal regions are among the top-performing territories nation-wide. Thus, the coastal zones of Russia not only occupy priority positions in the innovation space of the country, but also act as a significant environment for building cross-border interactions in the science and innovation sphere.

DISCUSSION

Formation of Russia, as noted by the eminent geographer, historian and ethnologist L.N. Gumilev (1989), is inextricably linked with the Eurasian continent, with its ethnocultural and socio-economic rhythm. This kind of «Eurasian determinant» of the development of our country was extremely bright and large-scale in the post-Soviet period, and especially from the «turning point» of 2013-2014, when the geopolitical confrontation in the Russia-West system was sharply manifested, and the southeast of the Eurasian continent began to move forward rapidly towards global proscenium (the new reality is increasingly being associated with the emergence of «Greater Eurasia»; Bordachev 2015), primarily China. It is this state (concentrating more than 18% of world GDP) that is currently the most important global actor in the field of port facilities and commercial shipping (Wang et al. 2018). The mega-project «One Belt - One Road» (一带一路) promoted by the PRC creates a format both to consolidate this leadership (Liu 2015) and to turn the transport and logistics projects of the «Middle State» into the main determinant of maritime activity, including for Russia, giving new development impulses for its coastal areas.

It is symptomatic that, according to the Strategy for the Spatial Development of the Russian Federation until 2025, all 47 border (including coastal) regions are classified as «geostrategic territories». Acquiring a multi-vector nature in its foreign policy (Druzhinin 2016b), Russia is simultaneously steadily «turning to the sea» (Druzhinin 2019), which is predetermined not only by the growing demand of Eurasian states for the use of resource and raw materials (including the shelf zone) and transport and transit (with an emphasis on the Northern Sea Route) potential of the Russian Federation, but also by the increasing geopolitical turbulence and the need for Russia and its corporations to build «flexible» logistics schemes and economic partnerships with the main Eurasian «centers of power» (the EU, China, India, etc.), relying, first of all, on the developed system of seaports.

The trend for the entire post-Soviet period (in the logic of «Westernization» of foreign economic and humanitarian relations of the country (Vardomskiy 2017), which still retains its dominance and inertia) is the development of port complexes (with the centers of the port industry) in the Baltic and in the Black Sea region – in new geopolitical and geo-economic context is complemented by an intensive (especially after 2014) deployment of military infrastructure in the regions of Russian geostrategic interests (Kaliningrad region, Crimea, Kuriles, the Arctic zone, etc.), as well as «spot» promotion of economic activity in the coastal zones in the format of Eurasian partnerships (including the creation of 7 «coastal» Advanced Development Territories (ADT) in the regions of Pacific Russia and the ADT «Kaspiysk» in Dagestan). The «driver» of further «marine orientation» of the country is also the implementation of large international (in terms of the scale and structure of investments) projects for the extraction of hydrocarbons on the shelf, as well as the localization of complexes for the liquefaction and transportation of natural gas in the coastal zones. As a result, while in the preceding post-Soviet decades, the dynamics of coastal zones was characterized mainly by a «shrinking concentration» on the main communication corridors (with an emphasis on the Baltic and Black Sea coast), in the modern time (increasingly clearly expressed) diversification of foreign economic relations of the Russian Federation and, accordingly, the inevitable activation of cross-border interactions throughout its external contour, is complemented by geopolitical motives and interests – creation of prerequisites for both social and economic «consolidation» of the coastal zones, and for their spatial «expansion», including the spread of their economic influence over the entire Russian coast. These processes correspond to the increasingly active inclusion of coastal zones in the processes of economic internationalization and cross-border regionalization.

The most complete and clear forms of aquatic transboundary regionalization are now found in the Baltic Sea (Druzhinin et al. 2018), where Russian geo-economic interests are focused, and the «Russian presence» in the coastal zone is equally significant (in the Baltic region as a whole, 16 million people are concentrated in the coastal metropolitan areas; 7.5 million – within the Russian Federation). In the post-Soviet period, its integration contour also revealed the Black Sea region (Dobransky 2013), characterized by socio-economic asymmetry (almost 70% of the population of the Black Sea cities now live in the «Turkish segment» of the coastal zone) and which became not only an acute geopolitical confrontation since 2014, but and the space for building economic interconnections that are significant for «Greater Eurasia» (mainly meridional orientation).

The megatrend of «marinization» of the spatial structure associated with aquatic regionogenesis will manifest itself, however, not only in the western border of the country but also practically along the entire perimeter of the Russian borders, and probably most dynamically in Pacific Russia (primarily within the Sea of Okhotsk and Japan; Baklanov 2018) and, of course, in the spaces of the Arctic (PRC symptomatically positions itself as a «near-arctic state»; Collins 2017). The potential of cross-border regionalization (with an emphasis on the energy resource sphere) also has the macro-region of the Caspian Sea «centered» on the Baku agglomeration.

CONCLUSIONS

In the structure of modern Russian space, coastal zones are among its significant and large-scale components. «Surrounding» the arrays of intracontinental territories (and in unity with them ensuring the integration of the country into world economic processes), creating prerequisites for the functioning of the «maritime economy» (and concentrating its most important service and production infrastructure elements), the coastal zones, at the same time, «attract» the population and predetermine the specifics of its settlement (spatial «pattern» of settlements, their functionality, dynamics), generating opportunities for cross-border and innovative activity.

The role of coastal zones (as a priority communication, resource, consumer and geostrategic territory) increases in the modern global and Eurasian context. The consonant with the growing «marine orientation» of Russia, the spatial «expansion» of the country's key segments of the coastal zones (and their spheres of influence), their economic and residential «condensation» (these processes are most noticeable in the Baltic and Black Sea regions, also manifesting in fragments in Pacific Russia and in the Arctic zone of the country) is combined with the ongoing intensive «stratification» of coastal areas in terms of the level and pace of socio-economic dynamics. The development of coastal zones, their perspective, acquires an increasingly pronounced polyfactorness, multi-vectorness, demonstrates visible, growing differences «from place to place», allowing to consider this phenomenon as a priority object of socio-geographical analysis.

ACKNOWLEDGEMENTS

The study was prepared with the support of the Russian Science Foundation, project no. 1918-00005 «Eurasian Vector of Russian Marine Economy: Regional Economy Perspective». Alexander Druzhinin expresses gratitude to the 5-100 competitiveness program of the Immanuel Kant Baltic Federal University for supporting his senior research fellowship.

REFERENCES

Alaev E.B. (1977). Economic-geographical terminology. Moscow: Mysl'. 1977. 189 (in Russian)

Amos C.L., Al Rashidi T., Rakha K., El-Gamily H. and Nicholls R. (2013). Sea surface temperature trends in the coastal ocean. Current development in geography, 6(1), 1-13.

Anikeev V.A. (2012). The town-planning aspects of settling maritime areas basing on the examples of port towns. FEFU: School of Engineering Bulletin, 2(11), 45-53.

Antipolis S. (2001). Urban Sprawl in the Mediterranean region. Mediterranean Blue Plan. [online] Available at: www.planbleu.org/sites/default/files/publications/urbsprawl_1.pdf (Accessed 28/05/2014)

Arzamastsev I.S. (2009). Nature management in coastal zones: basic concepts, zoning and management problems. Tamozhennaya politika Rossii na Dal'nem Vostoke, 4(49), 76-89. (in Russian).

Arzamastsev I.S. and Sorokin P.S. (2008). How to divide the coastal zone? Rybnoye khozyaystvo, 6, 39-41.

Aybulatov N.A. (1989). Geoecology of the coastal zone of the sea. Problemy razvitiya morskikh beregov. Moscow. 81-87.

Baklanov P.Ya. (2018). Marine Spatial Planning: Theoretical Aspects. Baltic region, 10(2), 76-85.

Beach D. (2002). Coastal Sprawl: the effects of urban design on aquatic ecosystems in the United States. Arlington, VA. Pew Oceans Commission.

Beatley T., Brower D.J. and Schwab A.K. (2002). An Introduction to Coastal Zone Management. Washington, D.C. Island Press.

Bondarenko V.S. (1981). Economic-geographical study of coastal zones. Vestnik MGU. Geografiya, 1, 36-41. (in Russian).

Bordachev T. (2015). New Eurasianism. Rossiya v globalnoy politike, 5, 32-41.

Bowen B., Frankic A. and Davis M. (2006). Human development and resources use in the coastal zone: influences on human health. Oceanography, 19(2), 62-71.

Bulleri F. and Chapman M.G. (2010). The Introduction of Coastal Infrastructure as a Driver of Change in Marine Environments. Journal of Applied Ecology, 47(1), 26-35.

Cetin M., Musaoglu N. and Tanik A. (2008). Multitemporal Assessment of Land-Use Change in a Rapidly Urbanizing Coastal Region in Turkey Using Remote Sensing. Environmental Engineering Science, 25(6), 917-928.

Cheung K.Y. and Lin P. (2004). Spillover effects of FDI on innovation in China: Evidence from the provincial data. China Economic Review, 15, 25-44.

Collins J.F. (2017). The Arctic in the Period of Geopolitical Changes: Estimates and Recommendations. Valdayskiye zapiski, 75, 15. (in Russian).

Costanza R., D'Arge R., DeGroot R., Farber S. et al. (1997). The value of the world's ecosystem services and natural capital. Nature, 387, 253-260.

Cracknell A.P. (1999) Remote sensing techniques in estuaries and coastal zones an update, International Journal of Remote Sensing, 20(3), 485-496.

Crowel M., Scott E., Kevin C. and Scott M. (2007). How Many People Live in Coastal Areas?, Journal of Coastal Research, 23(5), 3-6.

Dergachev V.A. (1980). Basics of economic zoning of the oceans. Kiyev. (in Russian)

Dobransky S. (2013). The formation of a Black sea community. Polis. Political Studies, 1, 177-181 (in Russian).

Doloreux D. and Melancon Y. (2008). On the dynamics of innovation in Quebec's coastal maritime industry. Technovation, 28, 231-243.

Druzhinin A.G. (2016a). «Maritime» component in the Russian human geography: Traditions and novations. Izvestiya Rossiiskaya Akademii Nauk, Seriya Geograficheskaya, 6, 7-16. (in Russian).

Druzhinin A.G. (2016b). Russia in modern Eurasia: The Vision of a Russian Geographer. Quaestiones Geographicae, 35(3), 31-39.

Druzhinin A.G. (2016c). Russia's Coastal Zone as a Social and Geographic Phenomenon: Conceptualisation and Delimitation. Baltic region, 8(2), 57-67.

Druzhinin A.G. (2016d). Socio-economic development of coastal areas of the European part of Russia: factors, trends, models. Rostov-on-Don, SFU publishing house. (in Russian).

Druzhinin A.G. (2017). The coastalisation of population in today's Russia: A sociogeographical explication. Baltic region, 9(2), 19-30.

Druzhinin A.G. (2019). The sea factor in the spatial and socio-economic dynamics of today's Russia. Quaestiones Geographicae, 38(2), 91-100.

Druzhinin A.G. et al. (2018). Coastal zones of Russia in the Baltic: factors, features, prospects and strategies for cross-border clustering. Moscow: «INFRA-M» (in Russian).

El-Sabh M., Demers S. and Lafontaine D. (1998). Coastal management and sustainable development: From Stockholm to Rimouski. Ocean & Coastal Management, 39, 1-24.

Fadeev S.A. (1998). Transformation of coastal zones in Western Europe and European Russia: delimitation and control. Doctoral thesis. Moscow. (in Russian)

Fedorov G.M. (2010). The Kaliningrad dilemma: a 'development corridor' or a 'double periphery? The geopolitical factor of the development of the Russian exclave on the Baltic Sea. Baltic region, 2(4), 4-12.

Fedorov G.M., Mikhailov A.S. and Kuznetsova T.Yu. (2017). The influence of the sea on the economic development and settlement structure in the Baltic Sea region. Baltic region, 9(2), 4-18.

Gogoberidze G.G. (2008). Structure and main features of marine economic potential of the coastal territory. IKBFU's Vestnik, 3, 75-81.

Gorkin A.P. (2013). Socio-economic geography: concepts and terms. Dictionary. Smolensk. (in Russian)

Green D.R (2009). Issues in coastal zone management. UK: Thomas Telford Press.

Gumilev L.N. (1989). Ancient Russia and the Great Steppe. Moscow: Mysl.

Hinrichsen, D. 1998. Coastal Waters of the World: Trends, Threats, and Strategies. Washington D.C. Island Press.

Ilyin I.A. (1934). About Russia. Three speeches. 1926-1933. Sofiya: Izd. «Za Rossiyu». (in Russian).

Jacobson C., Carter R.W., Thomsen D.C. and Smith T.F. (2014). Monitoring and evaluation for adaptive coastal management. Ocean & Coastal Management, 89, 51-57.

Kolosov V.A. (2018). Russian frontier: neighborhood challenges. M.: Matushkina.

Kurt S. (2016). Analysis of Temporal Change Taking Place at the Coastline and Coastal Area of the South Coast of the Marmara Sea. Gaziantep University Journal of Social Sciences, 15(3), 899-924.

Lam N.S.-N., Arenas H., Li Z. and Liu K.-B. (2009). An Estimate of Population Impacted by Climate Change Along the U.S. Coast. Journal of Coastal Research, 2(56), 1522-1526.

Liu W.D. (2015). Scientific understanding of the Belt and Road Initiative of China and related research themes. Progress in Geography, 35(5), 538-544.

Lyth A., Holbrook N.J. and Beggs P.J. (2005). Climate, urbanisation and vulnerability to vector-borne disease in subtropical coastal Australia: Sustainable policy for a changing environment. Environmental Hazards, 6(4), 189-200.

Makhnovsky D.E. (2014). The Coastal Regions of Europe: Economic Development at the Turn of the 20th Century. Baltic region, 4(22), 50-66.

Mikhaylov A.S., Mikhaylova A.A. and Kuznetsova T.Yu. (2018). Coastalization effect and spatial divergence: segregation of European regions. Ocean & Coastal Management, 161, 57-65.

Mikhaylova A.A. (2019). The dimension of innovation in the economic security of Russian regions. European Journal of Geography, 9(4), 88-104.

Mimura N., Nurse L., McLean R., et al. (Eds.): Climate Change 2007: Impacts, Adaptations and Vulnerability, 687-716.

Morrissey K. and O'Donoghue C. (2012). The Irish marine economy and regional development. Marine Policy, 36, 358-64.

Ngoile M. and Horrill C. (1993). Coastal ecosystems, productivity and ecosystem protection: Coastal ecosystem management. Ambio, 22, 461-467.

OECD (2003). Coastal zone management: integrated policies. Paris: OECD.

Pak A. and Majd F. (2011). Integrated coastal management plan in free trade zones, a case study. Ocean and Coastal Management, 54, 129-136.

Pak A. and Farajzadeh M. (2007). Iran's Integrated Coastal Management plan: Persian Gulf, Oman Sea, and southern Caspian Sea coastline. Ocean & Coastal Management, 50, 754-773.

Pernetta, J.C. and Elder D.L. (1992). Climate, sea-level rise and the coastal zone: Management and planning for global changes. Ocean and Coastal Management, 18, 113-160.

Pokshishevskiy V.V. (1982). Economic geography of the World Ocean. Moscow.

Qasim S.Z., Sengupta R. and Kureishy T.W. 1988. Pollution of the seas around India. Proceedings of Indian Academy of Science (Animal Science), 97(2), 117-131.

Radvanyi J. (2017). Qu and Vladimir Poutine se fait géographe… Hérodote. Géopolitique de la Russie, 166-167, 113-132.

Reed D. (2010). Understanding the Effects of Sea-Level Rise on Coastal Wetlands: The Human Dimension. Geophysical Research Abstracts, 12.

Rodoman B.B. (1999). Territorial ranges and networks. Moscow: Oykumena.

Salnikov S.S. (1984). The economic geography of the ocean – a new promising direction of economic and social geography. Sovetskaya geografiya. L.: Nauka. (in Russian)

Savitskiy P.N. (1997). Continent-Ocean (Russia and the world market). Kontinent-Yevraziya. Moscow: Agraf. 398-418. (in Russian)

Semenov Tyan Shanskiy V. P. (1915). On the powerful territorial possession in relation to Russia. Petrograd: Tipogr. M.M. Stasyulevicha (in Russian).

Siry H.Y. (2007). Making Decentralized Coastal Zone Management Work for the Southeast Asian Region: Comparative perspectives office of Legal Affairs: division for Ocean Affairs and the Law of the Sea. The United Nations.

Slevich S.B. (1988). Ocean: resources and economy. L. (in Russian).

Small C. and Nicholls R.J. (2003). A global analysis of human settlement in coastal zones. Journal of Coastal Research, 19(3), 584-599.

Susanta K.C. (2013). Interactions of environmental variables determining the biodiversity of coastal-mangrove ecosystem of west Bengal, India. The Ecoscan: An International Quarterly Journal of Environmental Sciences, 3, 251-265.

Tookwinas S. (1999). Coastal Planning of Shrimp Farming: Carrying Capacity, Zoning and Integrated Planning in Thailand. 108-109.

Turner R.K., Subak S., and Adger W.N. (1996). Pressures, trends, and impacts in coastal zones: interactions between socioeconomic and natural systems. Environmental Management, 20(2), 159-173.

Vardomskiy L.B. (2017). Post-Soviet Integration and Economic Growth of the New Borderland of Russia in 2005-2015. Prostranstvennaya Ekonomika, 4, 23-40.

Wang C., Wang M.J. and Wang Y. (2018). Identification of Chinese key nodes in the shipping network of the 21st Century Maritime Silk Road. Progress in Geography, 37, 1485-498.

Wang X.H., Pearson S.G., Morrison R.J. et al. (2011). Integrated coastal zone management research in Australia and China. Labour and Management in Development Journal, 11, 1-17.

Wescott G. (2009). Stimulating Vertical Integration in Coastal Management in a Federated Nation: The Case of Australian Coastal Policy Reform, Coastal Management, 37(6), 501-513.

Zalogin B.S. (1984). Economic geography of the World Ocean. Moscow: Izd-vo MGU (in Russian).

Received on July 3th, 2019 Accepted on January 21th, 2020

Alexander N. Pilyasov[1*]
[1] Institute of Regional Consulting, Moscow, Russia
***Corresponding author:** pelyasov@mail.ru

ARCTIC ENTREPRENEURSHIP DEVELOPMENT FACTORS

ABSTRACT. When studying Arctic entrepreneurship, the researcher encounters many paradoxes. Against the background of a powerfully developed topic of entrepreneurship and entrepreneurship in regional science, the layer of work on the entrepreneur in the Arctic is extremely thin. What is even worse – well-known mainstream theories turns out to be unsuitable for the study of Arctic entrepreneurship.

Under these conditions, the only way out for a researcher is to attach to a zonal paradigm and recognize the Arctic as an anti-mainland, and Arctic entrepreneur as the full antipode of a continental fellow. The adherence to the zonal paradigm removes contradictions and equips the researcher with a comparative method to elaborate all facets of the Arctic exceptionalism in the figure of the entrepreneur. The entrepreneur here is understood as a close «relative» to the indigenous peoples and the entire Arctic economy, the Arctic environment, and their specific behavioral traits and adaptation strategies to natural and economic extremes.

The most important factors in the development of entrepreneurship in the Arctic, which create a mosaic picture of situations in the Arctic territories, are: the geographical location – an island or quasi-mainland position; the presence of large resource corporations and a specific stage of their exploitation of the resource field; the institutional structure of the local economy in terms of the degree of nationalization. The cumulative impact of these factors creates a multi-faceted picture, when the Murmansk oblast is at one pole of the extreme favorable conditions for business, and the Chukotka autonomous okrug is at the other extreme hardness.

However, the favorable factors themselves do not guarantee the active development of entrepreneurship. The example of the Murmansk oblast demonstrates the opposite: the conditions for entrepreneurs here are so comfortable that they prefer to slip to the south, where there is no need to bear the burden of northern guarantees and compensations for the employee, and from there to conduct their business activities.

KEY WORDS: Arctic entrepreneurship, factors of development, Arctic economy, marine issues

CITATION: Alexander N. Pilyasov (2020) Arctic Entrepreneurship Development Factors. Geography, Environment, Sustainability, Vol.13, No 1, p. 46-56
DOI-10.24057/2071-9388-2019-91

INTRODUCTION

In the national project «Small and Medium Entrepreneurship and Support for an Individual Entrepreneurial Initiative», exclusively ambitious tasks are set to reach 32.5% of GDP and 25 million jobs by the end of 2024 owing to entrepreneurial activity (National project 2019). Given the enormous diversity of the natural and socio-economic conditions of our country, it is impossible to solve these problems without carefully taking into account local specifics.

The Arctic zone is precisely one of those territories of Russia that are maximally excluded from the general standard of unified federal approaches. And the development of entrepreneurship here is also very specific. So far, the federal policy of support, tuned to the common «arshin»[1], cannot reverse the negative trends that have developed here in recent years: in 2018 alone, the number of small businesses has decreased by 4,000 and workers by 25,000 people[2].

When you compare the flow of literature on entrepreneurship within regional science and the flow of literature on Arctic entrepreneurship, bewilderment arises. On one hand, there is a powerful and multi-subject layer of work on entrepreneurship, which for several decades has been developed within the world social science. On the other hand, there is nothing similar inside the Arctic studies: despite the fact that the number of works on Arctic entrepreneurship has been increasing in recent years, very narrow plots are developed, absolutely incomparable in breadth and scope with the work of colleagues in the social science.

Separate articles on individual facets do not capture the general phenomenon of an entrepreneur in the Arctic. But then a natural question arises, on which theoretical platform is a new, holistic view on the problem of entrepreneurship in the Arctic possible? It seems that such a comprehensive view provides a zonal paradigm, that is, a geographical, rather than an anthropological, sociocultural or economic view of an entrepreneur in the Arctic.

Arctic is a special zone of the earth's surface (The Arctic Frontier 1966; UNESCO 2009; Russian Arctic 2014; The New Arctic 2015; Vinokurova et al. 2016; Petrov et al., 2017), so different from the rest, first of all from the temperate zone, densely populated and intellectually dominant in all textbooks, in all theoretical models and constructions, that here it is needed from geography, from natural climatic and socio-economic features to explore the nature, the phenomenon of the Arctic entrepreneur. That is precisely here, in the Arctic zone, geographers can lead specialists from other social disciplines, as nowhere else.

On one hand, low transport accessibility and remoteness of the Arctic regions, small dispersed markets, lack of

[1] Russian measure of length, approximately 0.7 meter.
[2] Arctic: territory of dialogue. Roundtable: Large potential of small business. Speech by Alexander Kalinin, «OPORA ROSSII». April 9, 2019.

investment, lack of human capital in the Arctic, high costs for heating and electricity, the need to implement northern guarantees and compensations for employees form the strongest barriers to business in the Arctic. On the other hand, the effects of a temporary local monopoly, which small entrepreneurs easily gain here, the unique resources of the Arctic, including its cultural heritage, including indigenous peoples and their traditional knowledge, reindeer herding and traditional hunting, create their comparative advantage against their «fellows» in temperate zone.

Both that radically distinguishes the Arctic businessman from the «mainland». Paradoxical from the point of view of established canons of world social science, Arctic entrepreneurship is much closer to adaptation strategies in the extreme environment of indigenous peoples of the North, with their colossal plasticity and «stretchability» of the size of traditional economy, their ability to constantly maneuver and migrate in a high-risk natural and economic environment, than to the behavior of their fellows in a temperate, densely populated zone of Europe and America.

The purpose of this work is to show the specifics of the Arctic entrepreneurship, to characterize the main factors of its development and to demonstrate its development on the example of a specific Arctic region – the Murmansk oblast.

MATERIALS AND METHODS

Entrepreneurship in the Russian Arctic is multicolored, from the near-mainland case in large Arctic urban agglomerations to near-rural (as in the rural periphery of central regions of Russia) in the most hard-to-reach areas of the Far North. Therefore, there is a huge temptation to start studying specific cases and not to see the general phenomenon of Arctic entrepreneurship, which is anti-mainland in its essence, behind the description of individual trees, with all the multitude of local versions and variants.

The only recipe for this danger is to remain loyal to the zonal paradigm (Pilyasov, 2009; Siberian economy through the lenses of the latitudinal zones, 1984). That is, to see above all the general effect of cold discomfort and distance on all other and more particular features of the portrait of an arctic entrepreneur. Not to lose this common is, in our opinion, the real most important task of research on Arctic entrepreneurship. Despite all the local versions, the entrepreneurial class of the Arctic is welded together and is organically close to each other so in the polar cities as on the nomadic sites of private reindeer herders. It is the general pressure of the extreme natural and economic environment that forges this common feature of Arctic entrepreneurs, no matter where they are.

Loyalty to the zonal paradigm immediately reveals the illegitimacy of the concepts, approaches, methodology for describing and studying entrepreneurship for the temperate zone in relation to the Arctic. Take, for example, statistics. It does not distinguish between the entrepreneur of the main settlement zone and the Arctic. Meanwhile, it is obvious that the same data on the development of entrepreneurship in Russia and the Arctic actually means quite different – in one case, the description of a very stationary repetitive phenomenon planted on a statistical reporting «pin»; in another case, suddenly caught up in a digital description of a gust of wind, which tomorrow will be different, and the day after tomorrow – the third. It is like a photograph of a static or impetuous human figure: formally, this is a single phenomenon, but in fact it's about very different states.

The approach to the Arctic as an anti-mainland, which became the basis of our study of Arctic entrepreneurship (see first attempt – Pilyasov, Zamyatina 2016) and which itself, in turn, ideologically flows from the zonal, geographical paradigm, determined the set of analytical tools that was in our hands. First of all, it is a comparative method of research, which was used in a variety of guises.

First, it is a comparison of entrepreneurship in the context of latitudinal zones – the tundra of the Arctic (for example, the Yamalo-Nenets autonomous okrug), the taiga North (Khanty-Mansi autonomous okrug-Yugra), the forest-steppe temperate zone (for example, the south of the Tyumen oblast). The Tyumen oblast in general is the most appreciative object of zonal analysis and comparison in the Russian Federation precisely because here in all three latitudinal zones there is a state regional statistics due to the fact that each of the three regions included in the Tyumen «matryoshka» is a subject of the Russian Federation.

Secondly, it is a comparison of the Arctic and the mainland, the Arctic and Russia, which becomes possible due to the emergence of Arctic statistics in the Russian Arctic as a whole, which Rosstat has been developing since 2017 (Rosstat 2016-2019), including information on individual, small and medium enterprises in the Arctic.

Thirdly, this is a comparison of entrepreneurship within the Arctic itself. Here there are specific difficulties that required the development of a special research position to overcome them.

The fact is that the status of nine regions, fully or partially included in the Russian Arctic, are significantly different. There are four entirely Arctic territories here – this is the Murmansk oblast and three autonomous okrugs – Nenets, Yamalo-Nenets and Chukotka.

There is a separate group of three arctic-northern regions, within which arctic territories fixed by presidential decree and districts of the Far North and / or equivalent to them are separated: Arkhangelsk region without Nenets autonomous okrug, Komi Republic (with Arctic Vorkuta), Karelia Republic with three coastal Arctic regions . There are two vast multi-latitude regions, including the arctic, northern, and even forest-steppe zones – Krasnoyarsk krai, which by its characteristics is often close to the average Russian region, and Republic of Sakha-Yakutia.

Fourthly, this is a comparison of the «quasi-mainland» areas of entrepreneurship within the Arctic itself, which have a permanent ground road connection with central Russia, and «island» areas of entrepreneurship with limited delivery times, which have only seasonal, sea, river, air connection with the «mother» Russia. It can be called a comparison of entrepreneurs «near» and «remote» Arctic. It is clear that in many of their characteristics they will be different: the first closer to the comfortable existence of the mainland entrepreneurs, and the latter – to the type of independent, frontier-type entrepreneurs who perceive their business as a lifestyle.

The main informational basis of the work has been the materials of state statistics on entrepreneurship, which have been considered since 2000s. During this period, the criteria for the small business had been changed several times, separate reporting for the micro-enterprises, for individual enterprises arose, several comprehensive surveys of Russian entrepreneurship had been carried out. But the comparative method allowed us to overcome these turbulences: we selected and fixed stable correlations (Arctic-Russia, Arctic-North and others), which were maintained during all the transformations. In addition to statistical information, we used materials from our surveys of businessmen in the Arctic during expeditionary surveys of the Nenets, Yamalo-Nenets autonomous okrugs, and the Norilsk industrial region.

RESULTS

Arctic entrepreneurship: challenge and response

In the Arctic, the level of entrepreneurial activity, as measured by the number of small businesses per 10,000 inhabitants, is significantly lower than the average for Russia and the North[1]. And this ratio, for example, among the three subjects of the Russian Federation of the Tyumen region is steadily reproduced for all the years of observation, despite numerous changes in the criteria who should be attributed to a small entrepreneur: a maximum in the south, average values in the Khanty-Mansiysk autonomous okrug-Yugra, minimum – on the Arctic Yamal.

But at the same time, and this is also confirmed for different years of observation, the average size of a small enterprise (not just a micro, but a small enterprise) in the Arctic is larger than the average for Russia, for example, in Yamal more than in the south of the Tyumen region. And this can be explained by the fact that small business on average in the Arctic is more industrial in nature than on the mainland – plus the fact that traditional non-productive areas of entrepreneurship – trade, personal services – are often sold in the Arctic by state and municipal enterprises. It turns out that the average small business in the Arctic is like a polar bear: there are fewer individuals here in the Arctic than brown bears in the temperate zone, but they are bigger.

The challenge of rising costs from cold and remoteness makes every entrepreneur in the Arctic look for his own creative answer to it. And this search for an answer is the common thing that unites all entrepreneurs of the Arctic zone.

Among all the answers, consider those that are, firstly, recorded by official statistics; secondly, which have a clear arctic specificity. (For this reason, we shall not consider going into illegal status, because, firstly, it is impossible to statistically evaluate this «answer»; secondly, it is difficult to assess the Arctic specificity here without detailed sociological methods).

Given that the small enterprises themselves in the Arctic are comparatively less than on the mainland, the proportion of employed under part-time contracts here is always higher. But you need to look even broader: it's not just part-timers. It is colossal organizational flexibility in all types – the combination of budget, corporate employment with entrepreneurship; active attraction of part-timers; strategy of complex specialization in a wide range of goods and services – that helps to neutralize the negative effect of the Northern extra-costs for the Arctic entrepreneur.

Another form of cost reduction is the emphasis to trade mainly of its own products and services – as a rule, to a greater extent than on the mainland and on average in Russia. For example, on average in Russia this share for small enterprises in 2018 was 39,8%, for the Chukotka autonomous okrug (CHAO) – 72.8, for the Yamalo-Nenets autonomous okrug (YANAO) – 55.4%, for the Nenets autonomous okrug (NAO) – 52.8%, for the Murmansk oblast – 50 ,4%[2]. Arctic entrepreneurs have less opportunity to sell other firms' products, because this requires better logistic schemes, better road infrastructure than in the Arctic. And here any movement of cargo is expensive and the risks are high.

It is not surprising that the Arctic autonomous okrugs, which are characterized by the most harsh natural and economic conditions, have the maximum share of part-timers and products/services of their own production in the total sales (Table 1).

Another creative response of the Arctic entrepreneur to the challenges of the northern costs is constant mobility. The Arctic business can be fully called a business «on its feet»: it is impossible to be successful here, constantly not moving in the space between the input supply base and numerous local markets.

But after all, a successful entrepreneur on the mainland is also mobile. The difference is that the comparative role and time to travel in the Arctic is higher: the «travel» costs here are simply organically integrated into the business. The phenomenon of mobility seems to be the same, but its role in business success is many times higher in the Arctic than on the «mainland».

In conditions of remoteness, the value of mobility for an entrepreneur goes far beyond the traditional interaction with suppliers and consumers. Both business trips and vacations work for the Arctic entrepreneurship: here, mobility at the same time provides for picking up new knowledge, new technologies and business schemes that come from the mainland and can be used constructively in the Arctic.

For the Arctic entrepreneur, success increasingly depends not only on labor and capital as the key factors of production before, but on their own ingenuity, innovation, which are impossible without extracting new knowledge from outside and relying on talent. And it is precisely mobility that helps

Table 1. Part-timers and the dominance of their own products in the total sales as mechanisms for reducing Arctic costs for entrepreneurs

	2011	2014	2018
The share of external part-timers in small enterprises (without micro-enterprises),%			
Arctic – YANAO	20	10	15
North – KHMAO-YUGRA	10	10	9
South of Tyumen oblast	10	10	8
The share of own production in total sales of small enterprises (without micro-enterprises),%			
Arctic – YANAO	60	50	55,4
North – KHMAO-YUGRA	60	60	57,7
South of Tyumen oblast	50	40	46,2

Source: collections of «Small and medium-sized enterprises in Russia» for 2011, 2014, 2018: https://www.gks.ru/folder/210/document/13223

[1] *Rosstat, 2016-2019. Available at: http://www.gks.ru/free_doc/new_site/region_stat/arc_zona.html Accessed 15 July 2019*
[2] *http://www.gks.ru/free_doc/new_site/region_stat/arc_zona.html*

an entrepreneur to be the first to deliver innovation to his community, to the local market, and to get his temporary monopoly on this.

And this is all the more important when there are no universities and research centers nearby, from which knowledge flows to the entrepreneur in developed regions for subsequent commercialization. Here, the search for a novelty is entirely the responsibility of the entrepreneur himself, he does not have either partners or assistants here. Apart from the Internet, which Arctic entrepreneurs use in their business (to find partners, to discover new technology, to search for personnel, etc.) much more active than their colleagues on the mainland (Rosstat 2016-2019)

And this is not surprising: after all, on the mainland, an entrepreneur is always on hand with alternative sources for obtaining important business information, which the Arctic entrepreneur is deprived of.

Our expeditionary surveys of businessmen in the Arctic confirm that there is a combination of Internet and «live» search: partners for Internet negotiations are often determined first, which are then checked during business trips. The Internet in no way reduces the significance of personal communications with business partners, which for an Arctic entrepreneur are associated with significantly more time consuming process than for an entrepreneur on the mainland. In the conditions of Arctic remoteness, paradoxically, the role of such a «temporary proximity» (Rallet, Torre 2009), formed as a result of business trips to the mainland of an Arctic entrepreneur, for picking up new business ideas, for the process of educating an entrepreneur, is simply unprecedented.

The economic landscape of many cities and settlements in the Arctic is formed by the supply bases for storage of goods brought into during the summer navigation. All subjects of the Arctic economy, both large corporations and small businesses, are doomed to create stocks of means of production and consumer goods due to the high cost and irregularity of supplies from the mainland. Working «from the wheels», as is customary for entrepreneurs in central Russia, is simply impossible here.

It is not surprising that the Arctic entrepreneur in the most widespread production activities – in construction and transport, in contrast to the northern and southern ones, has much larger areas of industrial premises. There are no such differences in trade and business services – here the Arctic entrepreneur has less storage space than the northern and southern ones (Table 2).

Previously, under the Soviet planned economy, state schemes for the provision of products to the North and the Arctic had been centralized and unified – from several supply base centers and under one transport scheme (to the Arctic mainly along the Northern Sea Route). Such a typical scheme of delivery clearly demonstrated the homogeneity of the space of the «state» Soviet Arctic.

But now the space of the «new» Arctic is highly differentiated, and therefore there is a triumph of private decisions instead of the former unified delivery of goods, which to the maximum extent take into account the peculiarities of the place, and not the Arctic as a whole. The effect on economy of scale in the case of icebreaking caravan pilotage along the Arctic coast (with entry into all main ports-entry bases) is replaced by the effect on economy on diversity.

Each entrepreneur chooses his own «capillary» (for example, along rivers and winter roads), an economically feasible supply scheme from the continental rear bases and relies on his own base of goods storage at the «entry» point. Creative logistics for an Arctic entrepreneur is the most important way to reduce northern production costs. The strength of an Arctic entrepreneur is precisely to take into account the peculiarities of his place and his markets and, through this consideration of specifics, to fight and conquer northern extra-costs. In this sense, he is doomed to be the genius of a place.

If we try to see the total in the infinitely varied schemes of curbing the high costs of doing business in the Arctic, which are used by local entrepreneurs, then this will be an extremely dynamic, constantly changing combinatorics of production relations and niche markets. The mainland entrepreneur, imitating the large structures in which he wants to grow, is prone to stability and settled position.

On the other hand, the Arctic entrepreneur, quite the opposite, sees his benefits in the constructive exploitation of mobility and instability, to which he is ready every day. In the end, it is they who give him a chance to become a local monopoly and get his margin on it! For most of them it makes no sense to grow in the local small markets, and often this is simply impossible. In Alaska, there is even such a concept of «lifestyle business» (business as a lifestyle) – this is when entrepreneurs are satisfied with the existing size of their company and do not strive for growth (Northern Opportunity 2017).

Arctic entrepreneur as a mirror of the Arctic economy

Any regional entrepreneurship is always a copy of the structure of the local economy, the settlement system, the characteristics of local communities and their cultural traditions. But for Arctic entrepreneurship because of the vivid

Table 2. The presence of specially equipped premises of small and medium-sized businesses – legal entities by type of economic activity (according to the results of complete observation for 2010)

Name	Per one enterprise having a specially equipped room		
	Total area of premises squ m		
	YANAO	KHMAO-YUGRA	South of Tyumen oblast
Total	593.8	563.5	737.1
Building construction	1219.5	883.6	749.6
Transport and communication	1144.3	643.9	547.3
Real estate transactions, rental and provision of services	477.8	401.6	555.3
Wholesale and retail trade; repair of motor vehicles, motorcycles, household goods and personal items	306.3	439.4	357.9

Source: Rosstat, results of a one-time survey of small businesses in 2010 specifics of the Arctic, this is doubly true. Entrepreneurship here really embodies the special features of the Arctic economy itself, its small size, dependence on life-supporting transport and energy infrastructure, instability, ethnicity, storage / reservation syndrome.

To fully understand the deep nature of Arctic entrepreneurship, it is necessary to see his organic connection with the indigenous peoples of the North living here for centuries and strategies of adapting of their traditional economy to this extreme environment. It is necessary to see in the entrepreneur's daily response to the challenges of the Arctic remoteness and cold discomfort relatedness to the behavior of the Arctic natural systems: technology monitored by the entrepreneur from nature is usually the most economical.

Entrepreneurs of the Arctic fully perceive the natural Arctic rhythm – super-fast deployment in a favorable short summer season, and when adverse conditions occur, the same rapid coagulation as in winter hibernation, to reappear in a new place at a favorable time. Let us call this property as the «mercury»-like behaviour of the businessmen of the Arctic. Indeed, the share of ultra-dynamic enterprises in terms of growth or extinction in the Arctic is higher than in the north and in the south (Table 3). The accelerated Arctic rhythm characteristic for natural systems here is also fully manifested in the Arctic entrepreneurship.

Researchers of the indigenous peoples of the North (Krupnik 1989 et al.) note the unparalleled plasticity of the internal structure (flexibility of countless recombinations) of traditional economy (reindeer herding and traditional crafts), which ensures its ability to withstand the incredible amplitudes of natural conditions and ultra-fast changes in the natural environment and climate without breakdowns. But to the same extent, this is also true for an entrepreneur in the Arctic: his legal or illegal enterprise is a really temporary, non-stationary coalition on a type of activity that has temporarily become economically attractive. Changing clothes from an entrepreneur to a public sector worker in the small and fine Arctic labor market is not a big deal on the go. Many do this many times during their working life.

To maintain its viability, this firm enters into very mobile and volatile coalitions (daily changing unions and associations) with other participants – corporate, government and other business structures. This associated character is «catched» by Arctic businessmen from the indigenous peoples.

In the Arctic, there is not enough economic density to create the conditions for competition of small businesses. But if there is no market, market competition, then what then creates the conditions for constructive selection, for selecting the most economically viable entrepreneur?

This selection is here due to the factors of the natural and socio-economic environment, which has the properties of colossal instability and variability. The rapid change of natural and economic conditions, economic configurations, unions, associations with a small number of economic entities themselves form a constructive pressure and cause selection among the participants of the local economy.

On the mainland, competition unfolds in the background of a relatively stationary natural and socio-economic environment. And in this context, it is constructive to select the best entrepreneurs. But in the Arctic, in a colossally unsteady environment, the intensification of the forces of competition is simply destructive. Therefore, there exists a mobile dialectic of local monopolies and non-stationary environment, which, on the one hand, supports the sustainability of business development; on the other hand, it provides a constructive selection pressure on economic actors.

How does this specifically look like? The number of births of new firms (for example, per 1000 residents) is relatively less here than on the mainland. But the dynamics of changes in their states, the transition from one phase to another is more powerful and steeper here than in the developed zone. And this very quick flashing of states, branching of trajectories, such an economic combinatorics of different types creates the conditions for the selection of the optimal state and the optimal solution. The actors themselves do not create diversity here, in which only the forces of selection work. They are too few. But a mobile change of states of the environment creates such conditions of the necessary diversity.

When they say that small business creates a competitive environment and contributes to a constructive selection, it is definitely not about the Arctic. Business here is small and can not create competition. The realities of the business of the Arctic – each monopoly sits in its garden bed and carefully protects it. What then protects against social sclerosis? Colossal variability and instability of the environment of activity of economic entities! Working in a highly unsteady environment and lack of competition are the main differences between the Arctic small business and the mainland.

The conditions of competition of the mainland entrepreneur are pushing him to choose a growth strategy, expansion as the only correct one. But for an Arctic entrepreneur working in the conditions of natural and economic extremes, but outside the harsh pressure of competition, growth values are not a priority. It is more important for him to be necessary for his community, to preserve the lifestyle of his ancestors, to ensure the standard of life for his family. Fulfillment of these tasks does not require expansion of the business to new niches and markets. But this principle of reasonable sufficiency is fully inherent in the traditional way of the indigenous people: it is important to be generous among partners, and not to be first among competitors.

Therefore, Arctic entrepreneurship is a very specific kind of human economic activity in the harsh conditions of low transport accessibility, remoteness, small dispersed markets, high costs for heating and electricity, and with specific features of mobility, local monopoly, traditional local knowledge which are all utilized to overcome the challenges of the extreme natural conditions.

Table 3. The proportion of «extreme» enterprises (by dynamics of the number of employees),%

	Fast growing		Fading away	
	2017	2018	2017	2018
YANAO	1.06	0.84 (59)	5.09	6.07
KHMAO-YUGRA	0.80	0.82 (215)	4.98	3.95
South of Tyumen oblast	0.56	0.54 (207)	4.14	3.23
Reference: Russia	0.72	0.58	3.89	6.35

Source: collections of «Small and medium-sized enterprises in Russia» for 2017, 2018: https://www.gks.ru/folder/210/document/13223

Factors of development of Arctic entrepreneurship

The results of our expeditionary and sociological surveys of Arctic entrepreneurs in single-industry towns of Yamal autonomous okrug convince us that the differences between Arctic and mainland entrepreneurs even of one type of activity in current economic behavior and business strategies turn out to be even greater than differences between entrepreneurs of different specializations, for example trade and manufacturing business. The most important reason is the zonal, geographical factor, which sharply separates the businessmen of the Arctic from all the others, even irrespective of the particular specialty chosen by him. The commonality of entrepreneurs on the Arctic conditions is stronger than industry differences.

Based on this, let us focus primarily on the specific Arctic factors of entrepreneurship development, which have a powerful effect on the daily life of local business[1]. The first in this series will be the transport distance, that is, the island or «mainland» position of a particular Arctic territory.

It is clear that the Arctic «islands» absolutely transport-isolated from mother Russia are like «double Arctic»: the negative effects of northern prices, remoteness from centers of state and municipal support, the benefits of a monopoly position on tiny local markets, the influence of the traditional way of indigenous peoples of the North and Russians old-timers are manifested here with increased force.

«The development of small and medium-sized businesses in a closed city is not an easy task. The main principle of market competition does not work here – that is, the strongest survives. On the mainland, the competitive struggle gives entrepreneurs the strongest impetus to development. Moreover, partly competition arises because of migrants. In Norilsk, this mechanism does not work because of its remoteness from the mainland» (Present and future of SME in Norilsk 2019).

On the other hand, the presence of a regular ground connection with the main settlement zone turns such areas of the Arctic into a «quasi-mainland». Here the layer of entrepreneurship is already more dense and the effects of competition are beginning to work, centers of state support are closer and it is easier to get it.

The analysis of two official lists – areas of the Arctic and areas with limited time for provision of goods – allows us to highlight three situations: complete coincidence of the lists, when all the Arctic areas are simultaneously transport-isolated, and the entrepreneurs here fully embody the «interior» Arctic specificity; partial coincidence when inside the Arctic area one can distinguish districts with limited time for delivery of the «genuine» Arctic in terms of all effects for business activities, and areas that are within the national road network, and therefore the conditions for business activities here bear features of both the Arctic zone and « the mainland «; and the situation when the Arctic regions and the seasonal navigation areas do not coincide at all – the Arctic entrepreneurs of these regions are affected by both Arctic and continental factors (Table 4).

Table 4. Two cases of Arctic entrepreneurship – in the «island» and «quasi-mainland» Arctic

	Arctic districts (Presidential Decree 2019)	Districts of Limited time of Delivery (The list of areas, 2016)
Complete coincidence		
NAO	All districts	All districts
CHAO	All districts	
Republic of Sakha-Yakutia	Allaikhovskiy, Anabarskiy, Bulunskiy, Nizhnekolymskiy i Ust'-Yanskiy; Abyyskiy, Verkhnekolymskiy, Verkhoyanskiy, Zhiganskiy, Momskiy, Olenekskiy, Srednekolymskiy i Eveno-Bytantayskiy districts	All districts and settlements excluding cities of Aldan, Tommot, settlements of Leninskii, Nizhnii Kuranakh of Aldan district and city of Neryungri
Krasnoyarsk krai	Territories of the urban district of the city of Norilsk, Taimyr Dolgan-Nenets Municipal District, Turukhansk District	Boguchansky, Yenisei, Kezhemsky, Motyginsky, North-Yeniseysky and Turukhansky districts; city of Igarka and Norilsk, Taimyr Dolgan-Nenets Municipal District
Partial coincidence		
Murmansk oblast	All districts	Chavanga, Chapoma, Tetrino and Pyalitsa villages of the Terek district; Kanevka, Krasnoshchele and Sosnovka villages of the Lovozero district; Ostrovnoy city, Lumbovka, Korabelnoye, Svyatoy Nos, Cape-Cherny, Mayak-Gorodetsky and Terek-Orlovsky Mayak of Ostrovnaya closed administrative territorial formation
Archangelsk oblast (without NAO)	City of Arkhangelsk, Mezensky District, Novaya Zemlya, City of Novodvinsk, Onega District, Primorsky District, City of Severodvinsk	Verkhnetoemsky, Lensky, Leshukonsky, Mezensky, Pinezhsky, Primorsky and Shenkursky districts
YANAO	All districts	all districts and settlements, with the exception of Labytnangi, Muravlenko, Novy Urengoi and Noyabrsk cities
No matches		
Republic of Karelia	Belomorsky, Loukhsky and Kemsky districts	Kalevala National District; Valaam village of Sortavalsky urban settlement
Republic of Komi	Vorkuta City District	Other districts

[1] At the same time, of course, we are well aware of the non-geographical factors for the development of entrepreneurship, to which dozens of works of our colleagues are devoted: the general level of education of the population, the presence of a university, the size of local demand, the culture of entrepreneurship, the level of crime and the strength of legislative protection of property rights, the share of migrants in the population etc., but here we shall not concentrate on them.

The dichotomy of the two Arctics – island and quasi-mainland in terms of the complex of natural and socio-economic conditions for entrepreneurship is so information-intensive that it fully captures the features of natural extremity: the Murmansk oblast, the Arkhangelsk oblast and three coastal regions of the Republic of Karelia, in which the Arctic territories are connected by year-round transport network with the «mother» Russia and are more comfortable in natural and climatic conditions. Polar Vorkuta and the major cities of the Yamalo-Nenets autonomous okrug occupy an intermediate position: there is a constantly working railway network, which partially mitigates the effects of natural discomfort factors. On the other hand, the Nenets, Chukotka autonomous districts, the Arctic of Yakutia and the Krasnoyarsk krai are territories of the ultimate embodiment of the Arctic specificity and the strongest operation of all Arctic factors (both negative and positive) in the development of entrepreneurship.

The second factor is the presence or absence of a large resource corporation that mines the hydrocarbon or mineral deposits. Of course, a specific role is also played by the concrete phase of exploitation of the resource object – the stage of «fountains», stabilization or extinction. It determines the real possibilities of the corporate structure to pick up the functions of local life support in the Arctic cities and towns of the company's presence.

The results of our surveys of single-industry cities of Yamal and Ugra, Norilsk, convince us in the duality of the impact of corporate structures of the Arctic on the development of local entrepreneurship. On the one hand, the level of wages in companies is often such that it leads to a total staff shortage in all other areas of activity, including business: energetic and entrepreneurial local talents are massively flowing into the corporate contour from the local budget and business sector. In a number of cases, a personnel «desert» arises next to a large company. In addition, often large companies that become structure-forming for the local economy of remote Arctic villages and cities, take on the functions of maintaining the local life-support systems – trade, personal services, transport, energy, etc., and therefore «take out» these areas of traditional entrepreneurial activity from local businesses.

On the other hand, large companies also create a field of opportunities for small businesses, and, what is important to emphasize, not only in the form of orders and contracts for small business (as follows from the classic courses in business schools), but also due to the fact that the employees in the resource corporations very often create their own «part-time» business, in which they realize their creative plans more fully and boldly than in their corporate company. There are many such examples, for example, in Norilsk, where many employees of the Norilsk Combine are at the same time businessmen in private carriages (taxis) or in the personal services sector.

The third factor in the development of Arctic entrepreneurship is institutional. To what extent are the traditional areas of entrepreneurial activity — trade, services, and others — «closed down» by state or municipal enterprises — in those cases where there are no large corporate structures nearby that could also take up these tasks?

The statistics confirm that while the role of state and municipal structures in general in the Arctic is higher than in Russia, specific variations of the situation are very strong (Table 5). At one extreme is Chukotka autonomous okrug, in which there is no large integrated corporate structure and therefore the state and municipalities are forced to accept many of the tasks of daily life-support for themselves, which is impossible for a business to carry out profitably. But this means that for entrepreneurship there remain narrower spheres of industrial activity, and construction or transport.

At the other extreme is the Murmansk oblast, in which relatively comfortable climatic and socio-economic conditions allow many daily life support functions to be performed by local businesses or business from neighboring regions of the temperate zone. It is not surprising that here the share of enterprises in state and municipal ownership is minimal.

It turns out that inside the «united» Arctic, in fact, the business is in very different conditions: from the Murmansk oblast, which has a permanent land connection with «mainland» Russia, proximity to large urban European centers, relatively comfortable climatic conditions, the presence of large resource corporations in the mining industry and a wide free field for the activities of entrepreneurs, and to the Chukotka autonomous okrug, which is extremely remote from the main settlement zone; has more severe climatic conditions; and very localized activity of resource corporations (due to their rotational scheme of development), the benefits of which are «spilled» only on the nearest national village; as well as a very narrow scope of activity that is attractive for entrepreneurs.

A separate factor in the development of entrepreneurship is state support measures. Sociological surveys of dozens of Yamal entrepreneurs, undertaken in 2016-2017 as part of work under grant 16-46-890363 r_a «Arctic entrepreneurship as a factor of sustainable development of the Yamalo-Nenets autonomous okrug» confirmed the presence of two groups of entrepreneurs: of the frontier type, who consciously refuse any state support, not wanting to lose independence in the management of their business (they are often located in more isolated and remote areas of the Arctic); and entrepreneurs who are very «creative» in adaptation of their activities under the current rules and norms of state and municipal support. According to the analysis of regulatory legal acts of nine Arctic regions and a dozen cities according to Consultant + database, it is possible to isolate unified measures that are indifferent to Arctic specifics, and measures that take into account the peculiarities of the Arctic economy. Let us dwell on the latter.

Table 5. The share of enterprises in state and municipal ownership in the total number of enterprises, 2016, %

Arctic	The share of enterprises in state ownership	The share of enterprises in municipal property
CHAO	20.3	20.4
NAO	19.3	8.8
YANAO	3.9	8.2
Murmansk oblast	4.0	6.0
Reference: Russia	2.3	4.3

Source: Rosstat, 2016-2019. Available at: http://www.gks.ru/free_doc/new_site/region_stat/arc_zona.html Accessed 15 July 2019

Dozens of «Arctic» measures of state and municipal entrepreneurship support programs can be grouped into five areas, which clearly reflect the features of economic activities in the Arctic and confirm our conclusion that entrepreneurs in the Arctic are flesh and blood of the Arctic economy itself.

Firstly, these are subsidies for organizing the northern provision of food and consumer goods to hard-to-reach localities in the form of compensation for a part of transport expenses to entrepreneurs, compensation for a part of expenses on paying interest on bank loans for organizing northern delivery. Within this direction, it is possible to allocate separately subsidies for the delivery of goods (including firewood) to trading posts, in the place of compact residence of the indigenous peoples of the North. Any merchandise movement in the Arctic is associated with increased friction due to very poor transport conditions and it is natural that entrepreneurs who are willing to work in the northern supply market (and this is primarily the European Arctic, because in the Asian part this field of activity is even less attractive for businesses and it is often performed by state, municipal, and corporate structures), and support measures are proposed.

Secondly, these are subsidies (partial reimbursement of transportation costs, etc.) for export, that is, support for the promotion of Arctic products (handicrafts, crafts, reindeer meat) to processing sites and to final markets. The most important limitation in the activity of an Arctic entrepreneur is a small local sales market, which does not allow for the effect of economies of scale on the volume of operations.

«The environment in the city is comfortable for starting a business, the Norilsk residents say. However, when a business crosses a two-year threshold, an entrepreneur does not always understand where to go next». (Present and future of SME in Norilsk 2019).

Even in Norilsk, large by the Arctic standards, local business is quickly sticking to the limits of growth due to limited demand. For many reasons, it is more difficult for an Arctic entrepreneur to move from the internal to the external market than for the mainland. They are more delineated here!

In developed areas, the internal and external markets often overlap, the first organically flows into the second: you have achieved recognition in the local market and become competitive in neighboring markets. But precisely because the «neighboring» markets in the Arctic are far away, to win in the local market, one need completely different qualities and skills than recognition in the outside world.

The domestic market of an Arctic entrepreneur consists of three sectors of the Arctic economy: corporate, state and traditional, on each he provides his services (for example, for the public sector structures or a resource company – services of trade, construction, transport), and inside the traditional one can develop independently in the form of private reindeer herding, for instance. Plus a variety of services to local households. Success requires brilliant implicit knowledge of local specifics, the local community, the local environment and resource potential.

On the other hand, success in neighboring markets requires a completely different knowledge of logistics, tastes of consumers in large cities, competitors' strategies, etc. And the «first» implicit knowledge here doesn't help succeed at all. Therefore, the role of state support for entrepreneurs in the Arctic, to be strengthened in external markets, is comparatively more significant for them than for mainland counterparts.

Transport and energy are two bottlenecks that drive a wave of high costs for an Arctic entrepreneur. Therefore, the third direction is to subsidize the cost of heat and power, including, for example, in greenhouses, marine and fishery bases, the operation of refrigeration units of processing points, as well as ptomoting energy efficiency of entrepreneurs. Unfortunately, until now the support measures do not encourage replacement of the supply of fuel with their local production (this project is described in detail in Zamyatina, Pilyasov 2019).

The Arctic entrepreneur, unlike his mainland counterpart, is super mobile. That is why the support measures include subsidies for the arrangement of new places of economic activity (the fourth direction) for hunters, for reindeer-herding brigades, for young representatives of small indigenous peoples, who start as entrepreneurs in traditional economic activities.

Finally, the fifth specific Arctic area of business support is grants in the form of subsidies for material and technical equipment and the development of the traditional economic activities of the indigenous minorities of the North. This assistance is aimed at ensuring that from purely subsistence reindeer herding and traditional activities become partially commercial, that is, they would find nearest local markets.

So far, a very small place is occupied by «intellectual subsidies» aimed at acquiring new, including specific Arctic, knowledge by local entrepreneurs: for example, subsidies for conducting geobotanical surveys of reindeer pastures and developing projects for on-farm land management of reindeer pastures; subsidies for energy audits at small businesses; compensation for the payment of training of representatives of small indigenous peoples for the safe handling of weapons, etc.

It is necessary to significantly more actively promote such knowledge spillovers between the Arctic entrepreneurs themselves, between the local and temporary research specialists of the scientific and educational departments and the Arctic entrepreneurs. World experience shows that today, every small business needs periodic «injections» of new knowledge to strengthen its resilience. And this is even more important than the support in the Arctic of creating objects of innovative infrastructure in the form of, for example, business incubators, which in the low-density environment of the Arctic can be the next «cathedrals in the desert».

Murmansk oblast: the contradictions of the expected and the real – how favorable factors close the business opportunities

For all factors of entrepreneurship development, the Murmansk oblast is the most favorable within the Arctic zone: a simple transport and logistics scheme for the delivery of goods from Central Russia, relatively comfortable environmental conditions for entrepreneurial activities (especially when compared with the regions of the Asian Arctic), comfortable economic conditions due to the many niches and markets, potentially attractive for entrepreneurship; neighbors advanced in terms of entrepreneurship development from the south (Leningrad oblast and St. Petersburg) and from the west (provinces of Norway and communes of Finland).

Therefore, let us consider this polar case (the other – the negative pole according to the conditions of business in the Arctic – is the Chukotka autonomous okrug) in more detail. Here are the largest small enterprises in the Arctic: in other polar regions there are relatively more contract workers, but the size of the company itself is smaller. Here, the largest cumulative turnover of products of individual, small and medium enterprises, however, their number per 1000 people, as a rule, is less than that of their neighbors in the Arctic (Table 6)

By the volume of investments in small business, more than half takes fishery. And this is not surprising – investors in the field of fish processing in coastal municipalities receive state preferences and support. The fish business, due to the very high wages of the employee, significantly raises the

Table 6. Comparison of the situation in the development of entrepreneurship in the Arctic, Arctic-northern and multi-latitude regions of Russia

Regions in the Arctic zone	The share of part-timers in small enterprises (without micro), %	Increase in the number of small and medium-sized businesses, % of the previous year		Number of individual entrepreneurs per 1000 residents – only the Arctic territories	Turnover of products (services) produced by small enterprises, including micro enterprises and individual entrepreneurs, bln roubles
	2018	2015	2016	2018	2017
Entirely Arctic					
Murmansk oblast	7	16.0	9.3	22.9	214.7
YANAO	15	-18.5	4.2	30.6	170.7
NAO	20	-4.7	-6.4	28.3	7.8
CHAO	15	33.2	-29.0	25.4	9.0
Arctic-Northern					
Archangelsk oblast without NAO	12	-9.7	-3.0	28.2	343.1
Republic of Komi	13	-4.7	-10.3	24.9	213.4
Republic of Karelia	9	14.5	-7.8	18.2	159.4
Multi-latitude					
Krasnoyarsk krai	5	-7.9	-3.7	26.3	1005.0
Republic of Sakha-Yakutia	18	-4.0	-9.6	32.4	317.0

Sources: Rosstat, 2016-2019. Available at: http://www.gks.ru/free_doc/new_site/region_stat/arc_zona.html Accessed 15 July 2019; collections of «Small and medium-sized enterprises in Russia» https://www.gks.ru/folder/210/document/13223

average earnings of the employee in the entrepreneurial sector throughout the Murmansk oblast.

If we consider the development of business in other sectors, it turns out that the Murmansk oblast, with all its favorable conditions, loses to its Arctic neighbors. Mining small business in the core for the whole oblast economy mineral resource complex practically does not develop, the number of firms here is just scanty and not comparable even with Chukotka autonomous okrug, which is in much more severe conditions.

In the small business of the processing profile, the Murmansk oblast is ahead of the neighboring Arkhangelsk oblast in terms of turnover, although it is inferior in its total number. Due to this segment on the «energy» of import substitution in 2015-2016 in the region there was a general increase in the number of small and medium enterprises. In the rest industrial production small business of the Murmansk oblast is inferior to the Yamalo-Nenets autonomous okrug and the Republic of Karelia.

A natural question arises: what caused the gap in the expected favorable conditions (the best in the Arctic) and the real situation with the development of local entrepreneurship? Indeed, the potential area for the development of entrepreneurship in the Murmansk region is extremely broad, especially in comparison with other Arctic regions, but the entrepreneurial activity itself is very moderate.

It seems that the reasons are enclosed in geography, this time not zonal (that is, physical geography), but economic geography. Paradoxically, it is the favorable factors for the development of entrepreneurship of the Murmansk oblast, that, on the contrary, it inhibits!

The Murmansk region is so comfortable compared to the rest of the Arctic neighbors that many entrepreneurs here prefer not to bear the burden of northern costs and make their lives even more comfortable by relocating their business to the south, to areas where there are no legally enshrined northern guarantees or compensations or where they are essentially more modest than in the Murmansk region. That is why many compensatory effects typical for the entire Arctic zone do not work in the Murmansk region: for example, the active involvement of employees under contracts. Business simply votes with its feet, legally relocated from the oblast to its neighbors to the south, while continuing to work in the markets of the Murmansk region.

Similar effects, only vice versa, were observed during the Soviet era, when new production enterprises were attracted to the southern edge of the North zone as a magnet, because it was easier to find workers from all over the Soviet Union due to attractive northern regional coefficients and seniority allowances. So now the effects of «the opposite» are observed in the entrepreneurship of the Murmansk region, when it is more profitable to legally be deployed to the south in order not to pay northern workers' compensation. And this effect of the economic and geographical situation determines the modern underdevelopment of Murmansk entrepreneurship.

DISCUSSION AND CONCLUSIONS

The very study of the topic of the Arctic entrepreneurship highlights the transition of the global Arctic from the managed to entrepreneurial economy. This trend, indicated by researchers another 15-20 years ago (Audretsch D. et al. 2001; Nijkamp P. et al. 2002; Audretsch D. et al. 2004; Baptista R. et al. 2007; Erdos Katalin et al. 2010; Feldman M. et al. 2012), is gradually affecting the Arctic.

Quite recently it seemed that entrepreneurship is not about the Arctic at all and that here this phenomenon is simply impossible due to numerous natural and social barriers. But after all, also in the industrial era, electrification and the conveyor gradually reached the tents of nomadic reindeer herders, although at its start it seemed that this was impossible.

We are on the verge of extensive and comprehensive research on the phenomenon of Arctic entrepreneurship. The need for them stems from at least two reasons: 1) there is an acute shortage of theoretical knowledge on how the development of entrepreneurship in the Arctic submits, to what laws and patterns. The practice here is far ahead of the theoretical understanding of an already phenomenon in play; 2) It is impossible to use research groundwork for the temperate zone here because Arctic entrepreneurship, by its nature, drivers, structure, differs significantly from the «mainland».

For example, what is called ethnic entrepreneurship in Europe and is a small business of recent immigrants, in the Arctic is a business; on the contrary, of the first settlers of this land that is, indigenous small peoples of the North. The textbook presentation is that small business is the most important agent of competition and is always under its own positive pressure. But in the Arctic, on the contrary, small business often uses the effects of a monopoly position, and competition in small and autonomous (dispersed) markets can be simply destructive for the local economy. In the classic works of colleagues from European countries (Audretsch D., Thurik A. 2001 and many others, above mentioned), the stabilizing role of small business is often mentioned. But in the Arctic, on the contrary, entrepreneurship exploits instability in its own interests and is itself the brightest agent of instability, which does not weaken it, but exaggerates it.

One can recognize the phenomenon of Arctic entrepreneurship as «anti-mainland» in nature. But on the platform of the zonal approach, the researcher will not be discouraged by these features and will be able to cope with them and constructively interpret them. This opens up opportunities for the leadership role of geographers in the interdisciplinary research teams studying Arctic entrepreneurs. A proven comparative method of analysis (business of the Arctic and the North, the Arctic and the mainland, etc.) can give here truly constructive results.

«Secret» research method let us understand the nature of the Arctic entrepreneur – to see its essential connection with the indigenous peoples of the North and their strategies to adapt to the extreme Arctic environment, recognize the commonality of the Arctic economy and Arctic entrepreneurship, make sure that taken from nature techniques and technologies are useful and efficient for Arctic entrepreneur.

Nine polar Russian regions can be differentiated into a group of entirely Arctic, Arctic-northern and multi-latitude regions, each of which, along with common features, has its own specific features of local small business development. But common to all of them is the entrepreneur's response to the challenges of northern costs, which consists primarily in unparalleled organizational flexibility, super-mobilty and a brilliant understanding of the place's properties – the local community, the local market, local resources, etc.

The main specific Arctic factors for the development of entrepreneurship, as confirmed by our expeditionary surveys of small businesses in Yamal, are in the geographical location of the main activity sites – island or quasi-mainland, the presence or absence of a large resource corporation and a specific stage of development of the main natural asset (growth, stabilization, decline), institutional structure of the local economy (to what extent are key daily services – trade, transport, household services nationalized) i.e. whether there is a niche for small businesses.

Within the Arctic zone of Russia itself, the situation with regard to the factors of entrepreneurship development is sharply heterogeneous. At one of the most favorable pole is the Murmansk oblast, on the other – the most uncomfortable – the Chukotka autonomous okrug. However, the presence of favorable prerequisites by itself does not guarantee the active development of entrepreneurship. For example, the Murmansk oblast, with all its favorable external factors, is paradoxically not the leader in the development of Arctic entrepreneurship. On the contrary, it is precisely for it that the maximum gap between potential and real state in the development of entrepreneurship is characteristic.

ACKNOWLEDGMENTS

This article is made possible under the generous support from the RSF grant «Eurasian Vectors of Russia's Maritime Activity: Regional Economic Projections» (project № 19-18-00005).

My gratitude to two anonymous reviewers who helped to improve the original version of this article.

REFERENCES

Audretsch D., Thurik A. (2001). What's new about the new economy? Sources of growth in the managed and entrepreneurial economies. Industrial and Corporate Change, 10(1), 267-315.

Audretsch D., Keilbach M. (2004). Entrepreneurship and regional growth: An evolutionary interpretation. Journal of Evolutionary Economics, 5, 605-616.

Baptista R., Preto M.T. (2007). New Firm Formation and Employment Growth: differences across Regions and Types of Start-ups. ERSA Congress 2007 in Paris.

Erdos Katalin, Varga Attila. (2010). The Academic Entrepreneur: myth or reality for increased regional growth in Europe? ERSA 2010 Congress in Jonkoping.

Feldman M., Zoller T. (2012). Dealmakers in place: Social capital connections in regional entrepreneurial economies. Regional Studies, 1(46), 23-37.

Krupnik I.I. (1989). Arctic ethnoecology. Moscow: Nauka. (In Russian).

National project: Small and medium entrepreneurship (2019). [online] Available at: www.government.ru/info/35563/ [Accessed 15 July, 2019]. (In Russian)

Nijkamp P., Stough R. (2002). Editorial introduction: Special issue on Entrepreneurship and Regional Economic Development. The Annals of Regional Science, 36, 369-371.

Northern Opportunity. (2017). Alaska's Economic Development Strategy 2017-2022. [online] Available at: www.northernopportunity.com/wp-content/uploads/2017/10/Final-Statewide-CEDS.pdf [Accessed 15 July 2019]

Petrov A.N., BurnSilver S., Chapin S.F., Fondahl G., et al. (2017). Arctic Sustainability Research: Past, Present and Future. Routledge.

Pilyasov A., Zamyatina N. (2016). Arctic entrepreneurship: conditions and opportunities for development. Arctic: environment, economy, energy, 4, 4-15. (In Russian).

Pilyasov A. (2009). Zonal factors of post-industrial transformation (the case of the Russian North). Arctic idea. April. 40-41. (In Russian).

Rallet Alain, Torre André. (2009). Temporary Geographical Proximity for Business and Work Coordination: When, How and Where? [online] Available at: www.researchgate.net/publication/253129792_Temporary_Geographical_Proximity_for_Business_and_Work_Coordination_When_How_and_Where [Accessed 15 July 2019].

Present and future of SME in Norilsk (2019). [online] Available at: www.ttelegraf.ru/news/title/v--norilske-predstavili-itogi-kompleksnogo-issledovaniya-malogo-i-srednego-biznesa- [Accessed 15 July 2019] (In Russian).

Presidential Decree of May 13, 2019, 220 «On Amendments to the Decree of the President of the Russian Federation of May 2, 2014 N 296» On Land Territories of the Arctic Zone of the Russian Federation» (In Russian).

Rosstat, statistics on the Arctic zone, 2016-2019. [online] Available at: www.gks.ru/free_doc/new_site/region_stat/arc_zona.html [Accessed 15 July 2019].

Russian Arctic: contemporary paradigm of development. (2014). Ed. by A. Tatarkin. Saint Peterburg. Nestor-istoriya. (In Russian).

Siberian economy through the lenses of the latitudinal zones. (1984). Ed. A. Granberg. Novosibirsk. Nauka. (In Russian).

Small and medium-sized enterprises in Russia. Statistic bulletin. 2010-2018. [online] Available at: www.gks.ru/folder/210/document/13223

The Arctic Frontier. (1966). Ed. by R. St. Macdonald. University of Toronto Press.

The list of areas of the Far North and equated localities with limited terms of delivery of goods (products) (approved by the Government of the Russian Federation on May 23, 2000, 402). With changes and additions from: 27.09.2001, 31.03.2009, 6.12.2016. (In Russian).

The New Arctic. (2015). Eds. Evengard B., Larsen J.N., Paasche O. Springer.

UNESCO. (2009). Climate change and Arctic Sustainable Development: scientific, social and educational challenges. UNESCO: Paris.

Vinokurova U., Yakovets Yu. (2016). Arctic circumpolar civilization. Novosibirsk: Nauka. (In Russian).

Zamyatina N., Pilyasov A. (2019). How do we «equip» the Arctic. Moscow. ANO IRC. (In Russian).

Received on July 17th, 2019 Accepted on January 21th, 2020

P.Ya. Baklanov[1*], M.T. Romanov[1]
[1] Pacific Geographical Institute of the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia
***Corresponding author:** pbaklanov@tigdvo.ru

TRANSBOUNDARY REGIONS OF EAST ASIA: GEOGRAPHICAL AND POLITICAL PRECONDITIONS AND LIMITATIONS TO LONG-TERM DEVELOPMENT

ABSTRACT. The contradictory integrity of globalization and regionalization processes in the modern world has been embodied in the formation of specific spatial areas – multiscale cross-border regions, whose functioning and development are determined both by the interactions between neighboring countries and by the totality of external geopolitical and geo-economic circumstances. The article is devoted to the factors and features of cross-border processes and socio-economic development within one of the largest and most dynamic structures of modern Eurasia – the Greater Macro-Region of East Asia, embracing the northeastern and eastern territories of Russia, eastern China, Japan, and both Korean republics, Vietnam, and a number of other countries facing the seas of the northwestern Pacific. The integrity of this vast and very heterogeneous macro-region ranging from Chukotka to the Philippine Sea is based on relatively stable cross-border relations, which, in turn, are one of the determinants of these territories' development. The role of geographical prerequisites (geographical location, climatic conditions, natural resources of land and sea) and geopolitical factors (geopolitical location and cross-border features) in the long-term development of this macro-region is assessed. It is shown that both favorable and negative prerequisites are associated with the cross-border nature of the integrated geosystems, including the marine ecosystems. Various types of cross-border regions with two-, three-, and four-link territorial segments belonging to different countries have been identified. The geopolitical potential of countries and regions is assessed, and the zones of geopolitical tension are revealed.

KEY WORDS: transboundary region, East Asia, geographical, geopolitical factors, a geopolitical position, contact structures, long-term development, territorial segments

CITATION: P.Ya. Baklanov and M.T. Romanov (2020) Transboundary Regions Of East Asia: Geographical And Political Preconditions And Limitations To Long-Term Development. Geography, Environment, Sustainability, Vol.13, No 1, p. 57-63
DOI-10.24057/2071-9388-2019-96

INTRODUCTION

East Asia is a large, meridionally extended macro-region situated at the junction of the largest continent (Eurasia) and two oceans (the Arctic and the Pacific). In literature, there is no strict definition of the East Asia composition. The macro-region is defined differently among the countries of China, Japan, Republic of Korea and DPRK (Pacific Russia 2017; Terada 2006). As such, the Russian Far East and Southeast Asian nations are not always included in this region. More often, Northeast Asia is identified as a composition of Northeast China, the Russian Far East, DPRK, Republic of Korea and Japan (Pacific Russia 2017; Tulokhonov 2014; Womarck Brantly 2014). A.G. Druzhinin proposes a new approach to structuring of Eurasia based on a geo-ethnocultural systems conception (Druzhinin 2016). The Asiatic group of such systems is identified without concretization of their composition.

At present, the largest region – the North Pacific, including the USA, is identified but its full configuration is not always consistent (Geopolitical potential 2010; Pacific Russia 2017). In our opinion, East Asia is a more integral region. Taking into account similarity of geographical positions, we included the territorial subjects of the Russian Federation (RF) adjacent to the Pacific Ocean, the eastern portions of China, and all countries entering the seas and the North Pacific in East Asia. This macro-region extends over nearly eight thousand kilometers from north to south (Fig. 1).

The specific features of East Asia include great differences and the contrasts of its particular territories and regions if considered from different perspectives. So, from the environmental standpoint, the macro-region houses virtually all natural zones of the Earth: from the ice Arctic deserts of North Siberia and Chukotka to the tropical forests of Indonesia. The macro-region is characterized by the most diverse combinations of the natural resources of the land and adjoining seas and oceans. In East Asia, the unique diversity of civilizations and cultures was formed including the unique Chinese, Korean, Japanese, and paleo-Asiatic ones, as well as the Russian northern indigenous nations represented by the Yakuts, Chukchi, Koryaks. Finally, in the countries of East Asia, the strong diverse political systems have emerged: presidential republics (Russia, Republic of Korea etc.), constitutional monarchies (Japan, Thailand), the countries of socialist orientation (China, Vietnam), and DPRK, a country with the communist political regime).

The differences in the levels of the social and economic development in East Asian countries are also considerable. Highly-developed countries (i.e. Japan, Singapore and Republic of Korea), dynamically developing ones (i.e. China, Malaysia, Vietnam and Russia) and less-developed states (i.e. Thailand, Laos, Myanmar, Indonesia and Philippines) are found side by side.

The critical characteristic unifying the countries and the regions of East Asia derives from their transboundary nature. Thus, a great part of the Russian Far East and Northeast China

Fig. 1. The transboundary regions of East Asia

are included into the transboundary basins of the Chukchi, Bering, Okhotsk, Japan, and Yellow seas. In addition, there are several transboundary river basins in the East-Asia region including the Amur, the Ussuri, and the Tumen. The basins of the East China, South China, as well as the Philippine seas are also transboundary. According to our estimates, the transboundary regions of East Asia include practically 100% of its territory. Therefore, the transboundary phenomenon is the most important factor for East Asia combining its regions and countries into one integrated transboundary space.

The environmental, social-economic, cultural, and political characteristics of the macro-region combine to form a pattern that is diverse and unique. The dynamics of the separate parts and structures of this space are determined in many respects by the effect of the geographical and geopolitical factors. These factors also have different impacts on the long-term development of the countries and regions of East Asia. Such separate factors create the favorable prerequisites for the regional development, while the others restrict them. As a whole, the combined impact of the geographical and geopolitical factors and their spatial differentiation on the long-term development of the East Asia macro-region has not been practically investigated. Some works were devoted to Northeast Asia (Baklanov 2000; Geopolitical potential 2010; Larin 2016; Nature management 2005; Tkachenko 2009; Womarck Brantly 2014), the Far-Eastern region of Russia (Baklanov et al. 2016; Baklanov et al. 2015; Larin 2016), Northeast China, Vietnam (Baklanov et al. 2017; Larin 2014; Nature management 2005), the countries of Southeast Asia (Asia-Pacific 2010; Baklanov et al. 2017; Pacific Russia 2017). However, the transboundary nature of the macro-region presents additional complexity to understanding its long-term development. Therefore, the task of this paper is to make an aggregated assessment of the geographical and geopolitical factors and their role in the long-term development of the macro-region.

RESULTS AND DISCUSSION

The main geographical and geopolitical factors

The geographical factors exerting a significant influence on the long-term development of the macro-region include their geographical and economic-geographical positions, the effect of the natural-climatic conditions due to their spatial differentiation, as well as a natural-resource potential in the form of the territorial and aqua-territorial combinations of natural resources.

The most important characteristic of the geographical position of East Asia resides in the fact that its greater part is a zone of the global contact among geographical structures (Baklanov 2000). Firstly, this is a junction between Northeast of Eurasia and the Pacific and Arctic Oceans resulting in different interrelations between land and the sea in the natural-ecological, natural-resource and economical spheres. Secondly, this is the place where Russia borders the world's largest economies (the USA, China and Japan). The contact structures serve as the basis for performing various functions, including utilization of different natural marine resources supported by inshore infrastructure in combination with terrestrial resources. The proximity and dynamic development of the countries of the macro-region generate the external-economic interactions through the joint partnership of land-based and marine transport systems, which in turn contribute to the extension of the market spaces.

The southern portions of the Russian Far East, the southeast portions of China as well as the areas of Vietnam, Singapore and Malaysia have the most favorable economic-geographical position from the viewpoint of interfacing with the external transportation routes and proximity to the developing markets.

Due to a variety of natural conditions, the most favorable opportunities for sustainable development are found in the southern areas of the Russian Far East and East China. Low temperatures, permafrost and other extreme conditions become the limiting factors in the central and northern portions of the Russian Far East. In the areas of East Indo-China, the tropical extreme natural processes, such as typhoons, floods and tsunami are frequent, although many of these regions are most favorable for cultivation of rice. Local populations have generally adapted to these natural-climatic disturbances (Baklanov et al. 2017).

For some areas of East Asia and adjacent seas, one can identify both land-based and aqua-terrestrial resources. Their similarity and differences for different areas, as well as their generalized natural-resource potential were estimated. The regions of the Russian Far East, Northeast China, Malaysia, Indonesia and Myanmar hold the highest potential. As a whole, one can note the great complementarity of land-based natural resources with marine natural resources, as well as some complementarity of natural resources in certain neighboring areas, like between the Russian Far East, Northeast China, Japan, Republic of Korea and DPRK (Nature management 2005; Tkachenko 2009). Such complementarity of land-based natural resources with those of the marine origin, as well as natural resources between neighboring countries, which is typical in the regions and countries of East Asia, create favorable conditions for the long-term sustainable development.

It should be noted that the regional contrasts within the transboundary space are generally large. Here, one finds neighboring countries vastly differing both by their size and development level (Table 1). The world's largest countries are China (18.7% of global GDP in 2018), Japan (4.1%) and Russia (3.1%). Smaller countries include Indonesia (2.6%), Republic of Korea (1.6%), Malaysia (0.74%), while Singapore is a small but highly-developed country.

As the Table shows, from 1990 to 2018, many countries of East Eurasia have demonstrated extremely high rates of development: China has increased its GDP by 2,162%, Myanmar by 1,727%, Laos by 1,013%, Vietnam by 1,005%, Malaysia by 709%, Indonesia by 542% and Philippines by 491%. This is in contrast to many other regions of the world, where increases in the GDP have been lower. For example, the GDP during this same period has increased in the EU by 227%, the USA by 244%, Russia by 236%, and for the world taken together – by 370% (gtmarket.ru 2019; fincan.ru 2019). Among the geopolitical factors essential in the long-term development of the macro-region, we consider the geopolitical position of East Asia as a whole and its separate regions, as well as its transboundary phenomenon to be an important specific characteristic of the geopolitical position of the macro-region.

The essential features of the geopolitical position of this macro-region are determined by Russia and China, the eastern areas of which belong directly to East Asia. Russia and China have different "weights" and political systems but, in recent years, the strategic partnership was established between them.

The crucial component of the geopolitical position of the macro-region is the proximate neighborhood with the USA and strategic partnership between the USA, Japan and Republic of Korea that enhances geopolitical contrasts. The USA is one of the three largest countries by its geopolitical potential in the world (15.2% of the global GDP). It neighbors Russia and East Asia across the Bering Strait. The essential

Table 1. Grouping of the East Eurasia countries and regions according to their economic potential and growth rate (according to data for 2018)

Groups of countries by GDP volumes	Names of countries	GDP by IMF estimate at PPP, billion $	Place occupied by country in the IMF list by GDP (of 192)	Growth of GDP (PPP) in the period of 1990 to 2018 by WB data (%)
1. Large	1. China	25,270	1	2162
	including its eastern parts	14,361	…	…
	2. Japan	5,594	4	127
	3. Russia	4,213	6	236
	including Pacific Russia	46.7	…	…
	4.Indonesia	3,495	7	542
	5.Republic of Korea	2,136	14	489
2. Medium	6. Thailand	1,320	19	441
	7. Taiwan	1,251	22	…
	8. Malaysia	999	26	709
	9. Philippines	953	28	491
	10. Vietnam	710	34	1005
	11. Singapore	566	37	728
	12. Myanmar	344	52	1727
3. Small	13 DPRK	66.7	…	…
	14. Cambodia	70.5	…	…
	15. Laos	53.7	…	1013
	16. East Timor	6.8	…	…

Compiled by: (fincan.ru 2019).

specific geopolitical role is performed by Japan. Being territorially a part of East Asia, this country has the closest and large economic and military-political ties with the USA. In this "team", Japan can be simultaneously considered an external geopolitical factor in East Asia. Its vicinity to the countries of Oceania and Australia, varying by political systems, is also of great geopolitical importance for East Asia.

In Eastern Asia, significant differences between countries and regions are reflected in their geopolitical potential (Table 2). On the one hand, China, Russia, and the USA are the largest countries of the world in terms of their geopolitical potential; and to perhaps a similar extent so is Japan and the Republic of Korea. However, there are a number of countries that have medium and small potential values.

Even greater contrast is reflected by the relative characteristics of the development of the territories including those which belong to the transboundary regions (Table 3). On the one hand, these estimates reflect a cumulative effect of the geographical factors on the development of the territories and, on the other hand, they reflect, perhaps to a larger extent, their geopolitical stability.

Generally, taking into consideration a transboundary phenomenon, the sea water area crossed by national frontiers, and the portions mainland of the territories adjacent to it to some extent become the influence zones of neighboring countries and the zones of intersection of their mutual geopolitical interests. The estimates of the natural-resource potential owned by one country (in the coastal areas and marine exclusive economic zone) and, especially, stability and efficiency of the national types of resource management in the neighboring countries become interrelated and interdependent. This refers to all transboundary regions. According to our research (Baklanov et al. 2008; Baklanov et al. 2015; Geopolitical potential 2010), a transboundary phenomenon creates both favorable and negative conditions for the long-term development. So, when reaching the high standards of natural resource management in a neighboring country, the space and efficiency of the renewable resources regeneration (fish, forest, land etc.) are extended.

Under otherwise equal conditions, the integrity of the transboundary geosystems (including the marine ecosystem) creates objective opportunities for the long-term international cooperation of the countries within the transboundary region. This contributes to the rise of a sustainable natural resources management system, extension of the market space, and, as a result, long-term viable development of the region. At the same time, a risk of transboundary transfer of technogenic pollutants from one country to another, regular violations of the natural resources management standards in one of the countries, and lack of the sufficient information can impair the achievement of sustainable development in the transboundary region as a whole. All of the above-mentioned points are of great relevance for the countries and groups of countries included in various transboundary regions of East Asia.

Depending on how many countries border a transboundary sea (river), one can carry out specific zoning programs. This will determine a overlapping of influence zones and geopolitical interests of two, three or more countries within one transboundary region. Thereupon, we have identified the following types of the transboundary

Table 2. Main characteristics reflecting the elements of the geopolitical potentials of the countries and regions of East Asia (according data for 2017)

Regions	Area, thou. km^2	Share of transboundary territories in total area, in %	Length of coastline, km	Population size, thou. people	GDP, billion US$
1 Pacific Russia	3,086.3	100	59,883	5,219	46.8
2 Eastern areas of China	2,207.3	100	30,017	770,327	14,61
3 Japan	377.8	100	29,020	126,824	5,443
4 DPRK	120.5	100	4,009	28,491	66.7
5 Republic of Korea	100.2	100	12,478	53,733	2,035
6 Taiwan	36.0	100	2,007	24,827	1,189
7 Singapore	0.78	100	268	5,470	528
8 Thailand	514.0	100	7,066	71,037	1,236
9 Laos	236.8	100	0	7,007	49.2
10 Cambodia	181.0	100	1,127	24,827	64.3
11 Vietnam	329.6	100	11,409	93,402	649
12 Philippines	299.8	100	33,900	107,143	877
13. Malaysia	329.8	100	9,323	41,700	993
14. Myanmar	679.5	100	1,930	54,045	344
15. Indonesia	1,904.5	100	95,181	269,479	3,250
16. East Timor	14.9	100	706	1,212	6.8
In all	10,418.8		298,324	1,684,743	31,138.8
Average value for country, region	651.2	100	18,645	105,296	1,946.2

Compiled by: (www.wri.org 2019; gtmarket.ru 2019; fincan.ru 2019; InfoTables.ru 2019; gtmarket.ru 2019; Pacific Russia 2017).

Table 3. Contrast of East Asia transboundary space according to the level of development (according to data 2018)

Names of countries	Population size, people/ km^2	Economic density, thous. US$ /km^2	GDP per capita, US$/person	Comparison to the country with maximum density (times) *
1. China	139	2,633	18,110	53 / 276
including its eastern parts	349	6,506	18,643	21 / 112
2. Japan	336	14,805	44,227	22 / 49
3. Russia	8.6	246	29,267	859 / 2,957
including Pacific Russia	12	48	9,817	6,157 / 15,218
4.Indonesia	127	1,835	13,230	58 / 396
5.Republic of Korea	494	21,315	41,351	15 / 34
6.Thailand	131	2,568	19,476	56 / 283
7.Taiwan	650	34,565	53,023	11 / 21
8.Malaysia	86	3,030	30,860	86 / 240
9.Philippines	333	3,179	8,936	22 / 229
10.Vietnam	272	2,154	7,511	27 / 338
11. Singapore	7389	727,506	100,345	1 / 1
12. DPRK	189	556	2,341	39 / 1,308
13. Myanmar	79	507	6,511	94 / 1,434
14.Cambodia	80	389	4,335	92 / 1,868
15.Laos	27	227	7,925	273 / 3,208
16.East Timor	77	447	5,242	96 / 1,627

Compiled by: (gtmarket.ru 2019; fincan.ru 2019; InfoTables.ru 2019; gtmarket.ru 2019).
In the numerator – differences in population density, in the denominator – in economic density.

Table 4. Basic zones of geopolitical tension in the transboundary regions of East Asia

Names of transboundary regions	Subject and zones of the geopolitical tension	Parties (countries) where the geopolitical tension exists
1. Chukchi Sea basin	Demarcation of the eastern part of the Arctic shelf	Russia – the USA
2. Sea of Okhotsk basin	Claims of Japan to the South Kurile Islands	Russia – Japan
3. Sea of Japan basin	Absence of peace treaties between countries Absence of official borders in the area of the Takeshima-Dokdo Islands	DPRK – Republic of Korea Japan – Republic of Korea
4. East China Sea basin	Ownership of the Senkaku Archipelago (8 small islands with a total area of 7 km^2)	Japan – Republic of China (Taiwan) – People's Republic of China (PRC)
5. East China Sea basin	Ownership of Taiwan Island. Affiliation of Taiwan. China considers the island of Taiwan its territory.	China – Taiwan
6. South China Sea basin	Ownership of the Spratly Islands (a key area in the context of regional presence and availability of oil and natural gas).	Vietnam – China– Taiwan – Malaysia–Philippines–Vietnam–Brunei
7. South China Sea basin	Affiliation of the Paracel Islands	China – Vietnam
8. Water area of the Sea of Japan	Naming of the sea (in the Republic of Korea and DPRK, this sea is called the Eastern Sea).	Republic of Korea, DPRK, Japan

Compiled by: (www.geopolitics.com 2019).

regions in East Asia: two-link regions facing the Chukchi, Okhotsk and Bering Seas; three-link zones facing the Yellow Sea; four-link territories (including the Amur transboundary region) facing the Sea of Japan; continent-island regions facing the East China and South China Seas and island regions facing the Philippine Sea. In the multilink transboundary regions, there is a concentration of national frontiers (land and maritime) of different countries which complicates their geopolitical relations in general (Baklanov 2000).

East Asia, mainly in its transboundary regions, is characterized by the presence of zones of international tension and geopolitical problems related to unresolved borders, some of which are remaining legacies of the World War II. In particular, there is a certain unsettledness with regard to differences in positions and related geopolitical tensions between DPRK and the Republic of Korea, between PRC and Taiwan, as well as a number of geopolitical problems associated with the ownership of individual Pacific islands (Table 4).

The existence of such zones of tension in the transboundary regions creates significant constraints for sustainable natural resources management and development. The search for compromises and ratification of long-term international treaties are necessary.

CONCLUSION

Many countries and regions of East Asia have significant geographical (including natural-resource, advantages of location and collocation etc.) and geopolitical (transboundary phenomenon etc.) opportunities for long-term development.

Therefore, availability of considerable land-based natural resources and maritime natural resources seas is a favorable geographical factor for nearly all countries of the macro-region. Many of these resources are renewable (biological, forest, land, water, hydro-power, etc.) and thereby foster the long-term sustainable development of the macro-region.
At the current stage of development, this macro-region has the highest and steady rates of economic growth in the world. Eastern Asia is characterized by a very high contrast of the key socio-economic indicators of cross-border spaces of neighboring countries: population density, GDP values, and economic density. A steady decrease in the contrast among neighboring countries of Eastern Asia can be an important factor in reducing the level of conflict and the growth of sustainable development.

The most significant geographical limitations to its sustainable development are evident for the northern and northeastern portions of the Russian Far East, as well as the Western Pacific seismic arc. Within other countries, these limitations are less pronounced. Geopolitical tensions, such as territorial disputes, should be removed by diplomatic means, if all countries are to achieve sustainable international cooperation. In general, the transboundary phenomenon that is typical for all regions and countries of East Asia presents favorable opportunities for the long-term development.

The paper was prepared as part of the project No. 18-05-04 under the "Far East" Program of the Far Eastern Branch of the Russian Academy of Sciences.

REFERENCES

Asia-Pacific Economic Cooperation: yesterday, today, tomorrow. (2010). Vladivostok: the Far-Eastern Federal University. (in Russian).

Baklanov P.Y. and Moshkov A.V. (2015). Amur Region of Russia – natural resources, population and economy. In: Environmental Change and the Social Response in the Amur River Basin. Tokyo; Heidelberg; New York: Springer, 1-35.

Baklanov P.Ya. (2000). Contact geographical structures and their functions in the Northeast Asia. Proceedings of RAS, Series geogr., 1, 31-39. (in Russian).

Baklanov P.Ya. and Ganzei S.S. (2008). Transboundary territories: problems of the sustainable nature management. Vladivostok: Dalnauka. (in Russian with English summary).

Baklanov P.Ya., Moshkov A.V. and Romanov M.T. (2016). Pacific Russia: Basic factors and lines of the long-term development. Problems of geography, 141, 595-618. (in Russian with English summary).

Baklanov P.Ya., Yermoshin V.V., Karakin V.P., Zharikov V.V., Nguyen Van Qu, Dao Dinh Cham. (2017). Coastal-maritime nature management in the Pacific Russia and North Vietnam: concepts, structural features and types. Geography and natural resources, 4, 29-37. (in Russian with English summary).

Druzhinin A.G. (2016). Russian in the multipolar Eurasia: opinion of the geographer-social scientist. Rostov-on-Don: the Southern Federal University. (in Russian with English summary).

fincan.ru, (2019). GDP per capita PPP in 2018. [online] Available at: fincan.ru/articles/79_vvp-po-pps-na-dushu-naselenija-2018-stran-mira/ [Accessed 27.09.2019].

Geopolitical potential of the transboundary cooperation of the Asiatic-Pacific region countries: joint monograph. (2010). Vladivostok: Dalnauka. (in Russian with English summary).

gtmarket.ru (2019). The ranking of the countries by their population. [online] Available at: gtmarket.ru/ratings/world-population/info/ [Accessed 27.09.2019].

gtmarket.ru, (2019). Ranking of the countries by their gross domestic product // The World Bank. [online] Available at: https://gtmarket.ru/ratings/rating-countries-gdp/rating-countries-gdp-info/ [Accessed 27.09.2019].

InfoTables.ru, (2019). The area of the countries of the world (table). [online] Available at: https://infotables,ru/strany-i-goroda/19-tablitsa-ploshchad-territoriy-stran-mira [Accessed 27.09.2019].

Kolosov V.A. (2018). Paradoxes of globalization and increasing barrier functions of the frontiers of the modern era. Regional studies, 3, 9-18. (in Russian with English summary).

Larin V.L. (2014). Realities of the new world and Russian-Chinese relations. In: Russia and China: experience and potential of the regional and near-border interaction. Vladivostok: Dalnauka, 9-19. (in Russian with English summary).

Larin V.L. (2016). Russian-Chinese trans-boundedness within the framework of the Eurasian integration projects. World economy and international relations, 12, 69-81. (in Russian with English summary).

Nature management of the Far East and Northeast Asia: potential of integration and sustainable development. (2005). Khabarovsk: FEBRAS. (in Russian).

Pacific Russia in the integrated space of the North Pacific in the early 21st century. Experience and potential of the regional and near-border interaction. (2017). Vladivostok: IHAE FEBRAS. (in Russian with English summary).

Terada T. (2006). Forming an East Asian Community: A Site for Japan-China Power Struggles. Japanese Studies, 26(1), 1-13.

Tkachenko G. G. (2009). Territorial differentiation of the natural-resource potential of a group of countries in North-East Asia. Geography and Natural Resources, 30(2), 107-112. (in Russian with English summary).

Tulokhonov A.K. (2014). Political geography of North Asia under conditions of the globalization. Ulan-Ude: EKOS. (in Russian with English summary).

Womarck Brantly. (2014). Northeast Asia in a Multinodal World. East Asia, 31(3), 180.

www.geopolitics.com, (2019). All disputed territories of the world. Geopolicy. [online] Available at: www.geopolitics.ru/Karta-sajta/spisok-spornyx-territorij/ [Accessed 27.09.2019].

www.wri.org, (2019). List of the countries by their coastline // World Resources Institute. [online] Available at: www.wri.org/text/coastal-marine/variable-61.html [Accessed 27/09/2019].

Received on July 3th, 2019 Accepted on January 21th, 2020

Maria V. Zotova[1*], Anton A. Gritsenko[1,2]
[1] Institute of Geography, Russian Academy of Sciences. Moscow, Russia
[2] Immanuel Kant Baltic Federal University. Kaliningrad, Russia
***Corresponding author:** zotovam@bk.ru

NEIGHBOURHOOD AND PERCEPTIONS IN SMALL CITIES ON DIFFERENT RUSSIAN BORDERS

ABSTRACT. National neighbourhood have a significant influence on the life of people living along the state borders. They shape human interactions across borders and border residents' attitude towards neighbours. Many concepts like 'neighbourhood', 'proximity', 'trust', '(un)familiarity', and 'otherness' are usually used to explain this processes in border studies. However, insufficient attention has been paid to the comparing of perceptions, life strategies and everyday life of borderland population depends on neighbouring policy, border regime and neighbourship. Here we focus on different Russian borders with Ukraine (the new contested border in Crimea), Kazakhstan (the EAEU`s internal border), and China (old international and contact border) using different sources of information, including expert interviews as well as field observations and focus groups conducted with locals. We find that people differentiate between the neighbors they know and the neighbouring state they do not trust. Significant differences between neighbouring territories, unfamiliarity, and otherness are not allowed to get in the way of contact, because it is this contact that allows local residents to make a living. In conclusion, our results suggest that while the objective differences between the various sections of Russian borders serve to diversify the neighbourhood situations, their subjective perceptions and social representations serve to unite them.

KEY WORDS: border, border regions, neighbourhood, everyday life, perceptions, Russian borderland

CITATION: Maria V. Zotova and Anton A. Gritsenko (2020) Neighbourhood And Perceptions In Small Cities On Different Russian Borders. Geography, Environment, Sustainability, Vol.13, No 1, p. 64-73
DOI-10.24057/2071-9388-2019-124

INTRODUCTION

In recent decades, state borders have attracted the close attention of scholars from around the world. Initially, such scholars' research was focused on interaction, cross-border cooperation, and the flow of people, goods, ideas, and information across borders, but it has subsequently seen a gradual shift from examining borders per se to considering the processes of bordering and othering (Brambilla 2015; Newman and Paasi 1998) — in other words, scholarship has turned its attention from territorial to social and political functions. Despite this shift, borderlands as contact zones — with intensive international exchanges, social and political encounters and contradictions, mutual transitions, and manifestations of differences — remain a key locus for the investigation of multidimensional and multiscale border phenomena. The interest in borderlands and border practices is thus oriented to (re)structuring social and political spaces, and shaping people's identities and everyday life, across multiple borders created by diverse actors. It is also bound up with the search for answers to a number of fundamental questions raised in various aspects by F. Bart, G. Simmel, M. Foucault, B. Anderson, and others, namely: how do social and political borders relate to each other, and what role do they play in our lives, activities, and social and political relations? The overlapping interests of nation-states and the inhabitants of borderlands results in an ambiguity surrounding the processes taking place at the border, where rivalry enforces cooperation, which proceeds 'over the barriers' (Vendina et al. 2007; Brazhalovich et al. 2017). This highlights the necessity of understanding what borders mean to people and how border-related practices affect societies in general. Recent debates on borders have devoted considerable attention to the everyday life of people living in borderlands, their personal motives for crossing borders, and their perceptions of neighbours (Van der Velde et al. 2008; Helleiner 2009; Aure 2011; Ghosh 2011; Balogh 2013; Stoklosa 2013; Domaniewski et al. 2016; Laine 2016). However, comparative questions concerning territorial differences and their consequences for local life mostly remain beyond the scope of such studies, despite the fact that various borders and border regimes create different challenges for people, their well-being, and the possibility of cross-border activities. This paper aims to fill this significant gap in research by exploring and discussing the social function of neighbourhood and borders in a variety of small shrinking cities at different points along the Russian border. First, we begin by sketching the theoretical framework underpinning our study. Then we will consider how people perceive neighbouring states and neighbours, and what they mean for them. Following this, we will turn analyse what people associate with border regimes, how often and why they cross the border, and what they face when they do. Finally, we will discuss the indirect effect of the border and neighbourhood by considering the opportunities, hopes, and expectations of people living at the border.

THEORETICAL FRAMEWORK

A variety of concepts in border studies have been used to explain human interactions across borders and border residents' attitude towards neighbours.

The concepts of **proximity and distance** addresses to the formation of social ties, solidarity, and identity upon the influence of interaction across the border (Szytniewski et al. 2017). As noted by (Trippl 2010), (Boschma 2005), (Torre and Rallet 2005), geographical proximity alone is neither necessary nor sufficient to facilitate the formation of social ties. Consequently, frequent social and cultural encounters can generate feelings of familiarity, recognition, and security (Van Houtum 1999). However, when cultural differences are too great, people may not be able to make sense of them when using existing knowledge and representations of otherness; they will eventually experience discomfort (Szytniewski et al. 2017). Proximity and distance are also associated with the character of functional connectivity and accessibility, related to cross-border institutes and activities, and not to the geographical location alone (Schack 2001).

The concept of **trust** is often used in the contexts of transnational relation, cooperation, and human interactions. Social capital is an important element of cross-border networking and regional integration (Koch 2018). Trust of neighbours is especially important both in cooperation and in establishing informal ties, because it contributes to overcoming differences including territorial tensions. On the grassroots level, social integration occurs easily between neighbours with higher levels of trust (Rippl et al. 2009). A good example of this is found in territories with long-term coexistence within a single country where the ability of people to communicate and understand each other is based upon common values, language, and identity, as well as upon similarities in physical surroundings and daily routine.

To better understand cross-border (im)mobility Spierings and van der Velde (2008) have proposed the concept of **(un) familiarity.** It suggests that borders could in fact promote mobility due to the functional, physical, and socio-cultural differences between places. Depending on how people perceive differences between countries, they either stay at home or visit the other side (Spierings and van der Velde 2013). On the one hand, dissimilarities between 'here' and 'here', and the uniqueness of foreign places could lead to border-crossing for a variety of purposes (tourism, leisure, shopping etc.) (Spierings and van der Velde 2008). When the 'unfamiliarity' of a place dissolves, its attractiveness to visitors from abroad may also disappear (Timothy 1995). On the other hand, large dissimilarities in the socio-cultural sense will result in mental borders that have a negative impact on cross-border interactions (Van Houtum 1999). When dissimilarities between places on each side of the border are too large, they could restrain people from making the cross-border trip (Szytniewski and Spierings 2014).

Another theoretical construction proposed by Dolińska and Niedźwiecka-Iwańczak (2017) is based on George Simmel's concept of **'strangeness'.** It references to the dichotomy of **'Insiders'** and **'Outsiders',** with the division of the latter into 'Strangers' and 'Others'. While a 'Stranger' is currently, or potentially, dangerous, and poses a threat to values that an individual holds dear, the 'Other' is one whom we do not understand and thus do not accept at all (Kozera 1999), as he/she is not an 'Insider'. In the process of structuring the social world, otherness may remain just otherness, but it may also turn into strangeness. The perception of individuals and groups as other, dissimilar, or different from us does not necessarily trigger any form of strangeness if it is limited to giving facts and does not engage in evaluation. The awareness of otherness turns into strangeness only when the perceived dissimilarity is combined with negative emotions and attitudes.

In our study, we apply all these concepts to interpret the setting of living on the border of a different neighbourhood. We proceed from the understanding of neighbourhood as not only top-down geopolitical imaginaries and everyday perceptions and representations (Scott et al. 2019) but also as differences between places, including socio-economical, cultural, and institutional contrasts, among others.

CASE STUDIES

To remove the need to consider territorial proximity and to clarify the role of other border-related factors constituting personal life strategies, cross-border practices and perceptions we decided to focus on the Russian border cities, because they function as important economic centers for border areas and in particularly for social activities and cross-border cooperation. In Russia, there are only 17 such cities. Some of them are rather big and diverse. We focused on such of them that have similar socio-economic conditions and problems of urban development that make them more available for comparative analysis and reflecting geopolitical and territorial factors of neighbourhood. We selected three sections of the Russian border that differ in age and origin, regimes and socio-economic contrasts across the border. The first one is new conflict international border in Northern Crimea, contested by Ukraine and unrecognized by the international community. It appeared after the annexation of Crimea to Russia in 2014, and is characterized by a strict border crossing regime. Before 1954 there was a border between Russian and Ukrainian Soviet Republics. Then it existed as administrative boundary of Crimean autonomy in Ukraine. The second one is the internal border of the Eurasian Economic Union (EAEU) between Russia and Kazakhstan. Before the collapse of the USSR, the border was also administrative and completely open as in Crimea, binding together different territories of a single state. Finally, the third section is old border between Russia and China, that was highly militarized and closed during the Soviet time. Today it is much more open and contact, while it is dividing too different cultural worlds.

Thus, we selected four small cities on these sections of Russian border. Despite of mentioned similarities they have certain specific of historical and cultural development that define some features of social, economic and cross-border activities. Two cities are in Northern Crimea, close to the new contested border: Armyansk (population 22,000, located 5 km from the border) and Dzhankoy (population 39,000, located 40 km from the border). Armyansk is a mono-industrial city. Its defining chemical enterprise 'Crimean Titan' provides employment for almost half the local population. Dzhankoy, by contrast, is ex-industrial and now a principally commercial city, having mutual links with surrounding agricultural areas. Before the border appeared, it was also a transport hub connecting Crimea with Ukraine and Russia by roads and railways. Many locals found jobs servicing traffic flows. Since 2014, however, Crimea has experienced the shock of huge changes caused by the transition to the Russian legal and institutional field, a break almost all its old supply chains and economic ties with Ukraine, an energy, water and transport blockade, and international sanctions. These have undermined the stability of many Crimean enterprises, some of which have had to close. 'Crimean Titan' began to experience interruptions in its supply of raw materials and was forced to cut staff. The Dzhankoy transport hub has become into a transport dead end (Fig. 1), and agricultural enterprises have been forced to change their specialization due to lack of water.

The next city, Troitsk (population 75,000, located 11 km from the border with Kazakhstan) like the border itself, has experienced frequent changes in its territorial functions. For a long time it was a prosperous trading and cultural center on the routes between Russia, China, and Central Asia. During the 19th century Troitsk became the jewel of architecture in the South Urals due to the construction of various attractive public buildings, banks, and trade houses (Fig. 2). In the Soviet period, it received a diverse industrial development, enjoyed extensive cooperative ties with Kazakhstan, and was the leading Russian centre for the Virgin Lands campaign. After the collapse of the USSR, 8 of the 12 large enterprises were closed and the city ended up on the state border, the ethno-cultural contrasts of which gradually grew with the national state-building agendas in Russia and Kazakhstan. De-bordering processes were launched only in the 2010s with the implementation of Eurasian integration initiatives. The city nowadays has a development deficit, and decline and outflow of population.

Fig. 1. Dzhankoy railway station lies deserted following the events of 2014

Fig. 2. Historical building of trade house in pseudo-Russian style (1870). Nowadays it is a branch of the Chelyabinsk State University in Troitsk

Finally, Zabaikalsk (population 13,000, located 5 km from the border) is a small town (Fig. 3) on the Chinese Eastern Railway station with container terminal nowadays serving 60% of the traffic between Russia and China, and the largest checkpoint, passing more than one million people annually. Changing the border regime entailed a restructuring of the local economy in Zabaikalsk towards management of the border, border crossings, and transit flows.

Peripherality, limited economic development, and negative demographic processes make the life of each of these cities highly dependent on external impulses, including changes of neighbourhood relations and cross-border interactions.

RESEARCH METHODS

We conducted our research into these Russian border cities using various sources of information and methods of analysis. On the one hand, we used official documents of territorial development and data (e.g., border crossing, economic, and demographic statistics); on the other hand, we relied on interviews with local and regional experts field observations, and focus groups with locals.

We considered focus groups as the most relevant method for gathering and analyzing grassroots information due to its flexibility and adaptability both in terms of the composition of groups and in terms of the non-standard conditions of communication. Between autumn 2017 and spring 2018, 13 focus groups were consulted: four in Dzhankoy, five in Troitsk, and two each in Armyansk and Zabaikalsk. The number of groups in each located varied due to population size and the poly/mono-functionality of the cities. The initial condition for participant selection was that we focus on the most representative groups of the urban population, not exclusively those people involved in cross-border activity. We proposed that even if a person does not actively participate in cross-border movement, exchange, and communication, the border nonetheless affects his life through his relatives, friends, shopping, fashion, ideas, fears, and expectations. Much attention was paid to the dominant cohorts among the economically active part of the population aged 30–50 years. Employment in different spheres of activity, and differences in income, social status, and ethnicity were also taken into account in the selection process.

The recruitment of respondents was carried out by professional recruiters using the 'snowball' method. The selection of focus group participants was conducted at the second stage, in which files of potential respondents were screened under the proposed selection criteria: age, professional activities and ethnic self-identification. All focus groups have included ethnic minorities that play a significant role in the cities and have a particular symbolic capital because they also represent title ethnic group of neighboring country (except for Chinese).

The composition of focus groups was the following: in Armyansk, the group consisted of factory workers and local entrepreneurs; in Dzhankoy, groups contained state employees, pensioners, local entrepreneurs, and a mixed group with local Ukrainians and Crimean Tatars; in Troitsk, the groups contained state employees, pensioners, local entrepreneurs, factory workers, and management staff; finally, in Zabaikalsk, the group consisted of state employees and local entrepreneurs. All group discussions were held using a single but flexible guide, which included the following research topics: images of Russia and the neighbouring state, current relations between states, border-related practices and cross-border mobility, interactions between people in a multi-ethnic environment and across the border, perceptions of the border, and welfare and specifics of everyday life. Projective questioning techniques (e.g., standard methods of associations, combinations, metaphorization, semantic attribution, etc.) were actively combined. However, this combination differed in our cases studies. For example, in Crimean cities we don`t use semantic test, when people gave characteristics for the most authoritative local ethnic groups, which they defined themselves. It was replaced by modelling of abstract dialogs, disclosing relationships between them.

RESULTS

Since state borders are both symbols of social institutions and power relations (Newman et al. 1998), the competition for the constitution of 'reality' and for the meanings of borders and neighbourhood occurs in the borderlands. Ideas about neighbours are defined from above through socialization, interstate relations, and management of public opinion, but they are also shaped in everyday border-related practices (Dolińska et al. 2017).

Images of neighbours.

Discussing neighbouring countries, participants distinguish two ideas: the first is that the neighbouring country is 'a state and government', and the second is that it is 'a country and people'. Concerning the first category, participants expressed various but mostly negative emotions, ranging from mistrust to hate and fear, while concerning the second category, participants tended to express positive emotions such as sympathy and trust.

Individual opinions concerning neighbouring countries and argumentation demonstrated a certain proximity to the official political discourse in Russia. Thus, in Northern Crimea, Ukraine is perceived as a hostile and threatening state with many internal problems and a bad international reputation. In Troitsk, Kazakhstan is considered a friendly country, which nonetheless suffers from ethnonationalism: respondents often noted that in Kazakhstan the rights of Russian-speaking citizens are regularly violated. In Zabaikalsk, the attitude towards neighbours is more ambivalent. China is perceived as a strong partner country. However, the participants did not have full confidence in China, and believed Russia should not rely on it. This corresponds with the skepticism to China widely found in mass media (Kolosov et al. 2019).

In contrast to this attitude towards the neighbouring states, attitudes towards the actual people living on the other side of the border were very different. Respondents associated them with cross-border practices and special trust in relationships that, according to recent studies (Zotova et al. 2018), are widespread in almost all sections of the Russian border. On the new post-Soviet borders, their existence is largely based on the long-term coexistence of people within a single country that ensured fluency in the Russian language, fostered ethno-cultural exchanges and mutual influences, and promoted common values, as well as shaping mixed or dual cultural identities on both sides of the border. Family ties, professional networks, and friendships between people from neighbouring regions have persisted to the present day. Today this allows people to communicate across the border and understand each other without any difficulty. Tests and abstract dialogues in which participants identified the characteristics for key ethnic groups of residents showed that people in Crimea hardly at all distinguish a difference between Russians and Ukrainians. Common opinion of respondents were the following: «We do not know who are Ukrainians there», «Nobody ask who is Ukrainian, and who is Russian», «We are not distinguish Russians and Ukrainians, we live there all together», «There is no difference, we are the same».

On the border of Kazakhstan people also hardly draw a distinction between Russians and Kazakhs. In fact, most of the characteristics ascribed to Russians and Kazakhs were absolutely the same (Table 1). Participants recognized that in many cases, local residents do not see a difference between people on either side of the border, and cannot identify, for example, those who came to the city from the nearest areas of the neighbouring state. However, they do feel some differences from people from other regions of Russia and

Fig. 3. Zabaikalsk with skyscrapers of neighbouring Manzhouli in the background

sometimes even from own region (from Southern Crimea in Armyansk and Dzhankoy; from the north of Chelyabinsk region in Troitsk; from Chita in Zabaikalsk). Difficulties in interacting and mutual understanding, as participants noted, only exist with migrants from South Caucasian and Central Asian countries. They are usually perceived as 'strangers' who pose a potential threat. In Crimea, Crimean Tatars often play such a role, while in Troitsk and Zabaikalsk it is Azerbaijanis, Uzbeks, etc, who are seen as 'strangers'. They were mostly endowed with negative characteristics. In certain situations they oppose them to dominant Russians. In all the cities self-images of Russians fully coincided with images conducted by mass opinion surveys of Russian leading sociological agencies.

At the Russian-Chinese border the situation appears differently. The limited list of characteristics given to Chinese suggests that citizens of the neighbouring state are poorly known; apparently they are excluded from the 'we-community' and play the role of 'others'. However, as shown in Table 2, less frequent encounters leads to a reduced familiarity. In Troitsk the image of Chinese people is more blurred; local experience of interaction was limited to the passive observation of Chinese workers employed in the construction of a new energy unit for the local power plant over two years. In Zabaikalsk, by contrast, local people communicate frequently and regularly with neighbours due to cross-border trips and shuttle trading. However, as almost no one in Zabaikalsk speaks Chinese such communication takes place in pidgin. It is enough for transactions, but not for genuine understanding. Therefore, knowledge about the neighbouring country and familiarity with its culture — including social norms and symbolic values — is reduced to stereotypes and is prevented from reaching any higher level of interaction.

Perceptions of the border and border regime

The dual perception of neighbouring countries as both 'a state and government' and 'a country and people' is reflected in the dual meaning of borders, which impact on the perception of border regimes. Thus, both the contact and barrier functions of a border can be viewed differently from each angle.

During focus groups we posed the question: 'What would happen if the border were to be fully opened/closed?' (Table 3). In all cases, the participants emphasized that the border with the neighbouring country is needed. Even if the relationship between states seems friendly for now, the border could not be fully opened because it provides peace and stability, and protects against negative influences from the outside (drug traffic, smuggling, etc.). The border is perceived as a symbol of the state and protection from 'chaos' and 'disorder', as well as a guarantee of personal security. On the other hand, the order and wellbeing of everyday life are also closely linked with the border regime. Nobody even could imagine the full closure of the border. Communication between relatives and friends, cross-border trade and other activities (like shopping, leisure, tourism, labor, education, etc.) motivated by territorial proximity to the neighbouring country and its markets were the most popular arguments for such a view. Thus, the attitude of the local community to the possible openness of the border explains their perception of state function: that the border is necessary for reasons of state stability and security. In turn, discussion of its full closure reflects their perception of their relationship to neighbours, from whom they could not imagine being fully separated.

If in Armyansk, Dzhankoy, and Troitsk, the respondents noted that the current state of the border mainly impedes upon family ties, in Zabaikalsk it helps them 'survive'. These varied opinions reflect differences in what local people value and confront on a daily basis. Our case studies allow us to understand what constitutes the differences and how cross-border communication is affected by differences in border regimes and neighbourships.

Border as obstacle

The contested status of the new international border in Crimea has hampered communication across it. Residents of Crimea may cross the border only by driving, in a car with Ukrainian number plates, and with a Ukrainian passport that they are allowed to keep. Border crossing by cars with Russian number plates is not permitted due to sanctions on the Ukrainian side. To cross the border by public transport is also impossible; regular buses and trains

Table 1. Characteristics given to Russians and Kazakhs in Troitsk

Positive		Negative	
Kazakhs	Russians	Kazakhs	Russians
Brave	Brave (4)	Arrogant (2)	Arrogant
Cultural (3)	Cultural	Retarded (2)	Retarded (2)
Generous	Generous (2)	Stupid	Stupid (3)
Hardworking	Hardworking	Evil	Evil (2)
Loving (3)	Loving (2)	Lazy	Lazy (2)
Sincere (2)	Sincere (2)	Secretive	Secretive
Open	Open (3)	Cowardly	Cowardly
Rational	Rational	Power-hungry	
Respectful to elders (2)	Respectful to elders		
Simple (2)	Simple (2)		
Smart	Smart (3)		
Traditional	Traditional		
Peace-loving (2)	Advanced (4)		
Hospitable (2)	Energetic		
	Good (4)		
	Patient (3)		
	Patriotic (4)		

Note: Numbers in brackets indicate the number of identical answers given by focus group participants during the test.

Table 2. Characteristics given to Chinese in Zabaikalsk and Troitsk

Positive		Negative	
Zabaikalsk	Troitsk	Zabaikalsk	Troitsk
Hardworking (2) Energetic Rational Loving	Hardworking (2) Clever	Arrogant Evil Stupid Secretive Cowardly	Brutal Underdeveloped

Note: Numbers in brackets indicate the number of identical answers given by focus group participants during the test.

connecting Crimea with Ukraine have been cancelled. At the border, meticulous, picky, and sometimes humiliating document and baggage checks denigrate individuals. Due to the established border regime, people have lost the opportunity to consume familiar Ukrainian products and goods. Russian customs officers confiscate all food and other Ukrainian goods prohibited for import. Moreover, inhabitants of Armyansk and Dzhankoy crossing the border often encounter misunderstanding and even hostility at Ukrainian checkpoints, as well as on the other side of the border, due to their support for the annexation of Crimea in the referendum in March 2014. Therefore, they significantly cut their trips and travel to Ukraine only in order to visit relatives. Citizens of neighbouring Ukrainian regions have also reduced the frequency of cross-border visits. During the period 2014–2017, the number of border crossing in the north of Crimea (checkpoints Kalanchak, Chaplinka and Chongar) fell from 3.6 to 2.5 million people. Ukrainians visit Crimea mostly for tourism and family visits. During these contacts, many people are careful not to mention political issues anymore, while others have simply fallen out of touch with friends and family across the border. In either case, relations have become tenser and less trusting.

The main barriers to crossing the Kazakhstan border are the large distances between cities and the undeveloped and overloaded checkpoints, rather than the border itself. For citizens of neighbouring countries, the border crossing procedure is simplified, and people are allowed to use internal identity documents. Six buses and two trains daily connect Troitsk with Kostanay, Zhitikara, and Rudnyy across the border. Every year, more than two million people cross the border in Troitsk, and a quarter of them are citizens of Kazakhstan. The traffic flow grew twice after the cancelation of customs in 2010. The most popular reasons for crossing are to visit relatives and friends, diversify consumption, and save on goods purchases (sausages, sweets, vodka, and other products are cheaper in Kazakhstan). Economic crises and devaluation of ruble in the end of 2014 contemporary increased flow of people across the border in 1.5 times due to migrants from Kazakhstan. However, the situation return to pre-crisis level up to 2017.

The Russian-Chinese border opened to the movement of people in the early 1990s. Those wishing to visit the neighbouring country needed to have received an appropriate visa through national consulates located in several large cities. The closest such consulate to Zabaikalsk was found in Khabarovsk, 2500 km away (1.5 to 2 days of travel). The impressive demand for cheap Chinese goods in Russia has intensified shuttle trade at the local level that has dramatically increased the flow of people and goods across the border. In the 2000s, China initiated a number of programs for the development of its northern border territories. As a result, Russian citizens received the opportunity to acquire visas at the border, and could easily visit certain adjacent cities in small groups. The traffic flow highly increased up to one million people per year. Manzhouli, located on the border just opposite Zabaikalsk, was one such city; nowadays the two cities are closely connected through transport links, with more than ten buses daily running between them.

Table 3. Responses to the question 'What would happen if the border were to be fully opened/closed?'

City	'If the border were to be fully opened'	'If the border were to be fully closed'
Armyansk	*«If they* [certain citizens of Ukraine] *come here, it will be worse than in Donbass»*	*«It is unacceptable. This will create discontent. We have family ties. We are deeply connected with each other»*
Dzhankoy	*«It will be a mess and chaos. Lawlessness»; «Certain groups will immediately come here from Ukraine, and this will lead to great bloodshed. They will twist our heads off»*	*«It's impossible»; «It cannot be closed»; «It complicates our relationships with relatives»*
Troitsk	*«The border is needed because we do not know what may happen tomorrow»; «We still need to keep the gunpowder dry. We cannot open the border absolutely»; «Kazakhstan, of course… friends, but all could turn. Today we are friends, but tomorrow?... With Ukraine, too, ... that is how it turned out...»*	*«It is impossible in any case. We need to cooperate, communicate, trade, we need a dialogue of cultures»; «It's scary because we have families, relatives, etc.»*
Zabaikalsk	*«It will be disorder», «All the Chinese will come here. They will take away all our land»; «If there is no border, there will be no Russians here. Only the Chinese will stay»*	*«Close the border and immediately there will be no one* [left] *here... Thanks to the border, people live here»; «We need the border to survive»*

Border as resource

A border also of course has several indirect effects on life in border regions that local people do not perceive and explicitly articulate.

The new contested border in Crimea has led to the sharp reduction of external economic relations, as well as property redistribution, reorganization, and instability of local enterprises (especially 'Crimean Titan' and 'Dzhankoy Railway Station'). At the same time, the hardship of transition was partly offset by new economic opportunities. A decrease in economic competition on the local market due to disappearing of Ukrainian products and distributers, for example, allowed local farms to emerged from the shadow and to be legalized, and intensify production. Russian social and economic transfers leaded to wage, pensions and social benefits growth, as well as infrastructural investments. The deployment of Russian military garrisons, customs, and border control provided new jobs, along with additional demand for housing, value-added goods, and services.

Thus, people and local enterprises have faced big challenges since the border emerged. But, on the other hand, they also have felt positive changes comparing current state with Ukrainian period when Northern Crimea did not received enough investments in infrastructure and city development. As noted by respondents, during that time local people feel themselves in some way as neglected by Ukrainian authorities.

The economic crisis of the 1990s, the establishment of the state border, and to related rupture of industrial ties with Northern Kazakhstan had resulted in the closure of eight large enterprises and a change in city functions. The emergence of customs and border services, together with cross-border trade and smuggling, can only partially compensate the loss of employment and income for many locals. Integration initiatives in the framework of the EAEU, creating a single labour market have not yet brought new employment prospects. On the contrary, some border services — for example, customs posts — have been abolished. The small contrast in living standards and prices could not stimulate cross-border mobility and shuttle trade. However, Troitsk was able to sustain its central function for Russians pushed from the neighbouring territories of Kazakhstan because of an active nation-building process there. Despite the growth of peripherality of Troitsk, it has retained its cultural capacity for surrounding territories, becoming one of the terminal centers for Kazakhstan citizens seeking a place to live, work, and pursue educational opportunities. If the immigration flow into Troitsk from the neighbouring country has declined significantly in recent years, the flow of other migrants has slightly increased.

The opening of the Russian-Chinese border was accompanied by its demilitarization, and, as a result, a reduction in military garrisons and the compensatory development of shuttle trade, as well as an increase in the flow of Chinese goods via the East China Railroad. As a result, Zabaikalsk — a formerly small, closed military town — has become the main entry point for Chinese goods into Russia. Most of the local citizens are employed in customs and border services, the railway sector, budget sector, and shuttle trade. Frequent trips to Manzhouli and servicing the movement of people and goods across the border enable local people to survive. China has become for them a place of work along with a place of rest, shopping, and leisure.

Life on 'the margin'

Border practices and neighbourhood obviously have a direct and indirect effect on the social and economic development of these small border cities, and bring diversity to their life, distinguishing them from cities further away from the border. Even those residents who are not directly involved in cross-border movements, exchanges, and communications receive their share in border-related profits, which are redistributed from commuters to the closest members of their social circle as well as to other spheres of the local economy. Indeed, by defining the daily activities, the border provides a range of opportunities as well as risks and costs. Economic and social instability probably tip balance between opportunities and costs, especially when people consider the future of their children when defining their life strategies. If the city's border location neither increases prosperity nor improves the quality of life, and the neighbourhood is perceived predominantly in terms of risks, then the incentive to live there is reduced.

The respondents drew much attention to the numerous socio-economic problems of their cities. Among the problems listed were lack of work and low incomes, poverty, the low quality of medical services, inefficient urban governance, and corruption. In Armyansk, Dzhankoy, and Troitsk, inhabitants were also worried about increasing peripheralization and the closure of old industrial enterprises. As noted above, in Crimea people have pointed out especially positive changes, due to new trends of development and expectations on Russian assistance. Semantic tests used to identify the associations people held with regards to their place of residence (Fig. 4) show the correspondence of the above problems and negative definitions ('undeveloped', 'backward', 'poor', etc.). Positive definitions given to participants' own cities (e.g., 'beautiful', 'native', 'green', 'ours', etc.), especially in the case of Troitsk, mostly indicate local patriotism and the attractiveness of the urban landscape. The prevalence of negative characteristics ascribed to Zabaikalsk appears to be related to the socio-economic contrasts across the border and a permanent comparison with neighbouring Manzhouli. While over the course of 25 years, the small town of Manzhouli has become a large city, with a population of 300,000, developed communications, skyscrapers, and night illumination, Zabaikalsk back on the Russian side of the border has remained a small peripheral town, separated from the regional capital by 500 km of uninhabited territory.

The massive exodus of young, active people in search of a better life correlates well with the majority share of negative definitions of Zabaikalsk, located far from the economically developed Russian centres. People in our case study cities choose mainly migrate to large and prosperous Russian cities, rather than neighbouring countries. This choice is indirectly confirmed by a positive perception of Russia (Fig. 5)[1], mainly based on statist patriotism and an emotional attachment to the Homeland.

The usage of projective methods also confirmed our thesis about the special attitude held by residents of small border cities to their country as a whole. People believe that 'the state and government' is exclusively responsible for the development of distant border territories. They feel entitled to border protection because they perceive their presence on the border as a defense agaisnt the territorial claims of their neighbours. Therefore, in their view, the state must provide them with a decent standard of living, to compensate for their burden of protecting the state border and enduring the risks associated with living in a distant border territory. Moreover, people believe that border areas need a special 'border status' that would better facilitate daily life and cross-border interactions with neighbours. However, locals also do not believe that they can actually affect decision-making by higher authorities regarding local development, foreign policy, international relations with neighbouring state, the border regime, and so on.

Societal attitudes, hopes, and expectations that shape everyday life in the local border community almost correspond with the public mood in Russian society as a whole. The structure of fears negatively reflects the structure of a community's values. Our research shows that residents of small border cities are most afraid of losing their relatives, illness, poverty, unemployment, deterioration of relations with their neighbouring country, and especially war. According to a survey of the Levada Center in 2017, Russians' worst fears were illness among close relatives, poverty, and warfare (levada.ru 2017). The fear of illness is not a fear of any particular disease, but an expression (in a negative form) of that which is considered most essential: health. Fear of poverty reveals an inverted feeling of social and economic helplessness, and is a reflection of social vulnerability and insecurity. While fear of poverty was not one of the greatest fears of those living in small border cities, it is reflected in the concerns by Zabaikalsk residents about the need for survival. Lev Gudkov (1999) classifies these fears as those of unarticulated, background, uncertain anxiety and panic modes, which are connected with the level of stability of social ties and institutes in society. The fact that such diffuse fears are widespread across all our case studies indicates the strong sense of vulnerability in borderland communities, and explains why people express their hope for improvements in the socio-economic situation, living standards, level of wages, and stability.

CONCLUSIONS

The studies carried out at different sectors of the Russian border allow us to compare cross-border relationships and interactions with neighbours at the local level and the patterns of perception of neighbours on borders with different regimes and functions, as well as their impact on everyday life and people's well-being in conditions of local development deficit. Studying cities located directly on the border, the research aims to provide insights for understanding the conjunctions of different feelings and forms of proximity and distance, based on (un)familiarity, (un)similarity and (un)connectivity, with a national neighbourhood that reflects the (geo)political situation, level of cooperation, and relations between neighbouring countries, as well as socio-economic contrasts across the border. Such a complex construction helps to bring us closer to an understanding about the mutual interweaving of the state and the everyday life of ordinary individuals, as well as grassroots challenges to the territoriality of national borders and power relations.

Russian reality sometimes runs counter to theory. The considerable differences encountered at the Russia-Chinese border do not restrain interactions. Yet the Chinese environment continues to be alien, despite the presence of Russian-language signs, the widespread use of pidgin, and intense interaction during cross-border shuttle trade. Russians simply do not feel comfortable in China. The different culture, lifestyle, and behavior of Chinese renders them 'Other' to those from neighbouring Russian cities. While the residents of border territories regularly cross the border, they participate in shuttle trade solely because they need to survive in the absence of other livelihood opportunities. Despite the numerous cases of fraud or deceit by Chinese dealers, residents of neighbouring territories still go to China in search of money.

The concepts of 'otherness' and 'strangeness' are confirmed by the situation in Northern Crimea; despite new trends, attitudes to people across the border have remained friendly, and such people are still perceived as 'Insiders'. On the border with Kazakhstan, although Kazakhs are still perceived as intimates, different nation-building processes and the long distances involved have increased the sense of difference and promoted the opinion of Kazakh neighbours as 'Strangers'.

Thus, institutional and social proximity on the border

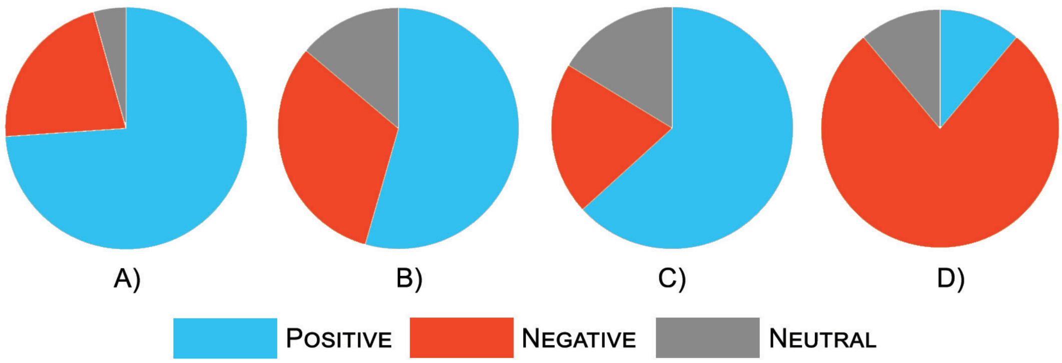


Fig. 4. Characteristics ascribed to participants' city of residence (A: Armyansk; B: Dzhankoy; C: Troitsk; D: Zabaikalsk)

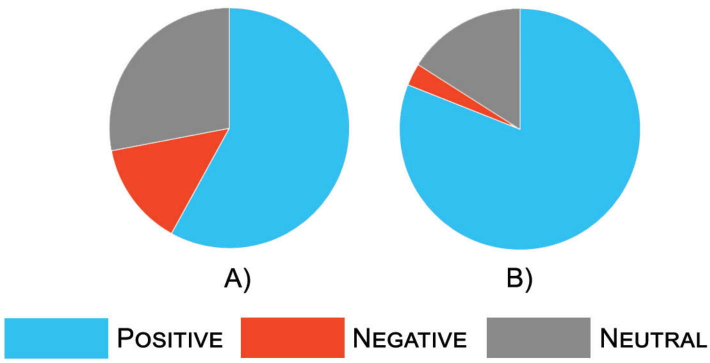


Fig. 5. Characteristics ascribed to Russia in general (A: Troitsk; B: Zabaikalsk)

[1] *Semantic tests to identify participants' associations with Russia were not conducted in Crimea.*

with Kazakhstan is neither necessary nor sufficient to the deepening of cross-border interaction. Despite the partnership and integration between countries, the intensity of border communications is less than expected. A low density of economic activities, poor infrastructure, and low population mobility still remain significant factors. Moreover, the simple lack of incentives to cross the boundary can play a much greater role than the obstacles that must be overcome to cross state borders. When motivation is weak, the abolition of barriers does not change anything (Zotova et al., 2018).

Despite lots of objective differences between the various sections of Russian borders, their subjective perceptions by people and social representations serve to unite them. Paying special attention to urban life during the focus groups has revealed that border towns are perceived primarily as a 'place of life'. Numerous everyday contacts with neighbours, including cross-border trips, force people to compare life here and there. The border is not considered anything extraordinary for the people living beside it, and any problems created by the border are perceived as secondary compared to the stagnation of urban development. Geographical distance and the low standard of living in small border cities promote the sense of abandonment, together with paternalism and a strong sense of frustration.

Local populations transfer the responsibility of their wellbeing to the government, because they believe that the state is required to provide them with a decent living for their important role as defenders of national borders. People expect that the Russian government will develop a well-balanced and comprehensive border policy; they hope border areas will receive a special status and the implementation of exclusive development programs. In the current socio-economic situation, in which 'the government does not provide an adequate standard of living', local populations are forced to use opportunities opened by the border and neighbouring state. That is why a potential border closure is perceived negatively as something that could undermine the last reason for living in the border city. Unfamiliarity, otherness, and significant differences between neighbouring territories are not allowed to get in the way of contact, because it is this contact that allows local residents to make a living.

ACKNOWLEDGMENTS

The research was supported by State Assignment IGRAS (0148-2019-0008, AAAA-A19-119022190170-1). Theoretical part of the article was prepared with partial support of the Immanuel Kant Baltic Federal University by the Russian Academic Excellence Project (5-100).

REFERENCES

Aure M. (2011). Borders of understanding: re-making frontiers in the Russian-Norwegian contact zone. Ethnopolitics, 2, 171-186.

Balogh P. (2013). Sleeping abroad but working at home: cross-border residential mobility between transnationalism and (re)bordering. Geografiska Annaler: Series B, Human Geography, 95(2), 189-204.

Boschma R. (2005). Proximity and innovation: A Critical assessment. Regional Studies, 39(1), 61-74.

Brambilla C. (2015). Exploring the critical potential of the borderscapes concept. Geopolitics, 1, 14-34.

Brazhalovich F.L., Kljuchnikov M.I., Kolosov V.A., Pavljuk S.G., Popov F.A. and Turov N.L. (2017). Life over the conflict borders: socio-economic aspects of cross-border relations within divided cities (on the example of Bender and Dubossary). Izvestiya Rossiiskoi Akademii Nauk. Seria Geograficheskaya, 3, 45-57 (in Russian).

Dolińska K. and Niedźwiecka-Iwańczak N. (2017) German neighbours from across the river – Insiders? Strangers? Others? Social Studies, 1, 95-116.

Domaniewski S. and Studzińska D. (2016) The small border traffic zone between Poland and Kaliningrad region (Russia): The impact of a local visa-free border regime. Geopolitics, 21(3), 538-555.

Ghosh S. (2011). Cross-border activities in everyday life: the Bengal borderland. Contemporary South Asia, 19, 49-60.

Gudkov L. (1999). Fear as a frame of comprehension. Monitoring public opinion 1999, 6 (44), 46-53.

Helleiner J. (2009). Young borderlanders, tourism work, and anti-Americanism in Canadian Niagara. Identities: global studies in culture and power, 16 (4), 438-462.

Houtum H. van (1999) The Geopolitics of Borders and Boundaries. Geopolitics 10(4), 672-679.

Koch K (2018). The spatiality of trust in EU external cross-border cooperation, European Planning Studies, 26(3), 591-610

Kolosov V.A. and Zotova M.V. (2019) China and Russian-Chinese relations in the mirror of Russian discourse. Problems of Geography, 148, 281-308 (in Russian).

Kozera B. (1999). «Stranger» and «Insider». Three reflections of an uninstrumental reason. In: Others insiders. Studies on ethnic issues. Katowice, 39-48.

Laine J. (2016). Threats, challenges, and Finnish-Russian cross-border security cooperation: a Finnish perspective tourism and geopolitics. In: Derek Hall, ed., Tourism and geopolitics: issues and concepts from central and Eastern Europe, Wallingford, CAB International, 178-191.

Levada.ru, (2015). Fears [online] Available at: www.levada.ru/2017/11/28/strahi-2/ [Accessed 20 Dec. 2019] (in Russian).

Newman D. and Paasi A. (1998). Fences and neighbours in the postmodern world: boundary narratives in political geography. Progress in Human Geography, 22 (2), 186-207.

Rippl S., Petrat A., Kindervater A. and Boehnke K. (2009). Transnational social capital: Are border regions a laboratory of social integration in Europe? Berliner Journal für Soziologie, 19, 79-103.

Schack M. (2001). Regional identity in border regions: The difference borders make, Journal of Borderlands Studies, 16 (2), 99-114.

Scott J.W., Celata F., Coletti R. (2019). Bordering imaginaries and the everyday construction of the Mediterranean neighbourhood. Introduction to the special issue European Urban and Regional Studies, 26(1), 3-8

Spierings B. and Velde M. van der (2013). Cross-Border Differences and Unfamiliarity: Shopping Mobility in the Dutch-German Rhine-Waal Euroregion. European Planning Studies, 1, 5-23

Stoklosa K. (2013). The border in the narratives of the inhabitants of the German-Polish border region. In: European Border Regions in Comparison, 257-274.

Szytniewski B. and Spierings B. (2014). Encounters with Otherness: Implications of (Un)familiarity for Daily Life in Borderlands. Journal of Borderlands Studies, 29(3), 339-351.

Szytniewski B., Spierings B. and Velde M. van der (2017). Socio-cultural proximity, daily life and shopping tourism in the Dutch-German border region. Tourism Geographies, 19(1), 63-77.

Timothy D. J. (1995). Political boundaries and tourism: Borders as tourist attractions, Tourism Management, 16(7), 525-532

Torre A. and Rallet A. (2005). Proximity and localization. Regional Studies, 39(1), 47-59.

Trippl M. (2010). Developing cross-border regional innovation systems: Key factors and chal- lenges. Tijdschrift voor Economische en Sociale Geografie, 101, 150-160.

Velde B. and van der Spierings B. (2008). Cross border shopping and the «bandwith of familiarity»: exploring the positive impact of national borders on consumer mobility in the Euregion Rhine, Nijmegen, Nijmegen School of Management.

Vendina O.I. and Kolosov V.A. (2007). Partnership to bypass barriers. In: Russia in global politics, 1, 142-154 (in Russian).

Zotova M., Gritsenko A. and Sebentsov A. (2018). Everyday Life in the Russian Borderlands: the motives and determinants of cross-border practices. Mir Rossii, 4, 56-77 (in Russian).

Received on September 17th, 2019 Accepted on January 21th, 2020

Alexander B. Sebentsov[1,2*]
[1] Institute of Geography, Russian Academy of Sciences, Moscow, Russia
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia.
***Corresponding author:** asebentsov@igras.ru

CROSS-BORDER COOPERATION ON THE EU-RUSSIAN BORDERS: RESULTS OF THE PROGRAM APPROACH

ABSTRACT. Cooperation across the Russia-EU border has been drawing much attention in recent years. The majority of studies point out programs' efficacy, high density of border institutions and resistance to geopolitical risks among other factors. These advancements can be explained by the theory of multilevel collaboration which implies that diverse and multiple cooperation institutions can effectively distinguish matters of high politics from practical issues concerning interests of those living along the external borders.

The article aims to analyze the impact of cross-border cooperation programs (CBC Programs) on the thematic, institutional and spatial structure of the cross-border relations.

The research is grounded in the overview and analysis of a large volume of empirical data including reports and descriptions of cross-border cooperation programs, data provided by the regional governing agencies, as well as 76 semi-structured interviews obtained from regional experts as part of several research expeditions by the Laboratory of Geopolitical Studies of the Institute of Geography RAS taken place over the period from 2011 to 2018.

Main characteristics and long-term trends of the cross-border program approach are examined as follows: growth in governmental coordination on various agency levels aimed at development and implementation of mutually beneficial partnerships, creation of joint program management bodies, development of uniform policies and joint funding sources for projects, and interest in maintaining an equal level of collaboration.

It is revealed that gradual rise of the programs' role in cross-border cooperation in the area contributed to the restructuring of its institutional systems, launching selection process for the existing border institutions (euro-regions, cross-border regional councils, and others), as well as triggering the synergy effect among them and the transborder forms of cooperation.

The main characteristics of spatial partnership structures are identified. These include those consisting of high concentration of project activities taking place within large urban centers along the external borders and those asymmetrical to cross-border interactions. The former is especially pertinent to the Russian side of the border where just a small number of such centers are involved in up to 70-80% of project activities. Even fewer number of Russian cities initiate their own collaborative projects. A gradual spatial shift of cooperative projects toward the areas immediately proximate to the borders, as well as the decrease in asymmetry of transborder cooperation are identified as the new trends by the author.

KEY WORDS: cross-border cooperation, European Neighborhood Policy, CBC Programs, Russian–EU relations, regionalization

CITATION: Alexander B. Sebentsov (2020) Cross-Border Cooperation On The Eu-Russian Borders: Results Of The Program Approach. Geography, Environment, Sustainability, Vol.13, No 1, p. 74-83
DOI-10.24057/2071-9388-2019-136

INTRODUCTION

In recent years, the experience of cross-border cooperation (CBC) with countries adjoining the Northwest border of Russia has been characterized by innovation and may serve as a model for other border areas of Russia. Over the past two decades, studies of different approaches to cooperation (Mezhevich 2009; Kuzneczov 2004; Sebentsov 2018, Scott 2015) showed that the European experience is utalized on other borders of Russia as well. Despite the noticeable cooling in relations with European neighbors and the EU as a whole, it is on the borders with the European Union where the highest density of various cooperation institutions have been achieved, which allowed L.B. Vardomsky (2008; 2009) to identify a special «European» or institutional cooperation type. Institutionalization of CBC in modern Russian and foreign studies is usually viewed as a necessary condition for effective cross-border interactions. A.S. Makarychev (2002) explains this by high «transaction costs» of border actors facing interstate differences in economic and legal regulations, institutional environment, etc. At the same time, numerous studies show that not all institutions are equally effective, and the development of some institutions is often accompanied by the decline of others. Moreover, formal institutions do not always work in the actual trans-border interactions practice, while numerous informal institutions have a great influence on cooperation development.

CBC Programs have become one of the most important institutions of CBC in the last twenty years. Recent studies show that CBC Programs determine the behavior and motivation of most actors of cooperation, and therefore radically change other formats of trans-border interactions (Korneevets et. al. 2010; Scott 2015; Gumenyuk 2018; Sebentsov et. al. 2018).

Thus, the purpose of this article is to show how CBC Programs affect cross-border cooperation with Russia's northwestern neighbors. This article describes a brief analysis of the program approach evolution in the EU neighborhood area, the main trends and patterns of CBC development within the programs framework, and the impact of these trends on other institutions and the spatial structure of cooperation.

MATERIALS AND METHODS

To address study goals, as described by this paper, required a synthesis of considerable material. This included numerous documents describing the regulatory and associated definitions related to CBC content at the border with the EU. Among them, it is necessary to highlight the actual texts of cooperation programs, which made it possible to analyze the evolution of the program approach. Other important sources of information were the reports on the CBC Programs implementation posted on the special internet portal called Knowledge and Expertise in European Programs (KEEP)[1]. Since different Territorial Cooperation Programs have different capabilities related to work with the data to be supplied to KEEP, the level of completeness and update of the data in KEEP varies considerably from program to program. Therefore, additional data were requested directly from CBC cooperators and associated governmental administrations. The collected materials were then subjected to geocoding, which allowed visualizing the cooperation projects data and their participants; territorial structure and networks of cooperation.

The third set of materials was represented by 66 expert semi-structured interviews, which were collected during expeditions of the laboratory of Geopolitical Studies of the Institute of Geography of the Russian Academy of Sciences from 2011 to 2018 in following regions: Kaliningrad (2011 – 12 interview, 2014 – 13), Karelia (2014 – 15), Pskov(2015 – 8), Leningrad (2015 – 10), Saint-Petersburg (2018 – 8). Interviews were conducted with representatives of federal, regional and local authorities, public organizations, businessmen, CBC participants and researchers. 10 interviews were conducted with profile departments of Ministry of Foreign Affairs (2018) and Ministry of Transport (2017–2019).

The evolution of the program approach at the borders with the EU

CBC programs first emerged in Europe. The first prototypes of such programs were several cross-border development plans initiated between 1972 and 1989 within the Euroregion (EUREGIO), located on the border of Germany and the Netherlands (Scott 1993; Perkmann and Sum 2002; Perkmann 2003). Financial support for these plans was provided at the request of EUREGIO by the European Commission, which then took into account the experience gained in the Interreg Program development. The main idea of the new program, launched in 1989, was to support cooperation between different territories within the European communities (Yarovoj 2007). These could be relatively compact border areas (Interreg A, cross-border cooperation), more extensive trans-border areas (Interreg B, trans-border cooperation), as well as regions for which immediate proximity was an optional feature (Interreg C, interregional cooperation).

The Interreg launch contributed to the growth of a number of Euroregions in the late 1980s and early 1990s, providing them not only with financial resources but also with proven practices and cooperation formats agreed at the regional, national and supranational levels (Perkmann 2007). Already in 1989, several regulations were adopted that allowed the implementation of Interreg Programs, as well as the use of funds of the European Regional Development Fund (ERDF) not only internally but also with respect to the external borders of the EU (McCall 2015).

For Russia, the experience of CBC became available after Finland and Sweden entered the EU (1995). Interreg IIA «Karelia» (with the participation of the Republic of Karelia) and «Southeast Finland – Russia» (with the Leningrad region and St. Petersburg) were implemented on the Russian – Finnish border. Cooperation with the Norwegian and Finnish border regions was carried out within the Barents Region Program framework. The atmosphere of mutual optimism in relations between Russia and European countries was a good background for the development of cross-border cooperation. A concrete manifestation of this was the 1997 Partnership and Cooperation Agreement (PCA), which created a political and legal framework for further cooperation. (The EU – Russia borderland… 2012) Despite high expectations, the actual, fully entitled participation for the Russian side in this program was small (Shlyamin 2002). The federal and regional authorities did not have sufficient resources to support cooperation, and the possibility of financing projects from EU funds was extremely limited (Karelia CBC... 2007).

The first significant funds for CBC were received within the framework of the Tacis CBC Small Project Facility Program which financed 146 projects from 1996 to 2003. A wide range of entities could apply for proposals under this program including regional authorities, local authorities, public institutions (hospitals, schools, universities, museums, non-profit organizations, etc.). However, despite the presence of a special Regional Support Bureau in St. Petersburg with two offices in Petrozavodsk and Kaliningrad, there was little progress in coordinating with Interreg Programs. (Pooling of financial resources of TACIS… 2001)

In the new program period of 2000-2006, the greatest success in coordination was achieved on the Russian-Finnish border, where, thanks to the Euroregion «Karelia» formed in February 2000, the first joint «Our common border 2001-2006» Program was prepared in October of the same year based on the Russian «Cross-border cooperation program of the Republic of Karelia, 2001-2006» and the European Interreg IIIA «Karelia» Program. This approach facilitated the identification of common objectives and the creation of a list of common projects. The cooperation results on other parts of the border were much more modest, including with the new EU members (Poland, Lithuania, Estonia, and Latvia), where two new Interreg Programs were also launched.

By 2004, the accumulated interaction experience showed that greater success in implementing joint activities would require better coordination between Tacis and Interreg instruments (Paving the for a New Neighborhood..., 2003; The EU's Eastern Neighbourhood…, 2016). As a result, existing CBC Programs were transformed into Neighborhood Programs. According to the new approach, the programs were to be developed with the participation of regional and local authorities on both sides of the border. Also, the achievement of a higher degree of coordination in project management and the use of various financing instruments was expected. Thus, Russian participants and their European partners had the opportunity to submit joint applications for project financing.

However, there was no unified approach to the review and approval of applications. For European partners, the competition procedures were held at the local managing authority of the Interreg Program, and for Russian participants

[1] *KEEP is a free database on Territorial Cooperation projects, project partners and programs (including CBC Programs within the scope of the Instrument for Pre-Accession and the European Neighbourhood and Partnership Instrument). It covers the financing periods starting in 2000. The KEEP database is Available at: https://www.keep.eu (Accessed: 31 Aug. 2019)*

– at the European Commission Delegation in Russia, which managed the TACIS Program funds (The KOLARCTIC ENPI… 2007). In addition, numerous inconsistencies between the financial and organizational procedures of the two programs led to persistent funding failures.

The new program period of 2007–2013 brought several significant changes to the cooperation programs. On January 1, 2007, the new European Neighborhood and Partnership Instrument (ENPI) came into effect providing the financial structure for the implementation of the new European Neighborhood Policy. ENPI allowed uniting all internal and external sources of funding from each of the countries in addition to participants' own financial resources. This ensured greater financial procedural uniformity and the timely nature of funding.

The principal changes were also intended to accommodate Russia's earlier reluctance to build relations with the EU within the framework of the European Neighborhood Policy. Instead, Russia had proposed an equal strategic partnership based on four common spaces as an alternative. In addition, Russia expressed its intention to participate in the development and financing of new cooperation programs. In 2009, a package of agreements between the Russian government and the European Union on financing and implementing five new CBC Programs from 2007 to 2013 was signed in Stockholm.

The organization of work needed to implement these programs took into account multilevel management principles, which engaged different levels of authority (supranational, federal, regional and municipal), as well as various actors – direct participants of cooperation. Another innovation was the creation of common governing bodies: the Joint Monitoring Committee (JMC), the Joint Technical Secretariat (JTC) and the Joint Project Selection Committee (JPSC). The JMC included representatives of central, regional and local authorities, and civil society representatives in some cases. Their tasks included development of the content aspects of the program, JTC creation, identification of project selection criteria, creation of the JPSC, and choosing the Joint Managing Authority (JMA). One of the executive power bodies of the EU member state participating in the Program often acted as the JMA. It was in charge for the program implementation, including technical assistance, operational and financial management. This approach made it possible to determine the basic rules of the game in advance by which numerous actors of cooperation were to act.

NUTS II and NUTS III[1] level provinces and municipalities adjacent to the shared EU borders and constituting the main territory of the programs could participate in the cooperation. Also, the possibility of indirect participation was provided to neighboring regions not adjacent to the border. In practice, the division of the program into the main and adjacent areas meant different opportunities for participation in projects. Only partners from the main program area could count on significant financial resources. In Russia, this formally involved huge areas covering the vast majority of the Northwestern portion of the country, but the border regions constituted its core (Fig. 1). The only exception was the «Kolarctic» Program. Portions of the Swedish, Norwegian and Russian areas (respectively Norrbotten, Troms and Nordland, Arkhangelsk region and Nenets Autonomous Okrug) do not have access to the land border. The explanation for this configuration of the program area is that «Kolarctic» is one of the key financial instruments of cooperation in the Barents Euro-Arctic region.

The development and launch period of new CBC Programs in 2014-2020 coincided with the geopolitical crisis in relations between Russia and its Western partners. Mutual sanctions and the rapid curtailment of bilateral ties between Russia and the EU gave reason to believe that CBC would also be frozen. However, the European Commission adopted a special decision not to apply sanctions to projects implemented under the new programs. After taking into account all the sanctions related risks, Russia also made the decision in favor of further cooperation (Conclusion of The State Duma Committee… 2018). Documents describing the implementation of these programs were agreed upon in December 2015. Intergovernmental agreements on funding and implementation of the programs were signed in 2016-2017, and the ratification of these agreements by the Russian side was completed by mid-November 2018. As a result, the financial agreements execution between the participants of specific cross-border projects was planned only for 2019, and the period of programs implementation is likely to be extended until 2024

The main innovation of the program period was the funding instrument reform. According to official documents, the European Neighborhood Instrument (ENI) provides for a more rapid and flexible financing of the European Neighborhood Policy, but it is not yet clear how this will specifically affect the work of the programs themselves (Programming of the European Neighborhood… 2013). Another innovation was the transition from multilateral to bilateral cooperation, which was informed by «political challenges» and «poor coordination of projects» by the regional and local authorities of two or more countries (Programming document for EU... 2013). The only exception was the «Kolarctic» Program, where the main cooperation area still covers the territories of four countries at once (The KOLARCTIC ENI CBC… 2015).

It was evident from interviews conducted by the author in 2014 and 2015 that this decision caused different reactions among the developers of new programs. The former developers of the «Estonia-Latvia-Russia» Program assessed the changes neutrally in general, believing that nothing would likely change. In contrast, for the Kaliningrad region, where the large «Poland-Lithuania-Russia» Program was divided into «Poland-Russia» and «Lithuania-Russia» Programs, most of the surveyed experts believed that the state of cooperation would deteriorate. This was predicated on the division of the program, which could lead to the lost opportunity for Russians to unite with one of the European partners for protection of their interests and promoting their objectives.

Evolution of cross-border cooperation within CBC Programs

Without exception, all CBC Programs were created to reduce costs related to the territory border status. These are their periphery, insufficient level of border infrastructure development, and the «soft security» issues. Other tasks of the Programs were the internal resources search for border region development and common natural and cultural heritage management. Thus, the borderland specifics largely determined the main cooperation directions. From the very beginning, however, the European Commission and Russia's neighboring countries had the pre-emptive right to establish cooperation priority areas.

Under the first TACIS Programs (1996-2003) cooperation was not systematic due to lack of coordination with foreign partners, joint management bodies and common priorities: funding was allocated on a competitive basis, and did not benefit the priorities developed for a specific program

[1] *NUTS or Nomenclature of Territorial Units for Statistics is a geocode EU standard for referencing the subdivisions of countries for statistical purposes. There are three levels of NUTS defined: NUTS I – national, NUTS II – regional, NUTS III – sub-regional.*

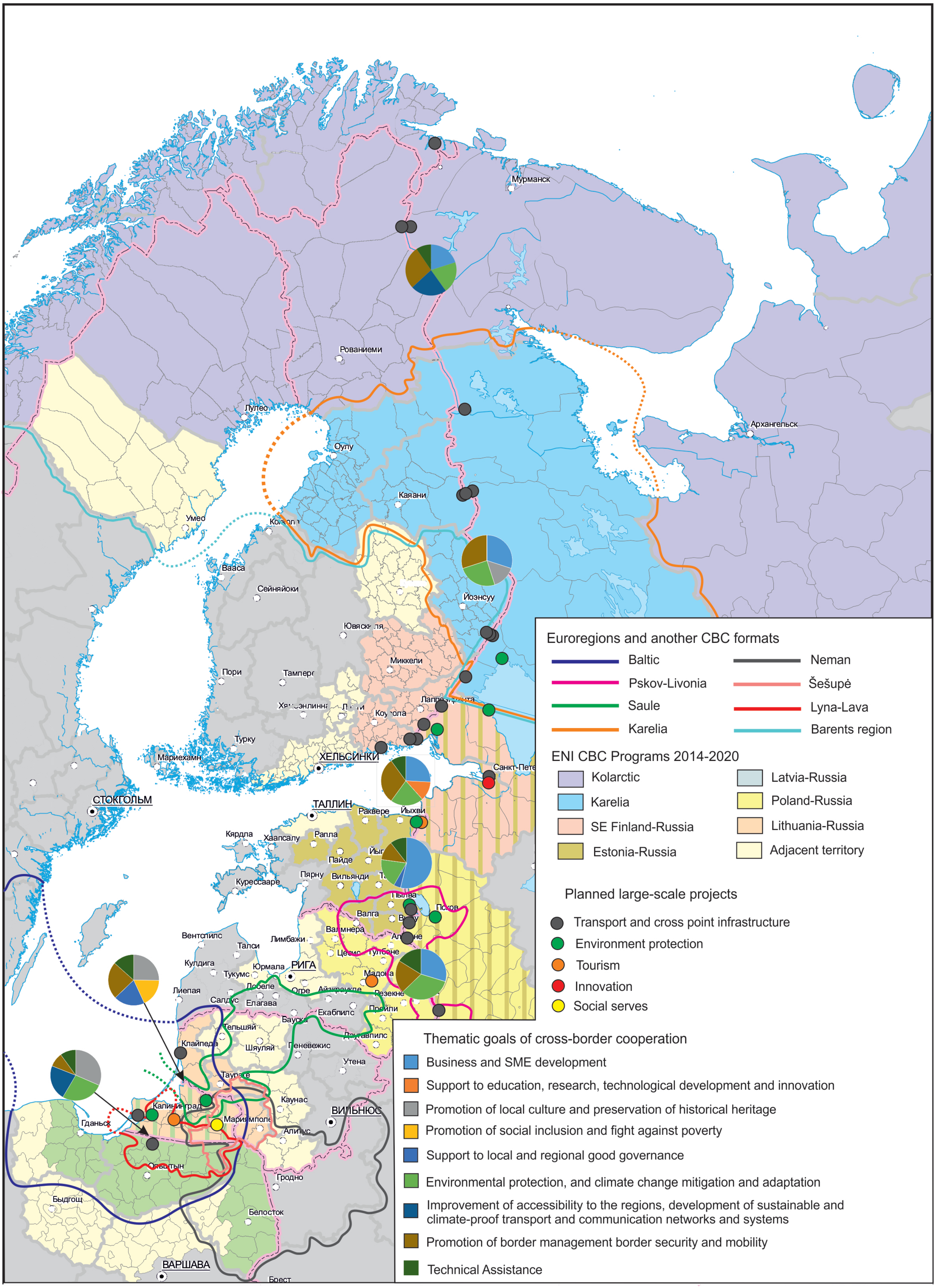


Fig. 1. CBC Programs 2014-2020 and other CBC formats on North-West borders of Russia

area. Approximately half of the funds were allocated for the infrastructure development, including construction at border crossing points. (Yarovoj and Belokurova 2012). Another 25% was spent on environmental protection and business support. These were primarily large projects with a budget exceeding 2 million euros. The remaining funds were spent on the realization of a wide variety of «small projects», the main beneficiaries of which were local and regional authorities. The funding amount for such projects was limited to 50 thousand euros, and the main implementation forms were various seminars, international exchanges, internships, and training programs. (Pooling of financial resources of TACIS…2001; TACIS cross-border cooperation… 2001)

In subsequent programs, in which regional and national authorities were more actively involved, cooperation became more systematic and varied in content. It is noted for the period 2004-2006 Russian opportunities to influence the content of TACIS/Interreg Programs were small. However, in 2007, after the start of Russian co-financing of ENPI Programs, federal and regional authorities took a more active position. Detailed CBC programming was carried out in 2007-2013 by development teams represented mainly by regional authorities (NUTS II territorial level). The European Commission approved the finished Programs, where they were checked for compliance with the priority tasks of the European Neighborhood Policy in the field of CBC.

The program developers had to take into account at least two of the four objectives set by the EU, but could also offer their own, focusing on the territory characteristics. The most popular objectives were those with limited funding, which could have the most visible effect. Thus, the need to solve the periphery problem, which was typical for almost all areas of the program, contributed to the popularity of the first objective which was the promotion of economic and social development within border regions. The second and fourth objectives (solving environmental problems and developing cross-border communications between people) also often became a part of cooperation programs. Only the third objective (ensuring efficient and secure borders) was not included in any of the programs under consideration, as it falls within the exclusive responsibility of central authorities. However, in reality, this objective was indirectly taken into account in all project proposals submitted under the first priority.

Nevertheless, the content side analysis of the implemented projects shows that the interaction trajectories remained the same. Thus, one of the key priorities of cooperation in both program periods was environment protection, which accounted for more than 20% of funds in TACIS/Interreg Programs of 2004-2006 and in ENPI Programs of 2007-2013. The main cooperation objects were transboundary water basins, where the central problems were related to the discharge of untreated industrial and domestic wastewater into the reservoirs. There were specific problems in the Barents Sea region, where various radiation safety issues proved to be the most immediate concern.

The second direction involved infrastructure development (about 18 and 43% respectively), including the roads reconstruction, checkpoints construction, street repair in border settlements, etc. For example, on the border with Lithuania, construction of the Panemune-Sovetsk bypass with a bridge over the Neman River was carried out, the Panemune and Kybartai checkpoints were modernized. Similar projects to expand the checkpoints capacity were carried out on other parts of the border.

Much attention was paid to the development of regional and municipal self-government (17.3 and 13.2%, respectively) which was supposed to be stimulated through cooperation in solving common problems or through joint spatial planning. Thus, on the Russian-Estonian border, the municipalities of Kohtla-Jarve (Estonia) and Slantsy (Leningrad region), which in Soviet times were part of a single complex of the Baltic shale basin, were engaged in the joint development of projects intended for the reconstruction of currently unused industrial facilities and the landscaping of the main city streets. Another example is the joint strategy for the development of twin cities – Finland's Imatra and Russia's Svetogorsk.

Tourism and cultural development projects were of particular importance in some parts of the borders with Finland, Poland and Lithuania. The issues of the common cultural heritage preservation, as well as the cross-border tourism development made such projects a natural priority in the area. Special emphasis in the Program was placed on the joint creation, preparation and development of the technical and economic basis for cross-border tourism products, joint activity on their promotion, classification and certification, as well as tour guide training.

In the 2007-2013 Programs, an important innovation was the so-called large-scale projects (LSP) with an investment component, for which up to 30% of the total budget could be spent. Their emergence was a response to requests by national and regional authorities, who hoped that large projects would have a more visible impact on the ordinary citizens' lives.

One of the key criteria for the implementation of these projects was geographical – they had to be implemented only in the main territory of the program, create a large-scale effect on the border areas, and have the unconditional support of national and regional authorities. The projects were also to be in compliance with national and regional strategies. The only exception to this rule are the two projects under the Kolarctic Program: reconstruction of the Arkhangelsk airport and the wind park in the Nenets Autonomous Okrug. These project locations are distant from the borders with Finland and the Scandinavian countries, at a distance of over 500 and 1000 km, respectively.

As a result, the vast majority of LSP was related to infrastructure development, which greatly affected the overall cost structure. For example, the Ivangorod-Narva and Pechora-Orava checkpoints were upgraded on the Estonian part of the border. This work included extending and repairing access roads, installation of new x-ray equipment for customs control, and construction of new terminals. Works similar in content and scale were carried out on the borders of Russia with Lithuania (Kybartai-Chernyshevskoye and Panemune-Sovetsk checkpoints) and Finland (Svetogorsk-Imatra, Brusnichnoye-Nuijamaa).

Considerable funds were invested in urban sewage systems in areas where large transboundary water basins cross borders. On the Russian-Latvian and Russian-Estonian borders, the projects were aimed at improving the environmental situation in the Peipsi lake reservoir and protecting the waters of the Narva river basin. As a result, sewage treatment facilities were modernized in several settlements of Estonia, the city of Gdov, and new sewage treatment facilities were built in Pskov. At the Finnish site, the largest was the «Clean Ladoga» project, in which the sewage treatment facilities were built in Sortavala. On the Lithuanian border, the main efforts were focused on improving the environmental situation in the Neman river basin which included construction of a sewage network and sewage treatment facilities in the towns of Neman and Slavsk (Kaliningrad region). Both sewage and water supply networks were developed for the villages of Pakalniškiai and Dumpiai (Klaipeda, Lithuania).

The only LSP in the tourism field in the Program «Estonia-Latvia-Russia» was associated with the development of a unique complex of Narva and Ivangorod fortresses as a single object of culture. Parts of the walls and buildings of both fortresses were reconstructed using the allocated 3.4 million euros. This work included improving access for disabled visitors and development of joint tourist routes.

Despite the fact that the new ENI Programs (as mentioned earlier) are only entering the stage of implementation for the first projects in 2019, it is possible to draw our first conclusions about the key cooperation areas of this period. The indicative financial plans analysis shows that the key cooperation topics will lie within border infrastructure development, which will draw about 25% of all allocated funds, support for entrepreneurship (about 23%) and environmental protection (about 21%). The traditional orientation of some segments of the borderland (Russian-Polish, Russian-Finnish, and partly Russian-Estonian) with regard to cooperation in the field of culture and preservation of the shared historical heritage is preserved. Innovation projects became a new phenomenon – it is planned to allocate about 15% of funds to such projects within the framework of the «South-Eastern Finland – Russia» program.

The trend towards an increase in the value of LSP, which emerged in the previous program period, remains. This is expressed in the direct selection of LSP at the stage of the program creation with the participation of developers and central authorities. Such approach should increase the number of projects implemented directly in the border area, as it is characterized by a special regime of visits and economic activities. Among the projects selected and approved for implementation, most (about 70%) are related to the transportation and border infrastructure development (e.g. road construction and the checkpoint reconstruction). About 20% of the projects are related to the resolution of environmental issues and involve construction of sewage systems in Pskov, Vyborg, and other locations. LSP in the field of culture and tourism is planned for the Kaliningrad region (development of a bicycle routes network and water tourism), as well as on the borders with the Baltic countries (development of joint tourist routes).

Thus, the wide participation of various actors in the development of priorities and implementation of concrete projects contributed to the thematic succession of programs. As a result, many projects started in one program period were continued in the next. This, in turn, led to the formation of long-term partner networks and spatial structures of cooperation.

CBC Programs and spatial structure of cross-border cooperation

The diversity of cooperation institutions is one of the key features of cross-border cooperation's spatial structure, distinguishing this Russian borders' part from others. There are three evolutionary stages of this cooperation in terms of spatial and institutional structure.

At the first stage, which is the period from 1991 to 1995, Russian regions and their neighbors signed numerous bilateral agreements on cross-border, trade and economic cooperation, mainly of a framework nature. In the same period, Intergovernmental Commissions on cross-border cooperation (e.g. with Latvia, Finland and Norway) or special Regional Councils (e.g. with Poland and Lithuania) started being established. Representatives of both regional and national authorities became their members on the basis of intergovernmental agreements. A special Council was established on Estonia's border, which, due to the difficult interstate relations with this country, worked under the regional authorities' auspices and transformed later in Euroregion «Pskov-Livonia». At the «old borders» with Norway, Poland and Finland, the Intergovernmental Commissions and Councils' main task was to establish checkpoints, discuss conditions of their work, regulate cross-border regime, and elaborate future cooperation's content. At the «new borders», pertinent issues of border delimitation and demarcation were added to these tasks.

At the second stage (from 1996 to 2004), the European Union became a new actor of CBC. Its active participation is manifested in the introduction of new forms and institutions of CBC, which had been already tested on other borders, including internal borders of EU. Additional effect is associated with the already mentioned TACIS border program, which provided these institutions with initial funding for their project activities (Yarovoj and Belokurova 2012). As a result, Euroregions began active development in this period (Fig. 1.). The Kaliningrad region was a pioneer in this matter, since it was literally «doomed» to CBC due to its exclave position, according to N. M. Mezhevich (2009, p.122).

Since 2004, after appearance of new Neighborhood Programs, the third stage of cooperation's spatial and institutional structure development arose. Its distinctive features are growth of coordination between the main actors (regional and national authorities), common program management bodies, coordination of joint priorities, and ensuring a high transparency with regard to competitive and financial procedures. Sustainability and large amounts of funding were especially attractive to Russian participants who could not find such funds within their country. As a result, the actors' rapid reorientation to Neighborhood Programs has had a significant impact on other CBC's formats, triggering selection processes among existing cooperation forms and institutions and also generating synergy effects between them.

Therefore, it can be assumed that the Neighborhood Programs contributed to the decrease in interest by regional and local actors to a whole number of Euroregions, particularly the small ones. Many of them initially had a narrow agenda (mainly environmental focus), and their territory covered only some municipalities in border regions. As a consequence of these features, the necessary financial and organizational resources for these cooperation projects were not provided. For example, the main activities in the Euroregion «Sheshupe» in recent years have been associated with organization of small events to improve the environmental situation in the Scheshupe River, as well as its fishing tournaments. The main project of the «Lyna-Lava» is an annual rafting on the Lyna-Lave River. The Euroregion «Saule» has not held meetings since June 2009. In the Euroregion «Neman», participation of Kaliningrad municipalities is almost imperceptible and is reduced to annual meetings.

On the other hand, the agenda of CBC Programs could be somewhat «narrow» for some of the Euroregions as well. For example, for the Euroregion «Baltic», with territory covering six Baltic countries, Kaliningrad border regional projects are quite local in relation to the goals and objectives.

A completely different situation is seen in the Euroregion «Karelia», that played the role of initiator of program approach in CBC. It participated in promoting its own interests in choosing cooperation priorities, actively using political, organizational and financial capabilities of regional and even national authorities for these purposes. In 2013, «Karelia» started its cooperation on program development for the period of 2014-2020. Later, in 2016, it adopted its own Strategy effectively resulting in the CBC ENI Program Karelia being considered a key financial tool for implementation. (Main Directions of Euregio Karelia… 2014)

An exceptional situation has been developing in the Barents region, where CBC Program «Kolarctic» became an important financial instrument for the Barents Regional

Council. As a result, this transborder institution shows typical Euroregion features in many areas of its activities. At the same time initially cross-border Program «Kolarctic» has acquired trans-border features. This conclusion is supported by continuing multilateral nature of cooperation, the enormous territory involved in cooperation, as well as considerable distance of many formally «border» projects from land borders.

Accumulating the vast majority of CBC projects, Neighborhood Programs began to determine its spatial structure's main features. First of all, attention should be paid to the **high concentration of cooperation projects in main centers of border regions (Fig. 2).** The main attraction points are the largest cities and administrative centers and capitals. This situation is seen on both sides of Northwest Russian state border, however, on the Russian side, the level of concentration is noticeably higher. Accordingly, a significantly smaller number of urban centers are involved in cooperation than on the neighbors' side. The absolute record belongs to St. Petersburg which was a partner of more than 70% of all CBC projects of the «South-East Finland – Russia» Program. The same city, being the most important Northwest center of Russia, attracted a significant part of the projects in other Programs (especially in «Estonia-Latvia-Russia» and «Karelia»).

The role of the administrative center in initiating cooperation projects is even higher. For example, in the first list of projects announced in CBC Program «Karelia» (2014-2020), all Russian projects without exception were initiated by actors from Petrozavodsk.

The capitals also become participants in CBC. In some cases (Finland, the Baltic States), the capitals are situated close to the border area or in adjacent territory of CBC Programs. As a result, they are involved in cross-border interactions (Fig. 2. B, E). Another reason for the participation of capitals in CBC is the involvement of certain central ministries and departments that are responsible for many contacts and projects in the border zone, operating of checkpoints, etc.

In addition, partners from national capitals are often involved as experts in implementing a number of complex technical projects (in the field of construction, environmental protection, etc.). Therefore, in the case of the «Kolarctic», «Karelia», and «SE Finland – Russia» programs, the involvement of capitals in cooperation projects remains noticeable, despite the great distance from the program territory (Fig. 2 A, C, D)

However, CBC has become «more cross-border». This is manifested in the fact that since 2007 CBC Programs are more concerned on border municipalities. As a result, the cooperation has partly shifted from large cities to less successful centers located directly near the border. These became platforms for the above-mentioned LSP (Fig 1, 2).

Another important feature of CBC spatial structure is its asymmetric nature. First of all, this is manifested by the results of CBC Programs that had been and remain much more visible on European partners' side **(result asymmetry).**

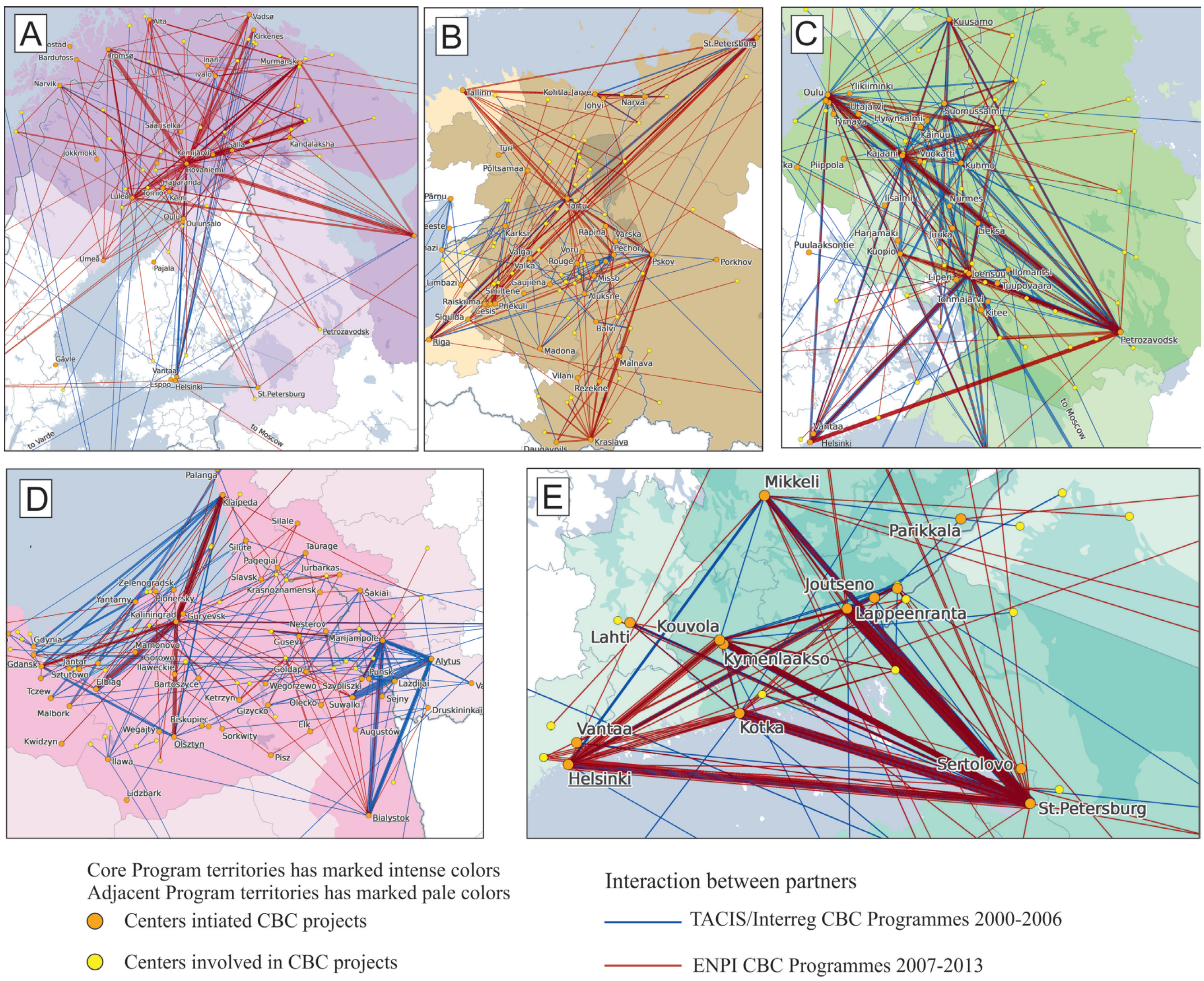


Fig. 2. Spatial structure of cooperation within the CBC Programs 2000-2013: Kolarcitc – A; Estonia-Latvia-Russia – B; Karelia – C; Lithuania-Poland-Russia – D; South-East Finland – Russia (complied by author using data from KEEP and regional administrations)

In 2004-2006 program period, an almost unbearable task for the majority of Russian partners was the requirement of co-financing of projects. Currency risks were also a factor. For example, Russian participants could plan their budgets only in rubles, so any jump in the exchange rate became a threat to the project. In addition, unlike European partners, Russian participants did not have supranational and national funding, which further enhanced **opportunities' asymmetry** (Lithuania-Poland-Russia… 2008).

Another set of serious obstacles for Russian partners was the language barrier, lack of experience in project activities, and lack of knowledge concerning European regulations. These circumstances were significant factors explaining why Russians didn't initiate the project **(initiative asymmetry**). For example, in case of CBC «Karelia» only Saint Petersburg and Sertolovo could be considered as initiators of cooperation. Eleven cities played the same role in Finish part of the borderland (Fig 2 E).

To be part of the CBC, Russian participants had to look for a «European partner» capable of preparing project documentation and taking on further management. In most cases, financing went through the European partners and they became the project's main beneficiaries. Russians often received small funds or became partners without financial participation. As a result, when building culture centers, roads or sports facilities, European partners received a ready-made project, while Russian partners received a standard design and rough estimate documentation (Sebentsov and Zotova 2018). For some participants interviewed by us (especially in Kaliningrad, Pskov and Leningrad regions), it was the sign of «unequal and asymmetric» cooperation, which was implemented primarily in the interests of neighbors and did not make a serious impact on the Russian regions and its socio-economic state. At the same time, the majority of experts agreed that cooperation can't be equal when the main financial obligations are assumed by the European side.

In 2007-2013 program period, the asymmetry decreased noticeably. This was related to the decision of the Russian Government to take part in co-financing of CBC Programs with the EU. Regional authorities also supported some of the CBC projects initiated by Russian participants. These measures increased Russian participants' opportunities to initiate cooperation. The experience gained by Russian participants in the previous program period also had an impact on increasing their project activity. However, the program still had a number of difficulties related to differences among institutions and legislation between countries. Thus, visa barriers (speed and cost of obtaining visas) were an obstacle for exchanges and seminars. The border regime, which restricts stay of foreigners in the five-kilometer zone along the border, had an impact on the projects in the Russian side of the border zone (Mezhevich 2009). Significant difficulties with interactions between authorities were related to the dissimilar distribution of competences between municipalities, regions and central authorities within the partnering nations. In interviews, foreign participants mentioned that the power structure in Russia was too centralized which led to Russian participants not having the ability to make independent decisions. Weak development of Russian civil society was also called a serious obstacle for cooperation. The small number of Russian NGOs, their isolation from external financial sources and the lack of internal funding made it difficult to find partners. In the financial sector, the major challenges were currency conversion, financial planning and reporting. Changing of these characteristics of institutional environment primarily concerns the Federal authorities' competence. Thus, it can be assumed that despite positive trends, the current features of CBC spatial structure will continue in the medium term.

CONCLUSIONS

CBC remains one of the most depoliticized areas in relations between Russia and the EU. Its accumulated experience of interaction can be considered as a model for other segments of Russian borders. This became possible due to the use of the program approach, the gradual development of which provided for the development of the agreed set of cooperation priorities, increased coordination of actions between different levels of government on both sides of the border, creation of joint program governance, and the development of unified rules for cooperative projects funding. By the middle of 2000s, such changes had made CBC Programs the most attractive tool for cross-border participants for implementation of their joint projects. This contributed to reformatting of the institutional cooperation structure on the Russian Northwestern borders.

Firstly, Programs launched selection processes for the existing cooperation institutions. Thus, there was a decrease in cooperation intensity in most of Euroregions. Due to their too narrow or too broad agenda, as well as small territorial coverage, some did not have sufficient organizational resources to participate in development of CBC Programs. In fact, only «Karelia» managed to use Russian-Finnish CBC Programs' resources to promote its own priorities and projects.

Secondly, a synergy between CBC institutions became a new force, which led, on the one hand, to this cooperation form's intensification, and, on the other hand, to the reduction of differences between them. This trend is most obvious in the Barents Euro-Arctic region, where activities are increasingly related to cross-border cooperation. CBC Program «Kolarctic», on the contrary, acquired features of trans-border cooperation expressed in the big distance between the program main area and the state borders and also in violation of direct neighborhood principle between the program's participants.

Thirdly, since most CBC projects are implemented within the framework of neighborhood programs, they determine cross-border cooperation spatial structure.

The main feature of CBC spatial structure is high concentration of project activities in the largest cities in the border areas and especially in regional centers. The level of concentration is especially high on the Russian side. That is why a significantly smaller number of centers are involved in cooperation than on the neighbors' part. Nevertheless, since the 2007-2013 program period, there has been a tendency to increase the «cross-border» nature of CBC. As a result, border actors started focusing directly on the border municipalities and small, unsuccessful border settlements. They became platforms for large-scale infrastructure projects, and in most cases aimed at improving transport accessibility of such settlements to the neighboring countries' territories. Another important priority was to reduce the negative impact of these settlements on the ecological status of transboundary river basins.

Another feature of CBC spatial structure is the asymmetry of cross-border interactions. It was also partially compensated by the latest trends of the period 2007-2013. The key factors for increasing the role of Russian participants in these projects were the Federal center's financial contribution to the neighborhood programs, the emergence of joint management bodies and practical cooperation experience gained by Russian participants. As a result, there is an increase in the number of projects initiated by Russian participants andthe number of Russian final beneficiaries is growing. In 2014-2020, about half of the projects were initiated by Russian partners during the application rounds.

According to interviews with participants of Programs and representatives from regional authorities, opportunities for further CBC are closely related to the Federal border policy's direction. This policy should include defined priorities of project activity and special financial tools for their implementation. Development of such programs within the Eurasian Economic Union could be more preferable, but difficult to implement in the short term.. On the internal Eurasian borders, such programs could be a part of the Eurasian cohesion policy (analogous to the European initiative Interreg), and on the external borders – of the European Neighborhood Policy (analogous to ENI).

ACKNOWLEDGEMENTS

The research is done as part of the state assignment (theme «Problems and Prospects of Territorial Development of Russia in the Conditions of its Unevenness and Global Instability» № 0148-2019-0008). The study's field phase, including interviews, in Saint-Petersburg, Leningrad region and Moscow was completed with support of Kone Foundation Project «Northern Dimensions of European Union Actorness. – The Case of Finland and Russia». Mapping and geocoding of CBC Projects were supported by the Russian Academic Excellence Project at the Immanuel Kant Baltic Federal University. The author also expresses gratitude to Stepan Barkov and Anastasia Malakhova – students of Moscow Pedagogical State University, who helped with collecting the data used in this study.

REFERENCES

Conclusion of The State Duma Committee on International Affairs on the draft Federal Law №. 55057-7 «Ratification of the Agreement on financing and implementation of the Cross-Border Cooperation Program» Russia-Lithuania « for the period 2014-2020» (2018). Database. System of maintenance of legislative activity [online] Available at: www.sozd.duma.gov.ru/ [Accessed 31 Aug. 2019]

ENPI CBC Joint Operational Programme Karelia (2007-2013). Programme Document Approved 19/12/2008 C(2008)8453 (2008) [online] Available at: www.cbcprojects.eu [Accessed 31 Aug. 2019]

Estonia – Latvia – Russia Cross-Border Cooperation Programme Within European Neighbourhood and Partnership Instrument 2007-2013 (2008). [online] Available at: www.europa.eu/regional_policy/sources/docoffic/official/communic/wider/wider_en [Accessed 31 Aug. 2019].

Gumenyuk I., Kuznetsova T. and Osmolovskaya L. (2016). Local border traffic as an efficient tool for developing cross-border cooperation, Baltic Region, 8(1), 67-82.

Gumenyuk I.S. (2018). Transport topics in cross-border cooperation programs: experience of the Kaliningrad region, Regional studies, 3(61), 84-91 (In Russian).

Karelia CBC. Joint Operational Programme 2014-2020. Programme Document Approved by the European Commission 18.12.2015 (2015). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

Kolosov V. A., Klemeshev A.P., Zotova M.V. and Sebentsov A. B. (2015). Russia-European Union borderlands: transboundary gradients, interactions, and current challenges, International Journal of Economics and Social Issues, 5(2S).

Korneevecz V.S. and Kropinova E.G. (2010). Cross-border cooperation program «Lithuania-Poland-Russia» 2007-2013 and formation of transborder tourist region and ensuring sustainable development of the territory, Bulletin of the Baltic Federal University. I. Kant. Series: Natural and medical Sciences, 7, 152-156 (In Russian).

Kostyukov A.N. and Semakina Yu.A. (2012). Problems of legal regulation of cross-border cooperation, Bulletin of Omsk University. Series «Right», 3, 51-56 (in Russian)

Kuzneczov A.B. (2004). Russian participation in Euroregions, Region of Cooperation. 4 (29). New spatial form of organization of economy, Kaliningrad: KGU, 5-19 (in Russian).

Kuzneczov A.B. (2008). Euroregions: half a century of «small» integration, Contemporary Europe, 2, 48-59 (in Russian).

Lithuania – Poland – Russia Cross-Border Co-Operation Programme 2007-2013 Final Version of Operational Programme after EC Comments. Document adopted by The European Commission on 17 December 2008 (2008). [online] Available at: www.ec.europa.eu/regional_policy/sources/docoffic/official/communic/wider/wider_en.pdf» [Accessed 31 Aug. 2019].

Lithuania – Russia Cross-Border Cooperation Programme 2014-2020. Approved by European Commission on 19 December 2016 (2016). [online] Available at: www.ec.europa.eu [Accessed 31 Aug. 2019].

Main Directions of Euregio Karelia 2020. Strategic Document of Euregio Karelia. Approved in Kostomuksha on 20th of May 2014. (2014). [online] Available at: www.euregiokarelia.com/en/toimintasuunnitelmat [Accessed 31 Aug. 2019].

Makarychev A.S. (2002). Spatial characteristics of transborder security. In: L.B. Vardomskij and S.V. Golunov, ed, Security and international cooperation in Russia's new border belt, Moscow: Volgograd, 8-40 (in Russian).

McCall C. (2015). State borders in Europe. In: Introduction to Border Studies, Vladivostok: Dalnauka.

Mezhevich N.M. (2009). Cross-border cooperation and practice of Euroregions' activity in the Russian North-West and in the Republic of Belarus: practical experience and legislative support, Saint-Petersburg: Levsha (in Russian).

Radvilavičius Š., Mezhevich N. (2010). The role of cross-border cooperation between The European Union and the Russian Federation in bilateral and regional programmes for economic development. ECO.

Paving the way for a New Neighbourhood: A New Framework for Relations with our Eastern and Southern Neighbours. Communication from the Commission (2003). [online] Available at: www.ec.europa.eu [Accessed 31 Aug. 2019].

Perkmann M. (2003). Cross-border regions in Europe: significance and drivers of regional cross-border co-operation, European Urban & Regional Studies, 10(2).

Perkmann M. (2007). Construction of New Territorial Scales: A Framework and Case Study of the EUREGIO Cross-border Region, Regional Studies. 2007. 41(2) 253-266.

Perkmann M. and Sum N. (2002). Globalization, regionalization, and cross-border region, Basingstoke: Palgrave Macmillan.

Poland – Russia Cross-Border Cooperation Programme 2014-2020. Approved by European Commission on 19 December 2016 (2016). [online] Available at: www.ec.europa.eu [Accessed 31 Aug. 2019].

Pooling of financial resources of TACIS and Interreg Programs (2001) Instruction. Luxemburg: Office for Official Publications of the European Communities (In Russian).

Programming document for EU support to ENI Cross-Border Cooperation. 2014-2020 (2013). Available at: www.ec.europa.eu/ [Accessed 31 Aug. 2019].

Programming of the European Neighbourhood Instrument (ENI) – 2014-2020. (2013). [online] Available at: www.eeas.europa.eu/ [Accessed 31 Aug. 2019].

Scott J.W. (1993) The Institutionalization of Transboundary Cooperation in Europe: Recent Development on the Dutch-German border, Journal of Borderland Studies, VIII(1).

Scott J.W. (2011) Reflections on EU Geopolitics: Consolidation, Neighbourhood and Civil Society in the Reordering of European Space, Geopolitics, 16, 146-175. DOI: 10.1080/14650045.2010.493781.

Scott J.W. (2015) Bordering, Border Politics and Cross-Border Cooperation in Europe. In: Neighbourhood Policy and the Construction of the European External Borders. Cham: Springer., 27-36.

Sebentsov A.B. (2018). Institutional dimension of cross-border cooperation on Russian borders, Regional Studies, 3(66), 66-75.

Sebentsov A.B. and Zotova M.V. (2018). The Kaliningrad Region: Challenges of the exclave position and ways to offset them, The Baltic Region, 10(1), 89-106.

Shlyamin V.A. (2002). Russia in «Northern Dimension», Petrozavodsk: Petrozavodsk State University (In Russian).

South-East Finland – Russia ENPI CBC Programme 2007–2013. Endorsed by the EU Commission on the 12/19/2008 (2008). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

South-East Finland – Russia. CBC 2014-2020 Joint Operational Programme. Endorsed by the European Commission on 18 December 2015 C(2015) 9184 (2015). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

South-East Finland–Russia Interreg III A/ Neighbourhood Programme 2000–2006. Final report. Approved by the Programme Monitoring Committee on 8 September 2010. (2010). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

TACIS cross-border cooperation micro-project assistance program. Instructions for applicants. (2001). World Data Centers. [online] Available at: www.wdcb.ru/grants/rus_guidelines_micro.rtf [Accessed 31 Aug. 2019].

The EU – Russia borderland: new contexts for regional co-operation (2012). ed. by Heikki Eskelinen, Ilkka Liikanen and James W. Scott. London; New York: Routledge.

The EU's eastern neighbourhood: migration, borders and regional stability (2016). edited by Ilkka Liikanen, James W. Scott and Tiina Sotkasiira [et al.]. London, New York: Routledge.

The KOLARCTIC ENI CBC Programme. 2014-2020. Programme Document Approved by the European Commission 18.12.2015 C(2015)9190 (2015). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

The KOLARCTIC ENPI CBC Programme. 2007-2013. Programme Document Approved 19/12/2008 C(2008)8453 (2008). [online] Available at: www.cbcprojects.eu/ [Accessed 31 Aug. 2019].

Vardomskij L.B. (2008). Cross-border cooperation at the «new and old» borders of Russia, Eurasian Economic Integration, 1, 3-16 (in Russian).

Vardomskij L.B. (2009). Russian borders in conditions of globalization. Moscow: LIBRCOM (in Russian).

Yarovoj G.O. (2007). Regionalism and cross-border cooperation in Europe (in Russian) Saint-Petersburg: Norma (in Russian).

Yarovoj G.O. and Belokurova E.V. (2012). European Union for regions: what Russian regions coi and should know about the EU, Saint-Petersburg: Norma (in Russian).

Received on October 29th, 2019 Accepted on January 21th, 2020

Vladimir A. Kolosov[1,2*], Maria V. Zotova[1]
[1] Institute of Geography, Russian Academy of Sciences, Moscow, Russia
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia
***Corresponding author:** kolosov@igras.ru

MULTIPLE BORDERS OF NAGORNO-KARABAKH

ABSTRACT. Though the agreement on ceasefire between Armenian and Azerbaijani troops in Nagorno-Karabakh was concluded more than 25 years ago, there is no progress in the negotiations between the sides. The conflict is intrinsically related to the partition of territory between the areas de facto controlled by the non-recognized Republic of Nagorno-Karabakh, boundaries of which do not match the administrative borders of Nagorno-Karabakh Autonomous Region in the Soviet period, and Azerbaijan. This paper considers the geopolitical situation of Nagorno-Karabakh through the lenses of its cross-border interactions and bordering. This notion widely used in contemporary border studies means not only border delimitation and management, but also the constant process of change in their functions, regime, and social importance. Such change can result, for instance, from the transformation of political strategies, shifts on the international arena and bilateral relations, currency exchange rates and global market prices, as well as in the course of the everyday practice and interactions. The authors analyzed first the existing pattern of borders in the context of security. Then they characterized de-bordering and interactions between Nagorno-Karabakh and its patron state, Armenia, describing the adaptation of the Karabakhi population and economy to the lack of international recognition. The demarcation line with Azerbaijan remains one of the rare cases of a completely closed border. One of the main and potentially long-term obstacles in finding a solution is the cultivation of the «image of the enemy» on both sides of this border.

KEY WORDS: Nagorno-Karabakh, bordering, conflict, geopolitical situation

CITATION: Vladimir A. Kolosov and Maria V. Zotova (2020) Multiple Borders Of Nagorno-Karabakh. Geography, Environment, Sustainability, Vol.13, No 1, p. 84-91 DOI-10.24057/2071-9388-2020-04

INTRODUCTION AND THE THEORETICAL FRAMEWORK

More than 25 years have passed since the ceasefire in Nagorno-Karabakh, achieved in May 1994 after a bloody armed conflict, but the prospects of its resolution are not visible in the medium term. The efforts of intermediaries are so far inconclusive. Azerbaijan firmly insists on the principle of the integrity of the state territory and the inviolability of the borders inherited from the Soviet period, and thus advocates the return of the former Nagorno-Karabakh Autonomous Region (NKAR). The Armenian side believes that the people of Nagorno-Karabakh will never abandon independence won in battles.

The Karabakh conflict is inextricably related to the territorial delimitation between the regions under the control of the Nagorno-Karabakh Republic (NKR, now named Republic of Artsakh in accordance with the Constitution adopted in 2017), Azerbaijan, and Armenia. Despite the quasi-invariance of their pattern after the ceasefire, the instability of the situation in the region is clearly reflected in the functions, regime and symbolic meaning of the external and internal borders in this part of the South Caucasus.

The delimitation or unification of Armenia and Artsakh, the configuration and functions of the de facto borders between Artsakh and Azerbaijan are determined by nation building associated with the internal political situation, shifts in the identity of citizens and their understanding of the place of their countries (regions) in the world, as well as the state of economy, changing gradients in the standard of living, and the development of transportation networks connecting the states of the South Caucasus with the outside world. Finally, new attempts to change existing borders by military force are not ruled out.

Most publications in both Russian and European languages on Nagorno-Karabakh are devoted to the history and stages of the conflict, interactions and the positions of the disputing parties (Markedonov 2012); as well as scenarios, prospects, models and settlement mechanisms, and the impact of the main external players (i.e. Russia , USA, EU, Turkey, and Iran) (Deriglazova et al. 2011; Markedonov 2018; Harutyunyan 2017). A number of authors compare the case of Nagorno-Karabakh with other post-Soviet conflicts. The significance of the Kosovo precedent for de facto state recognition is considered (Ozan 2008; Bayramov 2016; Tokarev 2017; Babayev et al. 2020). Many authors analyze domestic political causes and consequences of the outbreak of hostilities between Artsakh and Azerbaijan in April of 2016 – the so-called Four-Day War and the factors contributing to its quick end thanks to the efforts of Russia and other countries of the Minsk Group. In this context, they seek an explanation for the lack of progress in resolving the conflict. Some see such interests in the lack of open communication on the part of the intermediary countries that stems for their interest in maintaining relations with both sides (Bayramov 2016; Branch 2018). Others are trying to connect the Karabakh conflict to the lack of universally recognized criteria for the international legitimization of new states emphasizing that its only difference from other states is non-recognition (Berg et al. 2018; Caspersen 2012, 2015; Iskandaryan 2019). They argue that, objectively, neither Armenia nor Azerbaijan can afford to start a new large-scale war that would inevitably lead to unacceptable material losses and a breakdown in interaction with many key players (Ozan 2008; Babayev et al. 2020). However, political elites are actively using the conflict to legitimize their power and marginalize opponents (Minasyan 2011; Ayunts et al. 2016). The Karabakh conflict is of paramount importance in the political life and evolution of the identity of the citizens of Armenia and Azerbaijan. Therefore, the President of Armenia, L. Ter-Petrosyan, had to leave his post on suspicions of a willingness to reach a compromise with Azerbaijan by returning a section of the occupied areas that did not belong to the NKAR. His successors, L. Kocharyan and S. Sargsyan, were the natives of Karabakh and former leaders of the struggle for its independence. The fundamental obstacle to conflict resolution is the demonization of the opposite side, assigning to them the full responsibility for the tragic events of 1988-1994 (Harutyunyan 2003; Myths and conflicts..., 2013). At the same time, some authors stressed that not all citizens shared official narratives and quite a few tended to blame the conflict on the Soviet Union and present Russia (Radnitz 2019). In 2009-2010 and 2014, J. O'Loughlin, J. Toal and V. Kolosov, with the help of local colleagues and leading Russian sociological agencies, conducted representative polls in Nagorno-Karabakh, as well as in Abkhazia, South Ossetia and Transnistria. These studies revealed the attitude of various social strata and ethnic groups to the patron and the parent states, and other external actors, the degree of support for the political regime, the views and prospects of state building, and possible ways to resolve the conflict, etc. (O'Loughlin et al. 2015; O'Loughlin et al. 2017).

Thus, available works on the current state of economy, identity, political life and social problems of Karabakh, its interactions with other countries, and internal differences in the republic, although small in territory and population, but very diverse in geographical conditions, are few. These are mainly studies of Armenian authors (see, for example, Dadayan 2006; Sujyan 2010; Shakhnazarian 2009, 2011; Mgdesyan et al. 2016).

The objective of this article is to analyze the geopolitical situation of Nagorno-Karabakh through the lens of cross-border interactions and changes in its borders – the process of bordering. This term, a key concept in modern border studies, means not only the formation and management of borders, but also the ongoing process of transformations in regime, functions, and social meaning – for example, as a result of change in political strategies, shifts in the international situation and bilateral relations, exchange rates and world market prices, as a result of daily activities of political institutions and the practices of cross-border interactions, etc. (Newman 2011). In accordance with the functional and structuralist approaches in border studies, one of their tasks is to study the strengthening or weakening of the barrier role of borders in different historical periods or in different areas – re-bordering and de-bordering (Kolosov et al. 2013). Sources of information for the authors were official statistics from Armenia, Nagorno-Karabakh and Azerbaijan, published data and semi-structured interviews taken from academic experts, officials and leaders of NGO in the course of the field research conducted in September of 2019.

In the first section, the authors provide a brief historical context to the Karabakh conflict and contemporary borders in the region, then they present the current de facto border pattern and perceptions by the sides related to security. The next two sections are devoted to re-bordering and de-bordering as a reflection of the development of the conflict and the uncertainty of the prospects for its resolution, internal political and economic processes in Nagorno-Karabakh.

CONFLICT IN NAGORNO-KARABAKH: REDRAWING POLITICAL BORDERS

The territory of Nagorno-Karabakh became part of the Russian Empire in 1813 under the Gulistan Peace Treaty with Persia. After the collapse of the empire in 1917, the Armenian population of Karabakh and Zangezur refused to obey the authorities of the newly created Democratic Republic of Azerbaijan, which led to an armed conflict that was suppressed by the Red Army. Initially, considering the national composition of the population, the new authorities, with the consent of the Azerbaijani side, included these territories into Armenia. However, in July of 1921, Karabakh was left within the borders of Azerbaijan with the provision of autonomy, presumably in the interests of rapprochement with Kemalist Turkey, which recognized Soviet power. In 1923, the NKAR was created. However, the administrative borders of the NKAR did not coincide with the ethnic boundaries: three districts with the predominance of the Armenian population – Shaumyan, Dashkesan and Khanlar, as well as Kedabay and Shamkhor districts with a significant Armenian minority remained outside the NKAR. Almost simultaneously, a Kurdistan district was formed, which separated NKAR from Armenia (back in the 17th century, the Persian authorities moved the Kurds here). On its territory, which later entered the Lachin district, there was a road between the administrative center of NKAR Stepanakert and Armenia (Lachin corridor). According to the 1926 census, Armenians made up 90% of the NKAR population, but by 1989 their share had dropped to 76%. Based on the opinion of the Armenian side, economic policy of Baku toward the autonomy and, in particular, the allocation of budgetary funds, as well as neglect of cultural needs of the region, the creation of artificial barriers for relations between Karabakh and Armenia were unfair and caused discontent among the Armenian population of the NKAR. This repeatedly led to protests.

In 1988, the Council of People's Deputies of the NKAR appealed to the Supreme Soviets of the USSR, Azerbaijan and the Armenian SSR with a request to transfer the region to Armenia. This official request was the first one to violate the monopoly of the central authorities on changing the territory and borders of the union republics and autonomies. This petition meant for the Kremlin the beginning of an acute conflict between the two republics. Its satisfaction could become a dangerous precedent and cause a chain of unforeseen consequences in many parts of the country and ultimately provoke its collapse (Markedonov 2018). The Kremlin tried to solve the Karabakh problem in the usual technocratic way by changing local Communist party leaders and allocating funds for the construction of housing, schools and hospitals, as well as providing access to Armenian television programs. Ultimately, Moscow sided with Azerbaijan, supporting its territorial integrity.

The decision of the regional council sharply aggravated the political situation in both Azerbaijan and Armenia resulting in a strong impetus to the long-standing discourse on the historical past, territories and borders. The main slogan in Karabakh and Armenia at that time was «miatsum» – unification. The turning point was the Armenian pogroms in the industrial suburb of Baku, Sumgait, and then other cities. In turn, these events caused ethnic cleansing and massive flow of Armenian refugees from Azerbaijan and Azerbaijani refugees from Armenia and Karabakh. This coincided with widespread outbreaks of violence that led to the death of dozens of people and escalated into a military conflict starting in 1993 turned into a real war using all types of armed forces. On September 2, 1991, NKR independence was declared. Its territory included the NKAR and the Shahumyan region of Azerbaijan, inhabited by Armenians. The Azerbaijani authorities blocked the delivery of goods from Armenia to the NKAR, and then to Armenia proper, and Armenia to Nakhichevan (the Azerbaijani territory separated from the main part of the country by the territory of Armenia). The ceasefire was signed only in May 1994.

The contemporary pattern of borders and security challenges

The result of the Armenian-Azerbaijani armed conflict was a radical change in the de facto political borders in the South Caucasus. 92.5% of the territory of the former NKAR and five districts of Azerbaijan, declared the «security zone», were comprised of Kelbajar and Lachinsky, previously separated from Armenia, Kubatlinsky, Dzhebrailsky and Zangelansky, as well as parts of the Agdam and Fizuli districts – all passed under the control of NKR. At the same time, small territories of Martuni and Martakert districts of the NKAR, as well as the Shahumyan district and in part the Khanlar district, which entered the NKR during the conflict, went to Azerbaijan, forming 15%of its newly redrawn territory.

Not recognizing the NKR's right to self-determination, the Azerbaijani side considers the borders established as a result of the 1994 armistice to be temporary. The Karabakh side does not consider them fair, since the territory of the former NKAR and the Shahumyan region are not fully included in the Republic of Artsakh. The basis of the peace talks between Armenia and Azerbaijan conducted through the mediation of the OSCE Minsk Group including Russia, the USA and France are still the Madrid principles put forward in 2007. They provide for the transfer by the Armenian side to Azerbaijan of those territories occupied by it at the time of the armistice in 1994 beyond the former NKAR, although with an important reservation about the Lachin corridor. However, the Armenian side is skeptical of possible international security guarantees and is convinced that control over the occupied territories is of strategic importance for Karabakh, since it provides it with a relatively short border with Azerbaijan allowing maneuvering along the inaccessible Mravsky Range. Over a period of two decades, NKR invested considerable resources in the border reinforcement, including about 250 km of concrete fortifications, minefields and wire barrels.

There is another reason for the reluctance of the NKR to leave the occupied territories – control over water resources. At the end of hostilities, Azerbaijan lost access to the Sarsang reservoir built in 1976 on the Terter River in the territory of the NKAR. It provided water to six districts, most of which are now controlled by the Armenian side. Outside of the occupied territories there remained a part of the Terter district with a population of 102 thousand people, mainly refugees from Armenia and the former NKAR. Now the Azerbaijani side protests against the unfair distribution of water by Karabakh. This is in contrast to the previous situation when the Armenian side was extremely dissatisfied with the fact that most of the water from sources in the NKAR was spent in the lowland areas of Azerbaijan. Of the 128 thousand hectares of agricultural land irrigated from the Sarsang reservoir, 110 thousand were outside the NKAR (Babayan 2019).

The regime of the two longest and symbolic borders of Nagorno-Karabakh is typical for non-recognized states: high permeability of the border with Armenia (the «patron» state) and the complete absence of interactions through the separation line with the Azerbaijan (the «parent» state). Cross-border shootings in 2012-2016 became more frequent. During the «Four-Day War» in April 2016, Azerbaijani troops tried to advance deep into the territory controlled by Nagorno-Karabakh. The configuration of the separation line has slightly changed in favour of Azerbaijan, which regained control over the 8 sq. kms of territory.

Azerbaijan's policy was to bring about an economic and political crisis in Artsakh through a combination of military-political pressure, moderated armed confrontation at the de facto border, and diplomatic activity (Minasyan 2016). The most important element of this policy has become a long-term transport blockade. Before the conflict, most of the cargo for Armenia went through Baku, not through Georgia, as it is now (it should be borne in mind that, due to the Georgian-Abkhaz conflict, access to Russia from Georgia by railway is closed). The only railway line connecting Karabakh with the outside world passed through Azerbaijan and was closed in 1991. Supporting Azerbaijan as its closest ally, in the midst of hostilities in Karabakh, Turkey closed the border with Armenia in 1993. Gyumri – Kars railway is still not functioning.

The issue of road usage in the NKAR with the Armenian SSR has always been difficult. The road network was built in such a way as to direct the flow of vehicles through areas outside the autonomous region. Soon after the start of full-scale hostilities, the Armenian side seized control of the Lachin and Kelbajar districts which separated Artsakh from Armenia. The Yerevan – Goris – Stepanakert road, 340 km long, passes through the so-called Lachin corridor, which has become «a lifeline» for Karabakh and for a long time remained practically the only transportation route that allowed the NKR to avoid complete isolation. It was immediately reconstructed. In Soviet times, due to the extremely poor condition of the Lachin section of the Stepanakert – Yerevan road, drivers heading to Armenia were forced to travel hundreds of kilometers through Azerbaijan. In 2017, another road along the coast of Sevan and further through Vardenis came into operation, which significantly improved the geopolitical position of the non-recognized republic. Carriers prefer to use this route, since it is shorter, it has fewer turns and better coverage. According to rough estimates, the passenger flow from Stepanakert to Yerevan and back is about 250-300 people per day[1]. On the border, two checking points (on Eghegnadzor highway and on Vardenis highway) control the entry and exit of foreigners (except those from the CIS countries other than Azerbaijan and Turkmenistan) and over import /export of goods. Foreign citizens are required to obtain visas in the representations of Nagorno-Karabakh in Yerevan or immediately upon entry. The visa fee for foreigners who stay in Artsakh for less than 21 days was canceled as of late April of 2019.

An important impact on the perception of the nation's security by the political elite and citizens is provided by the demographic situation, which, in turn, is both a reflection and factor of territorial conflicts. Natural movement of population and migrations can change the ratio between major ethnic groups relatively quickly.

Ethnopolitical processes in the NKR and neighbouring regions have a long and complicated history. As noted above, according to the 1926 census, Armenians made up 90% of the population in the former NKAO, but by 1989 their share had fallen to 77%. According to the Armenian side, the Azerbaijani leadership pursued a deliberate policy of changing the ethnic structure of Karabakh, encouraging the resettlement of Azerbaijanis and Kurds from the densely populated plain territories to new settlements created around the cities of the NKAR. Moreover, among the Azerbaijani population, the birth rate was significantly higher than among the Armenians. The population dynamics of the NKAR was also influenced by demographic losses during the World War Two (35% of the population was drafted into the Soviet Army). Migration outflow to other regions, approximately equal to the natural increase, is much greater than from Azerbaijan and Armenia. Though the birth rate remained relatively high and mortality was low, which was due to the predominance of rural residents and a young sex

[1] *In 2012, the Stepanakert airport was reconstructed. However, it is closed due to the lack of a license from the Azerbaijani State Civil Aviation Administration. The Azerbaijani side is threatening to bring down planes that try to land in Stepanakert.*

and age structure, in 1926-1989 the population of the NKAR increased by a third – significantly less than in Azerbaijan. Shifts in the composition of the population were one of the main causes of the conflict.

By the time of the conclusion of the ceasefire agreement in May 1994, the number of residents of Artsakh had decreased by almost 50 thousand compared to 1989. Such a decline was the result of losses in the war with Azerbaijan and the «exchange» of the population with it. Almost all Azerbaijanis were forced to leave the NKR (about 40,000 left the former NKAR itself, and even more abandoned the seven regions of Azerbaijan annexed as a result of the war). At the same time, according to various estimates, 30-40 thousand Armenian refugees from Baku and other cities of Azerbaijan now live in the non-recognized republic. Consequently, the population of Artsakh has become ethnically homogeneous (99.7% are Armenians) (Averyanov 2014). The Azerbaijani side denies the legality of all elections in Karabakh, including on the grounds that refugees from the NKAR did not participate in them.

As of January 1, 2019, there were 148 thousand inhabitants in the republic. Its demographic situation is stable and favourable. Since the second half of the 1990s, despite demographic waves, the birth rate has been at the level of 15-20, the mortality rate is 8-9, the natural increase is 7-10 per 1000 inhabitants. Migration is now relatively small: it is estimated that about 4 thousand people are employed or for other reasons live outside the republic – as a percentage of the active population it is much smaller than in Azerbaijan (Sujyan 2010; Socio-economic situation…, 2019).

The Azerbaijani side accuses the Armenian side of deliberately settling Armenians in the districts that were not previously part of the NKAR in order to consolidate their belonging to the NKR. They provide various privileges to the migrants, while social infrastructure is being created. During the war in Syria, Armenia accepted about 17,000 Armenian refugees from Syria, some of whom, as rural residents, actually chose to live in Karabakh and are now engaged in agriculture. In the process of cognitive appropriation of these territories in the official discourse of the NKR, a gradual change of terminology is observed: from the «temporarily occupied territories» to the «security zone», then to the «liberated territories» and, finally, to the «territory of Artsakh». In the expert circles of Armenia itself these territories are ironically called «temporarily occupied – forever liberated.» Another confirmation of this process is the termination of the discussion about a possible «exchange» with Azerbaijan of all or part of these territories for recognition.

As a result, a significant demographic gradient has formed along the borders between Karabakh and both of its main neighbours, Armenia and Azerbaijan: the population density in Artsakh is much lower. Along the de facto border between Artsakh and Azerbaijan, this gradient also has a pronounced ethnic colouring, sharply dividing the areas with the Armenian and Azerbaijani (Muslim) populations.

DE-BORDERING: HOW «TRANSPARENT» IS THE BORDER WITH ARMENIA?

The high contact functions («transparency») of the border between Armenia and Nagorno-Karabakh reflects their deep integration in all areas. For Armenia, ensuring NKR security is inseparable from guarantees of its own security. Although formally there is a separate Defense Army of Nagorno-Karabakh, in fact, it is integrated into a single security system with the Republic of Armenia. According to the 1994 interstate agreement between Armenia and the NKR on joint defense, military-eligible citizens of Armenia must undergo part of the active military service on the separation line with Azerbaijan[1]. According to interviewed experts, *«everyone in Armenia is interested in having strong positions on the Azerbaijan border»*.

Income and expenses of the NKR budget are balanced only thanks to the subsidies provided by Armenia, officially called interstate credit. Although they are reduced annually, in 2018 their share in the Artsakh budget amounted to about 50%[2] (in 2010 – about 66%) (Grigoryan 2017). More than 80 agreements were concluded between Armenia and the NKR and, in fact, a single legal space has developed. NKR formally issued local passports, but since these passports do not allow leaving Armenia, all citizens of Karabakh have passports of Armenia to be able to go abroad. Thus, there was no real need in internal passports and their release was suspended. Foreign economic relations are also carried out through Armenia, where the enterprises of exporters are registered, so the NKR products are exported under the guise of Armenia (Nemtsova 2014). Unified customs legislation is in force in Armenia and NKR, which made it possible to enter the Eurasian Economic Union market with its products, though isolation of the NKR, separated from Armenia by mountain ranges, increases transport and administrative expenses caused by the need to pass customs formalities in Yerevan. Restrictive norms of the Eurasian Economic Union also apply to Artsakh – for example, quotas allocated to Armenia for the purchase of certain types of agricultural machinery in third countries.

Some sectors of Artsakh's economy are highly dependent on exports. So, one third of the products of a traditional branch of Artsakh's specialization, the agricultural sector, accounts for about 11% of GDP and is exported abroad. In particular, canned fruits and vegetables produced by the companies *Artsakh Fruit, Artsakh Berry*, and *Artsakh Bio* are exported to Russia, Ukraine, France and other countries (Feshchenko 2014).

Production cooperation between Armenian and Karabakh is developing. Under the government program «Grapes», NKR farmers receive up to 70 thousand cuttings of phylloxera-resistant grape varieties. Under the relevant agreement, Yerevan Brandy Company purchases Karabakh grapes (Beglaryan A., 2013). The Stepanakert construction materials factory (now Karin CJSC) was acquired by the *Sirkap-Armenia* company.

Largely due to close integration with Armenia and blurring the border with it, the electric power industry is becoming a new sector of NKR specialization and now accounts for 7.5% of GDP. After the collapse of the Soviet Union, only the Sarsang hydroelectric power station with the capacity of 50 MW was operating on the territory of the republic. However, over the past 10 years, 15 small hydropower plants with a total capacity of 52 MW have been built. At first, the task was to achieve self-sufficiency and security of energy supply. In 2018, 385.1 million kWh of electricity was produced in Artsakh. At the same time, the total potential makes it possible to generate annually up to 700 million kWh (Karabakh 2011; Grigoryan 2017-2). In 2017, surplus electricity from Artsakh was already exported to Armenia (Mgdesyan 2019).

The close integration of Armenia and Artsakh is also predetermined by the actual unity of their credit and banking system. Formally, it is regulated by *Artsakhbank*,

[1] According to the official data, the number of the NKR armed forces is about 20 thousand people, which is 36% of all the NKR employment. Military service in NKR is prestigious, as it provides stable income (almost 1.5 times higher than the national average).
[2] However, half of this amount is the customs income of Armenia from the NKR foreign trade.

which performs a number of functions of the NKR Central Bank. However, it is registered in Yerevan and has a general license of the Central Bank of Armenia, which allows to circumvent restrictions related to the non-recognition of the NKR independence. *Artsakhbank* is an affiliate member of the Armenian national payment system Ar*Ca*, the international payment system MasterCard International, and the money transfer system SWIFT; it provides banking services throughout Armenia. Although in 2004 the NKR established its own currency, the Karabakh dram, the means of payment is the Armenian dram, to which the Karabakh dram is equal in value.

Thus, after the beginning of the current conflict between Armenia (Karabakh) and Azerbaijan and the collapse of the Soviet Union, the border between the two countries, which turned into the line between non-recognized Karabakh and independent Armenia, experienced a process of sharp weakening of its barrier function, reflecting the possibility of a complete merger of the two polities.

At the same time, the de facto political border between Armenia and Karabakh still remains an economic and cultural barrier, although it is relatively weak. The Karabakh dialect is so different from the normative Armenian language that it is difficult for primary school pupils to use it in Karabakh schools. For a long period, Karabakh was ruled as a separate territory. The different historical past left a mark on regional identity, although the main identity of the Karabakh people is Armenian with all its cultural and historical markers, representations about past events, prominent political leaders and cultural figures, etc. The long-standing conflict with Azerbaijan, at first latent and then acute, also predetermined some features of the identity and political culture of the Karabakh people: consolidation with the authorities in face of an existential threat (however, combined with distrust of them), hope for paternalism of the state, and egalitarianism. The social stratification in Karabakh is less evident than in Armenia, in particular, thanks to state benefits and social policy. Thus, in Armenia, within the framework of «optimization», the number of small-class rural schools is being reduced, and in the NKR they are being preserved.

One of the factors determining the specificity of the Karabakh regional identity is Artsakh, where the brand of «an island of Christian culture in the Muslim world» is willingly cultivated. This has long been a powerful source of migrants who have maintained ties with their homeland. Unlike Armenia, they migrated to Baku, other industrial centers of Soviet Azerbaijan, to Russia, and other union republics. The Karabakh people were always better at speaking Russian than the residents of Armenia. In Artsakh, they are proud of the fact that all four Soviet marshals of Armenian descent are from the region.

There are also differences between Armenia and Karabakh in economic policy. The influx of investments from the diaspora, along with other factors, formed the basis for the programs established by the NKR government for the development of certain regions and priority sectors through changes in the tax structure. Taxes were reduced (revenue tax from 15% to 5%, individual income tax from 30% to 5%, land tax from 15% to 6%). Significant volumes of foreign investments, economic and humanitarian aid through various channels lobbied by the Armenian diaspora (including the 3.5 million to 7 million dollars aid per year from in the US state budget) began to flow into Karabakh. As a result, new enterprises and even entire industries were created.

The mining industry, one of the main sources of tax revenue, began to develop from scratch. For 25 years, the *Base Metals* mining company, owned by Moscow-based businessman of Armenian background V. Melumyan, has been operating in the republic. This company built a copper concentrate factory in Drmbon and is currently constructing a plant at the Kashen (Tsakhkashen) copper-molybdenum mine in the Martakert district.

Based on the former Karabakh silk factory in Stepanakert, sewing manufactures have been established that receive contracts from large Italian fashion houses (*Versace, Moschino, Prada, Armani*) and export their goods through Armenia.

Diaspora representatives provide support (equipment supply, training in new technologies) to small and medium-sized Karabakh entrepreneurs-winemakers. A well-known businessman in Armenia and Russia, G. Oganyan, invested $5 million into Karabakh wine-making. Now 13 wine-making companies operate in the NKR *(Domaine Avetissyan, Stepanakert Brandy* CJSC, *Artsakh Alco, Artsakh Brandy Company* CJSC, *etc.*), exporting wines, brandy, fruit and berry vodka to Russia, some countries of Europe and North America.

In the Martakert region, thanks to the investments of the Swiss businessman V. Sermakesh, the production of black caviar was launched. The Anivyan family from the United States invested in the production of dairy products. Iranian Armenians invested in the production of polyethylene pipes in Shusha. One of the most modern hospitals in the South Caucasus operates in Stepanakert, built with funding provided by S. Karapetyan, whose name is on the Forbes list of Russia.

As a result, Artsakh's GDP began to grow in the 2010s by 10-11% per year (Nemtsova 2014), much faster than in Armenia. However, this growth did not lead to a significant increase in population income due to the large contribution to the economic growth of big (considering the scale of the NKR) mining enterprises, their profits exported outside the region. Nevertheless, the per capita GRP and average salary in the NKR, although lower than in Yerevan, is significantly higher than in two of the three neighboring peripheral marzes (provinces) of Armenia (Table 1), with the exception of Syunik known for its mining-metallurgical industry.

The institutional factor also influences the preservation of the barrier function of the Armenian-Karabakh border. The only local fixed and mobile communications operator «Karabakh Telecom» in Artsakh belongs to the structures close to R. Kocharyan, the first president of NKR and the second president of Armenia. Although in 2018 the NKR authorities announced the monopolization of the telecommunications market, no changes have occurred. High prices for mobile communication with NKR cause dissatisfaction in Armenia, as many citizens maintain family and business contacts with the Karabakh.

RE-BORDERING: ONE OF THE MOST CLOSED BORDERLINES IN THE WORLD

The de facto border (the separation line) between Nagorno-Karabakh and Azerbaijan, which has not changed much since the day of ceasefire, is the most striking and rare example of a new closed «frontal» border. The complete absence of cross-border interactions differs even from the borders of Abkhazia and South Ossetia with Georgia. The deployment of military units along the separation line, the special regime of the border zone on both sides, constant skirmishes, and the destruction during the war and immediately after it of a number of cities and other settlements turned the border territories into an economic desert. Occupied territories were one of the main agricultural

Table 1. The main socio-economic indicators of NKR and neighbouring marzs (provinces) of Armenia, 2018

	NKR	Syunik	Wayotsdzor	Gegarkunik	Yerevan	Armenia
Population, '000	147.0	138.4	49.6	229.7	1 081.8	2 972.7
Industrial production, mln drams	125 006	323 503	27 641	64 613	735 190	1 664 279
GRP, mln drams	310 253	377 200	64 900	197 800	3 508 500	6 005 100
Per capita GRP	2 110 563	2 725 434	1 308 468	861 123	3 243 206	2 020 083
Average salary, drams	161 181	234 601	118 280	109 133	194 754	172 727
Retail trade turnover, mln drams	107 353	22 717	8 264	27 367	1 027 812	1 410 774
Per capita retail trade turnover, drams	730 293	164 142	166 605	119 141	950 094	474 577

Source: Socio-Economic Situation of Nagorno-Karabakh Republic in January-December 2018, Stepanakert, 2019. 140 p.

areas of Azerbaijan; 70% of summer pastures used to be ocated there. No matter how the configuration of the border changes, it remains highly likely a strong barrier for many years.

The per capita GRP in Artsakh for 2018 calculated at purchasing power parity (PPP) is almost half that of Azerbaijan, which receives significant income from production and export of oil and natural gas: 10,730 and 17,940 USD, respectively. However, if we take the belt of districts adjacent to the separation line[1], then according to official statistics, the average salary in the non-recognized republic (Statistical... 2019), recalculated by PPP, is slightly higher than in neighbouring regions of Azerbaijan, with the exception of Dashkesan, a territory with productive agriculture and developed mining industry. There is no pronounced economic gradient along the border.

The main factor of re-bordering is the «image of the enemy», cultivated for many years in the internal political struggle of South Caucasian societies in protracted conflicts (Kvarcheliya, 2013). The terms «Azerbaijanophobia» and «Armenophobia», used to denote ethnic resentment, fear, hostility or other negative feelings towards Azerbaijanis and Armenians, have been noticeably disseminated in the literature, media, and scientific research of the two countries (Aldricht 2002; Suny 1993). In both Armenia and Azerbaijan, the growth of spontaneous nationalism is based on the selective interpretation of history, myths, symbols, and religious images (Shnirelman 2003; Yunusov 2018).

At the same time, cross-conflict studies in both societies are largely taboo. An exception is a few works that consider the creation and translation of myths and dominant narratives in the South Caucasus (Myths and conflicts ..., 2013; Mikaelyan et al. 2011; Crombach 2019). They note that myths related to conflicts increase resistance to their resolution (Kvarcheliya 2013). As a result, the feeling of being a «victim» and the search for a «saviour» are used to manipulate public opinion (Hovhannisyan 2013).

According to many political scientists, the victim complex has spread through the Armenian society having arisen after the ethnic cleansing of 1915 at the end of the Ottoman Empire. The fear of becoming a victim again is being actively used today to justify control over Nagorno-Karabakh and hatred of Azerbaijanis (Tsiganok 2007). The ethnic prejudices of Armenians are based on deep-rooted stereotypes that identify Azerbaijanis with Turks, and therefore, associate with them the potential for a new genocide. In turn, Azerbaijanis also have a sense of persecution by the Armenians, and they use almost the same language, including the term «genocide» to refer to the crimes of Armenians. Both sides portray each other as an aggressor who attacks innocent civilian population. Extremely negative representations of Armenians are spreading in Azerbaijan through television, media and history books.

Studies of Azerbaijani textbooks emphasize the use of nationalist discourse, which excludes the understanding of history as a narrative. Instead, essentialist models of historical realities and the current state of affairs are offered (Adibekyan, Elibegova, 2013). «Friendship between nations» has given way to revised national stories that are hostile to the other and offered as «truths» that children should remember, which leads to the consolidation of the image of enemy in the psychology of the nation (Myths and conflicts ..., 2013). In response, Azerbaijani scholars emphasize that school books in Azerbaijan framed by the Karabakh conflict were called upon to «educate patriots who are ready, if necessary, to take part in the next conflict» (Yunusov 2011).

The Karabakh conflict is a painful issue for Azerbaijani society, on which one can observe the unity of opinion of almost all politicians. In the Azerbaijani political environment, the opposition is endowed with the image of a marginalized and small group directly linked by financial interests to the «historical enemy» – Armenians and the Armenian lobby (Abbasov 2013).

Thus, a quarter century after the ceasefire, the positions of the parties not only have not come close, but on the contrary, the situation has become even more severe. This is also confirmed by opinion polls. The vast majority of Azerbaijanis and Armenians (about 80-90%) consider each other enemies of their country.

CONCLUSION

The geographical isolation, the transport blockade by Azerbaijan, the diversion of significant funds to confront it, and especially the lack of international recognition, are slowing down the economic development of Nagorno-Karabakh. However, as in other non-recognized (partially recognized) states in the post-Soviet space, residents and households have successfully adapted to the situation thanks to the constant assistance of the «patron» state (Armenia in the case of NKR). Integration with it and control over the border with Lachin and Kelbajar regions leads to the erasing of the border (de-bordering). Having received an Armenian passport, citizens of Nagorno-Karabakh have international mobility opportunities. The non-recognized republic conducts foreign trade through Armenia, receives foreign investments,

[1] *Only areas completely controlled by the central government have been used, since data on the districts partially controlled by Stepanakert are not reliable.*

and develops new sectors of the economy. Nevertheless, this border is preserved, and its future depends on the prospects for resolving the Armenian-Azerbaijani conflict. Although the hypothetical international recognition of Artsakh is not strictly connected with the resolution of the conflict and/or unification with Armenia, it is hardly possible without the other. It is not yet possible to exclude the feasible return of the sides to the formula «territory in exchange for status». Despite the lack of progress in the negotiations between them, there are solutions to territorial problems, especially since Azerbaijan is interested in safe communication with Nakhichevan through the territories inhabited by Armenians (Okunev 2019; The Nagorno-Karabakh Deadlock 2020). In any case, the resolution of the conflict will affect not only the functions, but also the configuration of the de facto NKR borders and will trigger a new wave of de-bordering and re-bordering.

The irreconcilable positions of the parties are, of course, associated with the need for both sides to politically mobilize the population in order to legitimize the authorities for combating the external threat. The ceasefire line separating the Azerbaijani and Armenian forces has turned into a hermetically closed border – situation which constantly creates the risk of renewed hostilities. One of the main and potentially long-term obstacles in finding a solution is the cultivation of the «image of the enemy» on both sides of the de facto border, fostered through the state education systems.

ACKNOWLEDGEMENTS

This study was performed within the framework of the project «Post-Soviet Non-Recognized States: Factors of Viability and Risks for Russia» supported by Russian Science Foundation (N 19-17-00232), 2019-2021. The theoretical part of the article was prepared with partial support of the Immanuel Kant Baltic Federal University by the Russian Academic Excellence Project (5-100) and by State Assignment IGRAS (0148-2019-0008, AAAA-A19-119022190170-1).

REFERENCES

Abbasov I. (2013). «Games of Patriots»: Practices of marginalization of the opponent in internal political struggle in modern Azerbaijan. Myths and conflicts in the South Caucasus, 2, 38-62. (in Russian with English summary).

Adibekyan A., Elibegova A. (2013). Armenophobia in Azerbaijan. Yerevan: «Center of public relations and information» under the Office of the President of Armenia, 332 (in Russian with English summary).

Aldrich C. (2002). The Aldrich Dictionary of Phobias and Other Word Families. Trafford Publishing.

Averyanov Yu. (2014). Ethno-demographic processes in Nagorno-Karabakh: history and contemporary situation. Nakhichevan-on-the Don, 1, January, 10.

Ayunts A., Zolyan M. and Zakaryan T. (2016) Nagorno-Karabakh conflict: prospects for conflict transformation. Nationalities Papers, 2016. DOI: 10.1080/00905992.2016.1157158.

Beglaryan A. (2013). Nagorno-Karabakh: a difficult path to economic self-sufficiency [online] Center of Lev Gumilev. [online] Available at: www.gumilev-center.ru/nagornyjj-karabakh-trudnyjj-put-k-ehkonomicheskojj-samodostatochnosti/ [Accessed 20 Dec. 2019] (in Russian).

Branch A.R. (2018). Armenia and the South Caucasus: A New Security Environment Connections QJ 17, 2, 47-60. DOI: 10.11610/Connections.17.2.04.

Babayan D.K. (2019) Geopolitics of the Armenian-Azerbaijani conflict. Moscow-Yerevan: Defacto, (in Russian with English summary).

Barseghyan H., Sultanova S. (2012). Armenian image in history textbooks of Azerbaijan. CRS Issue 631, 27 Feb. Available at: http://iwpr.net/report-news/history-lessons-armenia-and-azerbaijan.

Bayramov A. (2016) Silencing the Nagorno-Karabakh Conflict and Challenges of the Four-Day War. Security and human rights, 27, 116-127.

Berg E., Wits K. (2018). Quest for Survival and Recognition: Insights into the Foreign Policy Endeavours of the Post-Soviet de facto States. Ethnopolitics, 17(4), 390-407.

Caspersen N. (2012). Unrecognized States. The Struggle for Sovereignty in the Modern International System. Cambridge: Polity Press.

Caspersen N. (2015). Degrees of legitimacy: Ensuring internal and external support in the absence of recognition. Geoforum, 184-192.

Crombach S. (2019). Ziia Buniiatov and the Invention of an Azerbaijani Past. Academisch Proefschrift. University of Amsterdam.

Dadayan D. S. (2006). Development and security Strategy of the Nagorno-Karabakh Republic as an unrecognized state (Abstract of PhD dissertation, political science). Moscow: Russian Institute of Strategic Studies (in Russian with English summary).

Deriglazova L., Minasyan S. (2011). Nagorno-Karabakh: paradoxes of strength and weakness in asymmetric conflict. Yerevan: Caucasus Institute (in Russian with English summary).

Feshchenko V. (2015). Made in Artsakh: how businessmen raised an unrecognized Republic from their knees [online] Available at: www.secretmag.ru/cases/stories/artsakh-epic.htm [Accessed 20 Dec. 2015] (in Russian).

Grigoryan A (2017). The budget of Nagorno-Karabakh for 2018 was adopted with a deficit [online] Kavkaz Uzel [online] Available at: www.kavkaz-uzel.eu/articles/314153/ [Accessed 20 Dec. 2019] (in Russian).

Grigoryan A. (2017). The authorities of Nagorno-Karabakh extended the program of subsidizing electricity tariffs [online] Kavkaz Uzel. Available at: www.kavkaz-uzel.eu/articles/312435/ [Accessed 20 Dec. 2019] (in Russian).

Harutyunyan B. (2003). Ethnic stereotype in intercultural communication: self-assessment and mutual assessment. Yerevan: Caucasus media Institute.

Harutyunyan A. (2017). Two state disputes and outside intervention: the case of Nagorno–Karabakh conflict. Eurasian Econ. Rev., 7, 69–93.

Hovhannisyan M. (2013). The use of images of an external friend and enemy in the internal political discourse of Armenia on the example of the analysis of the speeches of the first President Levon Ter-Petrosyan (September 2007 – May 2011). Myths and conflicts in the South Caucasus, 2, 62-82 (in Russian).

Gazazyan A. (2017). Troops of Nagorno-Karabakh: the active army of the unrecognized Republic. Deutsche Welle. [online] Available at: www.dw.com/ru/войска-нагорного-карабаха-действующая-армия-непризнанной-республики/a-6281785 [Accessed 20 Dec. 2019] (in Russian).

Karabakh plans to increase copper production and electricity production by several times (2011) [online] REGNUM. Available at: www.regnum.ru/news/economy/1485331.html [Accessed 20 Dec. 2015] (in Russian).

Kolosov V. and Scott J. Selected Conceptual Issues in Border Studies (2013) Belgeo, 4, 9-21.

Kvarchelia L. (2013). Instrumentalization of the conflict theme in political discourse. Myths and conflicts in the South Caucasus. 2, 13-16 (in Russian with English summary).

Markedonov S. (2012). De facto formation of the post-soviet space: twenty years of state construction. Yerevan: Caucasus Institute, 257 (in Russian with English summary).

Markedonov S.M. (2018). Thirty years of the Nagorno-Karabakh conflict: main stages and prospects of settlement post-Soviet studies. 1. (2), 23-41 (in Russian with English summary).

Mgdesyan V.M., Ryabovol I.V. (2016). Nagorno-Karabakh Republic: raising the economy through reforms and investments. Modern trends in the development of science and technology, 9-3, 98-102 (in Russian).

Mgdesyan A. (2019). Armenia is able to export cheap electricity Karabakh: interview [online] EurAsia Daily [online] Available at: www.eadaily.com/ru/news/2019/07/11/armeniya-mozhet-eksportirovat-deshevuyu-karabahskuyu-elektroenergiyu-intervyu [Accessed 20 Dec. 2019] (in Russian).

Mikaelyan G., Harutyunyan A., Melikyan D., Galanteryan L. (2011). Nationalist discourse in Armenia. A study supported by the Bureau of the Heinrich bell Foundation in the South Caucasus, URL: https://regnum.ru/news/polit/1413303.html [Accessed 20 Dec. 2015] (in Russian).

Minasyan S. (2011). Armenia in Karabakh, Karabakh in Armenia: living with a conflict identities, ideologies and institutions. A Decade of Insight into the Caucasus: 2001-2011. Yerevan: Caucasus Institute, 142-152 (in Russian with English summary).

Minasyan S. (2016). Deterrence in the Karabakh conflict. Yerevan: Caucasus Institute (in Russian).

Mythes and conflicts in the South Caucases (2013), 2. Instrumentalisation of the conflict's theme in political discourse. Javakhishvili J. and L.Kvarchelia, eds. International Alert.

The Nagorno-Karabakh Deadlock. Insights from Successful Conflict Settlements (2020). Babayev A., Schoch B., Spanger H.-J., eds. Wiesbaden: Springer.

Nemtsova Zh. (2014). RBC: Nagorno-Karabakh can be called the « Transcaucasian tiger» [online] Panorama.am. [online] Available at: www.panorama.am/ru/news/2014/07/19/melik-shakhnazarov/240578 [Accessed 15 Nov. 2019]. (in Russian).

Newman D. (2011). Contemporary Research Agendas in Border Studies: An Overview, in Wastl-Water D., ed. Ashgate Research Companion to Border Studies. Farnham: Ashgate Publishers, 33-47.

Okunev I. (2019). Ten ways to solve an unsolvable territorial dispute. Russia in global politics, 2019, 2 [online] Available at: www.globalaffairs.ru/number/Bylo-by-zhelanie-20040 [Accessed 20 Dec. 2015] (in Russian).

O'Loughlin J., Toal G., Kolosov V. (2016). Inside the post-Soviet de facto states: a comparison of attitudes in Abkhazia, Nagorny Karabakh, South Ossetia, and Transnistria. Eurasian Geography and Economics, 56(1), 1-34. DOI: 10.1080/15387216.2015.1012644.

O'Loughlin J. and Kolosov V. (2017). Building identities in post-Soviet «de facto states»: cultural and political icons in Nagorno-Karabakh, South Ossetia, Transdniestria, and Abkhazia. Eurasian Geography and Economy, 58(6), 691-715. DOI: 10.1080/15387216.2018.1468793.

Ozkan B. (2008). Who gains from the «No War No peace» situation? A Critical Analysis of the Nagorno-Karabakh Conflict. Geopolitics, 14(3), 572-599.

Radnitz S. (2019). Reinterpreting the enemy: Geopolitical beliefs and the attribution of blame in the Nagorno-Karabakh conflict. Political Geography, 70, 64-73.

Shahnazarian N. (2010). National Ideologies, Survival Strategies and Gender Identity in the Political and Symbolic Contexts of Karabakh War». In Cultural Paradigms and Political Change in the Caucasus. Lap Lambert Academic Publishing.

Shahnazarian N. (2011). In the close embrace of tradition: Patriarchy and War. Saint Petersburg: Aleteia (in Russian).

Shahnazaryan N., Sukhashvili T, and Banu Z. (2019). Stories of Help and Rescue: the Georgian-Ossetian and Nagorno-Karabakh Conflicts. Caucasus Edition: Journal of Conflict Transformation, 4, 15-23.

Shnirelman V.A. (2003). Memory wars. Myths, identity and politics in Transcaucasia. Moscow: Akademkniga, 453 (in Russian).

Socio-Economic Situation of Nagorno-Karabakh Republic in January-December 2018, Stepanakert, 2019.

Statistical Yearbook of Azerbaijan (2019). Baku: State Statistical Committee of the Republic of Azerbaijan.

Sujyan S. (2010). Demographic situation in the Nagorno-Karabakh Republic. 21st Century, 1 (13), 32-58. (in Russian).

Suny R.G. (1993). Looking Toward Ararat: Armenia in Modern History. Indiana University Press.

Tokarev A. A. (2017). Comparative analysis of secessions in the post-communist space: quantification of influence factors. Polis. Political Studies, 4, 106-117 (in Russian with English summary).

Yunusov A. (2018). Stereotypes and «image of the enemy» in Azerbaijan. [online] Available at: www.publicdialogues.info/node/874 [Accessed 20 Dec. 2015] (in Russian).

Received on January 11th, 2020 Accepted on January 21th, 2020

Elena A. Grigorieva[1*], Anna B. Sukhoveeva[1]
[1] Institute for Complex Analysis of Regional Problems Far Eastern Branch Russian Academy of Sciences, Birobidzhan, Russia
***Corresponding author:** eagrigor@yandex.ru

QUALITY OF LIFE, REPRODUCTIVE HEALTH AND SOCIAL SECURITY: MEDICAL AND SOCIAL ENVIRONMENT AT THE RUSSIAN FAR EAST

ABSTRACT. Medical and social environment is discussed for the southern part of the Russian Far East, in system "Quality of life and reproductive health" at different hierarchical levels; that are at the meso-level – Khabarovsk Krai and the Jewish Autonomous Region (JAR); at the micro-level – Smidovichsky District in JAR and Nanai District in Khabarovsk Krai; at the local level – municipal settlements in urban and rural areas. The aim of the research is to identify the features of the social and medical environment affecting the quality of life, with an emphasis on the health of indigenous and non-indigenous population of reproductive age as the main criterion of quality of life. For subjective estimation of their health, well-being and quality of life, sociological surveys of women of reproductive age and pregnant women was conducted using a special questionnaire "Medical and social passport of future parents". The region is characterized by low indicators of health and reproduction of the population, weak social infrastructure. The analysis of the quality of life in the region requires the establishment of priority groups of risk factors to improve the efficiency of medical and social control to minimize their impact.

KEY WORDS: quality of life, reproductive health, medical and social environment, indigenous people, Russian Far East

CITATION: Elena A. Grigorieva, Anna B. Sukhoveeva (2020). Quality Of Life, Reproductive Health and Social Security: Medical and Social Environment at the Russian Far East. Geography, Environment, Sustainability,Vol.13, No 1, p. 92-98
DOI-10.24057/2071-9388-2019-52

INTRODUCTION

Increasing the birth rate and life expectancy of the population, reducing mortality, preserving the reproductive health of women and reducing perinatal losses of the population and, consequently, improving the quality of life (QOL) are the main requirements for the consolidation of the population, as a part of the High priority National Programs in Russia. The transformation processes affecting all spheres of life in the Russian society have had a significant impact on public well-being, the level of public health and the quality of life. The material, social, demographic, political and religious environment of different social groups, including pregnant women and women in reproductive age, their needs in various spheres of life and the possibility of their satisfaction, have changed in all regions of the country, including Russian Far East, particularly in Khabarovsk Krai and the Jewish Autonomous Region (JAR). Social well-being of these groups of women, their subjective perception and assessment of living conditions are modifying, indicating changes in quality of life.

The consequences of a prolonged series of crises in Russia have increasingly affected the financial situation of pregnant women as a poorly protected social category. In this atmosphere a worthwhile QOL of women in reproductive age in general, and expectant mothers in particular, is an important criterion for increasing fertility in the region, and an assessment of the social efficiency of health care and management systems.

The problems related to the quality of life have long been at the forefront of state social policy in the most developed countries of the world. QOL of society as a whole and its various social groups can serve as an integral assessment of the governance effectiveness. In Russia, where state and municipal authorities were primarily concerned about the standard of living of individual social groups or territorial societies, the scientific study of the quality of life of pregnant women and women of reproductive age is of great importance for the understanding of social and managerial practices. At the same time, the sociological survey of factors affecting the quality of life of women of reproductive age and expectant mothers, as well as their adaptation strategies that change the level, image and quality of their life to ensure the implementation of their life plans, well-being, is of particular importance.

Sociological study of the transforming needs of these groups of women, aimed at changing their lives and their perception, allows providing modern municipal management practices with the necessary scientific knowledge, and making municipal management socially effective to further improve the QOL. At the same time, the change in the quality of life is a problem solved not only at the state or municipal levels, but also at the individual (personal) level, where the quality of life directly depends on the ability to form new models of social–economic behavior. Adaptation strategies that help to adapt to rapid changes in personal life situations and socio-economic transformations in the country can serve as a basis for these models.

Quality of Life

Quality of life (QOL) is one of the most important indicators of social well-being and health status; it is widespread in the Western world and is becoming increasingly used in Russia in recent years. This term was used first in the late 1950-ies in the works of J. Galbraith (1973) and D. Risman et al. (1961).

Literature review shows three approaches to the concept of "quality of life". In the 50-60-ies in the United States it was developed as a parallel to the American way of life in terms of social and cultural spheres of living. The meaning of the concept was revealed as the possibility of consumption of goods and services that characterize social reality through economic indicators (working conditions, housing and other material goods). Afterwards it was realized that spiritual goods and needs should also be included in the system of indicators of QOL.

Another interpretation of the quality of life was proposed by the American sociologist A. Toffler (1970), who interpreted QOL in three aspects: environmental, economic and social as a transition from the satisfaction with basic material needs to the stage of satisfaction of refined, modified personal needs of the consumer.

The attempts of Western sociologists to limit themselves only to subjective criteria (the degree of life satisfaction), and biologists – only to objective ones (the biopsychological status of a person) were not successful. Since the 1960s, there have been more and more studies in the West, where QOL was considered as an interdisciplinary category, with its main feature – the presence of objective conditions of formation and subjective perception of the population (Lydick and Epstein 1993, Peters 2019). It became obvious that this category should be both subjective and objective, matching the assessment of the quality of the environment as a whole. It should include living environment directly (structure of individual environment), personal biopsychic and social (individual potential) and the degree of satisfaction with life (Peters 2019). These issues were considered from different positions in the studies by A. Campbell, V. Roger, A. Sen, J. Galbraith, A. Toffler, M. Newell, J. Forrester, R. Bauer, G. Kahn, J. Fourastie, etc. Despite the fact that there were no clear formulations of the QOL, the concept was discussed as economic, and later – sociological, socio-geographical, environmental category.

In domestic research of the Soviet Union period, the increased interest to the problem of QOL appeared due to other reasons than in the Western countries, and was associated with the beginning of political and socio-economic reforms of the 1980-1990-ies, where QOL was considered from critical positions as the ideology of the capitalist way of life, and QOL differences were determined by the living conditions. S.I. Popov (1977) argued that the concept of “quality of life” must be put on a par with the “standard of living”, “way of life”. His idea is in a partial contradiction with discussion of the similarities and differences among the concepts of “quality of life”, “well-being” and “comfort”, by Pinto and co-authors (2017). They argued that QOL is mainly related to individual perception of satisfaction with life, while well-being reflects the psycho-spiritual dimensions (Pinto et al. 2017).

In modern Russian studies, the development of ideas about the QOL concept has a universal scientific character, and is used by a wide range of natural and human sciences, including mathematics and medicine, each of which offers its own interpretation of the concept, giving it specific features. Today, the problems of living standards, employment, and poor health have come to the fore in QOL studies, and the existing level of income in the region is becoming the most polarizing element affecting other QOL characteristics. Questions of theory and practice of the study of QOL and related categories, such as standard of living, lifestyle, living conditions, medical, demographic and social development were considered by many well-known Russian scientists: A.I. Alekseev, E.G. Animits, T.I. Zaslavskaya, N.I. Zubarevich, G.M. Lappo, V.V. Pokshishevsky, B.B. Prokhorov, N.M. Rimashevskaya, S.V. Ryashchenko, N.A. Shchietova, S.A. Aivazyan, etc. The QOL concept is multifactorial, i.e. it includes many components, a set of indicators suitable for any comparison. Its characteristics reflect in different proportions living conditions and quality of the population. The theoretical basis of modern studies of QOL is a systematic approach that considers objective and subjective indicators as equivalent.

Thus, literature review resumes three main approaches to the concept of “quality of life”:

1. Assessment of social reality through economic indicators;

2. Synthesis of environmental, economic and social aspects of QOL to characterize both ecological and economic societal needs, and social particular citizen’s needs;

3. Integral subjective-objective assessment of the environment quality as a whole, the immediate living environment (individual environment structure), the biopsychic and social condition of the individual (individual capacity) and life satisfaction.

The following main criteria necessary for the integrated assessment of QOL can be summarized: income, poverty and inequality, labor use and unemployment, demographic processes, education and training, health, food and nutrition, urban infrastructure and communication, security (social and political), culture and social relations, natural environment, political and civil institutions. Fig. 1. presents the list of criteria of for integrated assessment of QOL including quantitative and qualitative estimations.

The system approach includes both quantitative (specific values, digital data) and qualitative indicators, the last one representing subjective estimates of a personal well-being. The specificity and complexity of the use of these criteria in one goal is based on heterogeneous character of indicators (demographic, social, environmental, medical), which must be expressed in one measurement system.

The most well-known integral indicator in “quantitative” approach is the human development index (HDI), used by UN experts, including estimates of life expectancy, income and education (Anand and Sen 1994; Qiu et al. 2018; Permanyer and Smits 2019; etc.). The “objective” block includes indicators characterizing the conditions for subjective state of a person or social group – natural and geographical, environmental, socio-economic, medical and hygienic, etc. “Qualitative” approach studies QOL through the subjective assessment of a personal well-being, satisfaction with life as a whole or its individual aspects (demands, values, etc.). The number of indicators in this block varies depending on the scale and depth of the study, and can be supplemented by estimates of satisfaction with health and medical care, level of education, family relations, financial situation, assessment of mental state, labor, life and leisure conditions, etc. Most researchers consider the associated analysis of quantitative and qualitative indicators as a necessary requirement for the implementation of an integrated approach in assessment of QOL.

Health and quality of life

There are different ways of dealing with the characteristics of health in terms of the quality of life (Fallowfield 1990). According to the UN, among other 12 parameters, health status should be the first in the social category of QOL. The UN Economic Commission for Europe has systematized eight groups of QOL indicators, with health status also being at the first place. Human health, including its reproductive component, as a complex criterion of quality of life, is considered as the most important prerequisite for the reproduction of the quality of labor and human potential as a whole. In this regard, studies focused on the development of medical, social, environmental and climatic indicators and criteria for assessing the health status of the population as the main indicator of QOL are of fundamental importance (Malkhazova et al. 2015; Tikunov and Chereshnya 2016; Malkhazova et al. 2018; etc.)

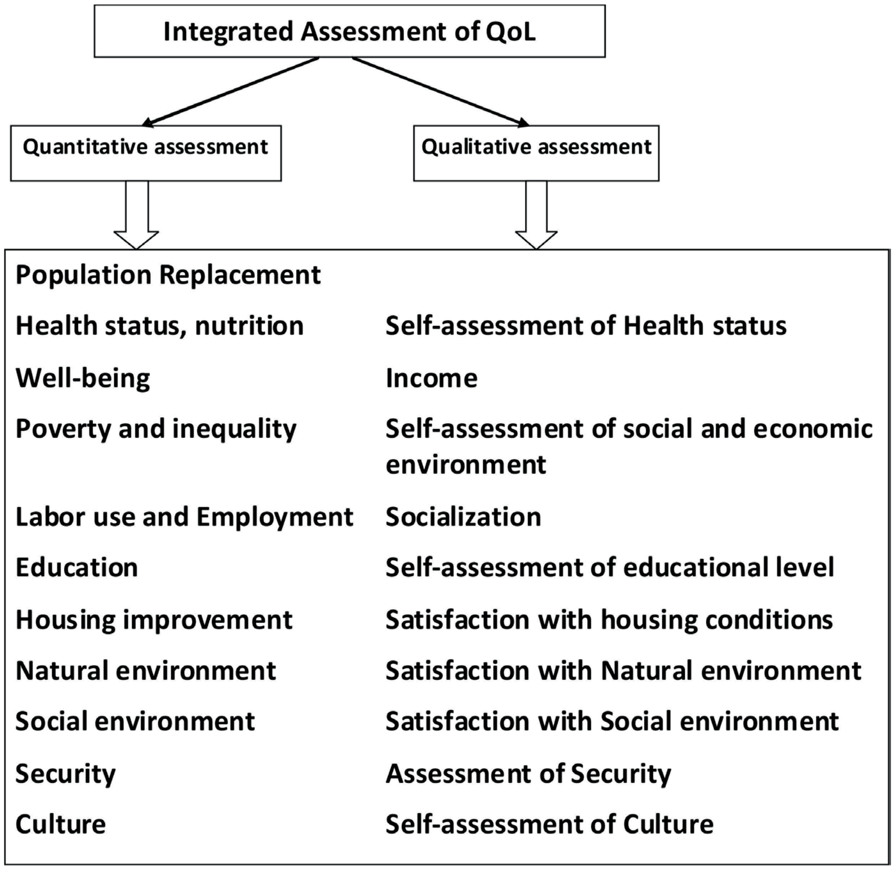


Fig. 1. The system of criteria for assessing the quality of life

This can be both an objective reflection of the features of natural-environmental and medical-social differentiation of living conditions, and subjective perception, showing the importance and individual characteristics of adaptation of a particular person (healthy or sick) through satisfaction with living conditions, external condition, the degree of personal and societal comfort.

Medicine offers a more complex, modern interpretation of QOL, calling it "health-related quality of life" and at the same time realizing that a good state of human health and well-being reflects not the absence of disease, but the satisfaction of demands and adaptation in the physical, psychological and social spheres (Guyatt et al. 1993). Thus, the quality of life associated with health is an integral characteristic of the physical, mental and social functioning of a healthy and sick person, based on his subjective perception. The quality of life associated with health, as a new methodology in medicine of the XXI century, makes it possible to assess the effectiveness of the health system through a subjective assessment by a human being of his physical, mental, social health and his role in the context of the impact on the overall quality of life (WHOQOL Group 1995).

Survey as a method for study of health status as an integral part of QOL

The profiles (the assessment of each component of QOL separately) and questionnaires (integrated evaluation) are the main instruments for QOL study in medical sector worldwide (Lin et al. 2013). Questionnaires are used as a research method to assess the level of individual health status to determine the prognosis of the disease and the effectiveness of treatment. Today, more than 50 questionnaires are used, which help to identify the subjective perception of health status and well-being in patients of different age groups, which cannot be achieved by using only traditional diagnostic procedures. Since 1995, the international non- profit organization for the study of QOL – MAPI Research Institute has been working in France (Mear and Giroudet 2012), coordinating the design of questionnaires and their cultural adaptation in different linguistic and economic environment.

In medicine, the assessment of patient's quality of life related to health is carried out in two ways: by another person, mostly by the doctor (objective approach); self-assessment (subjective approach). Development of subjective approaches was motivated by the results of studies that showed a lack of understanding by the doctor of adaptive or rehabilitation needs of the person (patient), when doctor's opinion may not correspond to the patient's settings. The most expedient is a combination of an objective approach, which reflects socially acceptable standards of life, and a subjective approach, that allows evaluating norms and preferences of a person (patient).

Next methods are widely used to measure various aspects of QOL related to health in medicine: Karnofsky scale (KPS) (Karnofsky and Burchenal 1949; Slevin et al. 1988); the index of well being by Campbell – Campbell's Index of Well-Being (Campbell et al. 1976), Nottingham health profile (NHP) (Hunt et al. 1981, 1985).

In Russia the most popular approach was proposed for the comprehensive integrated assessment of QOL by the questionnaire WHO-QOL-100 (The WHOQOL group 1998), and questionnaire Short Form – 36 (MOS-SF-36) (http://www.sf-36.com), both developed in accordance with the principles of evidence-based medicine and the requirements of Good Clinical Practice (GCP) which helped to expand the idea of the doctor on the condition of the person (patient) as a whole.

WHO-QOL-100 is a self-completion questionnaire. Its questions relate to the individual's perception of various aspects of his life with an assessment of the six major areas of QOL: physical functions, psychological functions, level of independence, social relations, environment and spiritual sphere, as well as directly measured by the respondent's perception of his QOL and health status as whole. The most popular is the Short Form – 36 questionnaire, which allows to cover persons in age categories from 14 years and older, consists of 36 questions grouped into eight scales, and involves a differentiated approach to the assessment of QOL depending on gender and age (Jenkinson et al. 1993).

The QOL in the Amur River basin in the southern part of the Russian Far East (RFE), hereinafter referred to as Priamurye is determined largely by the natural (geographical) environment, the history of population, economic and geographical development of the territory, which largely determined the types of economic activity, working and living culture of labor skills (Grigorieva and Sukhoveeva 2019).

The aim of this work is to identify the patterns of the social and medical environment affecting the quality of life at the RFE, with an emphasis on the health of indigenous and non-indigenous population of reproductive age as the main criterion of quality of life. The social and medical environment is considered as a factor in the formation of quality of life and evaluated in terms of its optimality for a healthy, harmonious, socially satisfied life of the population in region. Assessment of the quality of the environment provides a description of all components that form it and at the same time act as prerequisites for the implementation of the life of the indigenous and non-indigenous population.

MATERIALS AND METHODS

Methodologically, the QOL assessment of the quality of life is based on the synthesis of the system analysis of quantitative and qualitative indicators, taking into account the relationships in local social systems with environment at different spatial levels.

The research was carried out for rural and urban areas of Priamurye on different hierarchical levels, from April to October, 2018: Khabarovsk Krai and the Jewish Autonomous Region (JAR) at meso-level; Smidovichsky District in JAR and Nanai District in Khabarovsk Krai at the micro-level; municipal settlements of Khabarovsk Krai (Khabarovsk, Troitskoe,) and the JAR (Smidovich, Nikolaevka, Volochaevka, Danilovka, Aur, Peschanoe) at the local level. 92 pregnant women, 16 of them indigenous Nanai women in Troitskoe, Khabarovsk Krai, and 246 women of reproductive age, including 27 Nanai women in Troitskoe, Khabarovsk Krai, were surveyed.

The assessment of medical and demographic factors at the meso- and micro-levels is based on the analysis of statistical reports (Demographic Yearbook 2017; Khabarovsk Krai 2017). Qualitative assessment of QOL at the local level is based on the analysis of information obtained in a sociological survey using "Medical and social passport of future parents"; two categories of the population were interviewed. First, women of reproductive age were surveyed on satisfaction with their QOL and its components (health, medical care, housing, reproductive behavior, etc.). This subjective assessment was supplemented by the self-assessment by pregnant women (regardless of age and period of pregnancy) of their "health-related quality of life" with a description of the subjective perception of their health status and well-being, readiness for motherhood.

Questionnaire "Medical and social passport of future parents" was developed for an assessment and self-assessment of health status, well-being, quality of life, taking into account domestic and foreign social and medical experience. The questionnaire for pregnant women was developed taking into account the questionnaire of the international program Medical Outcomes Study Short Form (SF-36), which includes closed, open, semi-closed and personal questions. Additionally, respondents were asked some questions concerning specific problems of their health (the presence of symptoms of depression, bad habits, etc.).

RESULTS AND DISCUSSION

Medical and demographic situation

Representatives of 17 indigenous small-numbered peoples of the North, Siberia and the Far East live in the national villages of Khabarovsk Krai with total population of about 23 thousand people (1.7% of the total population); most of them (about 11 thousand) are Nanai people. Nanai villages are located on both banks of the Amur River. The number of various indigenous small-numbered peoples in Nanai District was 4752 in 2016, or 23% of the total population, of which 94% were Nanai. The population of Nanai District is rural, living in 42 national communities. During 2012-2016 the number of the permanent population decreased by 1.1 times (885 people), mainly due to migration. Nanai population is characterized by a positive rate of natural increase (0.3 to 1.4‰).

It's worth to mention, that today almost all indigenous peoples at the Russian Far East live in rural area, which is consistent with the Aborigines in the Arctic regions of Russia (Popova 2019). Let say, in contrast, 50% of native peoples in Canada live in urban areas (Wilson and Young, 2008).

The patterns of morbidity and mortality of indigenous minorities have their own characteristics. Mental disorders, diseases of the nervous system (including alcoholism and drug addiction) and infectious (tuberculosis) diseases dominate in the structure of morbidity. Alcoholism is 1.5 times higher than for the whole population of the region, and tuberculosis is twice higher. Higher levels of morbidity and mortality in indigenous compared to non-indigenous populations is noted, that is known worldwide (Rix et al. 2018). Injuries and poisoning, circulatory diseases and respiratory diseases prevail in the structure of mortality. We may suppose the reasons for this situation are the destruction of the traditional way of life and family relations; low standard of living; professional activities, not peculiar to the peoples of the North; the influence of the non-indigenous population. Attachment to traditional cultures and lifestyles cannot hamper achieving of economic goals. These findings are in consistence with the view of those such as Dockery (2010) in his research for Australian aboriginals that "continuity of traditional indigenous culture provides a degree of protection against those underlying causal factors", such as relationship between stronger cultural attachment and improved socio-economic outcomes (Dockery 2010, p. 330).

4686 people (11% of them are Nanai) lived in Troitskoe village of Nanai District in 2016 (13% in 2010). Rate of natural increase was 3.7‰ in 2016 (7‰ in 2010), birth rate was 19.4‰ (17.8‰ in 2010); mortality rate was 15.7 deaths per 1000 (14.6‰ in 2010).

In 01.01.2016, the population of the Jewish Autonomous Region (JAR) was 164.2 thousand (urban population – 68.6%). The entire population of the JAR is non-indigenous, unlike other regions of the Russian Far East, where the indigenous small-numbered peoples live.

The rate of natural increase (decline) of the population in the Jewish Autonomous Region is negative since 1992. It was –1.8‰ in 2016 due to the constant significant excess of the mortality rate over the birth rate. The mortality rate in JAR was 15.2 per 1000, exceeding indicator for the whole Far Eastern Federal District by 1.2 times (12.6 per 1000). The birth rate was 13.4‰. Predominance of the elderly population and the early mortality of men are the main reasons for the increase in mortality, resulting in high indicator of "sexual dimorphism of mortality" (1.4–1.6 times): a value showing how the mortality rate of men exceeds the mortality rate of women. Respiratory system, circulatory system and digestive system diseases are the main in the structure of morbidity, making 28.1, 12.7 and 7.0%, consequently.

The permanent population of Smidovichsky District in JAR reached 24 thousand people in 2018, decreasing by 4 thousand people during 2010-2018. For the period 2010-2015, mortality rate was 15.3–17.0‰; birth rate was 11.5 to 12.6‰; natural population decline was –2.6... –5.0 per 1,000. 65.84% of the population lives urban areas (Volochaevka, Nikolaevka, Priamursky and Smidovich).

In Smidovich the population was 4.3 thousand people in 2018, decreasing by 17% during 2010-2018. The main reasons are depopulation and migration, mainly interregional movements. During the period 2010-2018, the rate of natural increase was consistently negative –6.4 ... –8.0‰, with 11.7–13.3‰ for birth rate. The mortality rate 18.2–20.3 per 1000 indicates a high level of mortality compared to other municipalities of JAR and Khabarovsk Krai.

Population in Nikolaevka was 6.5 thousand people in 2018, 18% lower than in 2010. Rate of natural increase was –3.8+0.4 per 1,000 live births in 2012–2016, the mortality rate was 12.2–16.0‰, the birth rate was 11.4–14.1‰.

The results of the survey

The majority of women have specialized secondary education (37% of indigenous women and 49% of non-indigenous women). Two-thirds of all respondents (65%) are employed. The ratio of non-working women (housewives, women on maternity leave, students and the unemployed) is 31% and 21% for Nanai and non-indigenous people respectively. As for our opinion, the main reason for the high unemployment of Nanai women is the traditional way of life and the current unfavorable economic climate in the rural area. The survey results show that more than a half of the respondents (62%) are satisfied with their health status, of which 47% are pregnant women, and 53% are women of reproductive age. Every fifth respondent rated his health as "good and very good" (27% and 13%, respectively) (Fig. 2). At the same time, the proportion of indigenous women in both categories rating their health as "bad and very bad" was 21.5% and 19.0%, respectively. The higher satisfaction of indigenous women with their health status compared to non-indigenous residents of Priamurye is in a good agreement with results by Popova (2019) in her research on self-assessment of health by aboriginal people and migrants in rural areas of Yamal at the Russian North. Social uncertainty indicates a decrease in perceived control over various aspects of women's lives and can indirectly provoke reduction of a health risk, for example, through unemployment and, as a consequence, poor nutrition. Thus, the impact of unemployment on health is related not only to its psychological consequences, but also to the financial problems it causes.

The problem of limited financial resources for a decent standard of living and maintenance of their health is less acute for Nanai women of reproductive age – 39.0% against 40.6% of non-indigenous women. The following reasons were mentioned as "other" additional factors: psychological problems (stress); lack of own housing (forced hiring of an apartment or poor housing conditions); lack of qualified medical care, remoteness of specialized medical centers and other social facilities, inability to get higher education or retrain for a change of profession.

The results of the survey show that the main reasons for the restriction of opportunities for the preservation and restoration of health in women of reproductive age and pregnant women is the decline in the level of well-being, the growth of stressful situations, lack of employment, dissatisfaction with the social and living conditions of the rural areas. Our findings are in a good agreement with the results from rural areas at the Midwestern United States, where growing poverty and income disparities are reported over the past 20 years, which leads to sizable population losses (Peters 2019).

The survey data revealed that 69% of pregnant indigenous women have income less than 10 thousand rubles (i.e. below the subsistence minimum), 42% of them are women with two or three children in the family (Fig. 3). 38% of indigenous women in reproductive age have income below the subsistence minimum, compared to 32% of non-

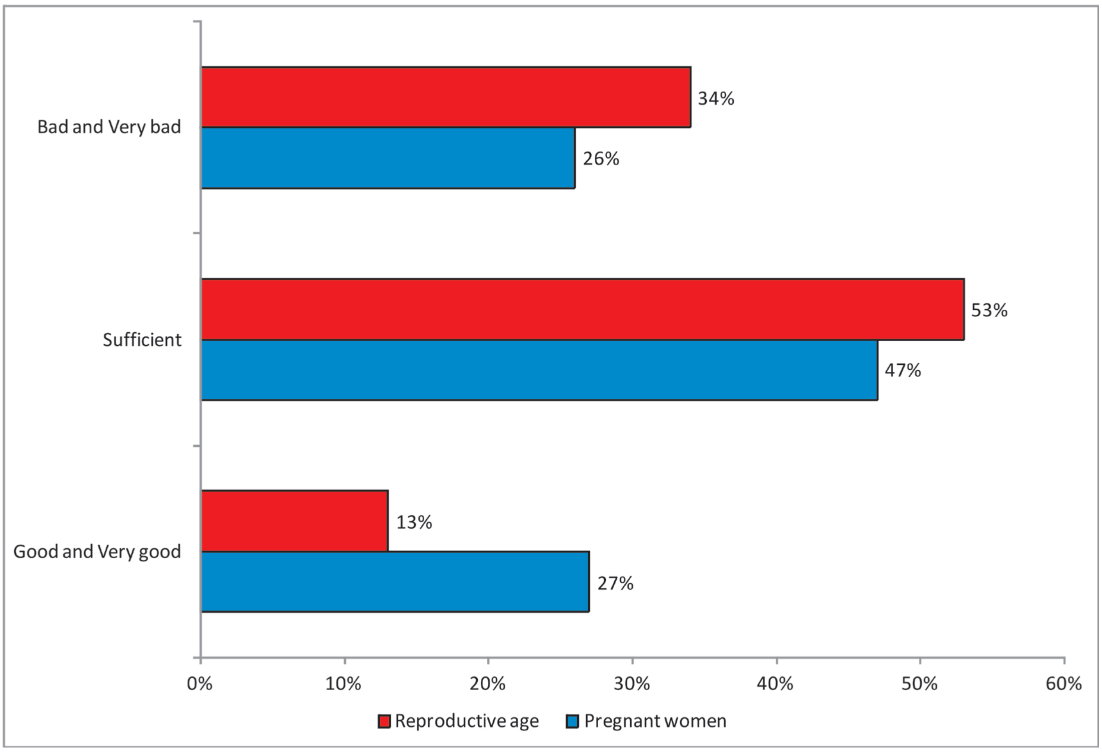


Fig. 2. Self-assessment of health status by women in Priamurye, %

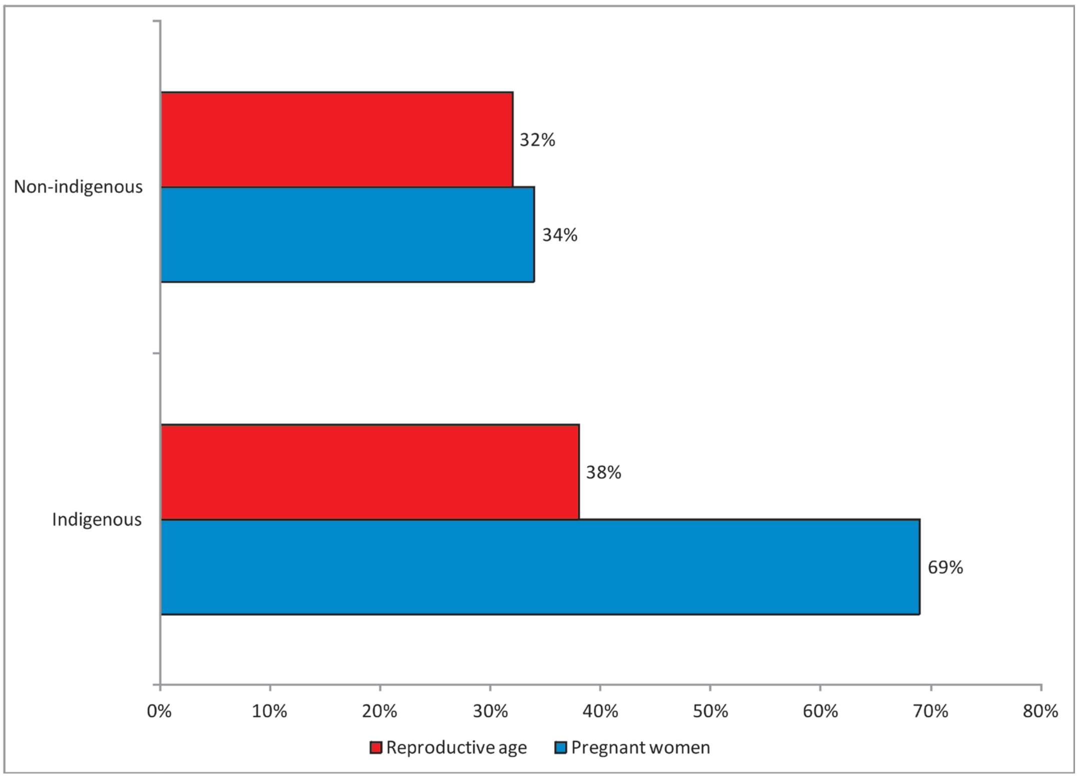


Fig. 3. Rate of women in Priamurye with income lower than 10000 Rub., %

indigenous women. The current average monthly income per family member (13,755 rubles) is insufficient to ensure a "decent" quality of life, to preserve and maintain health. It makes only 45.8% of the minimum limit (20 to 30 thousand rubles), which 72% of pregnant women and 37.5% of women of reproductive age consider necessary for a "wholesome" life.

To overcome demographic conditions unfavorable for women in reproductive age, particularly severe in rural areas, special social program, independent of outside socioeconomic or political forces, should be developed to reduce poverty and income inequality. Another proposal can be aimed at investments to economy development, which can lead to growing employment in industries. As Peters (2019) is pointing, its limitations are connected with requirement of sizable financing and long-term planning, when rural areas are powerless to make these major economic changes.

CONCLUSIONS

The differences and peculiarities in the structures of morbidity and mortality in the indigenous population (Nanai people) of Nanai District of the Khabarovsk Krai and non-indigenous population of Smidovichsky District, Jewish Autonomous Region; the features of the medical-demographic situation (processes of reproduction, mortality, health indicators), are identified. Nanai population is characterized by a positive rate of natural growth (0.3–1.4‰), in contrast to the non-indigenous population of the JAR. The structure of morbidity is dominated by mental disorders, diseases of the nervous system; injuries and poisoning dominate in the structure of mortality. The reasons are the destruction of the traditional way of life and family relations among the Nanai people, low standard of living, professional activities, not peculiar to the indigenous small-numbered peoples of the North, Siberia and the Far East; the influence of the non-indigenous population.

More than half of the respondents (62%) are satisfied with their health status (71% of pregnant women and 53% of women of reproductive age); every fifth respondent rates his health as "good" and "very good" (40% and 23%, respectively). About 40% of women, regardless of nationality, argue that high incomes, while not automatically guarantee of a good health, can provide a range of benefits and services necessary to maintain and strengthen their health potential. The social policy of the region and Russia as a whole, needs measures and actions to improve the reproductive health of women and increase the birth rate, to provide individual opportunities for women, such as guaranteed income, availability of employment, availability of comfortable housing and qualified medical care, investment in education, etc.

The decrease in the level of individual components and the overall assessment of QOL during pregnancy suggests that the pregnant woman is among the most vulnerable group of the population in terms of the impact of adverse factors and health conditions. Therefore, special focus is required to women in reproductive age, especially in unfavorable demographic environment at the Russian Far East.

As a result, important areas for future research can be pointed out. First, more attention needs to be directed towards subjective measures of improvements in medical staff or upgrading medical facilities. Second, future work should incorporate both objective and subjective measures of health status and income disparities in integral assessments of QOL. More research is needed to determine whether findings of the QOL measures from survey data are consistent across various economic and cultural contexts in both urban and rural areas.

ACKNOWLEDGEMENTS

The reported study was funded by RFBR according to the research project № 18-013-00923.

REFERENCES

Anand S., Sen A.K. (1994). Human Development Index: methodology and measurement. New York: UNDP Human Development Office.

Campbell A., Converse P.E., Rodgers W.L. (1976). The quality of American life: Perceptions, evaluations, and satisfactions. New York: Russell Sage Foundation.

Demographic Yearbook of the Jewish Autonomous Region (2017). Demographic Yearbook. Birobidzhan: Khabarovsk statistics, 81 p. (in Russian).

Dockery A.M. (2010). Culture and Wellbeing: The Case of Indigenous Australians Soc Indic Res, 99. 315–332. DOI: 10.1007/s11205-010-9582-y.

Fallowfield L. (1990). Quality of life. The missing measurement in health care. London: Souvenir Press, Human Horizon Series.

Galbraith J.K. (1973). Economics and the Public Purpose. New York: Signet.

Grigorieva E.A., Sukhoveeva A.B. (2019). Environmental and medical and social factors in the quality of life, reproductive health of the population: Priamurye as a case study. IOP Conf. Ser.: Earth Environ. Sci., 381, 012031. DOI: 10.1088/1755-1315/381/1/012031.

Guyatt G.H., Feeny D.H., Patrick D.L. (1993). Measuring Health-Related Quality of Life. Ann. Intern. Med. 118(8):622–629. DOI: 10.7326/0003-4819-118-8-199304150-00009.

Hunt S.M., McKenna S.P., McEwen J., Williams J., Papp E. (1981). The Nottingham health profile: Subjective health status and medical consultations. Social Science and Medicine. Part A: Medical Psychology & Medical Sociology, 15(3). 221–229. DOI: 10.1016/0271-7123(81)90005-5.

Hunt S.M., McEwen J., McKenna SP. (1985). Measuring health status: a new tool for clinicians and epidemiologists. J. R. Coll. Gen. Pract., 35. 185–188.

Jenkinson C., Coulter A., Wright L. (1993). Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. BMJ, 306(6890), pp. 1437–1440. DOI: 10.1136/bmj.306.6890.1437.

Karnofsky D.A., Burchenal J.H. (1949). The clinical evaluation of chemotherapeutic agents in cancer. New York, NY: Columbia University Press.

Khabarovsk Krai in numbers (2017). Short State Proceedings. Khabarovsk: Khabarovsk statistics, (in Russian).

Lin X.-J., Lin I.-M., Fan S.-Y. (2013). Methodological issues in measuring health-related quality of life. Tzu Chi Medical Journal, 25(1), 8–12. DOI: 10.1016/j.tcmj.2012.09.002.

Lydick E., Epstein R. S. (1993) Interpretation of quality of life changes. Quality of Life Research, 2(3), 221–226. DOI: 10.1007/BF00435226.

Malkhazova S.M., Mironova V.A., Orlov D.S., Adishcheva O.S. (2018). Influence of Climatic Factor on Naturally Determined Diseases in a Regional Context. Geography, Environment, Sustainability, 11(1), 157–170. DOI: 10.24057/2071-9388-2018-11-1-157-170.

Malkhazova S.M., Shartova N.V., Timonin S.A. (2015). Spatial patterns of public health in Russia. Geography, Environment, Sustainability, 08(04), 7–17. DOI: 10.24057/2071-9388-2015-8-4-7-17.

Mear I., Giroudet C. (2012). Linguistic Validation Procedures. In: Acquadro C, Conway K, Giroudet C, Mear I, eds. Linguistic Validation Manual for Health Outcome Assessments. Lyon, France: Mapi Institute, 15-117.

Permanyer I., Smits J. (2019). Inequality in human development across the globe. SocArXiv, DOI: 10.31235/osf.io/hskue.

Peters D.J. (2019). Community Resiliency in Declining Small Towns: Impact of Population Loss on Quality of Life over 20 Years. Rural Sociology. DOI: 10.1111/ruso.12261.

Pinto S., Fumincelli L., Mazzo A., Caldeira S., Martins J. C. (2017). Comfort, well-being and quality of life: Discussion of the differences and similarities among the concepts. Porto Biomedical Journal, 2(1), 6–12. DOI: 10.1016/j.pbj.2016.11.003.

Popov S.I. (1977). Problem of quality of life in the modern ideological struggle. Moscow: Politizdat, 280 p.

Popova T.L. (2019). Social and psychological adaptation of rural residents in Yamal. Vestn. Ural. Med. Akad. Nauki. = Journal of Ural Medical Academic Science, 16(2), 177-182. DOI: 10.22138/2500-0918-2019-16-2-177-182 (In Russ).

Rimashevskaya N.M. (2005). Social doctrine of Russia. Standards of living of Population in Russian Regions, 8/9, 63-93. (in Russian).

Risman D., Glazer N., Denney R. (1961) Lonely Crowd: A Study of Changing American Character. New Haven CT: Yale University Press.

Rix E.F., Wilson S., Sheehan N., Tujague N. (2018). Indigenist and Decolonizing Research Methodology. Handbook of Research Methods in Health Social Sciences, 1–5. DOI: 0.1007/978-981-10-2779-6_69-1.

Slevin M., Plant H., Lynch D., Drinkwater J., Gregory W. (1988). Who should measure quality of life, the doctor or the patient? British Journal of Cancer, 57(1), 109-112. DOI: 10.1038/bjc.1988.20.

Tikunov V., Chereshnya O. (2016). Public health index in the Russian Federation from 1990 to 2013 in comparison with other countries. Geography, Environment, Sustainability, 09(01), 94–108. DOI: 10.15356/2071-9388_01v09_2016_08.

Toffler A. (1970). Future Shock. Bantam Books, 1984.

Qiu Q., Sung J., Davis W., Tchernis R. (2018). Using spatial factor analysis to measure human development. Journal of Development Economics, 132, 130-149. DOI:10.1016/j.jdeveco.2017.12.007.

WHOQOL Group (1995). The World Health Organization Quality of Life assessment (WHOQOL): Position paper from the World Health Organization. Social Science and Medicine, 41, 1403–1409. DOI: 10.1016/0277-9536(95)00112-K.

WHOQOL group (1998). Development of the world health organization WHOQOL-BREF quality of life assessment. Psychol. Med., 28 (3), 551–558.

Wilson K., Young K. (2008). An overview of Aboriginal health research in the social sciences: current trends and future directions. International Journal of Circumpolar Health, 67(2-3), 179–189. DOI:10.3402/ijch.v67i2-3.18260.

Received on April 30th, 2019 Accepted on January 21th, 2020

Yelena A. Ogneva-Himmelberger[1*], Madeline Haynes[1]
[1] Clark University, 950 Main St., Worcester, MA 01610, USA
***Corresponding author:** yogneva@clarku.edu

SPATIAL PATTERNS OF ADVERSE BIRTH OUTCOMES AMONG BLACK AND WHITE WOMEN IN MASSACHUSETTS – THE ROLE OF POPULATION-LEVEL AND INDIVIDUAL-LEVEL FACTORS

ABSTRACT. This study explores spatial distribution of adverse birth outcomes (ABO), defined as low birth weight (<=2500 g) and preterm deliveries (gestational age <37 weeks), in black and white mothers in the state of Massachusetts, USA. It uses 817877 individual birth records from 2000-2014 aggregated to census tracts (census enumeration unit with population of approximately 4500 people). To account for small numbers of births in some tracts, an Empirical Bayes smoother algorithm is used to calculate ABO rates. The study applies ordinary least squares (OLS) and spatial regression to examine the relationship between ABO rates, seven individual-level factors from birth certificates and nine population-level factors (income level, education level, race) from census data. Explanatory power of these factors varies between the two races. In models based only on individual-level factors, all seven factors were significant (p<0.05) in the black mothers' model while only three were significant in the white mothers' model. Models based only on population-level variables produced better results for the white mothers than for black mothers. Models that included both individual and population-level variables explained 40% and 29% of ABO variance for black and white women respectively. The findings from this study give health-care providers and health-care policy-makers important information regarding ABO rates and the contributing factors at a local level, thus enabling them to isolate specific areas with the highest need for targeted interventions.

KEY WORDS: health; birth outcome; regression; GIS; USA

CITATION: Yelena A. Ogneva-Himmelberger, Madeline Haynes (2020) Spatial Patterns Of Adverse Birth Outcomes Among Black And White Women In Massachusetts – The Role Of Population-Level And Individual-Level Factors. Geography, Environment, Sustainability, Vol.13, No 1, p. 99-106 DOI-10.24057/2071-9388-2019-51

INTRODUCTION

Adverse birth outcomes (ABO), which include preterm deliveries (gestational age <37 weeks) and low birth weight deliveries (birth weight <=2,500 g) are a complex public health issue in the entire world. According to the World Health Organization, about 15 million babies are born preterm, and this number is on the rise, including in some high-income countries (World Health Organization 2018). In the U.S., the rate of preterm deliveries (PTD) in 2017 was 9.93% (up from 9.57% in 2014). For low birthweight (LBW) deliveries, the rate has also increased since 2014, rising from 8.00% to 8.28% in 2017 (Martin et al. 2018).

While the underlying cause of this increased prevalence of ABO is uncertain, potential hypotheses include a cultural transition to older age of women conceiving, increased use of assisted conception methods, and elevated prevalence of Cesarean sections and induced labor methods (American College of Obstetricians and Gynecologists 2016; Institute of Medicine 2007).

In the United States, the rate of PTD and LBW is higher in non-Hispanic black women than in non-Hispanic white women; rates of PTD and LBW in 2017 in black women were 13.93% and 13.89% while for white women they were 9.05% and 7.00% respectively (Martin et al. 2018). Despite many studies, the underlying causes of this disparity are not well understood (Burris and Hacker 2017; Kent et al. 2013; Lu and Halfon 2003; Manuck 2017).

Health risks associated with an ABO are severe and include an increased likelihood of respiratory problems, brain hemorrhage, heart complications, cerebral palsy, learning disabilities, and delayed motor and social skills (Centers for Disease Control and Prevention 2016; Clark et al. 2009; March of Dimes 2013; Rosenthal and Lobel 2011). Both PTD and LBW are also associated with increased infant mortality. According to the Center for Disease Control, PTD and LBW accounted for about 17% of infant deaths in 2015 (https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm).

Previous studies have investigated the role that an individual woman's health status may have on birth outcomes. Their results showed that diabetes, hypertension, substance use, previous ABO, and lower socioeconomic status increase the risk of an ABO (American College of Obstetricians and Gynecologists 2016; Berghella 2007; Goldenberg and Culhane 2007; Honein et al. 2009). It is important to consider the individual mother's health circumstances, though modeling ABO through these variables alone often do not capture the full risk present during pregnancy.

Several studies used various regression techniques to analyze the relationship between ABO rates, mother-level health characteristics and population-level characteristics (income, poverty, education, racial composition, population density, and environmental exposures). The geographical scope and unit of analysis varied from counties for the entire U.S. (Carmichael et al. 2014; DeFranco et al. 2008) to just metropolitan areas (Kramer and Hogue 2018), to zip (postal) codes and regional units within a particular state or province (Insaf and Talbot 2016; Kent et al. 2013; (Meng, Thompson et al. 2013). Researchers found that among mother-level characteristics, previous PTD, chronic hypertension, low pre-pregnancy weight, diabetes, maternal smoking during pregnancy, and elevated maternal age at delivery were associated with the likelihood of an ABO. Among population-level characteristics, percent population in poverty, percent with low education level, racial composition, and racial segregation were found to be significantly correlated with PTD and LBW (DeFranco et al. 2008; Insaf and Talbot 2016; Kent et al. 2013; Kramer and Hogue 2008).

The mechanisms, or pathways, through which population-level factors are transferred to individual risk factors, are complex and not fully understood. Research suggests that psycho-social stressors, associated with low socio-economic status (stressful work and living environment, reduced levels of social and financial support, deprivation, low access to health care facilities, and exposure to physical hazards, etc.), have impact on individual feelings and lead to depression and to unhealthy behaviors, such as smoking, drinking, substance abuse, delayed prenatal care and poor diet. These stressors also cause changes in neuroendocrine and immunological processes, increasing the risk of adverse birth outcomes (Meng, Thompson et al. 2013).

Although previous studies have investigated potential correlations between both socioeconomic and health related variables and birth outcomes, none evaluated correlations using more than a decade of individual birth data for an entire state at a detailed spatial scale (census tract). Census tract is the smallest geographical unit for which detailed socio-economic information is available from the Census Bureau.

To address these gaps in previous research, this study aims to: (1) to analyze geographic variability of ABO among black and white women in the state of Massachusetts, and (2) to examine the relationship between individual, area-level socio-demographic, and health-related factors and ABO rates at census tract level. Massachusetts ranges from densely populated metropolitan areas (Boston) and their suburbs to sparsely populated rural areas in the west and presents a wide variety of environmental and socio-demographic conditions. The state is divided into 14 counties, consisting of 39 cities and 312 towns. In Massachusetts, the distinction between a city and a town is based on the form of government chosen by the residents (https://www.sec.state.ma.us/cis/cislevelsofgov/ciscitytown.htm).

MATERIALS AND METHODS

We obtained individual birth data for 2000-2014 from the Massachusetts Department of Public Health and selected only singleton live births to non-Hispanic white and non-Hispanic black mothers for the analysis. Birth data were geocoded by the Department of Public Health to the census block level (the smallest enumeration unit in the U.S. Census). Six percent of births lacked census block information and were excluded from the analysis. Our final dataset included 725,582 births to white mothers, and 92,295 births to black mothers.

In order to facilitate the analysis of associations with socio-economic and demographic data, individual birth data was aggregated to census tracts. Census tract boundaries were obtained from the Office of Geographic Information, Commonwealth of Massachusetts (www.mass.gov/mgis/massgis.htm). There are 1472 census tracts with an average population of 4500 people in each census tract. The tracts that did not have any singleton live births to non-Hispanic white or non-Hispanic black mothers during the entire 15 years of study were excluded from the analysis, leaving 1467 census tracts for the analysis of births to white mothers, and 1449 tracts for the analysis of births to black mothers.

Each birth was assigned to a category based on birth weight and gestational age as follows: low birth weight (weight <= 2500 g) or normal birth weight (weight > 2500 g), and a full-term birth (gestational age >= 37 weeks) or preterm birth (gestational age < 37 weeks). A birth that was either preterm or low birth weight was considered an ABO. Total number of births and the number of ABOs for each census tract were calculated for the entire period (all 15 years combined).

Each birth record also contained mother-level data, such as mother's age, smoking during pregnancy, presence of gestational diabetes, gestational hypertension, chronic hypertension, and previous preterm delivery. Using this data, we calculated for each census tract the percentage of mothers

Table 1. Mean and standard deviation (SD) values for population-level and individual-level variables

Variable	Description	Mean (SD) for White mothers	Mean (SD) for Black mothers
Socioeconomic and demographic census variables			
High Education	Percent of individuals who possess an education level of a Bachelor's degree or higher	37.25 (20.43)	
Household Income	The median household income (US dollars)	67909.50 (29364.16)	
Low Education	Percent of individuals who possess an education level less than a high school diploma (or equivalent)	14.17 (13.02)	
Poverty	Percent of individuals living below poverty	11.77 (11.77)	
Per Capita Income	The average per capita income (US dollars)	33616.47 (14764.74)	
Unemployment	Percent of individuals who are unemployed	7.73 (4.92)	
Median Earnings	The median earnings for full time female employees	45764.20 (12775.34)	
Population Density	Population density (per square mile)	7269.29 (10649.15)	
Racial Composition*	Percent of individuals belonging to the corresponding racial group (either white or black)	80.48 (20.85)	6.50 (12.48)

Individual-level health variables			
Diabetes*	Percent of mothers with gestational diabetes	4.00 (1.64)	3.47 (2.91)
Pregnancy-Related Hypertension*	Percent of mothers with pregnancy-related hypertension	3.90 (1.6)	2.67 (2.53)
Chronic Hypertension*	Percent of mothers with chronic hypertension	1.36 (0.67)	1.92 (1.85)
Previous Preterm Infant*	Percent of mothers who delivered a previous infant who was preterm (<37 weeks gestational age)	0.96 (0.76)	0.79 (1.30)
Cigarette Use*	Percent of mothers who smoked cigarettes at any point during the pregnancy	8.67 (6.39)	4.42 (5.04)
Older Mom*	Percent of mothers who delivered a baby at an age greater than or equal to 35 years	20.28 (5.55)	11.30 (6.85)
Teen Mom*	Percent of mothers who delivered a baby at an age less than 20 years	4.10 (2.06)	5.25 (4.81)

These variables were calculated for the specific racial group

who had these health conditions and percent of teenage (younger than 20 years) and older (older than 35 years) mothers, separately for non-Hispanic black and non-Hispanic white mothers. This data is summarized in Table 1.

To explore geographic variation of ABO outcomes in more detail, we obtained boundaries of urban, suburban, towns, and rural locales from the National Center for Education Statistics (https://nces.ed.gov/programs/edge/Geographic/LocaleBoundaries) and overlaid them with census tract boundaries. Urban locales corresponds to principal cities with population over 100 thousand people; suburban locales have population between 50 and 100 thousand people and are located within the urbanized area adjacent to principal cities; towns are locales with population between 2.5 and 50 thousand people, located outside an urbanized area; and rural locales are all remaining territories. For the purposes of this research, we included towns into rural category because in Massachusetts both are similar in population density and types of land uses. Thus, we designated each census tracts as either primarily urban, suburban, or rural.

Population-level socioeconomic and demographic factors relevant to this study – education level, income, race, population density – were obtained from the 2006-2010 American Community Survey and the 2010 Census at a census tract level (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml). Table 1 provides detailed information about each variable. These variables were selected based on the findings of previous studies of ABO (Carmichael et al. 2014; DeFranco et al. 2008; Insaf and Talbot 2016; Kent et al. 2013; Kramer and Hogue 2008).

Raw ABO rates for each census tract were calculated for each census tract (dividing the number of ABO births by the total number of births), separately for both races. This approach produced potentially unreliable rates in areas with a small number of births. For example, if there were only two births in a census tract, and one of them was low birth weight, then the resulting ABO rate was 50%. This is often referred to as a "small numbers problem" or variance instability. One common approach to addressing this problem is to calculate adjusted, smoothed rates using Bayesian statistics. Using this approach, an estimate is obtained by combining the raw rates with "prior" information, such as the overall mean for the entire study area, i.e. an entire state (Anselin et al. 2006a). This smoothing method adjusts rates toward the overall mean, reduces variance instability, and produces robust and reliable rate estimates even for small samples (Kang et al. 2016; Mollalo et al. 2017). Adjusted ABO rates using Empirical Bayes smoother algorithm were calculated in the GeoDa software (Anselin et al. 2006b) for each census tract and these adjusted ABO rates were used in our analyses (Figure 1).

To contextualize ABO rates further, we used spatial selection tools in GIS and calculated ABO rates separately for urban, suburban and rural environments. A similar study found that urban areas in the state of Alabama had higher ABO rates

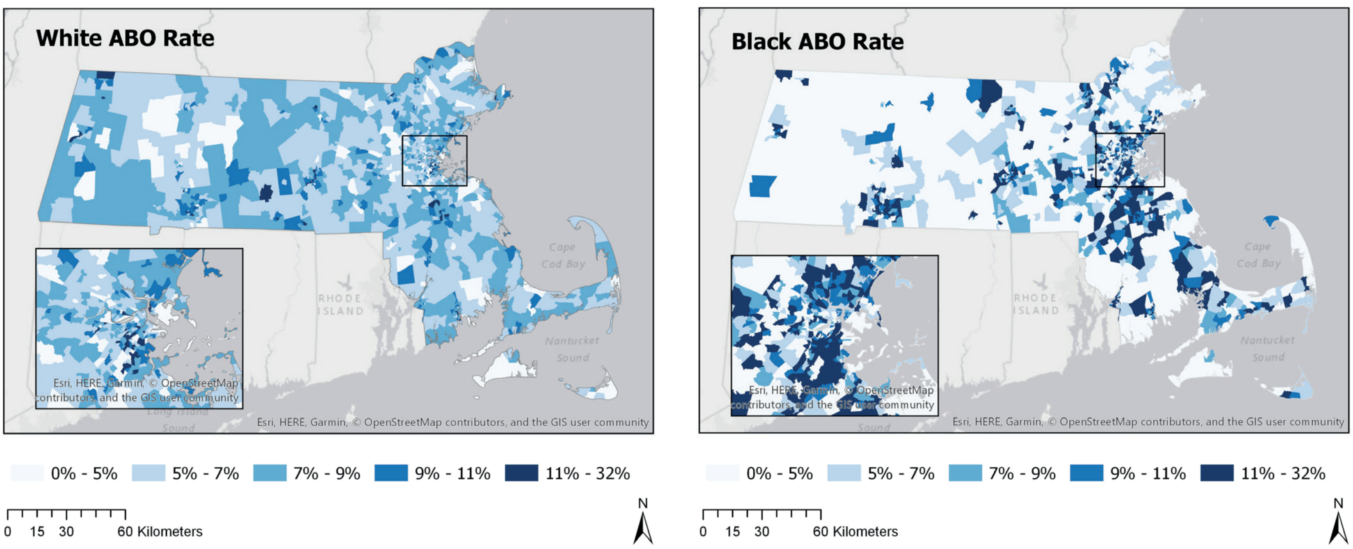


Fig. 1. ABO rates for white and black mothers

(Kent et al. 2013), and we wanted to see if the same is true in our state.

To characterize ABO rates' spatial pattern, a Global Moran's Index for both races was calculated. This index classifies the spatial pattern of a measured value (i.e. ABO rate) as random, clustered or dispersed, based on the index value and corresponding Z score. If the pattern is clustered or dispersed, it indicates that the observed pattern is not due to a random chance and that an underlying spatial process leads to a particular spatial pattern (Mitchell, 2005).

To determine the strength and the nature of the relationship between ABO rates and population-level and individual-level factors, we used regressions techniques. We applied multivariate ordinary least squares (OLS) regressions with the average rate of ABO for all 15 years as the dependent variable. Three separate regressions were run for each race: regressions with only individual-level variables (health-related data from birth certificates), with only population-level variables (Census data), and with all variables together. First run of each regression identified statistically significant variables (at 95% confidence level), and then only these variables were included in the final run of each regression. Independent variables included in the final OLS models are shown in Figure 2.

After each regression run, a Moran's I was calculated to test for spatial autocorrelation of the residuals. Z-scores were significant in all six regressions, indicating that residuals were not randomly distributed and suggesting a model misspecification. To address this problem, the two best-fit OLS regressions were selected (one for black and one for white births) and the same variables were used as the input into a spatial regression model. We followed Anselin's (2005) process for selecting the appropriate spatial regression model. This process compares multiple test statistics calculated in GeoDa software and indicates what model – spatial lag or spatial error – would be the best choice. Both models include an additional variable that explicitly captures spatial relationships in the data. Spatial lag model includes spatially lagged dependent variable as an additional independent variable in the analysis, and the spatial error model includes a spatial autoregressive error term (Anselin 2005).

RESULTS

Similar to previous studies, we found that there was a racial disparity in ABO rates between white and black women; statewide raw 15-year average ABO rate for white mothers was 7%, and for black mothers – 12%. When analyzed at a census tract level, maximum ABO rates are also very different between the two races, with the maximum rate for white women (18%) being much lower than for black women (32%). When stratified by urban-suburban-rural locations, the tract-level ABO rates for both white and black mothers showed the highest values in areas designated as urban and the lowest values in rural environments (Table 2).

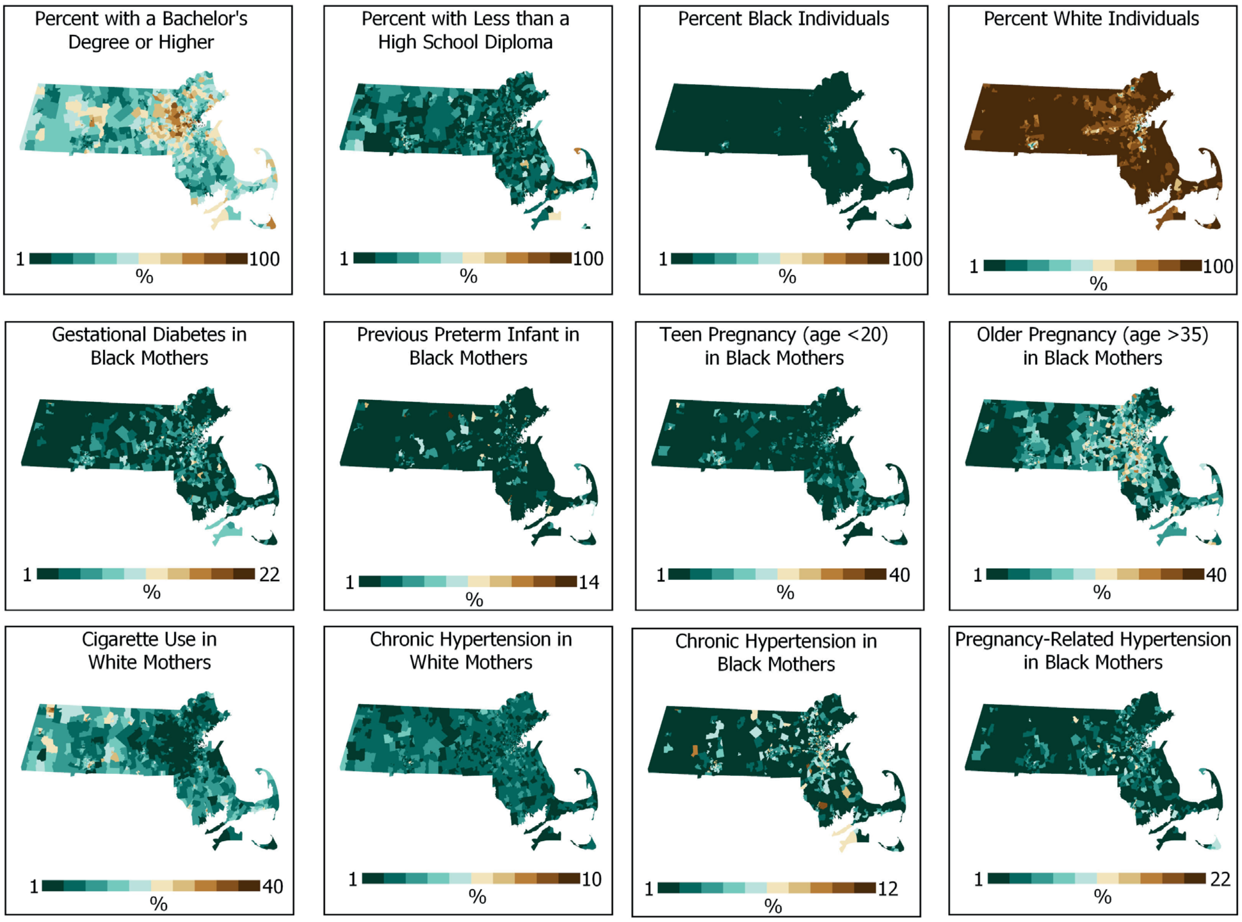


Fig. 2. Individual-level and population-level variables included in final regressions

Table 2. Mean ABO rate (standard deviation) per census tract for black and white mothers in urban, suburban and rural locations

	urban	suburban	rural
Black mothers	10.08 (5.64)	7.36 (5.44)	3.65 (4.30)
White mothers	7.98 (2.65)	7.45 (1.75)	6.92 (1.47)

Results of global Moran's I analysis indicate statistically significant clustering of ABO rates for both races (white Moran's I z-score = 4.86; black Moran's I z-score = 17.33). Maps of ABO rates confirm this finding; rates do not appear to be randomly distributed and there are several clusters of high values for both races located in different parts of the state (Figure 1).

Regression models based on individual-level variables explained 39 % of the variability in ABO rates for black women and considerably lower amount (23%) for white women (Table 3). Models that only contained population-level variables, explained similar percentages of the variance for black (21%) and white women (24%). Mixed models, containing both individual-level and population-level variables, explained 40% and 29% respectively, for black and white women.

Of the seven variables related to mothers' health and age, all were statistically significant at 95% confidence level in individual-level models for black mothers, and three – for white mothers. Percent teenage mothers has the largest effect on the ABO rates in black mothers (standardized coefficient 0.235), followed by the percent with chronic hypertension (standardized coefficient 0.182). For white mothers, the percent of mothers smoking and percent teenage mothers are the two strongest predictors (standardized coefficients are 0.361 and 0.143 respectively).

Of the nine population-level variables, education and income-related variables were statistically significant for both races at 95% confidence level, as well as percent corresponding race in the census tract. Unemployment was never a significant variable in any models.

When individual-level and population-level variables were combined in the mixed model, some of the variables remained statistically significant while others did not. For example, percent race and percent with chronic hypertension remained significant for both white and black mothers. For black mothers, all individual-level variables, except for one (percent smoking) remained significant. On the other hand, only two individual-level variables remained significant for white mothers in this mixed model. For population-level variables, an opposite pattern is present in the mixed model; while four variables remained significant for the white mothers' model (two education variables, median household income and percent white), only one variable (percent black) remained significant for black mothers' model.

The signs of the regressions coefficients were mostly in agreement with what we expected (e.g., higher percent college-educated and higher median household income were associated with lower ABO rates for both black and white mothers). Per capita income was statistically significant for white mothers, but its coefficient was the opposite of what

Table 3. Ordinary least squares regression results: Adjusted R^2, Akaike Information Criterion, and standardized coefficients

	Individual-level model		Population-level model		Mixed model	
	Blacks	Whites	Blacks	Whites	Blacks	Whites
Adjusted R^2	0.393	0.236	0.214	0.238	0.401	0.293
AICc	-4906.16	-7711.88	-4515.37	-7720.41	-4913.91	-7833.96
Population-level variables						
Percent with a Bachelor's Degree or Higher	n/a	n/a	-0.075	-0.301	n/a	-0.099
Median Household Income	n/a	n/a	-0.108	-0.176	n/a	-0.152
Percent Race	n/a	n/a	0.241	-0.173	0.120	-0.230
Percent with Less than High School Diploma	n/a	n/a	n/a	-0.054*	n/a	-0.080
Percent Below Poverty	n/a	n/a	0.143	0.108*	n/a	n/a
Per Capita Income	n/a	n/a	n/a	0.116	n/a	n/a
Median Earnings, Fulltime Female Employees	n/a	n/a	n/a	n/a	n/a	0.107
Population Density	n/a	n/a	0.062*	n/a	n/a	n/a
Individual-level variables						
Diabetes	0.113	n/a	n/a	n/a	0.115	n/a
Pregnancy-Related Hypertension	0.152	n/a	n/a	n/a	0.120	n/a
Chronic Hypertension	0.182	0.138	n/a	n/a	0.150	0.080
Previous Preterm Infant	0.091	-0.044*	n/a	n/a	0.110	n/a
Cigarette Use	0.064	0.361	n/a	n/a	n/a	0.338
Older Mom	0.155	n/a	n/a	n/a	0.137	n/a
Teen Mom	0.235	0.143	n/a	n/a	0.232	n/a

*All coefficients are significant at 95% confidence level; those marked with * are significant at 90% confidence level. Variables with an "n/a" were not included in the corresponding models.*

was expected (i.e., higher per capita income was associated with higher ABO rate). Percent race was a significant variable in all models, but had different signs – positive sign for black mothers, and negative sign for white mothers.

We selected two models with the highest R^2 (the mixed models) and used GeoDa software to calculate diagnostic statistics for spatial dependence. These tests showed that residuals were spatially autocorrelated in both models (z score for Moran's I for black model was 2.46; for white model – 2.48). Following model selection decision rule outlined by Anselin (2005), a spatial lag model was developed for black mothers, and spatial error model – for white mothers. To create spatial term in the equations, we experimented with different weights configurations and selected the weights that produced the best fitting regressions. We applied queen first order contiguity weights to black model, and queen second order contiguity weights – to the white model. In first order queen contiguity, census tracts that share common edges and corners are considered neighbors.

In both spatial regression models, all input variables remained statistically significant, and their coefficient signs were the same as in the OLS models. Spatial terms in both regressions (spatial lag term in the black model and spatial autoregressive error term in white model) had statistically significant coefficients with positive signs, representing spatial influence of the neighboring census tracts on ABO rates (Table 4).

Spatial regression models produced a psedo-R2 which is not directly comparable to the R2 from OLS models (Anselin 2005), so we used Log-Likelihood and AIC as measures of fit to compare these models to OLS models. In both spatial regressions, AIC value decreased and the Log-Likelihood value increased, suggesting an improvement of fit for the spatial models (Table 4).

After running spatial models, spatial autocorrelation was no longer present in the residuals (residuals Moran's I z-score was -0.1566 for black model; for white model 1.6653). This means that including the spatially lagged dependent variable term in the black model and spatially autoregressive error term in the white model has successfully eliminated all spatial autocorrelation.

DISCUSSION

In this study, we found that the ABO rates in Massachusetts varied considerably across census tracts and their distributions were very distinct for white and black mothers. Urban locations had higher ABO rates than suburban and rural locations. ABO rates with similar values were clustered in both races, but stronger clustering was observed in black mothers, as evidenced by their much higher z-score for Moran's I (17.33 vs. 4.86).

Most previous studies conducted the analysis at the scale of counties, metropolitan statistical areas or zip codes. Our study used census tract as the unit of analysis, because census tracts are small enough to allow for modeling of local variation in ABO rates. This spatial scale also facilitates the linking of the individual-level factors with census data and provides enough spatial detail to design a meaningful intervention or develop policy at regional or city/town level (Insaf and Talbot 2016).

When taken together, the selected socio-economic, demographic and health-related factors explained close to 30% of variability in ABO rates in white, and 40% in black mothers. When analyzed separately, individual-level and population-level factors explained the same amount of variability in ABO rates for white mothers (24%). For black mothers, individual-level factors explained almost twice the amount of ABO variability explained by the population-level factors (39% vs. 21%).

Among individual-level factors, smoking was the strongest predictor for white mothers and percent teenage mothers for black mothers (both in individual-level model, and in mixed model). To illustrate how these important findings can be useful to the health care providers and policy makers, the top 10% of tracts with the highest smoking rates for white mothers were selected, and 34 towns and cities that contain these census tracts were identified. The same process was repeated, and 31 towns and cities with the highest percent of teenage mothers for black mothers were identified. Twenty towns and cities were included in both lists, meaning that these locations have the highest smoking rates among white mothers during pregnancy and the highest rates of teen births to black mothers (Table 5). Health care providers and policy-makers in these towns and cities, armed with the findings from this study, could design targeted public outreach programs aimed at reducing smoking, especially among women, and delaying pregnancy among teenage women.

Our study had several limitations related to data sources and methodology. While birth data from the state department of Public Health is very detailed, we could not verify its quality and reliability. This data did not have information about father's health. The definition of some variables changed in the middle of our study period, rendering them unusable in the analysis (e.g., marital status and mother's education). We included population level variables measuring education level, but using mother's education level data would have been more relevant. Another limitation of the study is its inability to include other factors, which could have influenced ABO rates, such as other health conditions of the mother, psychosocial factors and potential environmental exposures. While many studies focus on either PTD or LBW, we combined them together in our analysis, recognizing that each may have a different, albeit overlapping, set of individual and population level factors. We combined them in order to increase the number of ABO outcomes in each census tract, thus alleviating the "small number problem" and increasing stability of our Empirical Bayes rates estimates.

These limitations notwithstanding, our study makes important contributions to the growing body of literature. It is the first to analyze ABO at the census tract level for an entire state, using 15 years of individual birth records. Additionally, this study is unique as it examines correlations of mother's health factors and socioeconomic factors separately as well as through a mixed model, which considers potential influences of both sets of characteristics.

Table 4. OLS and spatial regressions: Measures of fit and spatial term coefficients

Measure of fit	OLS model – Black	Spatial lag model – Black	OLS model – White	Spatial error model – White
Log likelihood	2466.02	2481.97	3926.04	3942.23
Akaike Info criterion	-4916.04	-4945.93	-7836.08	-7826.24
Spatial term coefficient and its z-value	-	0.1900 (5.64)	-	0.3614 (6.06)

Table 5. Cities and towns containing the top 10% of census tracts with highest rates of cigarette smoking among white expectant mothers and highest rates of teen pregnancy among black women

NAME	COUNTY	TYPE	POPULATION 2010
Attleboro	BRISTOL	City	43593
Brockton	PLYMOUTH	City	93810
Boston	SUFFOLK	City	617594
Barnstable	BARNSTABLE	Town	45193
Chicopee	HAMPDEN	City	55298
Fall River	BRISTOL	City	88857
Fitchburg	WORCESTER	City	40318
Gardner	WORCESTER	City	20228
Greenfield	FRANKLIN	Town	17456
Haverhill	ESSEX	City	60879
Lowell	MIDDLESEX	City	106519
Lynn	ESSEX	City	90329
Lawrence	ESSEX	City	76377
Montague	FRANKLIN	Town	8437
New Bedford	BRISTOL	City	95072
North Adams	BERKSHIRE	City	13708
Orange	FRANKLIN	Town	7839
Pittsfield	BERKSHIRE	City	44737
Southbridge	WORCESTER	Town	16719
Springfield	HAMPDEN	City	153060

The findings from our study provide Massachusetts health care providers and health-care policy makers with information regarding ABO rates and the contributing factors at a local level, giving them the ability to isolate specific areas with the highest need for targeted interventions. Examples of community-oriented public health interventions include improving access to healthy food and to primary prenatal care in low socio-economic areas, improving quality of health care accessible to expectant mothers, and increased social support in local communities (Lorch and Enlow 2016). These interventions would help alleviate the impacts of some psycho-social stressors and reduce the risk of adverse birth outcomes. In support of these interventions, the analysis of annual ABO rates thought time at census-tract level would also be very useful.

ACKNOWLEDGEMENTS

The work in this manuscript was supported by The Leir Foundation, in memory of Henry J. and Erna D. Leir.

REFERENCES

American College of Obstetricians and Gynecologists (2016). Preterm (Premature) Labor and Birth. Available at: https://www.acog.org/Patients/FAQs/Preterm-Premature-Labor-and-Birth. [Accessed 25 Apr.2019].

Anselin L. (2005). Exploring Spatial Data with GeoDaTM: A Workbook. Spatial Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign, Urbana, IL 61801.

Anselin L., Ibnu S., & Youngihn K. (2006). GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis, 38(1), 5-22.

Anselin L., Lozano N., Koschinsky J. (2006). Rate Transformations and Smoothing. Retrieved from Spatial Analysis Laboratory, Department of Geography, University of Illinois Urbana-Champaign. Available at: https://s3.amazonaws.com/geoda/software/docs/geodaworkbook.pdf [Accessed 25Apr.2019].

Berghella V. (2007). Prevention of recurrent fetal growth restriction. Obstetrics and Gynecology, 110(4), 904-912. DOI:10.1097/01.AOG.0000267203.55718.aa.

Burris H., & Hacker M. (2017). Birth outcome racial disparities: A result of intersecting social and environmental factors. Seminnars in Perinatology, 41(6), 360-366. DOI:10.1053/j.semperi.2017.07.002.

Carmichael S., Cullen M., Mayo J., Gould J., Loftus P., Stevenson D., Shaw G. (2014). Population-level correlates of preterm delivery among black and white women in the U.S. PLoS One, 9(4), e94153. DOI:10.1371/journal.pone.0094153.

Clark S., Miller D., Belfort M., Dildy G., Frye D., Meyers J. (2009). Neonatal and maternal outcomes associated with elective term delivery. American Journal of Obstetrics and Gynecology, 200(2), 156, 151-154. DOI:10.1016/j.ajog.2008.08.068

DeFranco E., Lian M., Muglia L., Schootman M. (2008). Area-level poverty and preterm birth risk: a population-based multilevel analysis. BMC Public Health, 8, p. 316. DOI:10.1186/1471-2458-8-316.

March of Dimes. (2013). Long-term health effects of premature birth. Available at: https://www.marchofdimes.org/complications/long-term-health-effects-of-prematu re-birth.aspx [Accessed 25 Apr.2019].

Goldenberg R., Culhane J. (2007). Low birth weight in the United States. The American Journal of Clinical Nutrition, 85(2), pp. 584s-590s. DOI:10.1093/ajcn/85.2.584S.

Honein M., Kirby R., Meyer R., Xing J., Skerrette N., Yuskiv N., Sever L. (2009). The association between major birth defects and preterm birth. Maternal and Child Health Journal, 13(2), 164-175. DOI:10.1007/s10995-008-0348-y.

Insaf T., Talbot T. (2016). Identifying areas at risk of low birth weight using spatial epidemiology: A small area surveillance study. Preventive Medicine, 88, 108-114. DOI:10.1016/j.ypmed.2016.03.019.

Institute of Medicine. (2007). Preterm Birth: Causes, Consequences, and Prevention. Washington, D.C.: The National Academies Press.

Kang S., Cramb S., White N., Ball S., Mengersen K. (2016). Making the most of spatial information in health: a tutorial in Bayesian disease mapping for areal data. Geospatial Health, 11(2), 428. DOI:10.4081/gh.2016.428.

Kent S., McClure L., Zaitchik B., & Gohlke J. (2013). Area-level risk factors for adverse birth outcomes: trends in urban and rural settings. BMC Pregnancy Childbirth, 13, 129. DOI:10.1186/1471-2393-13-129.

Kramer M., Hogue C. (2008). Place matters: variation in the black/white very preterm birth rate across U.S. metropolitan areas, 2002-2004. Public Health Reports, 123(5), 576-585. DOI:10.1177/003335490812300507.

Lorch S.A., Enlow E. (2016). The role of social determinants in explaining racial/ethnic disparities in perinatal outcomes. Pediatric Research 79(1): 141-147.

Lu M., Halfon N. (2003). Racial and ethnic disparities in birth outcomes: a life-course perspective. Maternal and Child Health Journal, 7(1), 13-30.

Manuck T. (2017). Racial and ethnic differences in preterm birth: A complex, multifactorial problem. Seminars in Perinatology, 41(8), 511-518. DOI:10.1053/j.semperi.2017.08.010.

Martin J., Hamilton B., Osterman M., Driscoll A., Drake P. (2018). Births: Final data for 2017. Centers for Disease Control and Prevention. National Center for Health Statistics. National Vital Statistics System. Available at: https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf [Accessed Apr. 25, 2019].

Meng G., Thompson M.E., Hall G.B. (2013). Pathways of neighbourhood-level socio-economic determinants of adverse birth outcomes. International Journal of Health Geographics 12, 32-32.

Mollalo A., Blackburn J., Morris L., & Glass G. (2017). A 24-year exploratory spatial data analysis of Lyme disease incidence rate in Connecticut, USA. Geospatial Health, 12(2), 588. DOI:10.4081/gh.2017.588.

Centers for Disease Control and Prevention (2016). Reproductive and Birth Outcomes. Available at: https://ephtracking.cdc.gov/showRbLBWGrowthRetardationEnv.action. [Accesed 25 Apr., 2019].

Rosenthal L., Lobel M. (2011). Explaining racial disparities in adverse birth outcomes: unique sources of stress for Black American women. Social Science & Medicine, 72(6), 977-983. DOI:10.1016/j.socscimed.2011.01.013.

World Health Organization. (2018). Preterm birth. Key facts. Available at: https://www.who.int/news-room/fact-sheets/detail/preterm-birth [Accessed Jan. 6, 2020].

Received on April 29th, 2019 Accepted on January 21th, 2020

Jaime Bosch[1], Irene Iglesias[1], Marta Martínez[1], Ana de la Torre[1*]
[1] Animal Health Research Center (INIA-CISA), Valdeolmos, Madrid, 28130, Spain
***Corresponding author:** torre@inia.es

CLIMATIC AND TOPOGRAPHIC TOLERANCE LIMITS OF WILD BOAR IN EURASIA: IMPLICATIONS FOR THEIR EXPANSION

ABSTRACT. Wild boar populations have continuously grown over the last century. This increase has led to various conflicts, including damage to agriculture and disturbed population equilibrium in natural areas, and it is a health threat due to animal and zoonotic infectious diseases, all with a high economic impact (e.g. Classical Swine Fever, African swine fever, tuberculosis or brucellosis). Addressing these problems requires understanding the geographic, climatic and topographic tolerance limits of wild boar. In this work, we determine these limits in Eurasia by spatially comparing the most widely accepted map on wild boar distribution (International Union for Conservation of Nature ,IUCN, 2008) with georeferenced records of wild boar presence (n = 34,233) gather from ecological and health sources. Results suggest a geographical expansion of the wild boar in the Eurasian zone outside the traditionally area described by the IUCN map. The specie has entered new biotopes and ecoregions, such as the equatorial region, where its presence is mainly associated with the large Asian plant monocultures. These results will support the development of population models, identification of permanent populations and habitats, and more effective decision-making about health and natural resource management.

KEY WORDS: wild boar; tolerance limits; distribution; population

CITATION: Jaime Bosch, Irene Iglesias, Marta Martínez, Ana de la Torre (2020) Climatic And Topographic Tolerance Limits Of Wild Boar In Eurasia: Implications For Their Expansion. Geography, Environment, Sustainability, Vol.13, No 1, p. 107-114
DOI-10.24057/2071-9388-2019-52

INTRODUCTION

The environmental tolerance of a species determines to a great extent its area of distribution, and is a result of the evolutionary process and environmental adaptations (Wiens and Graham 2005). The physiological limitations of a species are directly influenced by environmental conditions, mainly temperature, precipitation and humidity. Environmental variables and population factors interact in complex ways to influence species establishment in certain areas, survival and reproduction rates (Wiens and Graham 2005).

A good example of this dynamism is the wild boar (*Sus scrofa*), a species of great ecological plasticity. During the last century, its demographic growth has been exponential and continuous in Europe and many parts of the Palearctic (Lucchini et al 2005; Bosch et al. 2016; Parchizadeh 2017; Markov et al 2018). Wild boar have colonized new biotopes, natural and human-transformed (Markov et al. 2018). The worldwide increase in wild boar populations has led to numerous conflicts due to agricultural damages, problems in the conservation of natural areas, and threats to animal health. Wild boar are also increasingly present in human environments, invading urban areas and generating conflicts such as traffic accidents, attacks on people and pets, and health problems (Massei et al. 2015). Proper management of this growing wild boar population requires knowing its geographical distribution.

Wild boar can act as a reservoir for many transboundary diseases, such as classical swine fever and African swine fever (ASF), as well as zoonotic diseases such as tuberculosis or brucellosis, all with a high economic impact (Malmsten et al. 2017). Since 2007, ASF has been spreading across nearly the entire Eurasian territory, affecting 10 countries in the European Union (Iglesias et al. 2018; OIE 2018) and generating large losses in the global pig sector. Health authorities have emphasized the need for proper population control of boar in order to manage ASF, yet surveillance efforts remain inadequate (ECA 2016; EFSA 2018).

Models of species distribution can help to improve population and health management of wild animal populations (Ehrlén and Morris 2015; Bosch et al. 2016). To be useful, these models must include accurate information about wild species and related diseases. Biological records about wild boar have begun to be collected in a more uniform way through initiatives such as EUROBOAR, GBIF and Enetwild. Most of these initiatives are based on the distribution of wild boar within the area described by the International Union for the Conservation of Nature (IUCN) (Oliver and Leus 2008). However, wild boar have recently been observed at many sites outside the IUCN-demarcated area (Bosch et al. 2016; Markov et al. 2018), suggesting the urgent need to update our understanding of wild boar distribution. In addition, the ranges and limits of environmental tolerance of this species have never been described on a global level.

Therefore, the objective of the present work was to describe the current limits of wild boar distribution in Eurasia, as well as identify the climatic and topographic tolerance limits and biogeographic scenarios that condition its habitat.

MATERIALS AND METHODS

The current wild boar distribution in Eurasia, based on georeferenced occurrence data, was compared with the standard distribution maps from IUCN (Oliver and Leus 2008). Comparison focused on altitudinal distribution of wild boar, based on tolerance intervals for temperature and precipitation usually employed in wildlife distribution models, as well as on concepts in habitat quality and ecoregions (Sales et al. 2017).

Wild boar presence

The georeferenced occurrence of wild boar described by Bosch et al (2016) was updated using data from the following ecological and health sources from 2018: Global Biodiversity

Information Facility (GBIF), World Organization for Animal Health, Veterinary European Transnational Network for Nursing Education and Training, and the Genbank of the US National Center for Biotechnology Information. A total of 37,655 instances of wild boar presence were reported from 1982 to 2018 at a spatial resolution ≤10 km (~97% from field data), and all were initially considered in the study. To reduce potential spatial autocorrelation, the density of points was reduced by "extracting" data located close together (<10 km) using the statistical software R (R Development Core Team, 2012). Consequently, 34,233 wild boar occurrences from the original 37,655 were ultimately included in the study.

Environmental variables

Data on annual precipitation, precipitation during the driest month, and minimum and maximum temperatures during the coldest and warmest months were obtained from WorldClim (1950-2000) at a spatial resolution of 5 arc-minutes (~10 km) (Hijmans et al. 2005). Data on altitude were gathered from a global digital elevation model (LP DAAC 2004) at a spatial resolution of 5 arc-minutes (~10 km). Quality of the available habitat (QAH) for wild boar was obtained from Bosch et al (2016), while data on terrestrial ecoregions were obtained from the digital Köppen-Geiger world map on climate classification (Rubel and Koottek 2006).

Wild boar geographical distribution

To describe the current limits of wild boar distribution in Eurasia, wild boar presence was compared with the IUCN map (Oliver and Leus 2008) using overlays generated with ArcGIS 10.2 ESRI® software. The percentages of wild boar occurrences that fell within the IUCN area or in four buffer zones extending 10 km, 100 km, 500 km and >500 km from the edge of the IUCN area. The buffer zones were generated using the proximity analysis tool in ArcGIS 10.2.

After layer overlay, wild boar presence at increasing altitude was determined. The altitude associated with each wild boar occurrence was extracted, and the results were classified into six groups (minimum, maximum, median, 5th percentile, 25th percentile, and 95th percentile) in order to describe the distribution of altitudes at which wild boar were present.

Environmental tolerance, QAH and ecoregions of wild boar distribution

To describe the climatic limits that wild boar in Eurasia can tolerate and therefore that determine their habitat, wild boar occurrences were compared with the climatic variables selected by layer overlay. The value of each climatic variable associated with wild boar presence was extracted inside the IUCN area and in each of the four buffers, and the results were classified in six groups as previously explained. A similar analysis was performed to describe the biogeography of the wild boar habitat, in terms of QAH and ecoregions.

RESULTS AND DISCUSSION

The comparison showed that 89.1% (n = 31,720) of wild boar occurrences fall within the IUCN area, while 1% (n = 364) lie inside the 10 km buffer. In this buffer, wild boar presence may be associated with metapopulation movements, given an average movement of 10.38 ± 2.84 km for animals at least 17 months old (Keuling et al. 2010). In fact, previous studies have shown that 75-90% of wild boars are recaptured within 10 km of where they were first captured (Keuling et al. 2010).

The other 10% of wild boar occurrences fall outside the 10 km buffer: 8.4% in the 100 km buffer, 0.5% in the 500 km buffer, and 1% in the >500 km buffer (Fig. 1). These results indicate that wild boars are expanding in the Eurasian zone outside the traditional area described by the IUCN map, through their own movements as well as anthropogenic reintroductions. Below we look in greater detail at some countries showing evidence of settled wild boar populations outside the IUCN distribution area (Fig. 1).

In Russia, the largest percentage of wild boar occurrences outside the IUCN distribution area is located within the 100 km buffer in the western part of the country, at a latitude

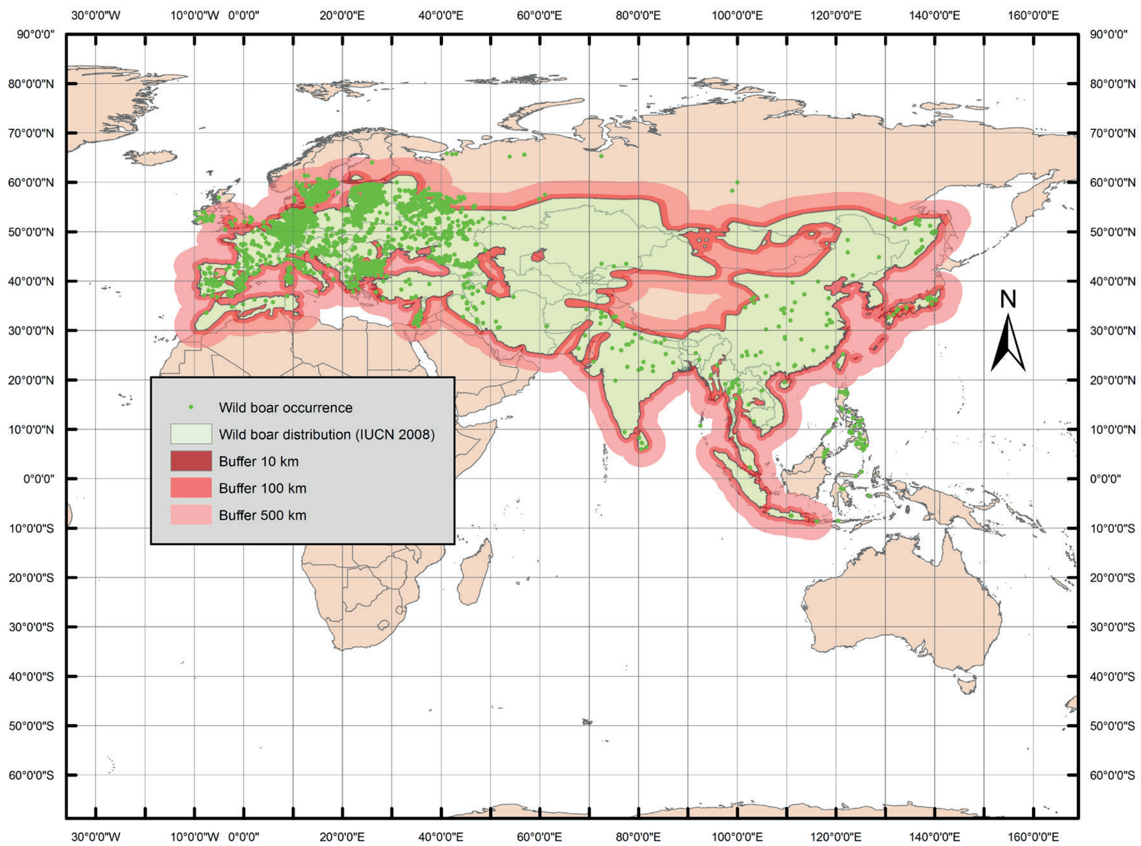


Fig. 1. Map showing wild boar distribution in Eurasia according to the International Union for Conservation of Nature (IUCN) area and buffers at 10, 100, 500, and >500 km

of 65° N (Fig. 1). Wild boar is also present at latitudes below 66° N, and permanently inhabited reproduction areas occur between 62° and 63° N; these areas are considered part of the geographical range of the species (Danilov et al. 2003; Markov et al. 2018). The natural expansion of the wild boar from the southwest to the north of Western Siberia may be due to the decrease in snow cover because of climate change, which can increase the availability of food (Bieber & Ruf 2005; Geisser and Reyer 2005; Melis et al. 2006; Spitz 1999; Powell 2004; Apollonio et al. 2010; Markov et al. 2018).

In Spain, wild boar occurrences outside the IUCN distribution area were found within the 100 km buffer. These expansion areas are distributed throughout the center, center-east, and south of the country, and they correspond mainly to valleys and river depressions used for agriculture. In Spain, as in other European countries, the growth of this species has been related to the following: reforestation (Servanty et al. 2011); disappearance of traditional agriculture and reduction of forestry activities; increase in shelter areas (Sáenz de Buruaga 1995); and increase in areas dedicated to certain crops, particularly corn, similarly to what has been observed in Poland, Sweden, Germany, and Switzerland (Baettig 1985; Fruzinski 1995; Keuling et al. 2009; Saïd et al. 2011). Another important factor is the absence of predators (Massei and Genov 2004; Fernández-Llario 2017).

In Sweden, wild boar occurrences outside the IUCN distribution lie mainly in the 100 and 500 km buffers, both corresponding to forests and agricultural areas (Fig. 1). Our results are in line with a study conducted in the southern and central part of Sweden in 2015, which showed an increase in the distribution and abundance of wild boar (Malmsten et al. 2017). This may be due to the high food availability in the wild and to environmental conditions favourable for reproduction.

In Belgium, the IUCN map covers only 2% of the country's surface. Wild boar occurrences are located in the 100 km buffer, and the distribution spreads throughout the country (Fig. 1). This country has recently been affected by ASF in wild boars, which creates risk of dispersion to neighbouring countries with large domestic pig production (OIE 2018). Our results suggest the need to include Belgium in analyses of wild boar distribution in order to improve population management, as well as improve efforts at ASF surveillance and control.

In Ireland, which is not included the IUCN distribution area, wild boar presences are located in the >500 km buffer, with distribution throughout the country, even at the southern border of Northern Ireland (Fig. 1) (NBDC 2018). Despite control programs implemented since 2008, the population of wild boar continues to expand (NBDC 2018).

Three other countries that are not included in the IUCN map but that show wild boar presence in the >500 km buffer in our analysis are the Philippines, Malaysia, and Indonesia, except the islands of Sumatra and Java (Fig. 1). In Southeast Asia, more taxa of wild pig populations exist than in other areas, including *Sus cebifrons* (Philippines), *Sus celebensis* (Sulawesi), and *Sus barbatus* (Indonesia and Malaysia). The most common taxon is the Eurasian boar (*Sus scrofa*) (Lucchini et al. 2005).

Traditionally, wild boar distribution has been described between 0 and 2400 m meters above sea level, although evidence of wild boar has been found in subalpine meadows above 2400 m in particular periods of the year (Markov 2018). Our results confirm the altitude range of 0-2400 m, but additionally show that the boar is present at altitudes below 0 and over 2400 m (Fig. 2). The highest altitudes above sea level showing the presence of wild boar are found in China and Southeast Tibet shrublands and meadows (3536 m), followed by altitudes above 2500 m in Iran (2728 m, Zagros mountains forest steppe), Philippines (2593 m, Mindalo montane rain forest), and Switzerland (2506 m, Alps conifers and mixed forest). Wild boar at altitudes below 0 m are found in Israel along the tectonic depression traversed by the Jordan River (-395 m) and in Russia. These results confirm wild boar presence in continental and coastal geographic depressions, as well as in great mountain ranges and high plateaus.

Maximum and minimum values for climatic variables in areas containing wild boar did not vary significantly between the IUCN area and the buffers (Table 2). The only exception

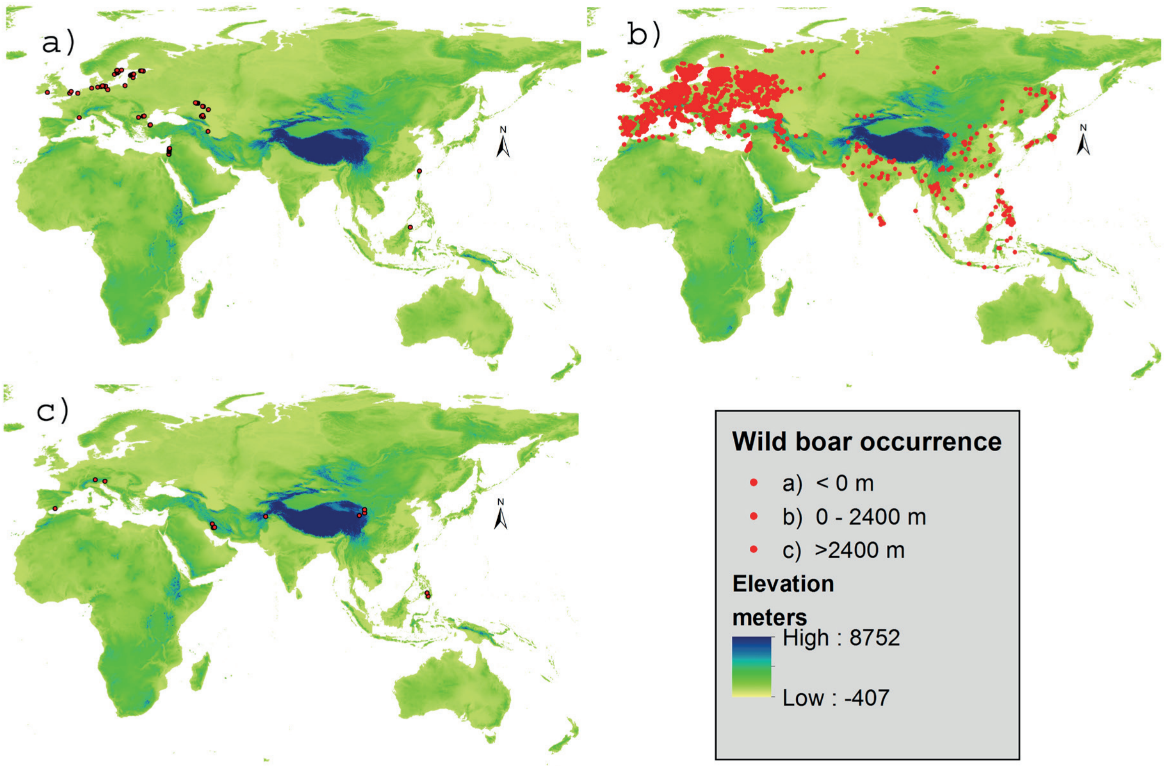


Fig. 2. Distribution of wild boar occurrences in Eurasia at altitudes lower than 0 m (a), between 0-2400 m (b), and higher than 2400 m (c)

was the Eurasia zone, mainly the >500 km buffer, where the minimum temperature in the coldest month was lower in the IUCN area (-7 °C) than in the >500 km buffer (19.3 °C), as was the maximum temperature in the warmest month (23.2 °C vs. 33.2 °C). These differences confirm the ability of wild boar to adapt to living outside the traditional IUCN area, including in northern Eurasia, Western Siberia (Markov et al. 2018), Southeast Asia (Lucchini et al. 2005), and desert zones such as Iran (Parchizadeh 2017; Rezaei et al. 2018). Climatic factors can also influence population density over time by affecting the availability of food and shelter as well as reproductive potential (Fig. 3) (Ehrlén and Morris 2015).

A similar trend was observed for QAH inside and outside the IUCN area (Table 1). Wild boar was present mostly in natural areas (QAH 1.5-2.0) accounting for 70% of the territory inside the IUCN area and 60% of the territory outside the IUCN area, and in agroforestry areas (QAH 1.75) accounting for 17% and 15% of the respective territories. Natural habitats suitable for wild boar (mainly QAH 1.5) have become more available due to warmer summers and milder winters in traditionally colder areas (Fernández-Llario 2017). This factor acts directly on the landscape connectivity, enhancing the migration of the species to areas such as the coniferous forests of northern Eurasia, and the Taiheiyo montane deciduous and evergreen forests in Japan (Bascompte and Solé 1996). Milder winters with little snow facilitate food access with lower energy expenditure, reducing mortality (Bieber and Ruf 2005; Rossi et al. 2011). Moreover, the increase in forest cover in Europe due to reforestation may have further facilitated the dispersal of wild boar (Servanty et al. 2011). The remaining wild boar occurrences are associated mostly with monocultures (QAH 1). Several factors have favoured the spread of wild boar and wild pigs (Lucchini et al. 2005) to previously unoccupied areas (Rosvold et al. 2008; Veeroja and Männil 2014; Massei et al. 2015): changes in land use (Servanty et al. 2011); shelter offered by some crops such as corn, sunflower, rice, wheat or rapeseed (Herrero et al. 2006; Keuling et al. 2009); and the decrease in predator populations in some regions (Fernández-Llario 2017, Jerina et al. 2014; Markov et al. 2019). As a result, wild boar are more abundant in European croplands (e.g. QAH 1 in the 500 km buffer), as well as in Southeast Asia.

The main ecoregions associated with wild boar presence are warm (present in 65% of boar occurrences) and snowy (present in 85%). Colonization of new arid ecoregions (27%) and new equatorial ecoregions (93%) is observed in 10 km

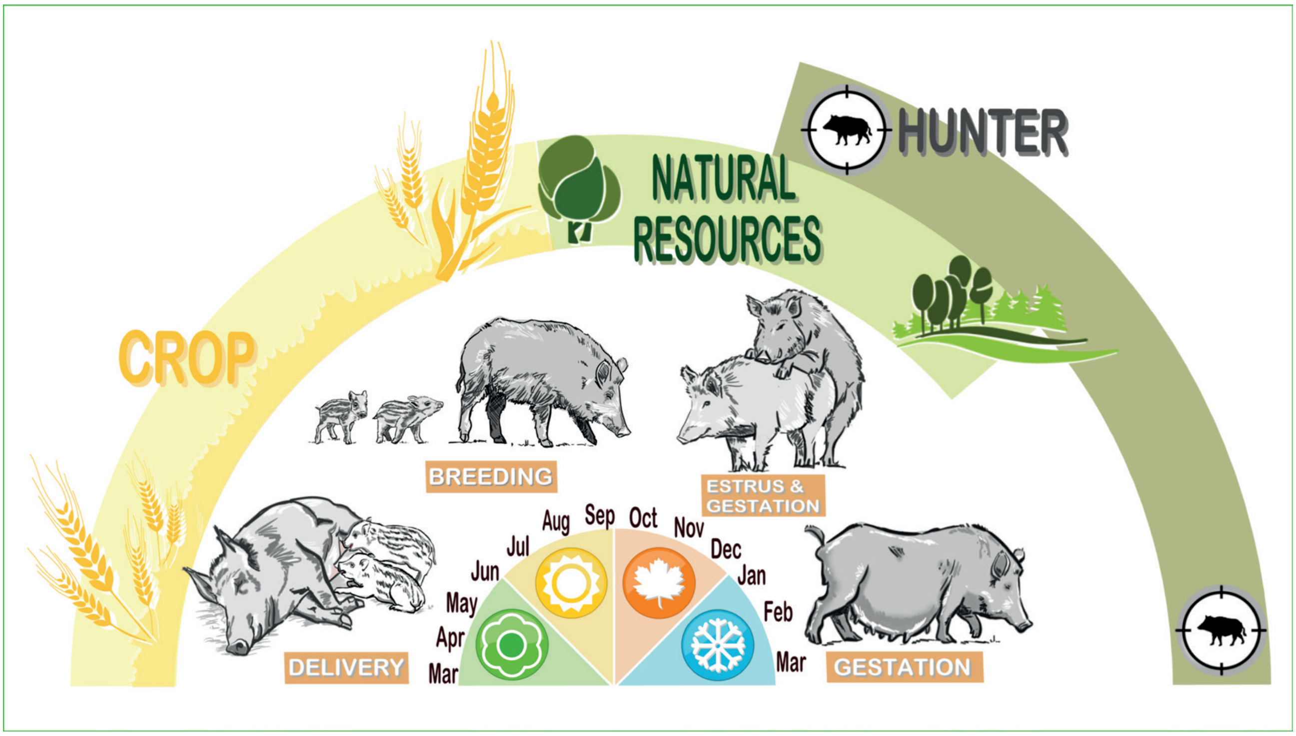


Fig. 3. Biological cycle of wild boar associated with food and shelter availability and hunting season

Table 1. Wild boar presence by quality of available habitats (QAH) or ecoregion inside and outside the International Union for Conservation of Nature (IUCN) area. Results are expressed as percentages of the total number of occurrences

Wild boar presence vs. IUCN area	Quality of available habitats (QAH)							Total occurrence (n)
	2	1.75	1.5	1	0.5	0.1	0	
IUCN+Buffer	61%	17.0%	7.0%	10.9%	0.2%	1.8%	2.2%	34,272
IUCN	61.8%	16.9%	6.9%	10.8%	0.2%	1.4%	2.1%	31,737
10 km	44.9%	26.4%	7.2%	16.3%	0.3%	2.8%	2.2%	363
100 km	52.3%	16.4%	9.2%	9.7%	0%	9.1%	3.3%	1,647
500 km	39.5%	11.4%	11.9%	27.6%	0%	2.7%	7%	185
>500 km	52.9%	25.0%	2.6%	16.8%	2.6%	0%	0%	340

	Ecoregion (according to Kottek et al, 2006)					
Wild boar presence vs. IUCN area	Equatorial	Arid	Warm	Snow	Polar	Total occurrence (n)
IUCN+Buffer	1.4	3.3	57.7	37.5	0.01	34272
IUCN	0.5	3	56.8	39.7	0.01	31737
10 km	0	27.3	69.7	3.0	0	363
100 km	0.3	6.4	83.2	10.1	0	1647
500 km	0	0	64.9	35.1	0	185
>500 km	93.5	0	4.1	2.4	0	340

and >500 km buffers (Table 1). The latter is especially true in Southeast Asia (>500 km buffer) (Fig. 4).

Our analysis does have some limitations. Predicting the distribution of wild boar, or of any wild species, reflects sampling bias because of sampling variations across the study area (Ferrier 2002; Varela et al. 2011, 2014). More georeferenced wild boar records are required for latitudinal and longitudinal gradients, which may become available after several years (Varela et al. 2011) through initiatives including GBIF, EUROBOAR, Enetwild, and the Food and Agriculture Organization.

CONCLUSIONS

Permanent populations of wild boar exist outside the IUCN distribution area, with nearly 70% of animal occurrences lying within the first 100 km of the distribution limit. Wild boar have also been found in the >500 km buffer. Wild boar presence extends below sea level on continental and coastal geographical depressions and to altitudes higher than 2400 m.a.s.l. on great mountain ranges and highlands; the two types of "extreme" areas show similar environmental characteristics and habitat quality. Our results further suggest that the wild boar has entered new biotopes and ecoregions,

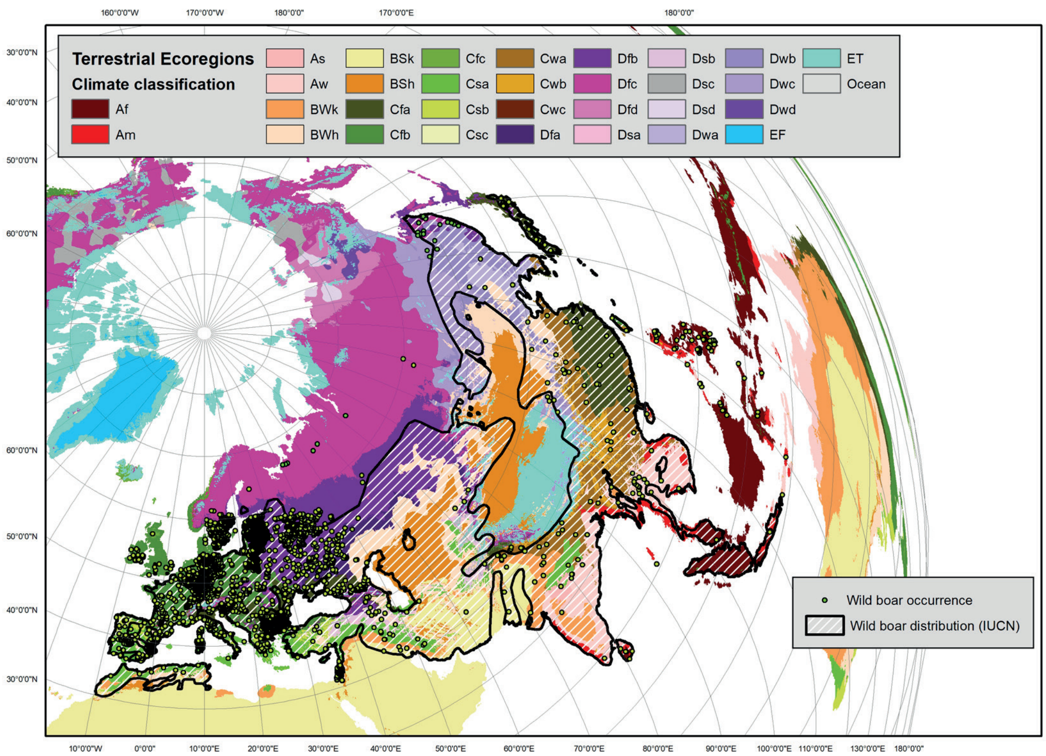


Fig. 4. Map of wild boar occurrences (dots) and wild boar distribution (lined area), by terrestrial ecoregions. Distribution was taken from the International Union for Conservation of Nature [Oliver and Leuss, 2008]. Ecoregions were defined based on the Köppen-Geiger climate classification [Kottek et al, 2006]

Abbreviations: Af (equatorial, fully humid), Am (equatorial, monsoonal), As (equatorial, summer dry), Aw (equatorial, winter dry), BWk (arid, desert, cold air), BWh (arid, desert, hot air), BSh (arid, steppe, hot air), Cfa (warm temperate, fully humid, hot summer), Cfb (warm temperate, fully humid, warm summer), Cfc (warm temperate, fully humid, cool summer), Csa (warm temperate, summer dry, hot summer), Csb (warm temperate, summer dry, warm summer), Csc (warm temperate, summer dry, cool summer), Cwa (warm temperate, winter dry, hot summer), Cwb (warm temperate, winter dry, warm summer), Cwc (warm temperate, winter dry, cool summer), Dfa (snow, fully humid, hot summer), Dfb (snow, fully humid, warm summer), Dfc (snow, fully humid, cool summer), Dfd (snow, fully humid, extremely continental), Dsa (snow, steppe, hot summer), Dsb (snow, steppe, warm summer), Dsc (snow, steppe, cool summer), Dsd (snow, steppe, extremely continental), Dwa (snow, desert, hot summer), Dwb (snow, desert, warm summer), Dwc (snow, desert, cool summer), Dwd (snow, desert, extremely continental), EF (polar, polar frost), ET (polar, polar tundra).

Table 2. Environmental variables in areas of wild board presence inside and outside the International Union for Conservation of Nature (IUCN) area (P=percentile)

	Minimum temperature during the coldest month (°C)						
Wild boar presence vs. IUCN area	min	P5	P25	median	P75	P95	max
IUCN+Buffer	-33.9	-10.2	-8.9	-7	-2.4	6.9	23.3
IUCN	-26.3	-10	-9	-7.2	-2.7	6.5	22.7
Buffer 10 km	-14	-5.2	-3.8	-2.3	0.7	3.8	10
Buffer 100 km	-19.3	-8.3	-7.4	-3.7	-2.2	2.7	19.4
Buffer 500 km	-18.9	-14.7	-13.9	-6.8	4.3	2.7	3.4
Buffer >500 km	-33.9	6.9	18.2	19.3	20.6	22.7	23.3
	Maximum temperature during the warmest month (°C)						
Wild boar presence vs. IUCN area	min	P5	P25	median	P75	P95	max
IUCN+Buffer	9.2	20.8	21.6	23.3	27.1	31.8	45.5
IUCN	9.2	20.9	21.6	23.2	26.6	31.4	45.5
Buffer 10 km	19.8	21.6	27.8	29	29.7	31	42.4
Buffer 100 km	19.5	20.5	20.9	27.8	29.2	30.8	40.9
Buffer 500 km	17.6	18.5	19.7	20.9	23.1	23.9	29.6
Buffer >500 km	17.4	19.1	31.4	32.5	33.2	33.3	34
	Annual precipitation (mm)						
Wild boar presence vs. IUCN area	min	P5	P25	median	P75	P95	max
IUCN+Buffer	53	469	566	624	658	991	5,495
IUCN	57	472	565	626	658	897	5,495
Buffer 10 km	132	386	498	599	612	688	1,596
Buffer 100 km	53	409	568	596	631	864	1,804
Buffer 500 km	352	526	568	605	638	1069	1,227
Buffer >500 km	355	1,520	2,159	2,163	2,529.5	2,817	3,399
	Precipitation during the driest month (mm)						
Wild boar presence vs. IUCN area	min	P5	P25	median	P75	P95	max
IUCN+Buffer	0	1	25	28	33	48	192
IUCN	0	0	25	28	32	45	121
Buffer 10 km	0	7	15	30	38	40	68
Buffer 100 km	0	8	26	28	30	41	76
Buffer 500 km	11	24	26	28	35	64	74
Buffer >500 km	8	25	48	51	64.5	134	192

such as the equatorial region, where its presence is mainly associated with the large Asian monocultures. The present study may more accurately define wild boar distribution in Eurasia than the conventional IUCN analysis, thereby helping to develop new models of species distribution, examine habitat selection, and identify permanent populations.

ACKNOWLEDGEMENTS

European Union's Horizon 2020 Research and Innovation Program under Grant Agreement 773830 (OHEJP), AT2015-002 and RTA2015-00033-C02-01.

REFERENCES

Apollonio M., Andersen R. and Putman R. (2010). European Ungulates and Their Management in the 21st Century. New York: Cambridge University Press.

Baettig J.L. and Sáez-Royuela C. (1985). L'evolution demographique du sanglier (Sus scrofa) en Espagne. Mammalia, 49, 195-202. (in French).

Bascompte J. and Solé R.V. (1996). Habitat fragmentation and extinction thresholds in spatially explicit models. J. Anim. Ecol., 65, 465-473.

Bieber C. and Ruf T. (2005). Population dynamics in wild boar Sus scrofa: ecology, elasticity of growth rate and implications for the management of pulsed resource consumers. J. Appl. Ecol., 42, 1203-1213.

Bosch J., Iglesias I., Muñoz M.J. and de la Torre A. (2016). A cartographic tool for managing African swine fever in Eurasia: mapping wild boar distribution based on the habitats. Transbound. Emerg. Dis., 64, 1720-1733.

Danilov P., Zimin V. and Ivanter E. (2003). Changes in the fauna and distribution dynamics of terrestrial vertebrates in the European north of Russia. Finnish Environ, 485, pp.175–181.

Ehrlén J. and Morris W.F. (2015). Predicting changes in the distribution and abundance of species under environmental change. Ecology letters, 18, 303-314.

European Court of Auditors, (2016). Special report no 06/2016: Eradication, control and monitoring programmes to contain animal diseases. Available at: https://www.eca.europa.eu/en/Pages/DocItem.aspx?did=36176 [Accessed 30 April 2019].

European Food Safety Authority (EFSA). (2018). Epidemiological analyses of African swine fever in the European Union (November 2017 until November 2018). EJ EFSA journal. doi: 10.2903/j.efsa.2018.5494.

Fernández-Llario P. (2017). Jabalí–Sus scrofa. In: Carrascal L.M., Salvador A. and Luque-Larena J.J., eds., Enciclopedia Virtual de los Vertebrados Españoles. Madrid: Museo Nacional de Ciencias Naturales. Available at: http://www.vertebradosibericos.org/ [Accessed 30 April 2019].

Ferrier S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? Syst. Biol., 51, 331-363.

Fruzinski B. (1995). Situation of wild boar populations in Western Poland. J. Mount. Ecol., 3, 186-187.

Geisser H. and Reyer H.U. (2005). The influence of food and temperature on population density of wild boar Sus scrofa in the Thurgau (Switzerland). J. Zool., 267, 89-96.

Herrero J., García–Serrano A., Couto S., Ortuño V.M. and García-González R. (2006). Diet of wild boar Sus scrofa L. and crop damage in an intensive agroecosystem. Eur. J. Willd Res., 52, 245-250.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G. and Jarvis A. (2005). Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol., 25(15), 1965-1978.

Iglesias I., Montes F., Martínez-Avilés M., Pérez A., Gogin A., Kolbasov D. and de la Torre A. (2018). Spatio-temporal kriging analysis to identify the role of wild boar in the spread of African swine fever in the Russian Federation. Spatial Statistics. DOI: 10.1016/j.spasta.2018.07.002.

Jerina K., Pokorny B. and Stergar M. (2014). First evidence of long–distance dispersal of adult female wild boar (Sus scrofa) with piglets. Eur. J. Willd Res., 60, 367-370.

Keuling O., Stier N. and Roth M. (2009). Commuting, shifting or remaining? Different spatial utilisation patterns of wild boar Sus scrofa L. in forest and field crops during summer. Mamm. Biol., 74, pp.145-152.

Keuling O., Lauterbach K., Stier N. And Roth M. (2010). Hunter feedback of individually marked wild boar Sus scrofa L.: dispersal and efficiency of hunting in northeastern Germany. Eur. J. Willd Res., 56, 159-167.

Lucchini V., Meijaard E., Diong C.H., Groves C.P. and Randi E. (2005). New phylogenetic perspectives among species of South-east Asian wild pig (Sus sp.) based on mtDNA sequences and morphometric data. J. Zool., 266, 25-35.

Malmsten A., Jansson G., Lundeheim N. and Dalin A.M. (2017). The reproductive pattern and potential of free ranging female wild boars (Sus scrofa) in Sweden. Acta Vet. Scand., 59, 52.

Markov N., Pankova N. and Filippov I. (2018). Wild boar (Sus scrofa L.) in the north of Western Siberia: history of expansion and modern distribution. Mamm. Res., 64, 1-9.

Markov N.I., Pankova N. L., Vasina A. L. and Pogodin N. L. (2019). Specific Features of Wild Boar, Sus scrofa, Rooting Activity at the Northern Boundary of Its Range in Western Siberia. Russian Journal of Ecology, 49(6), 584-587.

Massei G. and Genov P. 2004. The environmental impact of wild boar. Galemys, 16, 135-145.

Massei G., Kindberg J., Licoppe A., Gačić D., Šprem N., Kamler J., Baubet E., Hohmann U., Monaco A., Ozoliņš J., Cellina S., Podgórski T., Fonseca C., Markov N., Pokorny B., Rosell C. and Náhlik A. (2015). Wild boar populations up, numbers of hunters down? A review of trends and implications for Europe. Pest Manag. Sci., 71, 492-500.

Melis C., Szafranska P., Jędrzejewska B. and Bartoń K. (2006). Biogeographical variation in the population density of wild boar (Sus scrofa) in western Eurasia. J. Biogeogr., 33, 803-811.

National Biodiversity Data Centre, Ireland. (2018). Wild Boar (Sus scrofa), terrestrial distribution map. Available at: https://maps.biodiversityireland.ie/Species/TerrestrialDistributionMapPrintSize/119291. [Accessed 15 September 2018].

OIE, World Organisation for Animal Health. (2018). WAHID (World Animal Health Information System). Available at: http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/reportarchive. [Accessed 30 April 2019].

Oliver W. and Leus K. (2008). Sus scrofa. The IUCN Red List of Threatened Species 2008: e.T41775A10559847. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T41775A10559847. Available at: www.iucnredlist.org. [Accessed 30 April 2019].

Parchizadeh J. (2017). Spatiotemporal distribution of large-and medium-sized mammals and humans in the Lar Protected Area, Iran. Willd Res., 44, 400-406.

Powell D.M. (2004). Pigs (Suidae). Mammals IV. In: D.G. Kleiman,V. Geist, and M. McDade, eds. Grzimek´s animal life encyclopedia, vol. 15. 2nd ed. Michigan: Gale Research Inc.

Rezaei S., Naderi S. and Karami P. (2018). An investigation on attitude of human local communities to wild boar (Sus scrofa, Linnaeus, 1758) in Jasb wildlife refuge, Markazi province, Iran. Available at: https://www.researchgate.net/publication/328424442_An_investigation_on_attitude_of_human_local_communities_to_wild_boar_Sus_scrofa_Linnaeus_1758_in_Jasb_wildlife_refuge_Markazi_province_Iran. [Accessed 30 April 2019]. (in Arabic with English summary).

Rossi S., Toigo C., Hars J., Pol F., Hamann J.L., Depner K. and Le Potier M.F. (2011). New insights on the management of wildlife diseases using mult–state recapture models: the case of classical swine fever in wild boar. PLoS ONE, 6, e24257.

Rosvold J. and Andersen R. (2008). Wild boar in Norway – is climate a limiting factor? – NTNU Vitenskapsmuseet Zoologisk Rapport, 1, 1-23.

Rubel F. and Kottek M. (2010). Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. Meteorologische Zeitschrift, 19, 135-141.

Sáenz de Buruaga M. (1995). Alimentación del jabalí (Sus scrofa castilianus) en el norte de España. Ecología, 9, 367-386. (in Spanish).

Saïd S., Tolon V., Brandt S. and Baubet E. (2011). Sex effect on habitat selection in response to hunting disturbance: the study of wild boar. Eur. J. Willd. Res., 58(1), 107-115.

Sales L.P., Ribeiro B.R., Hayward M.W., Paglia A, Passamani M. and Loyola R. (2017). Niche conservatism and the invasive potential of the wild boar. J. Anim. Ecol., 86, 1214-1223.

Servanty S., Gaillard J.M., Ronchi F., Focardi S., Baubet E. and Gimenez O. (2011). Influence of harvesting pressure on demographic tactics: implications for wildlife management. J. Appl. Ecol., 48, 835-843.

Spitz F. (1999). Sus scrofa (Linnaeus, 1758). In: T. Mitchell-Jones, G. Amori, W. Bogdanowicz, B. Krystufek, P.J.H. Reijnders, F. Spitzenverger, M. Stubbe, J. Thissen, V. Vohralíki, and J. Zima, eds., The Atlas of European Mammals, 2nd. Ed. London: T & AD Poyser.

Varela S., Lobo J.M. and Hortal J. (2011). Using species distribution models in paleobiogeography: a matter of data, predictors and concepts. Palaeogeogr. Palaeoclimatol. Palaeoecol., 310, 451-463.

Varela S., Anderson R.P., García-Valdés R. and Fernández-González F. (2014). Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. Ecography, 37, 1084-1091.

Veeroja R. and Männil P. (2014). Population development and reproduction of wild boar (Sus scrofa) in Estonia. Wild. Biol. Pract., 10, 17-21.

Wiens J.J. and Graham C.H. (2005). Niche conservatism: integrating evolution, ecology, and conservation biology. Annu. Rev. Ecol. Evol. S., 36, 519-539.

Received on April 30th, 2019 Accepted on January 21th, 2020

Svetlana M. Malkhazova[1*], Polina V. Pestina[1], Vladimir S. Tikunov[1]
[1] Lomonosov Moscow State University, Moscow, Russia.
***Corresponding author:** sveta_geo@mail.ru

EMERGING AND RE-EMERGING NATURAL FOCAL DISEASES OF EUROPEAN RUSSIA (TYPOLOGICAL CLASSIFICATION OF NOSOLOGICAL PROFILES AND DYNAMICS OF INCIDENCE)

ABSTRACT. This study considers an automated typological classification version by using the extensive factual material in analysis of emerging and re-emerging natural focal diseases of European Russia.

The typological classification of nosological profile (a set of diseases) and the incidence dynamics for five nosological forms (hemorrhagic fever with renal syndrome, ixodic tick-borne borreliosis, tick-borne encephalitis, tularemia, and leptospirosis) was created using the formal methods of mathematical-cartographical modeling. This classification of the incidence in 1997–2015 yielded five types of the nosological profiles. These types vary by years, which is associated with the dependence of the incidence on climatic conditions in each specific year and on extent of deratization and preventive measures. The results obtained can be used to forecast potential epidemiological outbreaks and to develop targeted and appropriate for each region measures.

KEY WORDS: mathematical-cartographic modeling, typological classification, emerging and re-emerging diseases, nosological profiles, dynamics of incidence

CITATION: Svetlana M. Malkhazova, Polina V. Pestina, Vladimir S. Tikunov (2020) Emerging And Re-Emerging Natural Focal Diseases Of European Russia (Typological Classification Of Nosological Profiles And Dynamics Of Incidence). Geography, Environment, Sustainability, Vol.13, No 1, p. 115-127
DOI-10.24057/2071-9388-2019-61

INTRODUCTION

The classification of phenomena and processes is inherent, in one way or another, in all scientific fields and represent one of the main research methods. The problem of automated creation of classifications and various modifications of methods has increasingly gained importance in modern geoinformatics. A fairly large number of classification methods allow for the construction of appropriate formal algorithms.

A system of such algorithms (called "automated classification algorithms") has been already widely adopted in geoinformatics (Kapralov et al. 2010). The algorithms of automated classification methods are currently included in a number of GIS packages. However, these algorithms are often created without specific requirements of individual scientific areas where they can be applied. Such versatility has a positive value, allowing for the use of algorithms already developed in applied mathematicians or in other fields. However, this may create difficulties associated with underestimation of specific needs in a particular area.

The use of classifications in geography, environmental science, and medical geography has a long history. Regionalization, typology, and evaluation of complexes, often with rendering of the obtained results on maps, represent, as a rule, not only the methods, but also the research objectives (Feldman 1977; Keller 1992; Tikunov 1997; Malkhazova 2001; Kurolap et al., 2015; Malkhazova et al., 2017; and others).

In classification of geographical complexes, one encounters various challenges; some of them are typical of many fields while others are only characteristic of individual scientific areas, for example, medical geography. One of the problems is that the existing algorithms are usually associated with statistical characteristics that only indirectly or not at all represent spatial position of the phenomena. At the present time, there are no reliable methods of quantitative assessment of the relative importance of geographical location. This necessitates an additional consideration of the territorial aspect of the modeled phenomena. Currently, there have been only few attempts to modify statistical processing of indicators to account for spatial position (Trofimov 1985). The statements of geographical problems and descriptions of phenomena allow for some subjectivism or double interpretation, at least at the present stage of research. Formalized multidimensional classification algorithms may not correspond exactly to the level of formalization and accuracy of the tasks themselves, which sometimes leads to the results that do not correspond to the essence and meaning of the studied phenomena. This problem may be addressed by using the concept of fuzzy sets and development of classification methods based on it (Zadeh 1965; Tikunov 1989).

The issue of the optimal system of base parameters is characteristic of many classifications; such a system must comprehensively (to the extent dictated by the essence of the task) describe the studied phenomena. Consideration of all available data can lead to their redundancy. The data should not duplicate each other, be derived from one another, etc. Otherwise, they may obscure the most significant properties and lead to a distortion of the final result. As a rule, it is difficult to find a specific criterion that determines the use of a given parameter of a geographical complex. A deep knowledge of the studied object is critical and allows establishing the optimal set of parameters. Alternatively, it may be possible to experimentally check the extent to which the input data influence the result.

Another difficulty is associated with the varying degrees of parameters' significance. Some of them are so important that it is not possible to simulate phenomena without their consideration, while others only complement and clarify the basic system. This requires "weighting" the parameters and assessing the degree of their influence on the final results. However, the definition of "weights" is an independent, complex, and largely unsolved, at the present time, task. There have been attempts to justify the system of "weights" by expert surveys of researchers on specific topics (Kapralov et al. 2010).

Most of the classification problems in geography are associated with quantitative and/or qualitative (for example, those that come from some other classifications) parameters. This imposes certain restrictions on the use of various methods of multidimensional classification. Since a significant part of the data that are taken into account in classifications is of a qualitative nature, the algorithms must be able to work with non-numeric characteristics. The creation of systems of such algorithms is a progressing field of research on multidimensional statistical analysis.

In all geographical studies, complexes should be treated as spatiotemporal formations. In this regard, the role of temporal and content-rich characteristics of the studied complexes in the classifications used in natural and social sciences (e.g., biology, geology, economics, history, etc.) should not be underestimated.

The present study considers the existing difficulties and the provisions outlined above and tests an automated typological classification version by using the extensive factual material in analysis of emerging and re-emerging natural focal diseases of European Russia. The problem of emerging and re-emerging diseases arose as a major world health care problem at the end of the XXth century. The World Health Organization (WHO) states that emerging and re-emerging infections are those that have recently appeared or whose incidence has increased over the past two decades or has a potential to increase in the near future. This term includes both diseases that spread to new territories and returning infectious diseases (Morse 1995; Jones et al. 2008; Malkhazova et al. 2016; Malkhazova and Mironova 2017).

The typological classification of nosological profile (a set of diseases) and the incidence dynamics for five nosological forms was created using the formal methods of mathematico-cartographical modeling (Tikunov 1997). A purposefully configured calculation algorithm was created specifically for this study. The advantage of the developed classification is its ability to analyze patterns of incidence variation as a result of grouping the administrative-territorial units (TUs) into taxa with homogeneous dynamics patterns. Such a typology allows studying not only the individual series, but also their groups, which are less susceptible to random fluctuations.

MATERIALS AND METHODS

The goal of the study presented herein was to use the typological classification method discussed in (Tikunov 1983) to obtain homogeneous groups for 48 TUs of European Russia, using the data on the 1997–2015 (19 yrs) incidence statistics on five emerging and re-emerging (herein-thereafter, referred to as "emerging") natural focal infections: hemorrhagic fever with renal syndrome (HFRS), ixodic tick-borne borreliosis (ITBB) (i.e., lyme disease), tick-borne encephalitis (TBE), tularemia, and leptospirosis (Malkhazova et.al. 2019).. The calculations produced a series of maps which characterize the types of the nosological profiles for individual years (1997, 2005, 2015) and for the entire studied period (1997–2015), as well as a series of maps of incidence dynamics types for each nosological form for the considered period. The arithmetic mean incidence values for the nosological forms are shown on the maps of the nosological profiles as bar-charts, where the y-axis is the incidence and the x-axis is a nosological form, and on the maps of the incidence dynamics types — as line-charts of the incidence by years.

In the applied typological algorithm, the entire set of parameters for any TU can be written as a M-dimensional vector-string $[x_1, x_2, x_3, \ldots x_m]$, and the entire set of TUs (with the number of N equal to 48 in this study) is denoted by $X=\{x_1,\ldots,x_N\}$, where x_i- ^{i} is -th TU. Spatially, TUs may be described by various parameters and metrics, the main of which are those that allow calculating the distance between TUs (i.e., the coefficients of "similarity" or "difference" between them). In addition to geographical space, the studied set of TUs also exists in the M-parameter space (19) where TUs lose their geographical attributes and, regardless of their original nature, become *M*-dimensional points. The result of this (Kapralov et al. 2010) is:

1. Representation of the original TUs in the form of "TU—parameter" matrix, reflecting the measurements of M-parameters for *N* TUs with *N* rows and *M* columns:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x^N \end{pmatrix} = (x^{(1)}, \ldots, x^{(M)}) = \begin{pmatrix} x_1^{(1)} \cdots x_1^{(j)} \cdots x_1^{(M)} \\ \cdots \; \cdots \; \cdots \; \cdots \; \cdots \\ x_{N-1}^{(1)} \cdots x_{N-1}^{(j)} \cdots \; x_{N-1}^{(M)} \\ x_N^{(1)} \cdots x_N^{(J)} \cdots x_N^{(M)} \end{pmatrix},$$

where

$x_i=(x_i^{(1)},\ldots,x_i^{(M)})$- ^{i} -th TU in -dimensional space of parameters,
$x^{(j)}$ - ^{j} -th parameter, $x^{(j)}=(x_1^{(j)},\ldots, x_N^{(j)})$,
$x_i^{(j)}$- value of ^{j} -th parameter of ^{i} -th TU,
$i \in \{1,\ldots,N\}$ $j \in \{1,\ldots,M\}$.

2. Representation of the original TUs in the form of "TU—TU" matrix, reflecting the result of the TUs' comparison in the parameters' or geographical space with *N* rows and columns:

$$\underset{\in}{A} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1N} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{N-11} & \in & a_{N-1j} & \cdots & a_{N-1N} \\ a_{N1} & \cdots & a_{Nj} & \cdots & a_{NN} \end{pmatrix},$$

where

a_{ij} – result of comparison of ^{i} -th and ^{j}-th TU i,j $\in \{1,\ldots,N\}$.

As a rule, a_{ij} means the measure of difference (or similarity) of the TUs. In the case of interpretation of a_{ij} as a measure of difference, the matrix *A* is symmetric, with zeros on the main diagonal. The transition from "TU—parameter" matrix to "TU—TU" matrix is performed by setting the metric *d* (the distance between the TUs).

In this case, the TUs are compared with each other by the method of pairwise comparison.

The next stage of the TUs' classification is their pretreatment, including normalization, weighing, and dimensionality reduction.

The normalization was carried out according to the variance and mathematical expectation.

The purpose of this normalization is to bring each incidence parameter for five types of infections, each for 19 years, to a standard form (as a result, the mathematical expectation of any parameter becomes zero, and the variance becomes one).

$$Let \; \bar{x}^{(l)} = \frac{1}{N}\sum_{i=1}^{N} x_j^{(j)}$$

be the assessment of mathematical expectation of j -th parameter,

$$var(x^{(j)}) = \frac{1}{N}\sum_{i=1}^{N}(x_j^{(j)} - \bar{x}^{(j)})$$

be the assessment of variance of j -th parameter.

Then the normalization means the recalculation

$$x_i^{(j)} = \frac{x_i^{(j)} - \bar{x}^{(j)}}{\sqrt{varx^{(j)}}} \text{A} j \in \{1,...,M\}, i \in \{1,...,N\},$$

i.e.,

$$\Delta = \bar{x}^{(j)}, \Delta_2 = \sqrt{var(x^{(j)})}$$

The next step includes the application of the principal component method. Principal component analysis, or component analysis, is one of the most frequently used methods for dimensionality reduction. This method solves the problem of finding, using the existing system of parameters that describe the TUs, a new system with the following properties:

– the new system's parameters are the linear combinations of the original system's parameters;

– the number of parameters in the new system does not generally exceed (and in practice is always less) the number of parameters in the original system;

– the new system's parameters are orthogonal, i.e. uncorrelated;

– the new system's parameters are ordered in descending order of variance;

– the new system's parameters carry as much information (or a predetermined percentage of information, for example, 90%) about the variability of objects, as the original parameters. "Information" refers to the variance of parameters.

The principal component method is applied to correct the original parameters' space distorted by mutual correlations, to reduce the volume of the stored data without losing a significant part of the information on the TUs, to visualize the TUs in the parameters' space (which is achieved, for example, by rendering the TUs as points on the plane of the first two main components), and to identify the latent (i.e., hidden, not clearly observable) parameters that reflect the essence of processes or phenomena.

In a matrix form, the result of the principal component method is written as

$Z=XL$ or $Z_{N\cdot m}=X_{N\cdot M}L_{M\cdot m}$, where

M – the number of the original parameters;

m – the number of the principle components obtained, $m \le M$;

$Z=Z_{N\cdot m}=(z^{(1)},...,z^{(m)})$ – matrix of the new parameters (as in the original matrix, the parameters are arranged in columns);

$X=X_{N*M}=(x_j^{(1)},\ldots,x_j^{(M)})$ - the original matrix "TU—parameters";

$L=L_{M\cdot m}=(l^{(1)},...,l^{(m)})$ – the calculated matrix of the componential loads.

The most simple is the geometric interpretation of the principal component method. In a multidimensional parameters' space, the TUs are considered as points, whose cloud's geometrical arrangement, in the case of the normal distribution, resembles a M-dimensional ellipsoid. The main axes of the imaginary ellipsoid are treated as the new parameters, sorted in the descending order of the TUs' dispersions along the axes.

Of course, component analysis is not the only method of dimensionality reduction. As examples of other common methods, we should note factor analysis, multidimensional scaling, and the method of extreme grouping of parameters. There are various methods for measuring distances in a multidimensional parameters' space. The parameters can be measured on various scales — quantitative, ordinal, or nominal, which produces various types of distances between the TUs. This, first of all, allows constructing different "TU-TU" proximity matrixes for geographical and parameters' spaces. In addition, they are used to construct the distances between classes and functionals of classification quality.

The most general relation is used to calculate the distance for M quantitative parameters. This relation is called the Mahalanobis-type metric. Special cases of the Mahalanobis-type distance are:

– ordinary Euclidean distance

$$d_e(x_i, x_j) = \sqrt{\sum_{x=1}^{M}(x_i^{(x)} - x_j^{(x)})^2}$$

and

Mahalanobis distance

$$d_{cb}(x_i, x_j) = \sum_{x=1}^{M}\left|x_i^{(x)} - x_j^{(x)}\right|.$$

The TUs correlation coefficient can also be used as a measure of the TUs' proximity in the space of numerical parameters.

The distances between the TUs in the space with ordinal parameters are most often based on various rank-order correlation coefficients. The main ones are the Spearman and Kendall rank-order coefficients.

The distance between the TUs characterized by the nominal parameters is usually calculated as the number of matches or discrepancies between the parameters' values for two TUs:

$$d_{cnt}(x_i, x_j) = \sum_{k=1}^{M}(x_i^k \ne x_j^k), d_{cnt}(x_i, x_j) \in \{0,..., M\},$$

where

$$I(x_i^k \ne x_j^k) = \begin{Bmatrix} 1, (x_i^k \ne x_j^k) \\ 0, (x_i^k \ne x_j^k) \end{Bmatrix},.$$

Relevant literature on data analysis methods can be used to expand the list of methods for determining distances between objects in the parameters' space (Ayvazyan et al., 1985, 1989).

The typological method, which entails normalization of the base-parameter system by variances, represents a matrix for calculating the Euclidean distances (d_{ik}), connecting each pair of the TUs and capturing their differences. From the obtained d_{ik} values, the greatest distance is chosen, and the two TUs that it links become the cores of the homogeneous territorial clusters. The clusters are formed by the distribution of the remaining TUs between the two cores based on the minimal Euclidean distances. If the number of the clusters is large, the third core (and all the subsequent ones) are isolated by testing each of the remaining TUs as cores, and the remaining TUs are distributed among the three cores by minimality of d_{ik}; the variant with the smallest intragroup differences is identified. The resulting number of groupings can be analyzed by the absolute and relative heterogeneity coefficients and, thus, we can choose the optimal number of clusters:

$$A_k = \frac{100\left\{\sum_{k=1}^{K}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\sum_{p=1}^{P}(x_{ip} - x_{jp})^2\right]^{1/2} I_{ik}I_{jk}\right\}}{\sum_{i=1}^{t_{\max}}\left[\sum_{p=1}^{P}(x_{ip} - x_{jp})^2\right]^{1/2}}$$

$k=t_{min},t_{min}+1,...,t_{max}$;

$$O_k = \frac{100\left\{\sum_{k=1}^{K}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\sum_{p=1}^{P}(x_{ip} - x_{jp})^2\right]^{1/2} I_{ik}I_{jk}\right\}}{\sum_{i=1}^{t_{\max}}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\sum_{p=1}^{P}(x_{ip} - x_{jp})^2\right]^{1/2} I_{ik}I_{jk}},$$

$k=t_{min},t_{min}+1,...,t_{max}-1$.

where K – the number of identified groups; P – the number of the orthogonalized coefficients to calculate distances; n – the number of TUs; t_{max} – the maximal number of groups; t_{min} – the minimal number of groups; I – indicator (binary), pointing to the presence (1) or absence (0) of TUi in group k.

A sharp increase in the absolute or relative coefficients of heterogeneity with a decrease in the number of the identifiable clusters indicates the increase in heterogeneity within the identified clusters, while, on the contrary, a smooth increase in the coefficients is a sign of its uniform increase. The threshold followed by a sharp increase in heterogeneity can be optimally taken as the final number of clusters. In our analysis, the isolation of five clusters (specific types of nosological profiles) in all calculations turned out to be optimal.

RESULTS AND DISCUSSION

Types of nosological profiles

In 1997, the first two out of five isolated clusters (Fig. 1), each characterizing a specific set of re-emerging natural focal diseases, have similar features. The incidence of diseases is relatively high — more than 100 cases per 100,000 population, on average. The first type includes the Republic of Bashkortostan with a very high incidence of HFRS and a very low (less than one) incidence of TBE and ITBB. The second includes the Republic of Udmurtia with a high incidence of HFRS and a medium incidence of TBE and ITBB. The incidence of leptospirosis and tularemia is close to zero.

The third type is associated with the most part of European Russia where all nosological forms are recorded. There are 8.4 and 3.4 cases and per 100,000 population of HFRS and ITBB respectively; the incidence of other infections is less than one.

The fourth type includes six TUs — the Vologda, Kaluga, Smolensk, Tver, and Yaroslavl regions, and the Republic of Kalmykia. It has a fairly high incidence of ITBB (5.7 cases per 100,000 population, on average), a lower incidence of HFRS (two cases per 100,000 population), and a very low incidence of other nosological forms (less than one).

The fifth type includes two regions — the Krasnodar Territory and the Republic of Adygea, with a fairly high incidence of leptospirosis (over 20 cases per 100,000 population) and an extremely low incidence of HFRS (0.02 cases per 100,000 population).

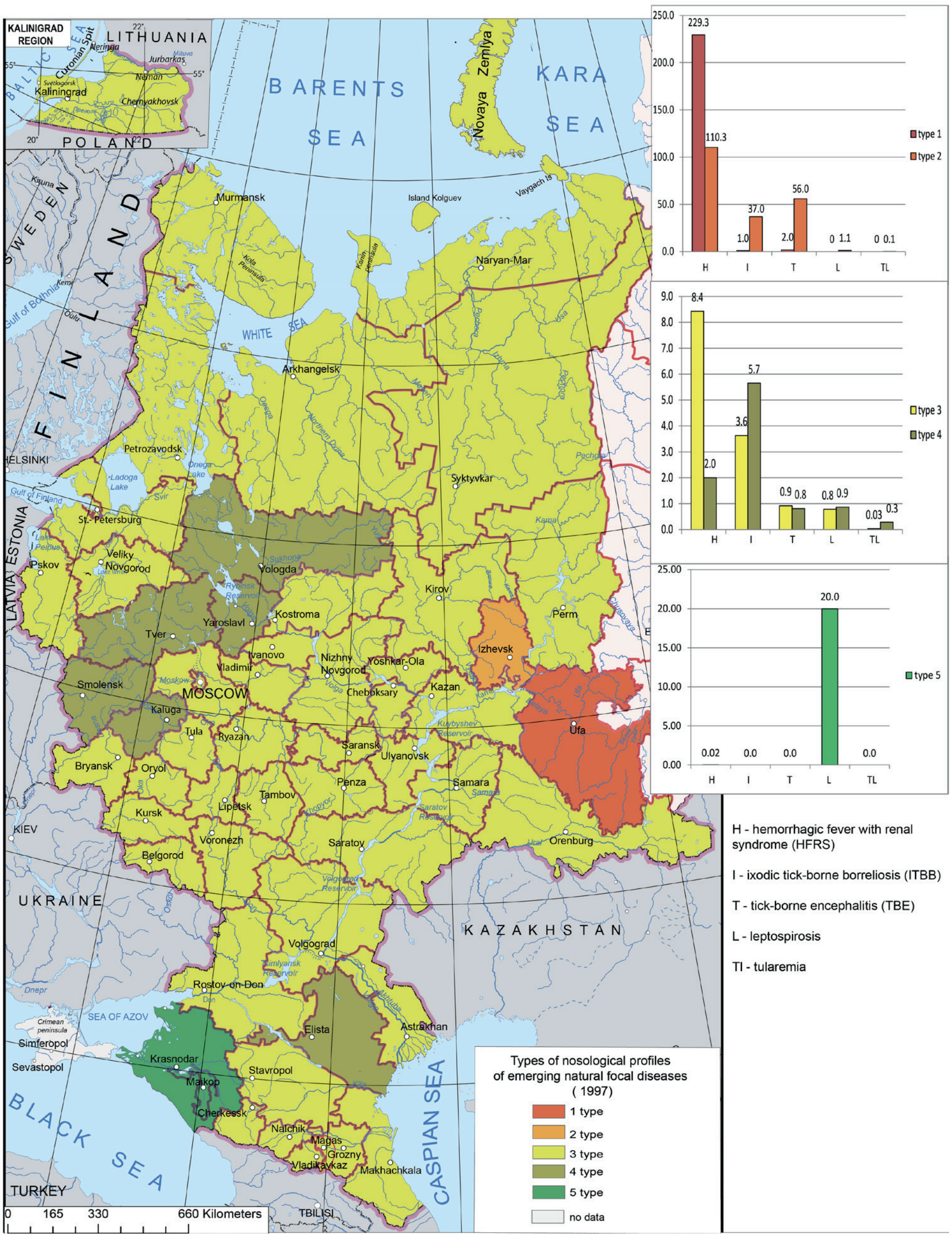


Fig. 1. The types of nosological profiles of emerging and re-emerging natural focal diseases in European Russia in 1997

In 2005, the situation differs from that in 1997 (Fig. 2). The first type, which includes the Republic of Udmurtia (Fig. 1), is characterized by a high incidence of ITBB and HFRS (36.7 and 28.3 cases per 100,000 population respectively) and a medium incidence of TBE (13.1 cases per 100,000 population). The incidence of leptospirosis is very low (two cases per 100,000 population), and there are no cases of tularemia.

The second type includes TUs of the Volga Federal District — the Orenburg, Penza, Samara, Saratov, and Ulyanovsk Regions, the Republics of Bashkortostan, Mari-El, Tatarstan, and Chuvashia; and of the Northwestern Federal District — the Pskov Region. There is a fairly high average incidence of HFRS (21.5 cases per 100,000 population) and a low incidence of TBE, ITBB, and leptospirosis. There are no tularemia cases.

The third type is typical of the Northwestern Federal District (the Arkhangelsk and Vologda Regions, and the Republic of Karelia), of the north of the Central Federal District (the Kostroma and Yaroslavl Regions) and of the north of the Volga Federal District (the Kirov Region and Perm Territory). The incidence of ITBB is relatively high (19.2 cases per 100,000 population). The incidence of other infections — TBE and HFRS, does not exceed seven cases per 100,000 population; and the incidence of leptospirosis and tularemia is close to zero.

In 2005, there was an outbreak of tularemia in the Ryazan Region, which, combined with a low incidence of HFRS and ITBB (about three cases per 100,000 population) and a very low incidence of leptospirosis, allowed us to classify this region as a separate, fourth type of the nosological profile.

The fifth type has some similarities with the fourth type isolated in 1997 (Fig. 1). As in 1997, it is characteristic of most European Russia (Fig. 2). The average incidence in the TUs of this cluster is quite low, with all of the analyzed nosological forms recorded.The nosological profiles of natural focal diseases in 2015 differ from those in 1997 and in 2005 (Fig. 3). The first type is characterized by a high incidence of HFRS (40 cases per 100,000 population), and a much lower incidence of TBE and ITBB (3.5 and 9.6 cases per 100,000 population, respectively). The incidence of leptospirosis is close to zero and tularemia was not recorded in this cluster in 2015. This type encompasses TUs of the Volga Federal District (the Republics of Bashkortostan, Mordovia, and Udmurtia, and the Perm Territory) and TUs of the Central Federal District (the Kostroma and Yaroslavl Regions).

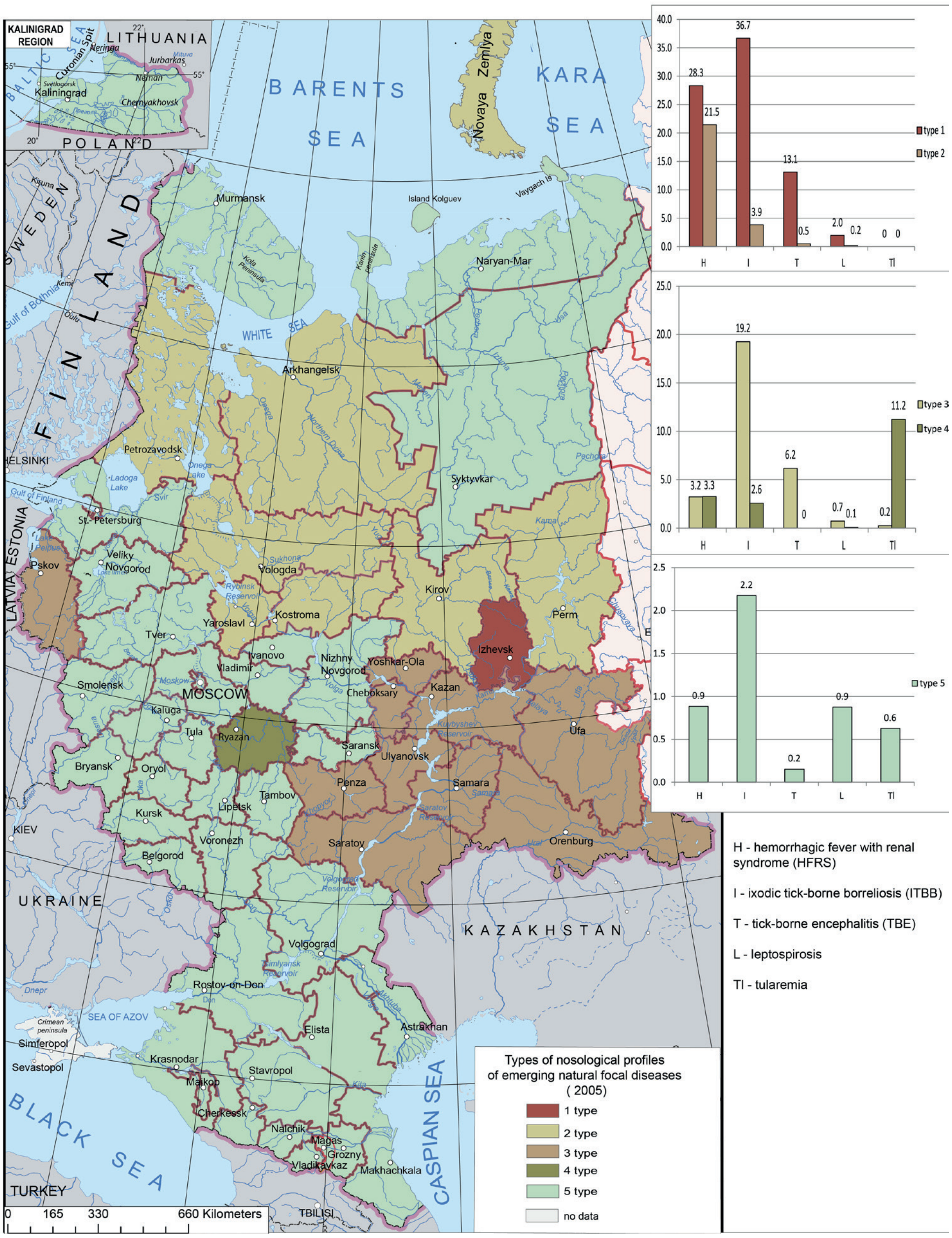


Fig. 2. The types of nosological profiles of emerging and re-emerging natural focal diseases in European Russia in 2005

The second type includes the Volga, Central, and Southern Federal Districts. It is characterized by a low incidence (the average number of cases of HFRS is 10.5 per 100,000) population; the incidence of TBE and ITBB is 0.1 and 2.5, respectively; the incidence of leptospirosis and tularemia is close to zero.

The third type includes the Vologda and Kirov Regions. It has a high (28.4 cases per 100,000 population) incidence of ITBB; the HFRS and TBE incidence is within 10 cases per 100,000 population each, while the incidence of leptospirosis and tularemia is close to zero.

The fourth type (Fig. 3) is similar to that in 1997 (Fig. 2) (less than seven cases per 100,000 population); the highest incidence is observed for ITBB and it is lower for HFRS (about two cases per 100,000 population). The incidence of TBE is less than 1.5 cases per 100,000 population and is close to zero for leptospirosis and tularemia. This type is typical of the Northwestern (the Arkhangelsk, Leningrad, Murmansk, Novgorod, and Pskov Regions; the Republic of Karelia and St. Petersburg) and the Central (the Belgorod, Vladimir, Lipetsk, and Ryazan Regions) Federal Districts.

The fifth type has a very low incidence of all five diseases (does not exceed one case per 100,000 population, on average). It includes the Nenets Autonomous District and the Komi Republic (the Northwestern Federal District); the Voronezh, Kursk, Moscow, Oryol, Tambov, and Tula Regions; Moscow (the Central Federal District); the Astrakhan, Volgograd, and Rostov Regions; the Republics of Kalmykia and Crimea (the Southern Federal District); and the Republics of the Northern Caucasus.

Considering the entire studied period (19 years), the first type includes only one region — the Republic of Udmurtia, with a high incidence of HFRS, TBE, and ITBB, and an extremely low incidence of leptospirosis and tularemia (Fig. 4).

The second type has a lower incidence of diseases in general. The TBE incidence is slightly lower than in the first type (about 21 cases per 100,000 population); the HFRS incidence is about five cases per 100,000 population. The incidence of leptospirosis and tularemia is close to zero. This type is typical of the TUs located in the north of the Volga Federal District (the Kirov Region and Perm Territory), the Northwestern Federal District (the Vologda, Murmansk, and Pskov Regions), and the north of the Central Federal District (the Kostroma and Yaroslavl Regions) (Fig. 4).

The third type is characterized by a low incidence of all five diseases in general (about seven cases per 100,000 population); at that, the incidence of HFRS is higher (six cases per 100,000 population) and it is lower for TBE and ITBB (below one and

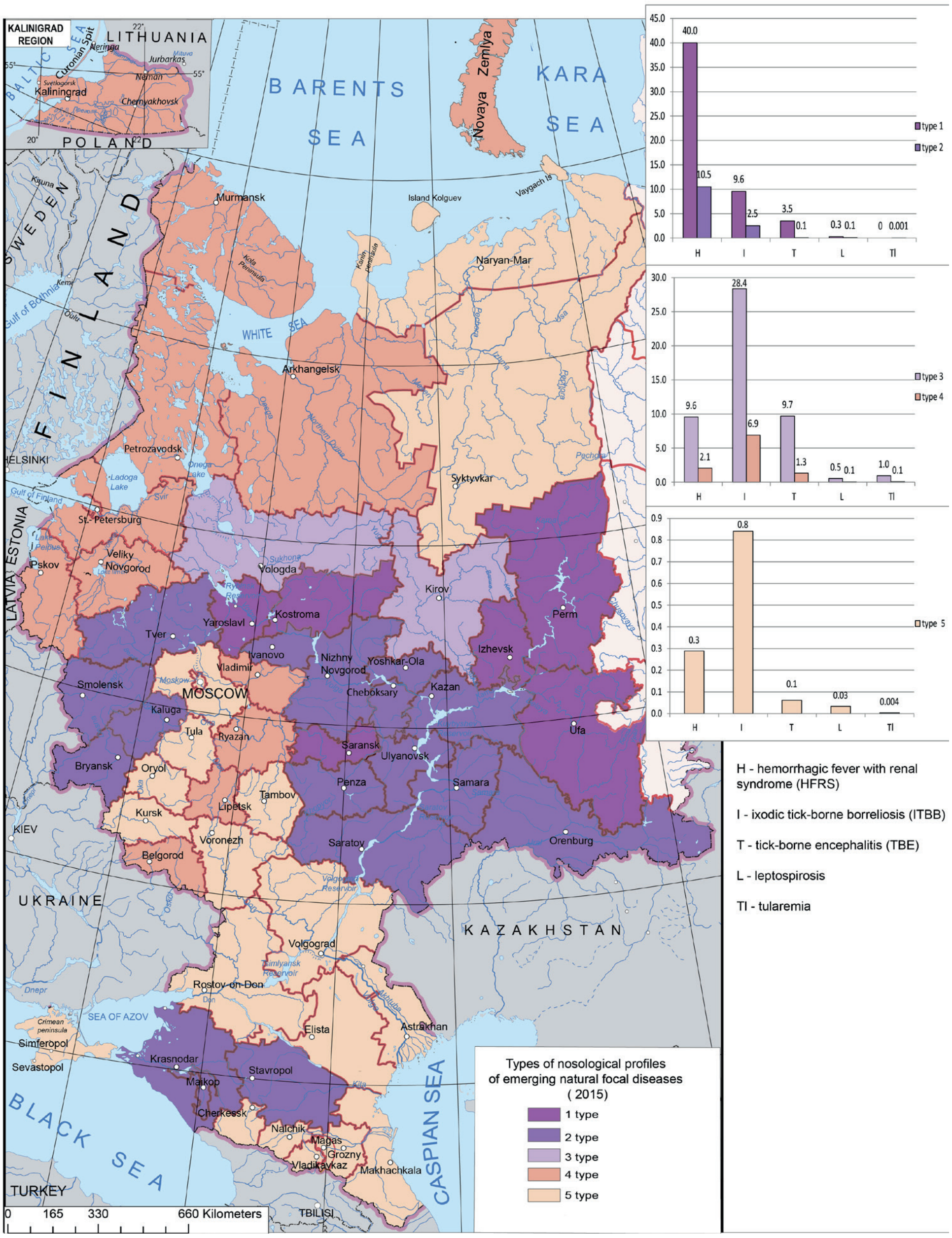


Fig. 3. The types of nosological profiles of emerging and re-emerging natural focal diseases in European Russia in 2015

about two cases per 100,000 population, respectively). The incidence of leptospirosis and tularemia is close to zero. This type includes the majority of European Russia's TUs in various Federal Districts.

The fourth type is also characterized by a generally low incidence of the diseases (about four cases per 100,000 population). The incidence of HFRS, ITBB, and leptospirosis is about three cases per 100,000 population; the incidence of TBE and tularemia is close to zero. The cluster includes the following TUs — the Kaliningrad Region (the Northwestern Federal District), the Kaluga, Smolensk and Tula Regions (the Central Federal District), the Republic of Mordovia (the Volga Federal District), the Krasnodar and Stavropol Territories, the Republic of Adygea and the Chechen Republic (the Southern and North Caucasian Federal Districts).

The fifth type of the nosological profile includes the TUs with a generally low (not exceeding 1.6 cases per 100,000 population) incidence of all five infections, but with a slightly higher incidence of TBE and ITBB. This type includes the Arkhangelsk Region and the Nenets Autonomous District (the Northwestern Federal District), the Ryazan Region (the Central Federal District), and the Republic of Dagestan (the North Caucasian Federal District).

Analysis has thus demonstrated that the nosological profiles of the emerging natural focal infectious diseases in the regions vary by year; the incidence depends on the climatic conditions in each particular year and the extent of deratization and other preventive measures. In general in the studied period, the third type was prevalent — the incidence of all considered infections was low with a somewhat higher incidence of HFRS. The exception is the Republic of Udmurtia (the first type prevailing) where there is a consistent high incidence of HFRS, TBE, and ITBB, as well as seven TUs located farther north and characterized by a high incidence of ITBB. The type with a very low incidence encompasses four TUs located in the north of European Russia (the Arkhangelsk Region and the Nenets Autonomous District), in the central part (the Ryazan Region), and in the south (the Republic of Dagestan) (Fig. 4).

Types of dynamics

The conducted typological classification of the incidence dynamics for all nosological forms considered, allowed isolation of five types of the TUs with specific incidence dynamics.

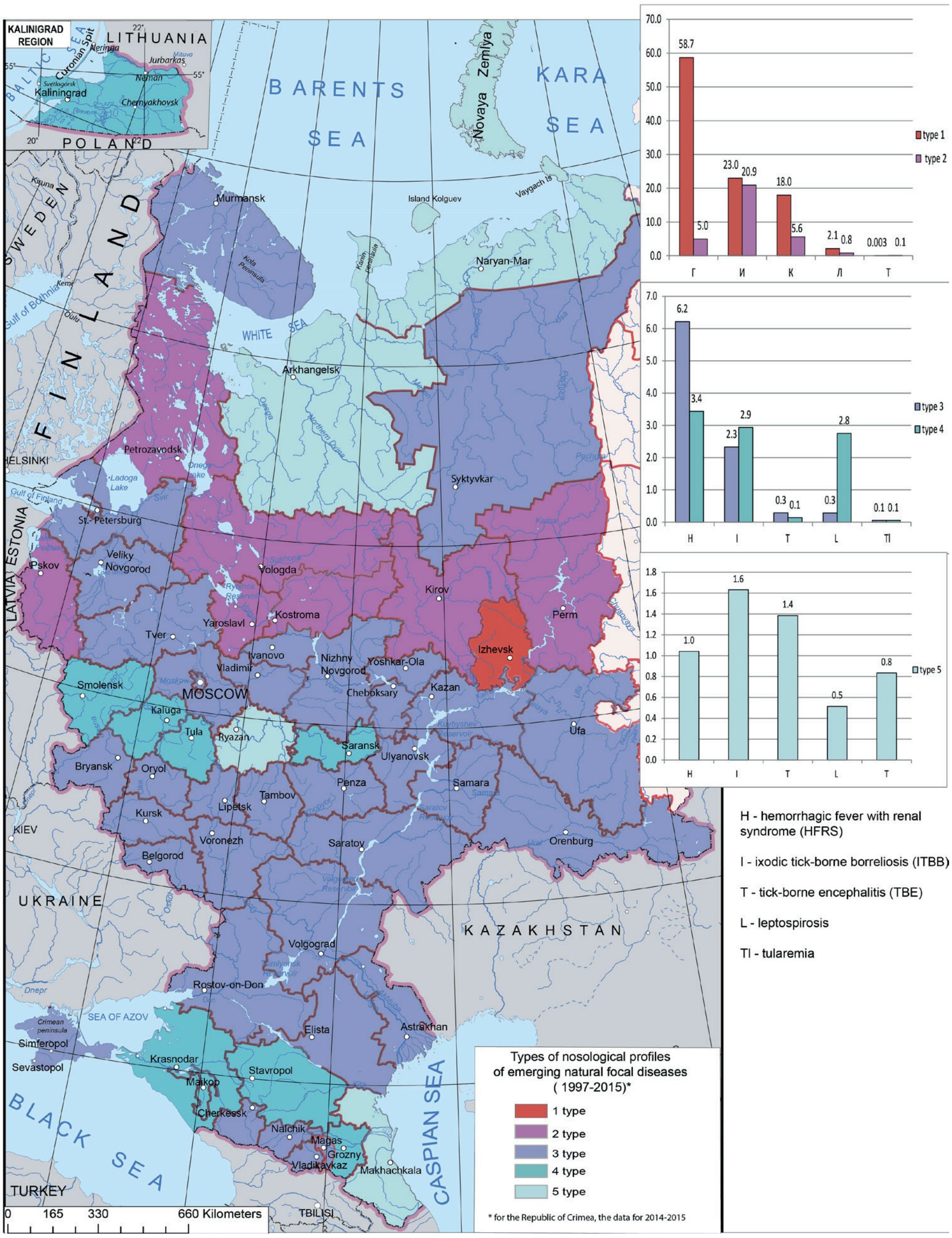


Fig. 4. The types of nosological profiles of emerging and re-emerging natural focal diseases in European Russia over 19 years (1997–2015)

Hemorrhagic fever with renal syndrome. The types of HFRS dynamics are shown in Fig. 5. The first type is characteristic of two regions — the Republics of Udmurtia and Bashkortostan. The highest incidence in these regions is recorded in 1997 (about 180 cases per 100,000 population). The incidence rises every two to four years and it is rather high in general (20 cases per 100,000 population). In 2015, the incidence is 80 cases per 100,000 population.

The second type is characteristic of some TUs of the Volga Federal District (the Orenburg, Penza, Samara, and Ulyanovsk Regions; the Republics of Mari El, Mordovia, Tatarstan, and Chuvashia), where incidence is slightly lower than in the first type — the maximum number of cases recorded in 1997 is about 40 per 100,000 population; the incidence rises every two to six years. The last maximum (25 cases per 100,000 population) is recorded in 2014.

The third type encompasses the rest of the Volga Federal District (the Kirov, Nizhny Novgorod, and Saratov Regions, and the Perm Territory) and some TUs of the Central Federal District (the Yaroslavl and Tula Regions). This type exhibits a relatively low incidence (not exceeding 20 cases per 100,000 population) with two peaks in 2004 and 2014. The incidence rises slightly every two to four years. Recently, there has been a trend towards an increase in the number of cases.

The fourth type, which includes one TU of the Northwestern Federal District (the Vologda Region) and part of the Central Federal District (the Bryansk, Ivanovo, Kaluga, Kostroma, Lipetsk, Ryazan, Smolensk, and Tver Regions), has a low incidence, with, however, a clear growing trend; for this type, a higher incidence is observed in 2004, 2008, and 2014–2015.

The fifth type includes most of the TUs of the Northwestern and Central Federal Districts, as well as all the TUs of the Southern and Northern Caucasus Federal Districts. The incidence is low (in some cases it is zero), but there are two peaks in 2007 and 2015 with a general growing trend.

Thus, the regions with a high incidence of HFRS and the peak in 1997 include the TUs of the Volga Federal District (the first and second types). In the regions where the incidence is minimal and the cases are recorded irregularly (the fourth and fifth types), two small outbreaks are observed in 2007 and 2015, and the incidence exhibits a clear growing trend. A relatively low incidence is associated with the third-type TUs; they are characterized by frequent incidence fluctuations, while maintaining a clear upward trend.

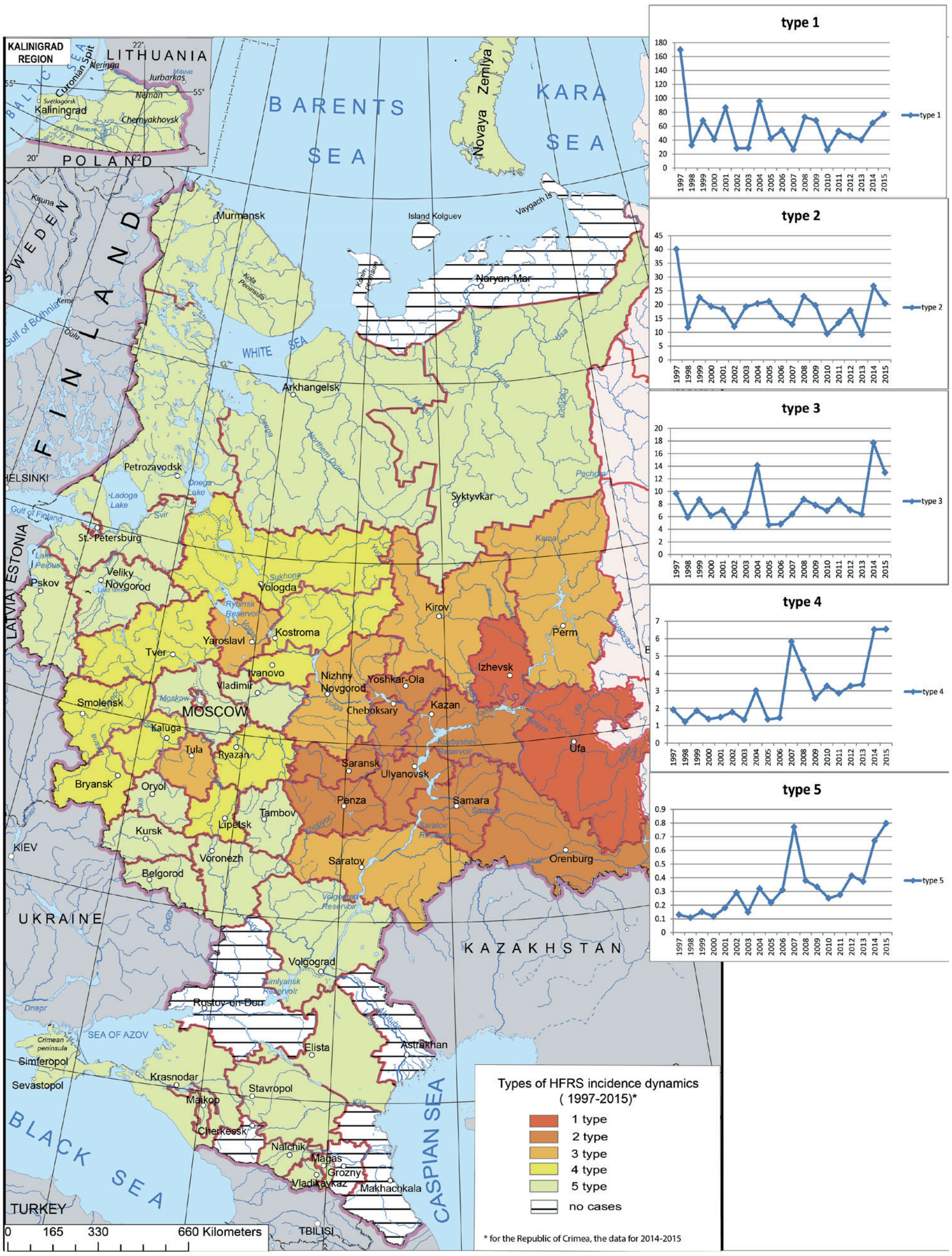


Fig. 5. The types of the HFRS incidence dynamics in European Russia

Tick-borne encephalitis. European Russia has TBE-free areas; cases are recorded mainly in the Northwestern, south of the Central, and the Volga Federal Districts (Fig. 6). The first type is characterized by a high incidence in 1997–1999 (up to 45 cases per 100,000 population), followed by a sharp decline, which continues to this day. In 2013-2015, the incidence did not exceed 10 cases per 100,000 population. Two TUs are in this type, namely, the Republic of Udmurtia and the Perm Territory.

The second type, which includes TUs of the Northwestern (the Arkhangelsk and Vologda Regions, and the Republic of Karelia), the Central (the Kostroma Region), and the Volga (the Kirov Region) Federal Districts, is characterized by a lower incidence — the number of cases does not exceed 12 per 100,000 population. The incidence undulates with a slight growing trend and two peaks in 2003 and 2009.

The third type has a low incidence and includes most of the TUs of the Northwestern (the Kaliningrad, Leningrad, Novgorod, and Pskov Regions, the Republic of Komi, and St. Petersburg) and the Central (the Yaroslavl Region) Federal Districts. The incidence is undulant with rises every three years. A small outbreak is recorded in 2003.

The fourth type encompasses the regions with a low incidence, generally not exceeding 0.5 cases per 100,000 population; the fluctuations are insignificant. This type includes the northern TUs of the Northwestern (the Murmansk Region, the Nenets Autonomous District), the Central (the Tver and Ivanovo Regions) Federal Districts, and the Volga (the Nizhny Novgorod, Orenburg, Samara, and Ulyanovsk Regions, the Republics of Bashkortostan, Mariy-El, and Tatarstan) Federal Districts. Over the 19-year period, there is a slight declining trend (Fig. 6).

The fifth type, which includes most of the TUs of the Central and some TUs of the Volga and Southern Federal Districts, has an extremely low incidence. TBE is not endemic in most of these regions and it is possible that the statistics includes a certain percentage of imported cases. Over the considered period, dynamics slighly increases and undulates; there are two pronounced peaks — in 2009 and 2014.

Thus, in the TUs associated with the first-type of dynamics the incidence decreases towards the end of the studied period. In other clusters, a slight growing trend is recorded in 1997–2015. The TBE outbreaks are registered in 1999, 2003, and 2009 in all TUs, except for those in the first-type cluster.

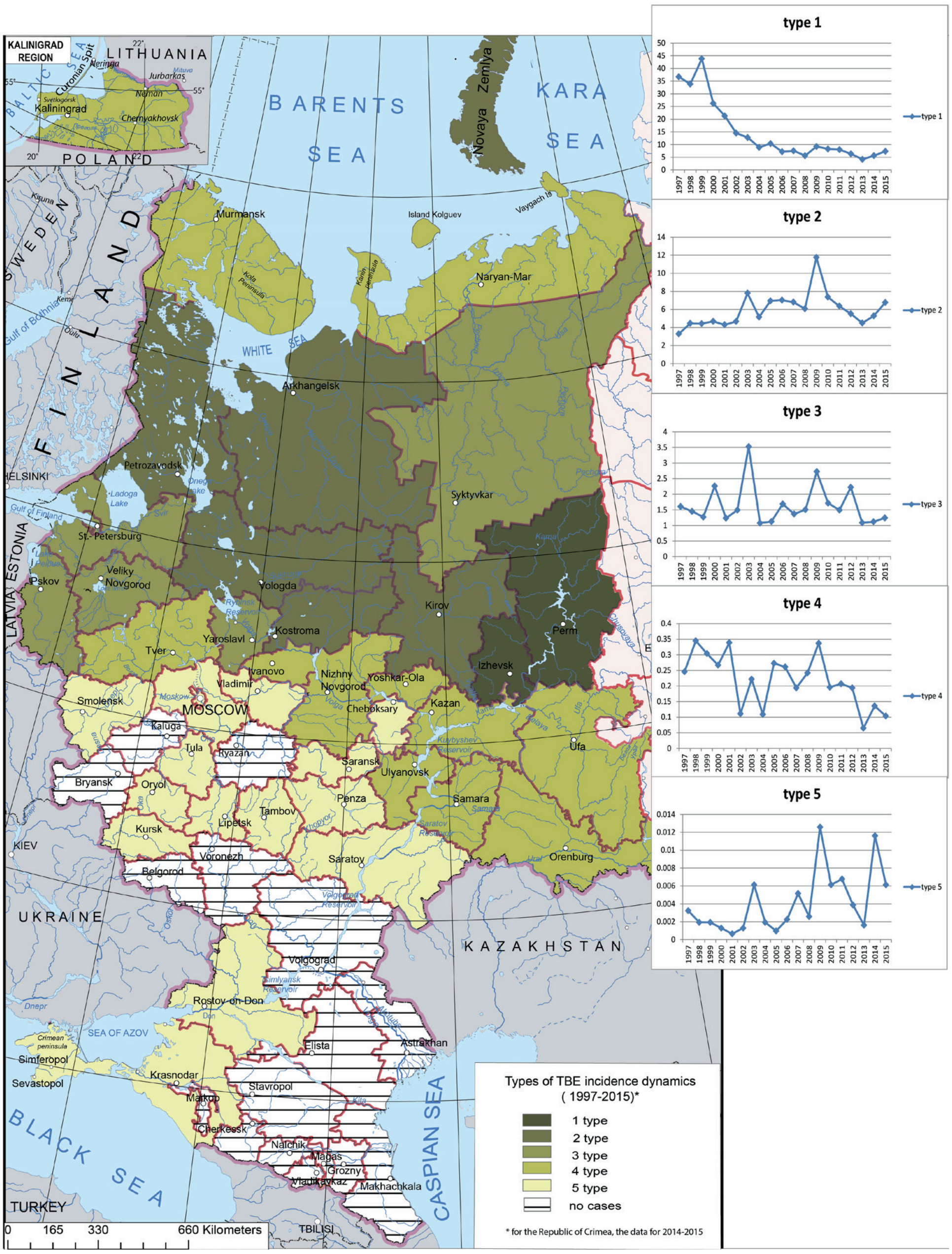


Fig. 6. The types of the TBE incidence dynamics in European Russia

Ixodic tick-borne borreliosis. The ITBB cases are recorded in most of the TUs of European Russia (Fig. 7). The first type is characteristic of the Vologda, Kirov, Kostroma, and Yaroslavl Regions and the Republic of Udmurtia. This type is characterized by a high incidence and rises every two to four years. In the past five years, there has been a slight downward trend in incidence with the last major increase in 2009.

The second type is characteristic of the Northwestern (the Kaliningrad, Leningrad, Novgorod, and Pskov Regions, the Republic of Karelia, and St. Petersburg), the Central (the Vladimir and Kaluga Regions), and the Volga (the Republics of Mariy-El and Chuvashia and the Perm Territory) Federal Districts. The incidence is generally low but rises every two to four years; the maximum is recorded in 2003. In the past few years, there has been a downward trend.

The third type is characteristic of the Northwestern Federal District (the Arkhangelsk Region), the Central Federal District (the Belgorod, Lipetsk, Moscow, Smolensk, and Tver Regions, and Moscow), and some TUs of the Volga Federal District (the Nizhny Novgorod, Penza, Saransk, and Ulyanovsk Regions). The incidence is also low with a noticeable upward trend over the 19-year period (from one case per 100,000 population in 1997 to five-six cases per 100,000 population in 2012–2015). Small rises and falls in incidence are observed every two to four years and the last minimum is registered in 2013.

The fourth type is characterized by an even lower incidence — it does not exceed three cases per 100,000 population, on average. This type has a growing trend with small ups and downs every two to four years. It includes a fairly large part of European Russia — the Komi Republic in the Northwestern Federal District, the TUs of the Central and Volga Federal Districts, and the Krasnodar and Stavropol Territories.

The fifth type is characterized by an extremely low, close to zero, incidence and includes the Murmansk Region and he Nenets Autonomous District in the north of European Russia. It also includes the Saratov, Volgograd, and Rostov Regions, and the Republics of Adygea, Kalmykia, Crimea, and Chechen in the central and southern parts of European Russia. The incidence generally increases over the studied period with ups and downs every two to five years.

Thus, the areas with the first two types of dynamics have a relatively high incidence with an upward trend; they are located predominantly in the Northwestern (except for its northern part) and in the northern part of the Volga and

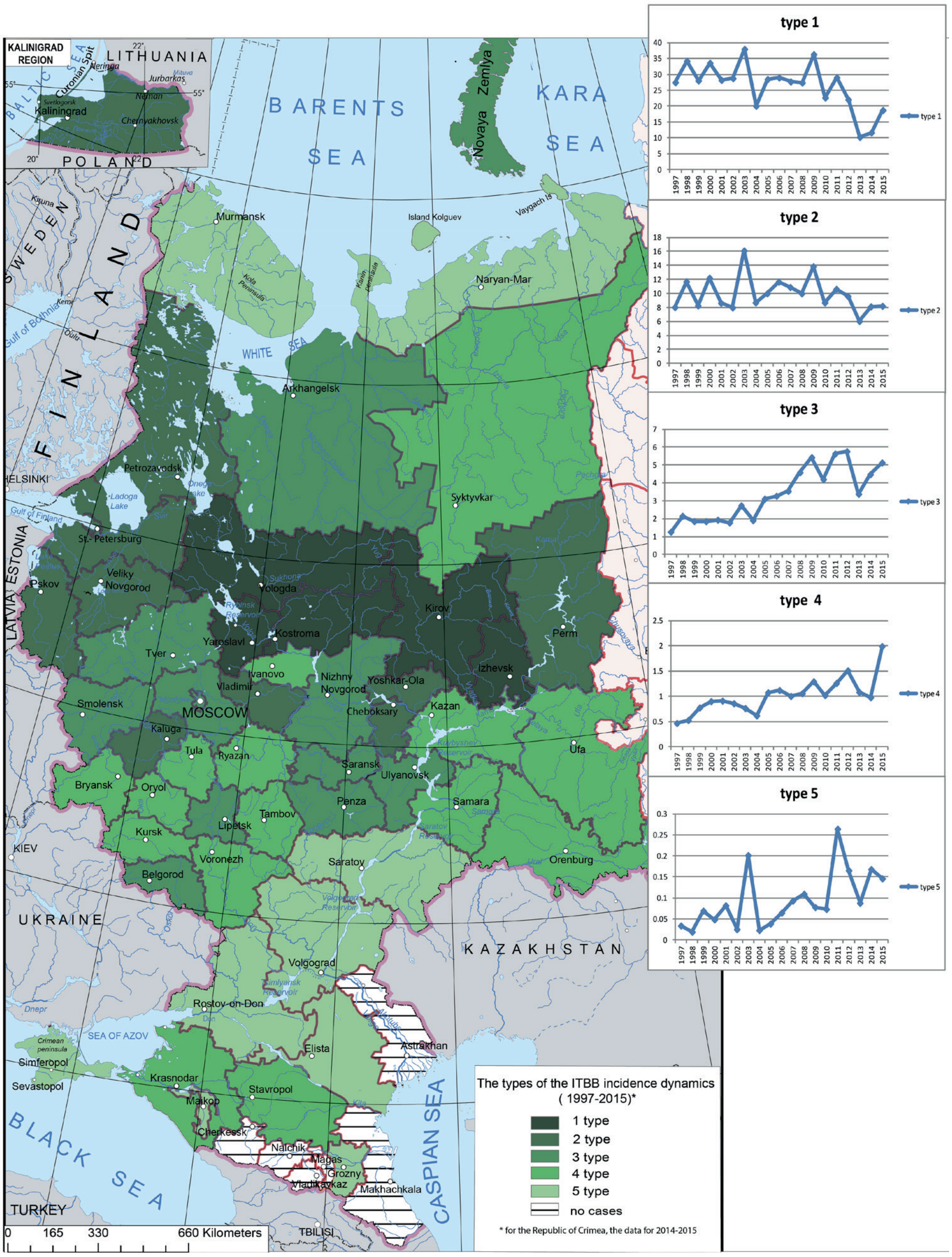


Fig. 7. The types of the ITBB incidence dynamics in European Russia

the Central Federal Districts. The areas with other types have low incidence that increases over the studied period. These regions are located mostly in the north of the Northwestern Federal District, in the Central Federal District, in the south of the Volga Federal District, and in the Southern and North Caucasian Federal Districts.

Tularemia. There have been tularemia outbreaks practically in all areas of European Russia (Fig. 8). The first type includes TUs with a medium incidence (about one case per 100,000 population). However, in 2005 in these TUs, there is a large outbreak with six cases per 100,000 population. This type includes the Ryazan Region and the Republic of Kalmykia

The second type is characteristic of the Arkhangelsk Region with a generally low incidence (close to zero) increasing after 2008. Small outbreaks are recorded in 2002, 2010, 2012 and 2014, with not more than 3.5 cases per 100,000 population.

The third type encompasses some TUs of the Northwestern (the Kaliningrad and Leningrad Regions, the Nenets Autonomous District, and St. Petersburg), most TUs of the Central Federal District, and the Krasnodar and Stavropol Territories. There, the incidence does not exceed one per 100,000 population. As a rule, it is within zero to 0.2. However, relative increases are recorded in 1998 and 2005, though the number of cases does not exceed one per 100,000 population.

The fourth type is characterized by the increasing incidence. This type includes the Vologda, Murmansk, Nizhny Novgorod, Kirov, Yaaroslavl Regions, and the Republics of Karelia and Crimea. The incidence is relatively low; however, there are two small outbreaks in 2005 and 2012 with the number of case not exceeding one per 100,000 population. The fifth type covers various TUs throughout European Russia and also includes the TUs with no cases in 1997–2015. The incidence is low with a small peak in 1999.

Thus, in 2005, there has been a marked rise in the incidence in the clusters of TUs with the first, third, and fourth types of dynamics and a small rise in the TUs with the fifth type. Despite the fact that within the entire European Russia, the incidence is generally low, relatively noticeable incidence rises, or "flares," varying in magnitude and frequency, are recorded for all types of dynamics.

Leptospirosis has been recorded practically in all parts of European Russia. However, the incidence is relatively low (Fig. 9). The first type encompasses TUs with the highest incidence

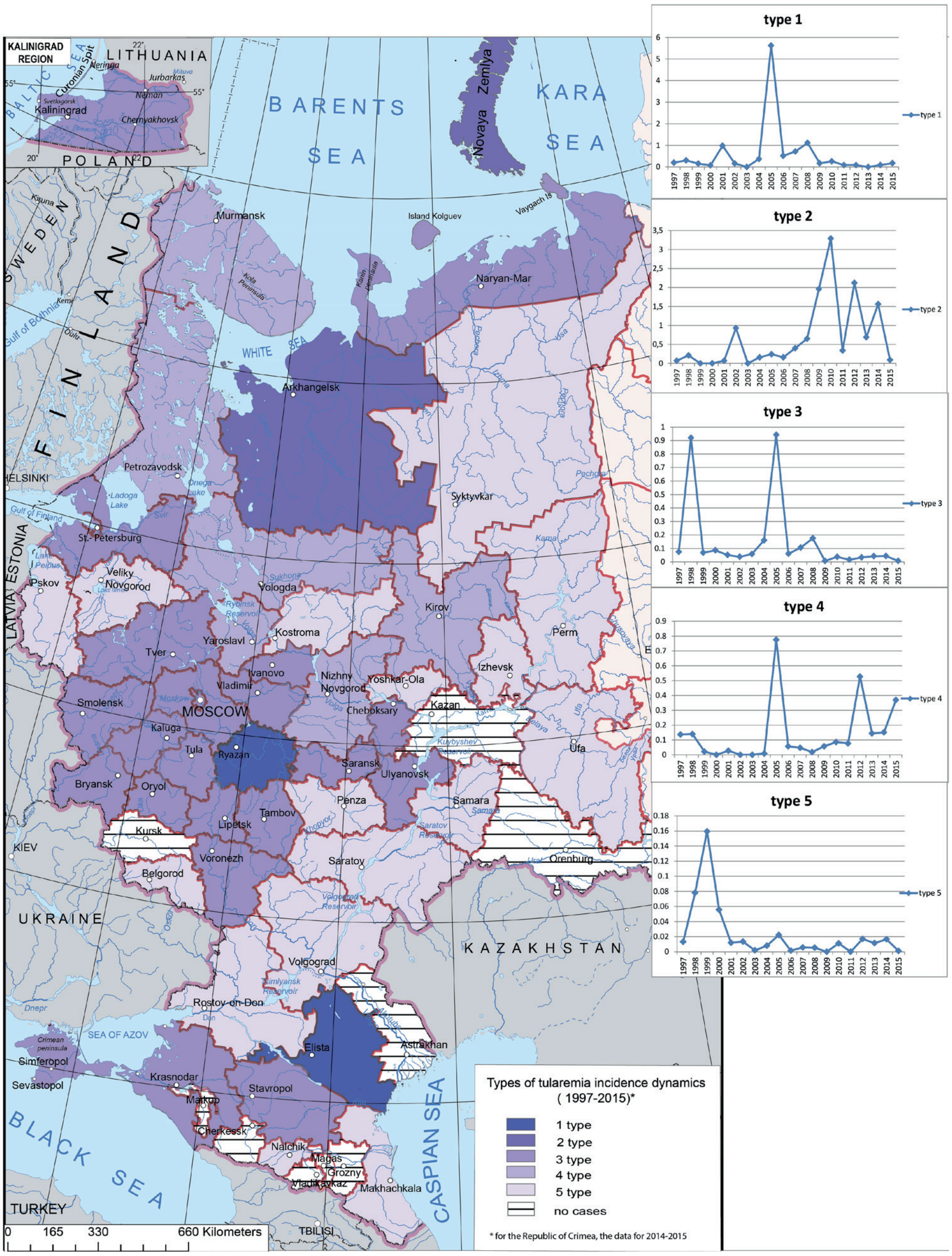


Fig. 8. The types of the tularemia incidence dynamics in European Russia

— in the Republic of Mordovia and the Krasnodarsky Territory. This type is characterized by three marked peaks in 1997, 2002 and 2004, attributed to the increase in the number of small rodents (the principle carriers of this infection). Over the 19-year period, the incidence is decreasing and in certain years, it is close to zero. Such a trend can be explained by the mainstream use of next-generation vaccines (Ananyana 2010).The incidence of the second type is lower than of the first. This type includes the Volga and Kaliningrad Regions (the Northwestern Federal District), the Kaluga, Smolensk, Yaroslavl, and Tula Regions (the Central Federal District), the Ulyanovsk Region, the Republic of Udmurtia, and the Perm Territory (the Volga Federal District) and the Republic of Adygeya (the Southern Federal District). In 1997-2015, the incidence is decreasing with one large outbreak in 2004.

The third type is found is all Federal Districts — the Archhangelsk, Voronezh, Kirov, Leningrad, Nizhny Novgorod, Novgorod, Orlov, and Ryazan Regions, the Republic of Crimea, and the Stavropol Territory. The general incidence in these regions is even lower (does not exceed 2.5 cases per 100,000 population). There are two clear rises in the incidence in 2001 and 2004. In the last years, the incidence is close to zero. Overall in the studied period, there is a clear downward trend.

The fourth and fifth types are registered in the most of European Russia. The fourth type includes the Belgorod, Bryansk, Vladimir, Volgograd, Ivanov, Murmansk, Moscow, Penza, Samara, and Tver Regions, and the Republics of Karachaevo-Cherkessia and North Ossetia. The incidence is lower than in the third type (one case per 100,000 population, on average). There are two small outbreaks in 2004 and 2008. Over the studied period, the incidence is decreasing and in the last years it approached zero, which is associated with mainstream vaccination (Ananyana 2010).

The fifth type that also covers all Federal Districts of European Russia is characterized by even lower incidence. There are two outbreaks in 2000 and 2004 (about 0.2 cases per 100,000 population) over the studied period. Between the outbreaks, the incidence is almost zero. It is possible that certain number of cases in this type of dynamics can be attributed to imported cases.

Despite the general decreasing incidence trend in all five types of dynamics, each type has characteristic outbreaks that differ in magnitude and frequency. In 2004, the leptospirosis incidence went up (an outbreak) in all dynamics clusters.

Thus, the typological classification of the multi-year temporal data series allowed isolating the clusters of TUs

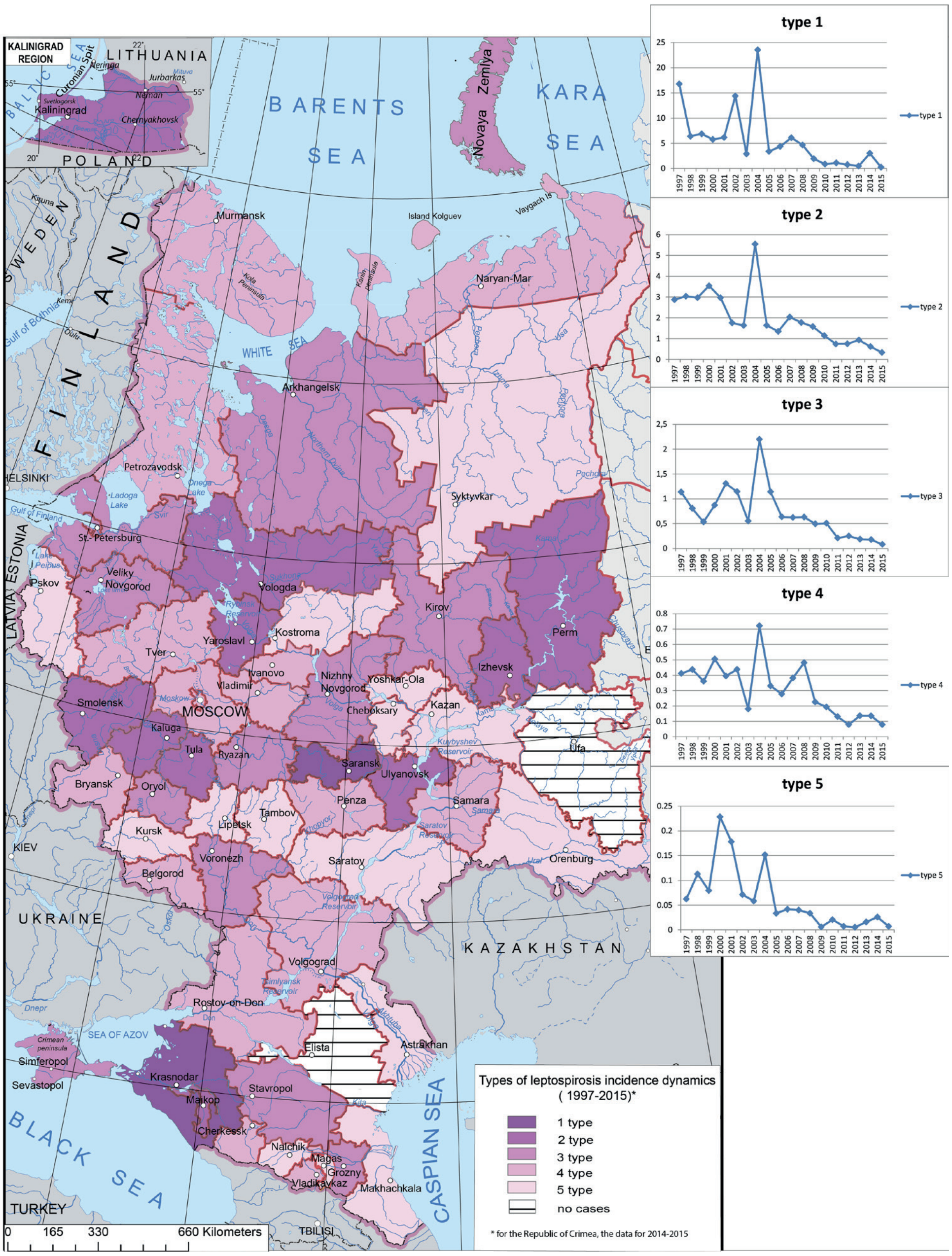


Fig. 9. The types of the leptospirosis incidence dynamics in European Russia

with similar dynamics patterns. The incidence of HFRS and ITBB exhibits a clear growing trend over the last years. The TBE incidence is lower than at the end of the XXth century, however, it increases in some TUs. There are outbreaks of leptospirosis and tularemia, but in general, their incidence is low.

CONCLUSION

The use of formal techniques of mathematical-cartographical modeling allowed us to implement the formal classification of the nosological profiles (i.e., sets of diseases) using the multi-year temporal data series for five emerging natural focal diseases of European Russia. This classification of the incidence in 1997–2015 yielded five types of the nosological profiles. These types vary by years, which is associated with the dependence of the incidence on climatic conditions in each specific year and on extent of deratization and preventive measures.

Furthermore, the classification allowed breaking the regions into the groups with similar patterns of dynamics. Five types of dynamics patterns were identified. Each type is characterized by periodic outbreaks varying in magnitude and frequency.

The results obtained can be used to forecast potential epidemiological outbreaks and to develop targeted and appropriate for each region measures. Automated classification algorithms, certain standard programs, and appropriate skills are crucial in modern research and practice.

ACKNOWLEDGEMENTS

The study was supported by the grants of the Russian Foundation for Basic Research (projects №18-05-00236 and 18-05-60037).

REFERENCES

Ayvazyan S.A., Yenyukov I.S., Meshalkin L.D. (1985). Applied statistics: A study of dependencies. Finance and Statistics

Ayvazyan S.A., Bukhshtaber V.M., Yenyukov I.S., Meshalkin L.D. (1989) Applied statistics: Classification and reduction of dimension. Finance and Statistics.

Ananyana Yu. V. (2010). Leptospirosis of people and animals: the spread of trends and the problems of prevention // Epidemiology and Preventive Vaccination, №. 2, pp. 13-16.

Feldman E.S. (1977). Medical and geographical study of the territory of Moldova. Kishenev, Shtiintsa.

Jones K.E., Patel N.G., Levy M.A., et al. (2008). Global trends in emerging infectious diseases // Nature, Vol. 451, №7181, pp. 990-993.

Kapralov E.G., Koshkarev A.V., Tikunov V.S., et al. (2010). Geoinformatics. In 2 books. University Texbook. V.S.Tikunov (Ed.). 3rd ed., Revised and corrected edition. Moscow: Academia.

Keller A.A. (1992). Medico-geographical approach to the study of regional health // Medico-geographical aspects of assessing the level of public health and the environment. SPb, pp. 37-45.

Kurolap S.A., Klepikov O.V., Vinogradov P.M. (2015). Integrated environmental assessment of the urban environment. Voronezh: Scientific book.

Malkhazova S.M. (2001). Medico-geographical analysis of territories: mapping, assessment, forecast. Moscow: Scientific world.

Malkhazova S.M., Mironova V.A. (2017). The problem of new and recurring infections: the task of medico-geographical study // Bulletin of Voronezh State University. Series: Geography. Geoecology, №1, pp. 21-31.

Malkhazova S.M., Mironova V.A., Pestina P.V., Orlov D.S. (2016). Emerging and re-emerging infections in Russia: the medico-geographical aspect // Moscow University Bulletin. Series 5: Geography, №5, pp. 24-32.

Malkhazova S.M., Pestina P.V., Shartova N.V. (2017). Public Health in the Regions of Russia at the Beginning of the XXI Century: Medico-Geographical Assessment // Reports of the Academy of Sciences, V. 475, №.3, pp. 329-332.

Malkhazova S. M. Mironova V.A., Shartova N.V., Orlov D.S. (2019). Mapping Russia's Natural Focal Diseases: History and Contemporary Approaches. Springer International Publishing

Morse S.S. (1995). Factors in Emergence of Infectious Diseases // Emerg. Inf. Dis., pp. 7-15

Tikunov V.S. (1983). Algorithm for modeling the thematic content of typological maps. – Vestn. Mosc. Univer., Ser. Geogr, №4, pp. 78-84.

Tikunov V.S. (1989). Classification and mapping of fuzzy geographic systems. – Vestn. Mosc. Univer. Ser.5. Geogr., №3, pp. 16-23.

Tikunov V.S. (1997). Classifications in geography: renaissance or fading?. Smolensk: SSU

Trofimov A.M., Zabotin Ya.I. Panasyuk M.B., Rubtsov V.A. (1985). Quantitative methods of regionalization and classification. Kazan, 120 p.

Zadeh L.A. (1965). Fuzzy sets. // Information and Control, 1965, V. 8, pp. 338-353.

Received on June 6th, 2019 Accepted on January 21th, 2020

Ricardo Almendra[1*], Paula Santana[1]
[1] CEGOT-Centre of Studies on Geography and Spatial Planning; Department of Geography and Tourism, Faculty of Arts and Humanities of the University of Coimbra
***Corresponding author:** ricardoalmendra85@gmail.com

MORTALITY FROM CARDIOVASCULAR DISEASES IN THE MUNICIPALITIES OF MAINLAND PORTUGAL: SPATIOTEMPORAL EVOLUTION BETWEEN 1991 AND 2017

ABSTRACT. During the last decades, important efforts have been taken to tackle cardiovascular diseases, which resulted in important mortality and disability decreases. Despite this, cardiovascular diseases are still one of the major causes of death in Portugal. Thus, the aim of this study is to analyse the evolution of the spatial pattern of deaths by cardiovascular diseases, between 1991 and 2017, identifying areas of high risk, and its variation, in the 278 municipalities of Continental Portugal.

Secondary data on annual resident population and deaths by cardiovascular diseases (International Classification of Diseases 10th revision: I00-I99) was collected from the Portuguese National Statistics for the municipalities of Portugal Mainland, from the period under analysis.

To identify areas with significant high and low risk of mortality by cardiovascular diseases, towards time and space, and areas with significantly high or low temporal trends, retrospective spatial-temporal cluster and a spatial variation in temporal trends analysis were conducted.

In the spatial-temporal analysis 3 clusters of high risk and two of low risk were identified; municipalities forming the clusters of high risk tend to have rural characteristics while the municipalities in the clusters of low risk are located in the two metropolitan areas. The majority of the municipalities forming the clusters of low risk also present higher decreasing trends than the country average.

The results presented can contribute to support the development of future interventions on cardiovascular mortality.

KEY WORDS: Cardiovascular diseases mortality; space–time clustering; Spatial variation in temporal trends; Portugal

CITATION: Ricardo Almendra, Paula Santana (2020) Mortality From Cardiovascular Diseases In The Municipalities Of Mainland Portugal: Spatiotemporal Evolution Between 1991 And 2017. Geography, Environment, Sustainability, Vol.13, No 1, p. 128-133
DOI-10.24057/2071-9388-2020-06

INTRODUCTION

Circulatory system diseases are still the most common cause of death in the world (Bhatnagar 2017). According to the WHO (2017), more people die annually from Cardiovascular Diseases (CVD) than from any other cause, being responsible for 31% of all deaths worldwide. CVD are also associated with important morbidity and disability burdens (Roth et al. 2017). The term "CVD" is used to represent a group of disorders of the heart and blood vessels and it includes coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic and congenital heart diseases and venous thromboembolism (Stewart et al. 2017).

The increasing international awareness about CVD (among other noncommunicable disease) has contributed to the growing evidence on CVD (Mahmood et al. 2014), to track, and benchmark, opportunities for prevention and efforts to reduce its burden (Roth et al. 2017). According to Bhatnagar (2017), CVD are caused by a set of chronic conditions resulting from a complex interplay between genetic predisposition, life style factors and environmental influences that lead to a progressive deterioration of the structure of cardiovascular tissues; this process, at least until a certain extent, could be prevented by maintaining a healthy lifestyle (even though genetic defects underlie some infrequent forms of heart disease).

Several authors have been providing evidence on the association between several behavioral, environmental and socioeconomic risk factors and CVD mortality and morbidity: unhealthy diet (Pekka et al. 2002, Åkesson et al. 2007; Jackson et al. 2019), insufficient physical activity (Matthews et al. 2012; Patterson et al. 2018), tobacco use (Huxley and Woodward 2011; Peters et al. 2013), harmful consumption of alcohol (Rehm et al. 2016; Ogunmoroti et al. 2019), air pollution (Brook et al. 2010; Hvidtfeldt et al. 2019), thermal discomfort (Vasconcelos et al. 2011, Almendra et al. 2017), socioeconomic deprivation (Chum and O'Campo 2015; Jimenez et al. 2019), unemployment (Naimi et al. 2009) and population density (Chaix et al. 2007), among others.

Environment can impact health in a direct way (e.g. through air quality) and also indirectly, through its influence on behaviors, and therefore on health (Malkhazova et al. 2014; Barton 2017; Mitsakou et al. 2019); for instance, living in municipalities with low job opportunities may lead to stress and poor mental health which in consequence may result in the adoption of unhealthier behaviors (Kamphuis et al. 2012; Loureiro et al. 2019). Studies assessing the impact of sudden and significant changes in the environment where people live, have been highlighting the consequences of environmental conditions on CVD (Bhatnagar 2017). Robertson et al (1977), found that the incidence of myocardial infarction and death from coronary heart disease was significantly higher among Japanese emigrants in the United States when compared to Japanese in Japan; material, behavioural and psychosocial factors were referred to as a major factor to explaining this inequality. More recently, Hedlund et al. (2007) studied pairs of twins from Finland where one twin had lived one year, or more, in Sweden and found a reduced prevalence of coronary heart disease among the twins that migrated from Finland to Sweden; they highlighted that the change in the environment where the migrant lived may have contributed to the adoption of new dietary habits and physical activity patterns.

Mortality due to CVD is currently decreasing in nearly all European countries (Roth et al. 2017). This decreasing trend results from the positive evolution of some behaviour risk factors such as smoking and alcohol consumption and also the improvements of the disease treatment (Wilkins et al. 2017). The same trend is found in Portugal (Almendra et al. 2015), mostly in result of the significant improvements in healthcare and life conditions (Santana 2014). Nonetheless, despite this important decrease CVD are still one of the main causes of morbidity and mortality in Portugal and should be a priority to public health and urban planning stakeholders.

Assessing the unequal distribution of health outcomes between regions with deferent environmental conditions, and its evolution, can contribute to the identification of risk factors to be addressed in the planning of suitable public health interventions (Vardoulakis et al. 2014). Thus, this study analyses the evolution of the spatial patterns of deaths by CVD, between 1991 and 2017, identifying areas of high risk, and its geographical variation, in the 278 municipalities of Continental Portugal.

MATERIALS AND METHODS

Portugal mainland (hereinafter referred as Portugal) is constituted of 278 municipalities. In 2017, Portugal resident population was near 10 million inhabitants, 70% of whom reside in urban spaces (Santana and Almendra 2018); according to the national statistics (2018), at municipal level, the population density (hab/km2) varies between 4 and 7,641 (with an average population density of 109 hab/km2). Lisbon and Porto are the two most populous cities; the two cities are the capital of the correspondent metropolitan areas, accounting together for nearly 4.5 million inhabitants (Fig. 1a).

To identify areas with significant high risk of mortality by CVD, towards time and space and areas with significantly unusual high or low temporal trends, secondary data was collected for the 278 municipalities of Portugal mainland, for 26 years (1991–2017) from the Portuguese National Statistics Institute (National Statistics): 1. annual resident population and 2. deaths due to CVD (International Classification of Diseases 10th revision: I00-I99).

The retrospective spatial-temporal method of clusterization, developed by Martin Kulldorff, (Kulldorff 1997), was applied to cluster the municipalities with significant higher or lower CVD mortality rates, when compared to the expected value, considering the period under analysis. It assumes a Poisson probability model, estimating the Relative Risk (RR), with significant levels of 5%, through the Monte Carlo method. The spatial structure of the model was defined considering: i) the centroid of each municipality, ii) a circular spatial window, iii) 20% of the population, as the maximum cluster size, iv) 2 years as the minimum temporal cluster dimension, v) the impossibility of clusters overlapping.

The analysis of the spatial variation in temporal trends identifies if the temporal trend within a group of municipalities is significantly different than the expected value (e.g. if a cluster of municipalities is decreasing at a slower pace than the reference). In this analysis, a significant cluster does not necessarily mean that the CVD mortality rate is higher or lower, it reflects significant differences between the rhythms of the evolution of CVD mortality.

SaTScan v9.6 was used to develop the spatiotemporal analysis and ArcMap 10.6 to map the results.

RESULTS

In 2017, the crude CVD mortality rate was of 312.4 deaths per 100,000 inhabitants in Portugal (was of 468.1 in 1991) (Fig 1b and 1c). The spatial-temporal analysis of the CVD mortality allowed the identification of 3 clusters with high risk (represented by warm colours in Fig. 2a) and 2 clusters of lower risk (represented by cold colours in Fig. 2a). The high-risk clusters count 26% of the total deaths analysed.

High risk municipality clusters cover the several regions of Portugal (Fig 2a): cluster E includes municipalities from the North East region; Cluster B includes the majority of the municipalities from the Centre and Southern regions; Cluster D is constituted uniquely by the municipality of Lisbon. Cluster D, lasting from 1991 until 2007, has the longest temporal frame and the highest RR (1.81) with an average CVD mortality rate of 668 deaths per 100,000 inhabitants (Table 1). No clusters of high risk were identified after 2008 (Clusters

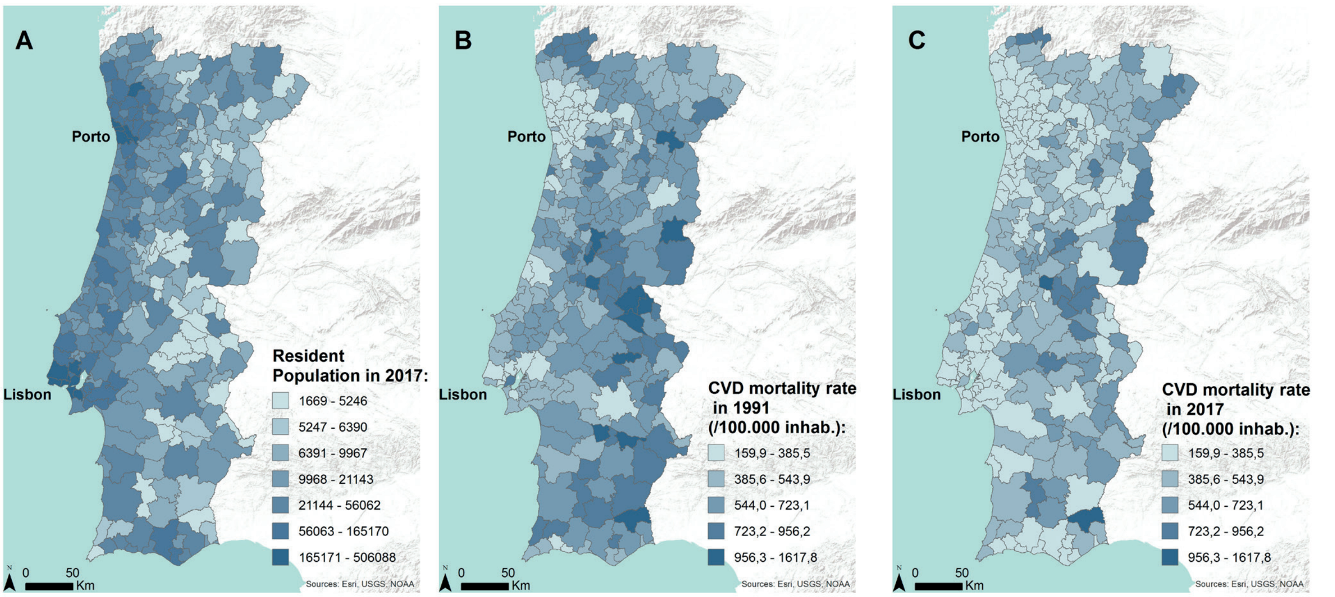


Fig. 1a. Resident population in 2017 **Fig. 1b. Cardiovascular mortality rate in 1991** **Fig. 1c. Cardiovascular mortality rate in 2017**

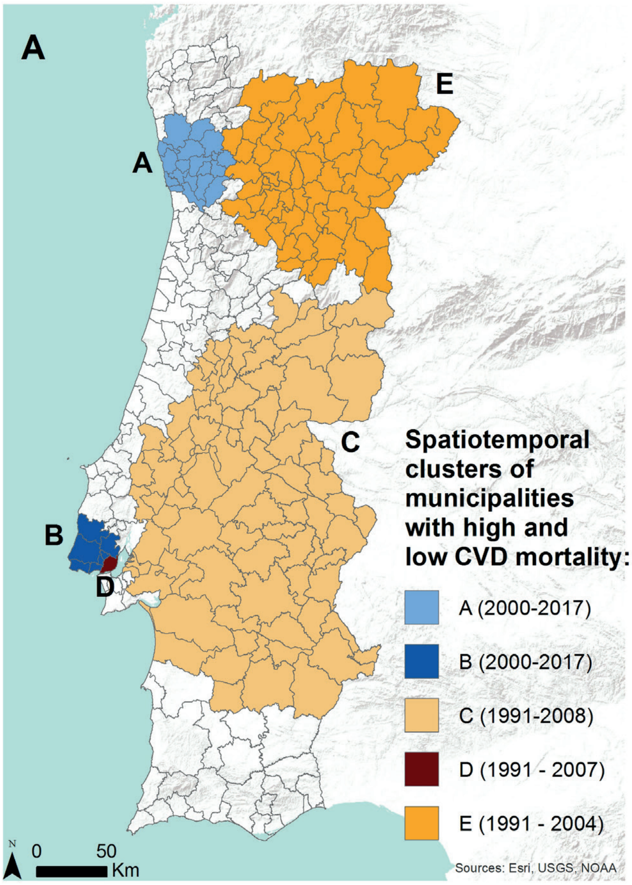


Fig. 2a. Spatiotemporal clusters of cardiovascular mortality

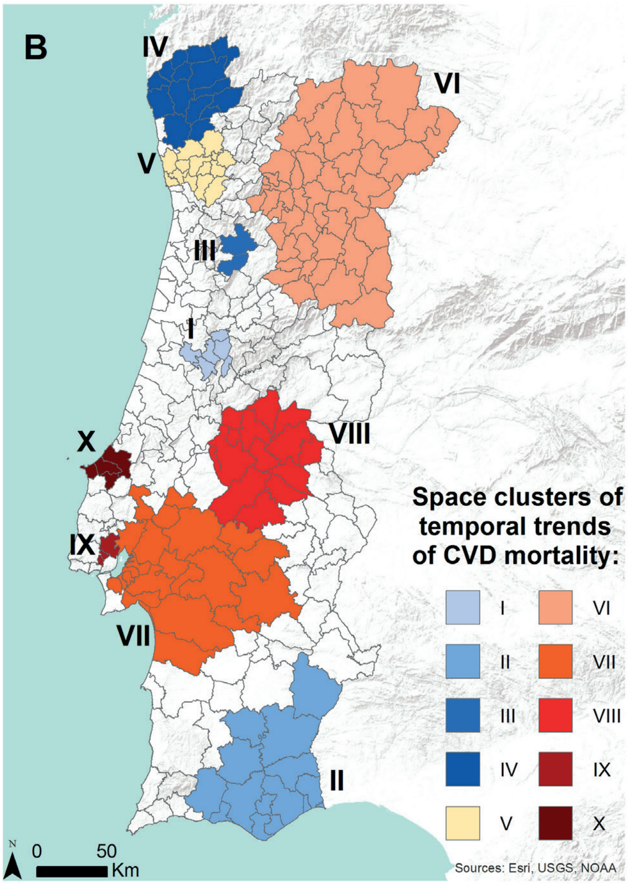


Fig. 2b. Space clusters of temporal trends of cardiovascular mortality

C, D and E last until 2008, 2007 and 2004, respectively) while clusters of low risk are identified from 2000 onwards, lasting until the end of the period under analysis.

Low risk clusters are mainly constituted by municipalities of the two metropolitan areas; Porto Metropolitan Area Cluster (Cluster A) includes the Porto municipality. On the opposite direction, Lisbon Metropolitan Area Cluster (Cluster B) does not include the municipality of Lisbon (which forms Cluster D of high risk).

Clusters A and Cluster B have the same temporal dimension (from 2000 until 2017) but the RR in cluster A is lower (0.58) when compared to cluster B (0.74).

In the temporal analysis of the spatial variation an annual average decrease of 1.8%, in the CVD mortality rate, was found between 1991 and 2017.

Table 2 presents the results of the spatial variation in temporal trend of CVD mortality. Despite the observed decline of CVD mortality in all municipalities, there are important inequalities in the pace of decrease. Clusters V to X are decreasing at a slower pace than what would be expected while clusters I to IV are decreasing faster than the rest of the country (Table 2, Fig 2b). In the municipalities constituting Cluster I, CVD mortality has declined faster than in any other municipality; despite this evolution the CVD mortality rate in this cluster is still higher than the country average.

DISCUSSION

This study analysed the evolution of the spatial patter of deaths by CVD, between 1991 and 2017. The methods adopted in this study combine space–time clustering and spatial variation in temporal trends, allowing the identification of municipalities (space or space–time) characterized by higher CVD mortality, and of municipalities with temporal trends different from the rest of the country. The spatiotemporal pattern of CVD mortality in Portugal has been changing between 1991 and 2017. A general decreasing trend was identified between 1991 and 2017 (annual average decrease of 1.8%); this trend is also represented in the evolution of the spatiotemporal clusters as no cluster of high risk is identified after 2008. Municipalities forming the lower risk clusters tend to be more urban when compared to the municipalities forming clusters of high

Table 1. Relative Risk of the spatiotemporal clusters of cardiovascular mortality

Clusters (p-value <0.05)	Temporal frame (years)	N.º of municipalities	Av. Year Population	Population density (hab./km^2)	RR	Cluster CVD mortality rate (per 100,000 inhab.)
A	2000-17	19	1,946,680	774.0	0.58	229
B	2000-17	7	1,284,042	1,321.4	0.74	288
C	1991-08	91	1,373,008	34.9	1.54	543
D	1991-07	1	562,593	5,622.9	1.81	668
E	1991-04	54	749,505	42.1	1.37	517

Table 2. Space clusters of temporal trends of cardiovascular mortality

Clusters (p-value <0.05)	N.º of municipalities	Av. Year Population	Population density (hab./km²)	Trend inside Cluster (%), 1991-2017	Trend outside Cluster (%), 1991-2017	CVD mortality rate in 1991	CVD mortality rate in 2017
I	6	61,519	89.2	-3.23	-1.79	677.8	336.6
II	14	334,389	45.4	-2.78	-1.77	594.5	350.0
III	2	29,314	54.0	-2.66	-1.80	826.9	471.2
IV	15	606,629	180.1	-2.23	-1.78	461.7	301.4
V	14	1,180,155	695.3	-1.28	-1.82	293.0	225.4
VI	47	602,960	36.4	-1.13	-1.84	513.7	416.8
VII	21	895,301	84.0	-1.07	-1.87	401.6	314.2
VIII	14	120,213	21.7	-0.91	-1.81	702.3	556.8
IX	3	517,334	2377.8	-0.70	-1.85	300.6	257.9
X	4	100,702	177.9	-0.52	-1.82	416.1	385.1

risk; it is noteworthy that Cluster D, with the higher risk, is constituted by the Lisbon Municipality alone. The majority of the municipalities in the rural clusters present a lower decreasing trend than what would be expected if they would be evolving at the same pace as the other municipalities.

A continuous and rapid decline in CVD mortality has been observed during the last years in the more developed regions (Roth et al. 2017). The decreasing trend found in this study is in line with previous works (Almendra et al. 2015; Santana, 2005; Santana and Almendra 2018). In Portugal, CVD mortality has been evolving positively for some decades as a result of improvements in living conditions in terms of basic sanitation, housing condition, education, for instance (Santana and Almendra 2018). The growing adoption of strategic and preventive measures and the improvements of diagnosis and treatment of stroke and myocardial infarction are also contributing to the trend identified (Direção-Geral da Saúde, 2017). In addition, important improvements in lifestyles determinants such as the decrease of tobacco and alcohol consumption have been registered during the last years and its impact must also be considered (Santana 2005; Santana and Almendra 2018). As an example, a set of public health initiatives were implemented to reduce salt intake which positively impacted the evolution of the number of CVD events, and eventually also impacted premature deaths (Abreu et al. 2018). Measures addressing excessive tobacco and alcohol consumption were implemented and may also have impacted the evolution of the CVD mortality.

The association between socioeconomic conditions and health is often assessed under the rationale that people living in more deprived areas people will present worst health outcomes (Nogueira 2010; Santana et al. 2015b). As referred before, the majority of the municipalities forming the Clusters E and C are located in the inland of the country which is characterized by higher material deprivation (Almendra et al. 2017), a more pronounced ageing pattern (Santana and Almendra, 2018) and lower geographical access to healthcare (Santana et al. 2015a). The conjugation of these factors may contribute to a higher mortality rate found in these municipalities. Santana (2005) and Correia et al. (2004) identified a similar pattern during the 1990', highlighting that CVD mortality was associated with rurality (e.g. more manual workers, lower education levels).

On the other hand, clusters with low risk are constituted by municipalities of the metropolitan areas, where the population density tend to be higher; Winkleby et al. (2007), in a study developed in Sweden, also found lower risk of coronary heart disease incidence in areas with higher population density. Malkhazova et al. (2014), refers that cities create the most favorable conditions for life: people living in urban spaces may benefit from better developed social and health structures, providing more opportunities for education and career choices.

Lisbon municipality was classified as a cluster of high risk, contrasting with the other municipalities of the metropolitan area. Possible explanations may be related with the high number of migrants living in Lisbon in worst health conditions (Harding et al. 2008) and with a high ageing population recorded in Lisbon; according to the national statistics (INE, 2019), in 2001, the ageing index in Lisbon was of 198, being the highest of the Lisbon metropolitan area (the country average was 105).

Roth et al. (2017) refer that political and social unrest may be in the origin of discontinuities in the evolution of CVD mortality, highlighting that further attention is needed to understand how CVD is influenced by rapid changes in material living conditions. In consequence of the "Great Recession", and the following implementation of the Economic Adjustment Programme, portuguese social and economic structures were shaken (e.g. such as strong unemployment increases, loss of purchasing power, important emigration flows) (Doetsch et al. 2017; Almendra et al. 2019). This may also have contributed to the observed spatial variation in temporal trend, since the decreasing trend of the CVD mortality may have been slowed down in the municipalities more affected by the consequences of the economic downturn.

Other factors influencing the spatial variation in temporal trend may be related to the different ageing rhythms recorded throughout the country as there is a positive association between CVD mortality and the ageing rate (Casper et al. 2016). The decline of CVD mortality may have been smoothed by the demographic evolution, mostly in municipalities where the ageing process has been more intense.

An important limitation of this study is related with the unavailability of deaths by age at municipal level, which did not allow the age-standardization of the results. This study applied spatiotemporal scan statistic to detect clusters in different space and periods but did not consider covariables that could help to understand these patterns. Further studies should build up from our results and include covariables on demographic and socioeconomic characteristics.

CONCLUSIONS

The spatiotemporal pattern of the CVD mortality in Portugal has been decreasing since 1991. Municipalities with rural characteristics tend to present higher CVD mortality risk and, simultaneously, a slower decreasing trend.

The joint analysis of the spatiotemporal clustering and spatial variation on temporal trends can contribute to support the development of future interventions to control this challenging problem. Strategies directed to the municipalities where the CVD mortality is higher and the decreasing trend is slower can have a high impact on the CVD burden. On the other hand, municipalities where the CVD decreased significantly faster than the country can be important study cases to identify conditions that contributed to the positive evolution.

ACKNOWLEDGEMENTS

Ricardo Almendra, and Paula Santana are members of CEGOT, which is supported by European Regional Development Funds through COMPETE 2020—Operational Programme "Competitiveness and Internationalization", under Grant POCI–01–0145– FEDER–006891, and by National Funds through the FCT under Grant UID/GEO/04084/2013.

REFERENCES

Abreu D., Sousa P., Matias-Dias C., Pinto F.J.(2018). Cardiovascular disease and high blood pressure trend analyses from 2002 to 2016: after the implementation of a salt reduction strategy. BMC Public Health 18, 722. DOI: 10.1186/s12889-018-5634-z.

Åkesson A., Weismayer C., Newby P.K., Wolk A. (2007). Combined Effect of Low-Risk Dietary and Lifestyle Behaviors in Primary Prevention of Myocardial Infarction in Women. Arch. Intern. Med. 167, 2122. DOI: 10.1001/archinte.167.19.2122.

Almendra R., Santana P., Vasconcelos J., Freire E. (2015). Seasonal mortality patterns due to diseases of the circulatory system in Portugal. Geogr. Environ. Sustain. 1, 71-77, DOI: 10.24057/2071-9388-2015-8-1-71-78.

Almendra R., Santana P., Vasconcelos J. (2017). Evidence of social deprivation on the spatial patterns of excess winter mortality. Int. J. Public Health 62, 849-856. DOI: 10.1007/s00038-017-0964-7.

Almendra R., Perelman J., Vasconcelos J., Santana P. (2019). Excess winter mortality and morbidity before, during, and after the Great Recession: the Portuguese case. Int. J. Biometeorol. 63, 873-883. DOI: 10.1007/s00484-019-01700-6.

Barton H. (2017). City of well-being : a radical guide to planning, Routledge. ed. New York.

Bhatnagar A. (2017). Environmental Determinants of Cardiovascular Disease. Circ. Res. 121, 162-180. DOI: 10.1161/CIRCRESAHA.117.306458.

Brook R.D., Rajagopalan S., Pope C.A., Brook J.R. et. al. (2010). Particulate Matter Air Pollution and Cardiovascular Disease. Circulation 121, 2331-2378. DOI: 10.1161/CIR.0b013e3181dbece1.

Casper M., Kramer M.R., Quick H., Schieb L.J. et al. (2016). Changes in the Geographic Patterns of Heart Disease Mortality in the United States: 1973 to 2010. Circulation 133, 1171-80. DOI: 10.1161/CIRCULATIONAHA.115.018663.

Chaix B., Rosvall M., Merlo J. (2007). Assessment of the magnitude of geographical variations and socioeconomic contextual effects on ischaemic heart disease mortality: a multilevel survival analysis of a large Swedish cohort. J. Epidemiol. Community Heal. 61, 349-355. DOI: 10.1136/jech.2006.047597.

Chum A., O'Campo P. (2015). Cross-sectional associations between residential environmental exposures and cardiovascular diseases. BMC Public Health 15, 438. DOI: 10.1186/s12889-015-1788-0.

Correia M., Silva M.R., Matos I., Magalhães R. et al. (2004). Prospective community-based study of stroke in Northern Portugal: Incidence and case fatality in rural and urban populations. Stroke 35, 2048-2053. DOI: 10.1161/01.STR.0000137606.34301.13.

Direção-Geral da Saúde (2017). Programa Nacional para as Doenças Cérebro-Cardiovasculares. Lisbon.

Doetsch J., Pilot E., Santana P., Krafft T. (2017). Potential barriers in healthcare access of the elderly population influenced by the economic crisis and the troika agreement: A qualitative case study in Lisbon, Portugal. Int. J. Equity Health 16, 184. DOI: 10.1186/s12939-017-0679-7.

Harding S., Teyhan A., Rosato M., Santana P. (2008). All cause and cardiovascular mortality in African migrants living in Portugal: evidence of large social inequalities. Eur. J. Cardiovasc. Prev. Rehabil. 15, 670-676, DOI: 10.1097/HJR.0b013e32830fe6ce.

Hedlund E., Kaprio J., Lange A., Koskenvuo M. et al. (2007). Migration and coronary heart disease: A study of Finnish twins living in Sweden and their co-twins residing in Finland. Scand. J. Public Health 35, 468-474, DOI: 10.1080/14034940701256875.

Huxley R.R., Woodward M. (2011). Cigarette smoking as a risk factor for coronary heart disease in women compared with men: A systematic review and meta-analysis of prospective cohort studies. Lancet 1297-1305. DOI: 10.1016/S0140-6736(11)60781-2.

Hvidtfeldt U.A., Geels C., Sørensen M., Ketzel M. et. al. (2019). Long-term residential exposure to PM2.5 constituents and mortality in a Danish cohort. Environ. Int. 133, 105268. DOI: 10.1016/j.envint.2019.105268.

Jackson J.K., MacDonald-Wicks L.K., McEvoy M.A., Forder P.M. et. al. (2019). Better diet quality scores are associated with a lower risk of hypertension and non-fatal CVD in middle-aged Australian women over 15 years of follow-up. Public Health Nutr. 1-12. DOI: 10.1017/S1368980019002842.

Jimenez M.P., Wellenius G.A., Subramanian S.V., Buka S. et. al. (2019). Longitudinal associations of neighborhood socioeconomic status with cardiovascular risk factors: A 46-year follow-up study. Soc. Sci. Med. 241, 112574. DOI: 10.1016/j.socscimed.2019.112574.

Kamphuis C.B.M., Turrell G., Giskes K., Mackenbach J.P., van Lenthe F.J. (2012). Socioeconomic inequalities in cardiovascular mortality and the role of childhood socioeconomic conditions and adulthood risk factors: a prospective cohort study with 17-years of follow up. BMC Public Health 12, 1045. DOI: 10.1186/1471-2458-12-1045

Kulldorff M. (1997). A spatial scan statistic. Commun. Stat. – Theory Methods 26, 1481-1496. DOI: 10.1080/03610929708831995.

Loureiro A., Santana P., Nunes C., Almendra R. (2019). The Role of Individual and Neighborhood Characteristics on Mental Health after a Period of Economic Crisis in the Lisbon Region (Portugal): A Multilevel Analysis. Int. J. Environ. Res. Public Health 16, 1-16.

Mahmood S.S., Levy D., Vasan R.S., Wang T.J. (2014). The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet 383, 999–1008. DOI: 10.1016/S0140-6736(13)61752-3.

Malkhazova S.M., Linsheng Y., Wuyi W., Orlov D.S., Shartova N.V. et. al. (2014). HEALTH OF URBAN POPULATION IN MOSCOW AND BEIJING AGGLOMERETIONS. Geogr. Environ. Sustain. 7, 41-53. DOI: 10.24057/2071-9388-2014-7-4-41-53.

Matthews C.E., George S.M., Moore S.C., Bowles H.R. et. al. (2012). Amount of time spent in sedentary behaviors and cause-specific mortality in US adults. Am. J. Clin. Nutr. 95, 437-445. DOI: 10.3945/ajcn.111.019620.

Mitsakou C., Dimitroulopoulou S., Heaviside C., Katsouyanni K. et. al. (2019). Environmental public health risks in European metropolitan areas within the EURO-HEALTHY project. Sci. Total Environ. 658, 1630-1639. DOI: 10.1016/J.SCITOTENV.2018.12.130.

Naimi A.I., Paquet C., Gauvin L., Daniel M. (2009). Associations between area-level unemployment, body mass index, and risk factors for cardiovascular disease in an urban area. Int. J. Environ. Res. Public Health 6, 3082-96, DOI: 10.3390/ijerph6123082.

Nogueira H.G. (2010). Deprivation amplification and health promoting resources in the context of a poor country. Soc. Sci. Med. 70, 1391-1395. DOI: 10.1016/j.socscimed.2010.01.011.

Ogunmoroti O., Osibogun O., McClelland R.L., Burke G.L. et al. (2019). Alcohol and ideal cardiovascular health: The Multi–Ethnic Study of Atherosclerosis. Clin. Cardiol. 42, 151-158. DOI: 10.1002/clc.23125.

Patterson R., McNamara E., Tainio M., de Sá T.H., Smith A.D. et. al. (2018). Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. Eur. J. Epidemiol. 33, 811-829. DOI: 10.1007/s10654-018-0380-1.

Pekka, P., Pirjo, P., Ulla, U. (2002). Influencing public nutrition for non-communicable disease prevention: from community intervention to national programme--experiences from Finland. Public Health Nutr. 5, 245-51.

Peters S.A.E., Huxley R.R., Woodward M. (2013). Smoking as a risk factor for stroke in women compared with men: A systematic review and meta-analysis of 81 cohorts, including 3 980 359 individuals and 42 401 strokes. Stroke 44, 2821-2828. DOI: 10.1161/STROKEAHA.113.002342.

Rehm J., Shield K.D., Roerecke M., Gmel G. (2016). Modelling the impact of alcohol consumption on cardiovascular disease mortality for comparative risk assessments: an overview. BMC Public Health 16, 363. DOI: 10.1186/s12889-016-3026-9

Robertson T.L., Kato H., Rhoads G.G., Kagan A. et. al. (1977). Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: Incidence of myocardial infarction and death from coronary heart disease. Am. J. Cardiol. 39, 239-243. DOI: 10.1016/S0002-9149(77)80197-5.

Roth G.A., Johnson C., Abajobir A., et al. (2017). Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015. J. Am. Coll. Cardiol. 70, 1-25. DOI: 10.1016/J.JACC.2017.04.052.

Santana P. (2005). Geografias da Saúde e do Desenvolvimento. Evolução e Tendências em Portugal, Almedina. ed. Coimbra.

Santana P. (2014). A saúde dos portugueses, in: Simões, J. (Ed.), 40 Anos de Abril Na Saúde. Almedina, Coimbra, 69-92.

Santana P., Costa C., Cardoso G., Loureiro A., Ferrão J. (2015a). Suicide in Portugal: Spatial determinants in a context of economic crisis. Heal. Place 35, 85-94. DOI: 10.1016/j.healthplace.2015.07.001.

Santana P., Costa C., Marí-Dell'Olmo M., Gotsens M., Borrell C. (2015b). Mortality, material deprivation and urbanization: exploring the social patterns of a metropolitan area. Int. J. Equity Health 14, p. 55. DOI: 10.1186/s12939-015-0182-y.

Santana P., Almendra R. (2018). The health of the Portuguese over the last four decades. Méditerranée. DOI: 10.4000/mediterranee.10348.

Stewart J., Manmathan G., Wilkinson P. (2017). Primary prevention of cardiovascular disease: A review of contemporary guidance and literature. JRSM Cardiovasc. Dis. 6. DOI: 10.1177/2048004016687211.

Vardoulakis S., Dear K., Hajat S., Heaviside C., Eggen B. (2014). Comparative Assessment of the Effects of Climate Change on Heat- and Cold-Related Mortality in the United Kingdom and Australia. Environ. Health Perspect. DOI: 10.1289/ehp.1307524.

Vasconcelos J., Freire E., Morais J., MacHado J.R., Santana P. (2011). The health impacts of poor housing conditions and thermal discomfort. Procedia Environ. Sci. 4, 158-164. DOI: 10.1016/j.proenv.2011.03.019.

WHO (2017). Fact sheet on Cardiovascular diseases [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)]. WHO Fact sheet.

Wilkins E., Wickramasinghe K. (2017). European Cardiovascular Disease Statistics 2017 edition, European Heart Network. Brussels.

Winkleby M., Sundquist K., Cubbin C. (2007). Inequities in CHD incidence and case fatality by neighborhood deprivation. Am. J. Prev. Med. 32, 97-106. DOI: 10.1016/j.amepre.2006.10.002.

Received on June 6th, 2019 Accepted on January 21th, 2020

Sarsenbay K. Abdrakhmanov[1*], Yersyn Y. Mukhanbetkaliyev[1], Fedor I. Korennoy[2], Kanatzhan K. Beisembayev[1], Ablaikhan S. Kadyrov[1], Anar M. Kabzhanova[1], Julie Adamchick[3], Gulzhan N. Yessembekova[1]
[1] S.Seifullin Kazakh Agro Technical University, Nur-Sultan (Astana), Republic of Kazakhstan
[2] Federal Center for Animal Health (FGBI ARRIAH), Vladimir, Russian Federation
[3] Department of Veterinary Population Medicine, University of Minnesota, St. Paul, USA
***Corresponding author:** s_abdrakhmanov@mail.ru

ZONING OF THE REPUBLIC OF KAZAKHSTAN AS TO THE RISK OF NATURAL FOCAL DISEASES IN ANIMALS: THE CASE OF RABIES AND ANTHRAX

ABSTRACT. Rabies and anthrax, being natural focal diseases, are characterized by the ability to persist in areas with a certain combination of environmental factors without human intervention. These infections annually cause sporadic outbreaks in domestic, livestock and wild animals in the Republic of Kazakhstan (RK) receiving close attention of the veterinary service. In particular, targeted mass vaccination and surveillance are conducted, which requires zoning of the country according to the exposure to the diseases.

This paper presents a zoning approach based on the estimation of suitability to the study diseases using the Environmental Niche Modelling method. Retrospective data on animal rabies outbreaks in the RK for 2003-2014, as well as data on anthrax burial sites for 1933-2014 were used. The following environmental factors were treated as potential explanatory variables: 1) a set of climate data derived variables BIOCLIM; 2) altitude above the sea level; 3) land cover type; 4) the maximum green vegetation fraction and 5) soil type.

The modelling outcomes for both diseases indicate elevated risks along the northern and southeastern borders of the RK that not only follows the distribution of historic disease cases, but also accounts for potentially suitable environmental conditions. To comply with the requirements of the veterinary service, gridded risk maps were converted into categorical maps by averaging risk values within municipal districts and ranking according to four categories: low, medium, high, and very high.

The maps obtained may be used as recommendations to the veterinary service as a basis for developing region-specific anti-epizootic measures.

KEY WORDS: Anthrax, rabies, zoning, Kazakhstan, suitability modelling, Maxent

CITATION: Sarsenbay K. Abdrakhmanov, Yersyn Y. Mukhanbetkaliyev, Fedor I. Korennoy, Kanatzhan K. Beisembayev, Ablaikhan S. Kadyrov, Anar M. Kabzhanova, Julie Adamchick, Gulzhan N. Yessembekova (2020) Zoning Of The Republic Of Kazakhstan As To The Risk Of Natural Focal Diseases In Animals: The Case Of Rabies And Anthrax. Geography, Environment, Sustainability, Vol.13, No 1, p. 134-144 DOI-10.24057/2071-9388-2020-10

INTRODUCTION

The prevention and elimination of zoonotic infections remains a priority and responsibility of veterinary science and practice. Rabies and anthrax are among the most significant zoonoses that form the epizootic and epidemic status of many countries and regions of the world, including the Republic of Kazakhstan (RK). The epidemic and epizootic status of these infections is specified according to the level of epidemic danger and the degree of activity of foci in the territory of the RK (Sanzybaev 2003; Antyuganov 2012).

Rabies is a highly dangerous zoonotic disease. This is a viral disease that affects all warm-blooded animals and humans. This disease is characterized by an acute course, signs of polyencephalomyelitis and absolute mortality in the absence of timely treatment, and causes deaths of more than 55 thousand people and more than 1 million animals annually. Today, rabies is registered in 113 countries of the world (Zavodskih 2007; Makarov et al. 2008; Smreczak et al. 2009, 2012; Orlowska et al. 2011; Youla et al. 2014). Direct economic losses from rabies are about 4 billion Euro per year (Nouvellet et al. 2013; Robardet et al. 2013). Over the recent years, the epizootic situation on rabies in the RK remains tense: the rabies incidence in animals (foxes, raccoon dogs, wolves, cats and cattle) tends to increase by an average of 7 % annually. Up to 700 head of farm animals die annually from rabies in the RK; more than 50% of them are cattle and up to 25% are sheep and goats (Abdrakhmanov et al. 2010).

In most regions of the RK, the epizootic status for rabies is extremely difficult: the incidence among various animal species has increased over the last decade; the disease also causes fatalities in humans (Bersagurov 2002; Zholshorinov 2004). Ongoing measures have failed to limit the spread of the rabies infection and completely eliminate the disease in the RK. This fact is associated with many factors and, in particular, the presence of natural foci of infection (Domsky 2002; Chubirko 2003; Dudnikov 2003). The main measures to control the rabies situation in the RK include: 1) oral vaccination of wild animals in epizootic outbreaks and in areas of potential infection spread; 2) forced and prophylactic vaccination of susceptible productive and domestic preying animals, the latter as a necessary measure of urban control of the rabies spread among people. Moreover, strict registration and control of the number of domestic and stray preying animals and outreach activities among the population have been implemented.

Anthrax may be considered "old" and well-studied disease, but the problems associated with its prevention in animals and humans are still unresolved. This is due both to epidemiological characteristics of the disease and to the ecology of the pathogen itself. Every year, from 2 000 to 20 000 cases of anthrax are registered in the world. The disease is widespread in many countries in Africa, Asia, South and Central America, the Middle East and the Caribbean (Adamovich and Nikonov 1970; Aikembayev et al. 2010).

Kazakhstan historically belongs to the category of anthrax endemic countries. The system of anti-anthrax measures presently used in the veterinary medicine of the RK has enabled reducing the intensity of the epizootic situation. However, the risk of new livestock animals becoming infected with anthrax in historically affected areas has been maintained for decades.

Maintaining the epizootological and epidemiological danger in the country is related to the presence of a large number of persistent anthrax locations, which comprises numerous burials of animal corpses fallen from anthrax (Lukhnova et al. 2013; Abdrakhmanov et al. 2014).

As of December 31, 2014, there are 4 058 historic anthrax foci in the republic since 1933, of which 1 800 sites have been registered and certified as permanent anthrax locations.

Against the background of geographically distributed epidemiological risk factors, the sporadic outbreaks in both humans and animals are still registered. As such, eight anthrax cases in livestock animals have been reported 2015–2019 in different regions of the country (FAO EMPRES-i 2019).

The wide spread of animal anthrax and rabies with the formation of favorable prerequisites for new epizootic complications drives the need to improve the measures of epizootological supervision while preventing these zoonoses in the RK.

Research and applied developments provide tools that are both relevant and required for ensuring the country's biological safety regarding the study infections. These tools allow the visualization, zoning, mathematical modelling, and predicting the risk of re-emerging outbreaks of the diseases. Among the geography-based methods, the most informative is zoning, i.e. categorizing of the country in accordance with the intensity of the epizootic situation and the risk of re-occurrence of the disease. Zoning is of great practical importance and allows focused attention of regional veterinary services to the territories (regions, districts, settlements), which have the highest level of zoonotic occurrence and the greatest potential for recurrent outbreaks, based on a combination of favorable landscape-climatic factors coupled with the presence of historically registered foci of diseases (Norstrøm 2001).

This paper was initiated to generalize the methodology of zoning of the RK according to the degree of risk of rabies and anthrax outbreaks (Abdrakhmanov et al. 2016, 2017). Given the natural focal nature of both studied diseases, the zoning was carried out using the mathematical-geographic method for modelling environmental niches (ENM) with optimization by the principle of maximum entropy (MaxEnt). Databases on historic outbreaks of rabies and anthrax in the RK were used as input presence data, while the set of ecological and climatic variables BIOCLIM along with some geographical parameters was used as explanatory factors.

MATERIALS AND METHODS

Rabies data

Data on rabies outbreaks in the RK for 2003-2014 were provided by the veterinary service of the administrative territories (region, district) and collected during field visits. The database consists of 762 rabies cases among various animals: cats, dogs, cows, foxes, camels, sheep, horses and wolves.

For the purposes of modelling, all animal species are divided into three categories: domestic, wild and farm (livestock) animals. Cats and dogs are classified as domestic animals; wolves and foxes are wild animals; horses, cows, sheep and camels are farm animals.

For each outbreak of rabies, the following data are available that are relevant for further modelling: geographical coordinates (latitude, longitude); date of the outbreak; the number and species of infected animals; the name of the settlement, district and region. Fig. 1 shows a map of the RK with mapped cases of rabies in three categories of animals.

Anthrax data

Data on anthrax outbreaks in the RK for 1933-2014 were provided by the veterinary service of administrative territories (regions, districts) and collected during field visits. The database includes anthrax outbreaks in farm animals. For each outbreak, the following attributes are indicated: geographical coordinates, diseased species, number of infected animals, and date of the outbreak registration. After exclusion of unreliable data (in particular, with no geographical coordinates or with erroneously indicated coordinates), the database has 4058 anthrax outbreaks, each of which reports from 1 to 851 infected animals. Humans were also infected in a number of outbreaks. The outbreak map is shown in Fig. 2.

It should be noted that in some cases several outbreaks were associated with the same location (settlement), i.e. identified by the same geographic coordinates. This resulted in 1798 unique locations, a lower number than the total number of outbreaks.

All data on outbreaks were provided for the cartographic presentation in the form of shape files. For modelling with MaxEnt data were converted into .csv files.

Climate and landscape data

The following geospatial variables were used:

1. A set of 19 bioclimatic variables BIO1 – BIO19 (hereinafter – BIOCLIM), which are derived from monthly averages of air temperature and precipitation obtained from meteorological stations worldwide (Hijmans et al. 2005). Data are available for download at http://worldclim.org/version1. The "current" data set for 1950–2000 was used. The variables are presented in Table 1;

2. Altitude above the sea level (ALT) (USGS 2019);

3. Maximum green vegetation fraction (MGVF), reflecting the presence and intensity of vegetation cover. Annual average data for 2001–2012 are used (Broxton et al. 2019a);

4. Land cover type presented by 17 categories dataset based on IGBP classification (Table 2) (Broxton et al. 2019b);

5. Soil type (SOIL) (Digital Soil Map of the World 2019). This variable was included in the model for anthrax only, since the influence of soil type on the survival of the anthrax pathogen is known, and, therefore, a correlation may be expected between the type of soil in a given area and the presence of anthrax outbreaks (Cherkassky 1999, Hugh-Jones and Blackburn 2009; Mullins et al. 2011). Table 3 summarizes the soil units presented in the RK.

All geospatial variables presented in raster format were corrected to a common resolution of 1x1 km, clipped by the shape of the RK, and converted to ASCII format as required for modelling by the MaxEnt software.

Method of risk identification: a description of the maximum entropy principle

To identify the prevailing tendency of the rabies and anthrax outbreaks in animals in areas with a certain combination of landscape-climatic conditions, the maximum

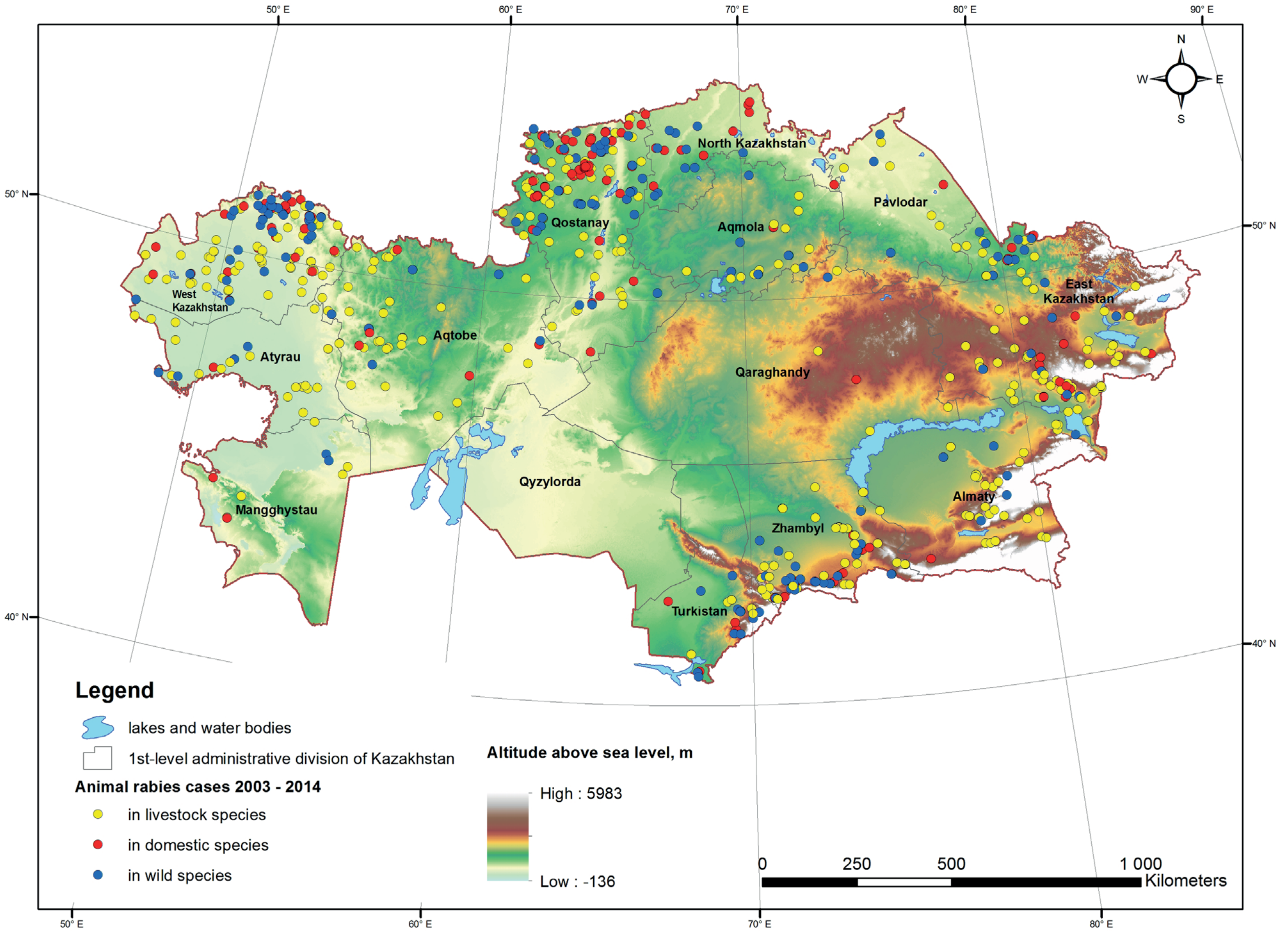


Fig. 1. Physical map of the Republic of Kazakhstan and rabies cases in three categories of animals, 2003-2014

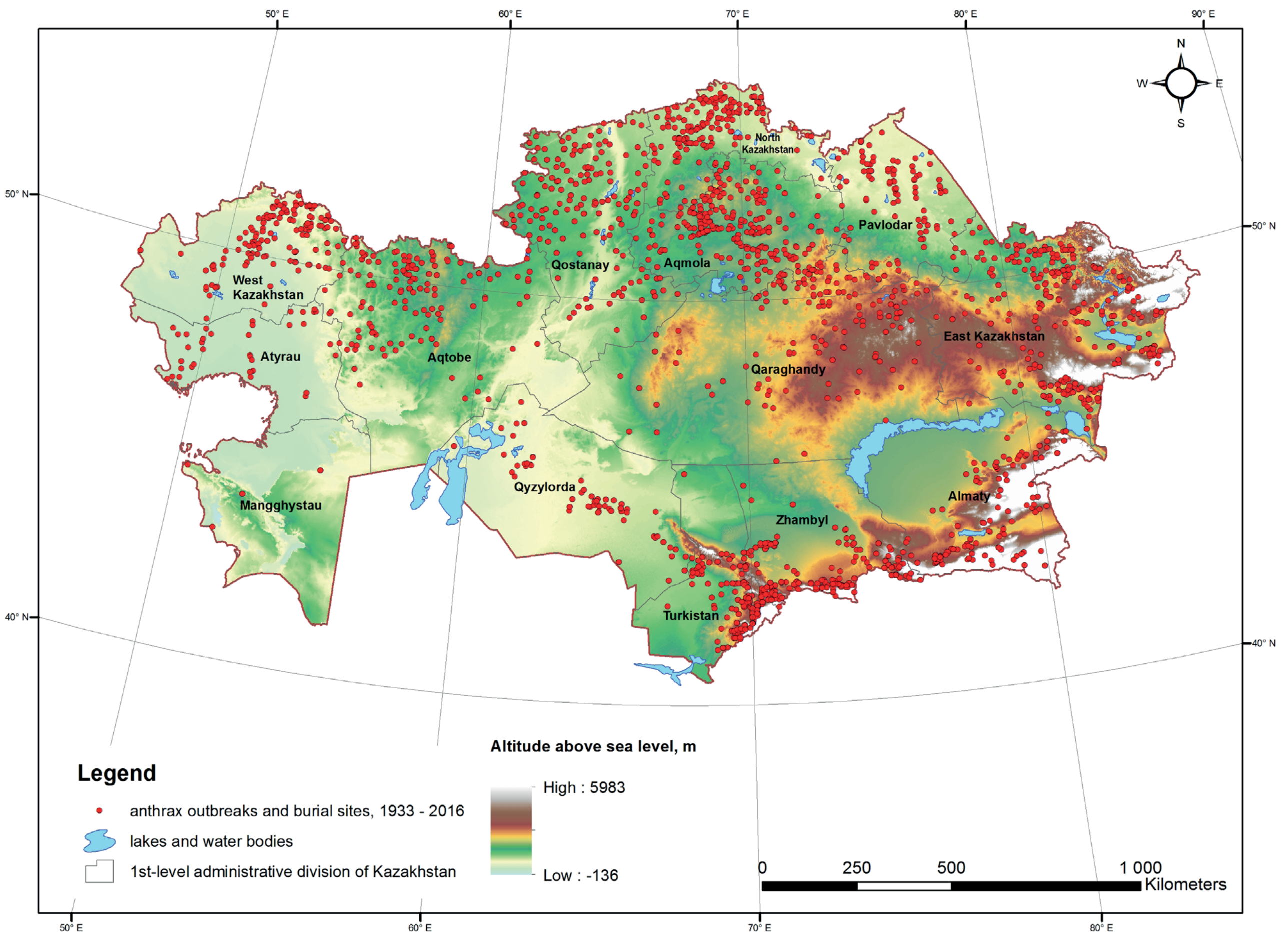


Fig. 2. Physical map of the Republic of Kazakhstan and anthrax burial sites, 1933-2014

Table 1. BIOCLIM variables description

Variable	Description
BIO1	the annual average temperature
BIO2	the average daily temperature
BIO3	isothermal (BIO2/BIO7 * 100)
BIO4	seasonal temperatures (standard deviation * 100)
BIO5	the maximum temperature of the warmest month
BIO6	the minimum temperature of the coldest month
BIO7	the annual temperature range (BIO5 – BIO6)
BIO8	the average temperature of the wettest quarter
BIO9	the average temperature of the driest quarter
BIO10	the average temperature of the warmest quarter
BIO11	the average temperature of the coldest quarter
BIO12	the annual average of precipitation
BIO13	precipitation in the wettest month
BIO14	precipitation of the driest month
BIO15	seasonal rainfall (coefficient of variation)
BIO16	precipitation of the wettest quarter
BIO17	precipitation of the driest quarter
BIO18	precipitation of the warmest quarter
BIO19	precipitation of the coldest quarter

Table 2. Categories of LAND COVER

Category of land cover	Land cover description
0	water
1	evergreen coniferous forests
2	evergreen broad-leaved forests
3	deciduous coniferous forests
4	deciduous broad-leaved forests
5	mixed forests
6	thick bushes
7	rare bushes
8	the wooded savannah
9	savannah
10	meadows
11	permanent wetlands
12	arable land
13	urban and built-up areas
14	mixed arable land and natural vegetation
15	the snow and ice
16	barren lands and wastelands

entropy method was used for modelling (MaxEnt) (Phillips et al. 2004). This method belongs to the class of Environmental Niche Modelling methods that require only "presence data", i.e. those locations where the studied phenomenon (in our case – outbreaks of diseases) are reliably registered. The method is widely used to model 1) the habitat of a particular species, as well as 2) potential range of infectious diseases caused by a specific pathogen. In the latter case, the application of the method is based on the assumption that the disease range is determined by the range of the causative pathogen. As explanatory variables, geospatial factors are normally used that describe the distribution of environmental and (sometimes)

Table 3. Soil units presented in the Republic of Kazakhstan

Category key	Name
Bc	Chromic Cambisols
Bk	Calcic Cambisols
Ch	Haplic Chernozems
Ck	Calcic Chernozems
Cl	Luvic Chernozems
Gc	Calcaric Gleysols
Ge	Eutric Gleysols
Gm	Mollic Gleysols
I	Lithosols
Jc	Calcaric Fluvisols
Je	Eutric Fluvisols
Kh	Haplic Kastanozems
Kk	Calcic Kastanozems
Kl	Luvic Kastanozems
Mo	Orthic Greyzems
Oe	Eutric Histosols
Qc	Cambic Arenosols
Sg	Gleyic Solonetz
Sm	Mollic Solonetz
So	Orthic Solonetz
Xh	Haplic Xerosols
Xk	Calcic Xerosols
Xl	Luvic Xerosols
Yh	Haplic Yermosols
Yk	Calcic Yermosols
Yt	Takyric Yermosols
Zg	Gleyic Solonchaks
Zo	Orthic Solonchaks
Zt	Takyric Solonchaks
RK	Rock debris
GL	Glaciers
DS	Dunes/Shifting sand
ST	Salt flats

socio-economic characteristics in the study area. The factors used can presumably contribute to the formation of conditions favorable for the existence of the studied pathogen and the spread of the disease it causes. The essence of the modelling is to select a probability distribution of the pathogen (or disease) presence in the study area, which is the most uniform (i.e., has the maximum information entropy) of all possible distributions, taking into account the observed actual data distribution. The resulting distribution indicates the suitability of each location for the presence of a pathogen (disease) in terms of the combination of environmental factors used.

The Maxent modelling is performed for rabies and anthrax. In rabies, the modelling was carried out separately for three categories of animals: domestic, wild and farm. The modelling was carried out in 10 replications, and in each replication, the model worked up to 500 iterations to select the most optimal distribution. In each iteration, 75% of all outbreaks were randomly selected for the training of the model (i.e., to identify the desired distribution), and the remaining 25% used for validation and testing of the resulting distribution.

In each iteration, the contribution of each variable to the total distribution was determined by the Jack-knife method. This procedure allows identification of the variables that are most significant in terms of their impact on the probability of outbreaks.

When modelling rabies, to compensate for the possible data bias caused by uneven diagnostics near settlements, the density grid of the main roads in the RK was used. It follows the assumption that cases of the disease were more likely recorded in close proximity to settlements and roads. To build the road density grid, we used road data available in the dataset Esri Data & Maps (https://www.esri.com/arcgis-blog/products/product/mapping/esri-data-maps/). The density grid has been made using the procedure Kernel Density from the software toolkit Spatial Analyst, ArcGIS.

The predictive ability of a MaxEnt model is usually estimated by the area under the ROC curve (AUC value), which indicates the probability that a randomly selected point of presence will be evaluated by the model as "suitable" rather than "unsuitable". It is usually considered that the value AUC = 0.5 indicates the absence of predictive ability of the model; AUC > 0.7 is considered a good indicator, and AUC > 0.8 is an excellent indicator (Araújo and Pearson 2005).

Generalization of risk levels within administrative areas

Upon obtaining risk maps, which are regular grid-based distribution of the probability of the rabies and anthrax outbreaks in the RK, the maps were summarized for zoning at municipal level in accordance with the practice of the veterinary service of the RK. For this end, the risk levels were integrated by municipal districts by determining the average value within each district (the maximum of three values was taken for rabies). Thereafter, the obtained values were ranked in accordance with Table 4 (Abdrakhmanov et al. 2016, 2017). The zoning results are presented the form of risk maps for each disease.

Software

Data geoprocessing and visualization were made using geographic information system ArcGIS version 10.6 (ESRI, USA) with custom add-in for Environmental Niche Modelling related data processing SDM Toolbox (Brown 2014). Maximum entropy modelling was performed using Maxent software (Maxent software… 2019). Additional statistical data processing and file conversion was done using Microsoft Office Excel (Microsoft, USA).

RESULTS

Rabies zoning of the Republic of Kazakhstan

The distribution of the average "suitability" of the RK territory was obtained for each of the categories "farm animals", "domestic animals" and "wild animals", and shown in Fig. 3 (a, b, c respectively).

The final generalized representation of zoning by municipal districts is shown in Fig. 4.

The predictive ability of the model (AUC value) is 0.782±0.031 for the category "farm animals"; 0.859±0,042 for the category "domestic animals" and 0.809±0,045 for the category "wild animals". Thus, the obtained probability distributions demonstrate sufficiently high degree of reliability in description of the distribution of existing rabies cases in the RK, depending on the combination of the chosen environmental factors.

The variables that make the most contribution to the model were defined as those for which the relative importance of the modelling results is not less than 10%. For the category of "farm animals", the variables BIO19 (precipitation level of the coldest quarter), LANDCOV (land cover type) and BIO1 (average annual temperature) have the most contribution.

For the category of "domestic animals", the variables LANDCOV, ALT (altitude above sea level), BIO12 (annual average precipitation) and BIO19 have the most contribution.

For the category of "wild animals", the variables LANDCOV, BIO19, ALT and BIO12 have the most contribution.

Anthrax zoning of the Republic of Kazakhstan

Upon Maxent modelling, the AUC value is 0.834 ± 0.005, which can be considered a good indicator of the predictive ability of the model.

The probability distribution of anthrax outbreaks, taking into account the geospatial factors analyzed, is shown as grid-based probability surface constructed in modelling using all variables (Fig. 5).

When processing geospatial variables, the following factors were found to most explain the distribution of anthrax outbreaks: MGVF (maximum green vegetation fraction), BIO12 (average annual precipitation), BIO13 (precipitation of the wettest month), BIO16 (precipitation of the wettest quarter) and SOILS (soil type). The result obtained allows us to confirm the strong dependence of the locations of anthrax outbreaks on climatic factors, in particular on humidity and precipitation (as demonstrated collectively by the variables BIO12, 13 and 16), on the presence and intensity of vegetation (MGVF) and on the type of soil. In the case of moisture and vegetation, an increase in the corresponding factor entails a corresponding increase in the outbreak probability. The most favorable soil types for the occurrence of anthrax are Bc – Chromic Cambisols, Ch – Haplic Chernozems, Ck – Calcic Chernozems, Je – Eutric Fluvisols and Xk – Calcic Xerosols.

Summarizing and ranking of risk levels by area in accordance with Table 4 gives the following pattern of RK zoning by the degree of biological safety regarding anthrax (Fig. 6).

The obtained probability distribution is in good agreement with the results of other authors that have used similar modelling methods (Mullins et al. 2013).

DISCUSSION

The current system of epizootic surveillance needs automation and the use of modern information technologies to improve its main function – deterrence, prevention and elimination of the infectious threat (Anderson and May 1992). Due to this condition, there is a need to perform zoning in order to assign a certain status to the regions of the RK by categories of biological safety with regard to rabies and anthrax.

In this paper, we perform zoning based on modern mathematical-cartographic methods that take into account not only the presence or absence of outbreaks in a particular territory, as it applies normally, but also the exposure of the territory to their occurrence based on a set of ecological and geographic features.

The maximum entropy method is often used to model the habitats of particular species based on 1) exactly known places where their presence was detected and 2) aggregates of explanatory environmental variables within a study area.

Table 4. The ranking of risks

Risk level (suitability)	Risk category
< 10%	Low risk
10 – 25 %	Average risk
25 – 50%	High risk
>50%	Very high risk

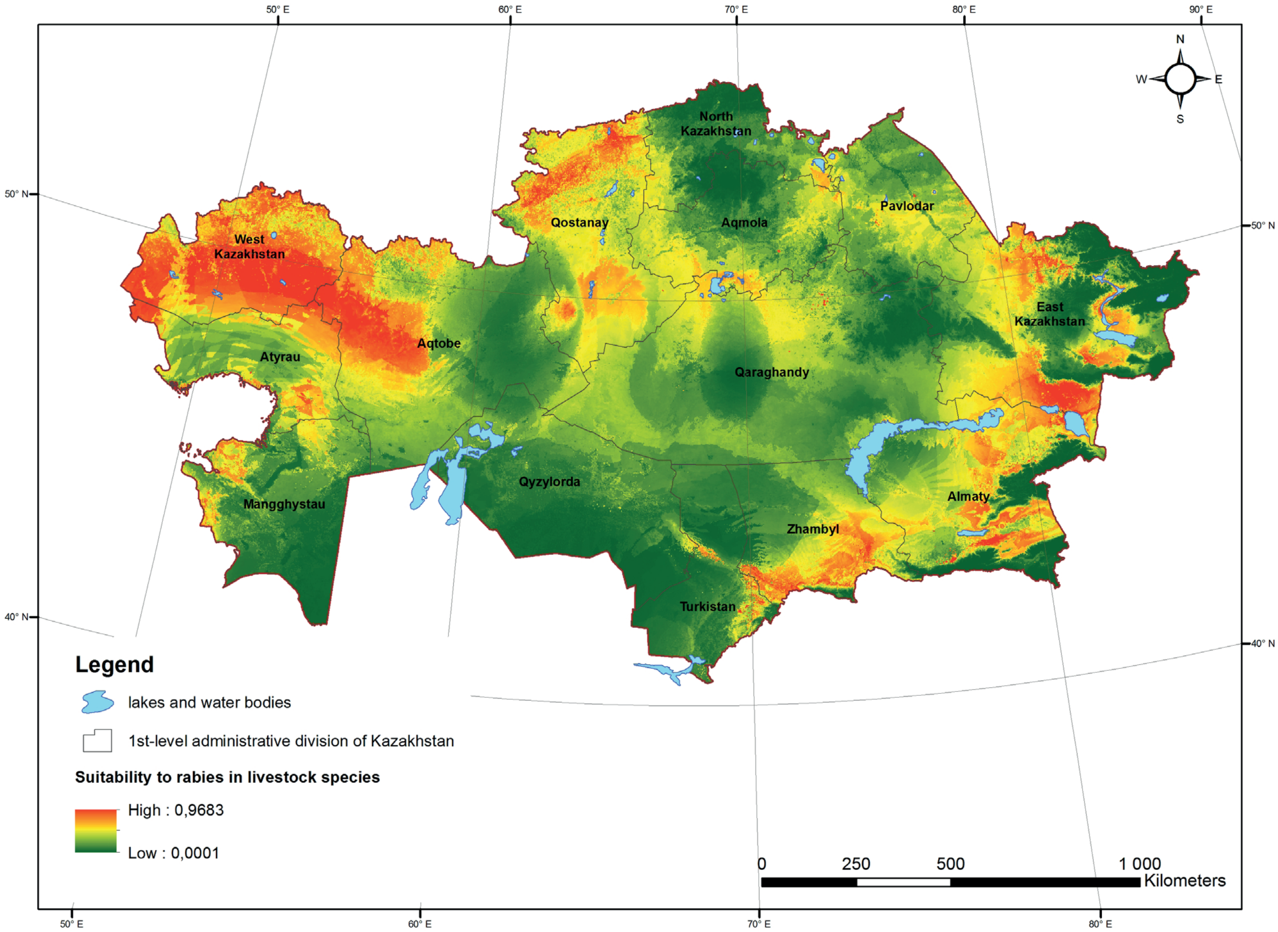


Fig. 3a. Gridded suitability of the RK territory to rabies in farm species

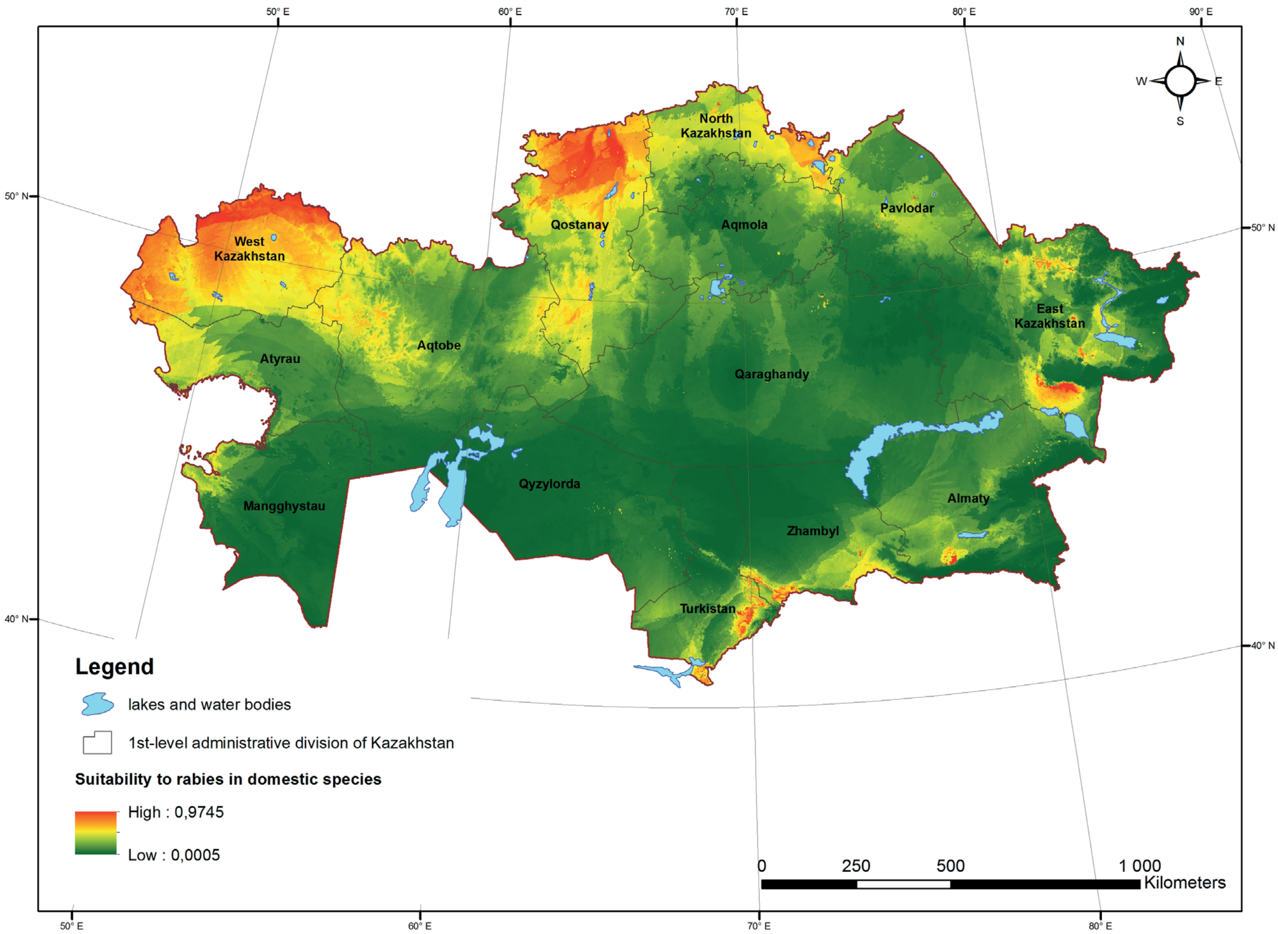


Fig. 3b. Griddded suitability of the RK territory to rabies in domestic species

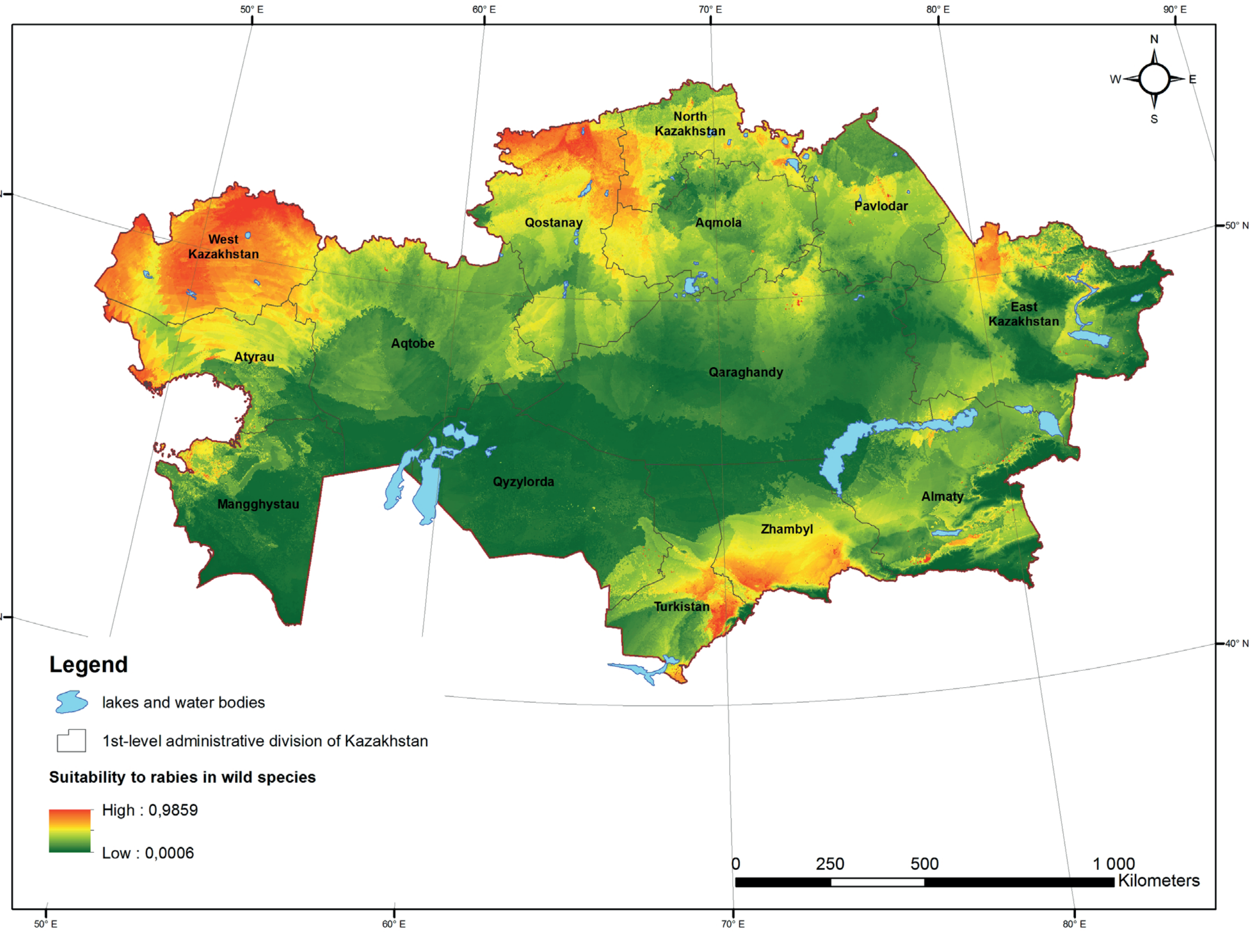


Fig. 3c. Gridded suitability of the RK territory to rabies in wild species

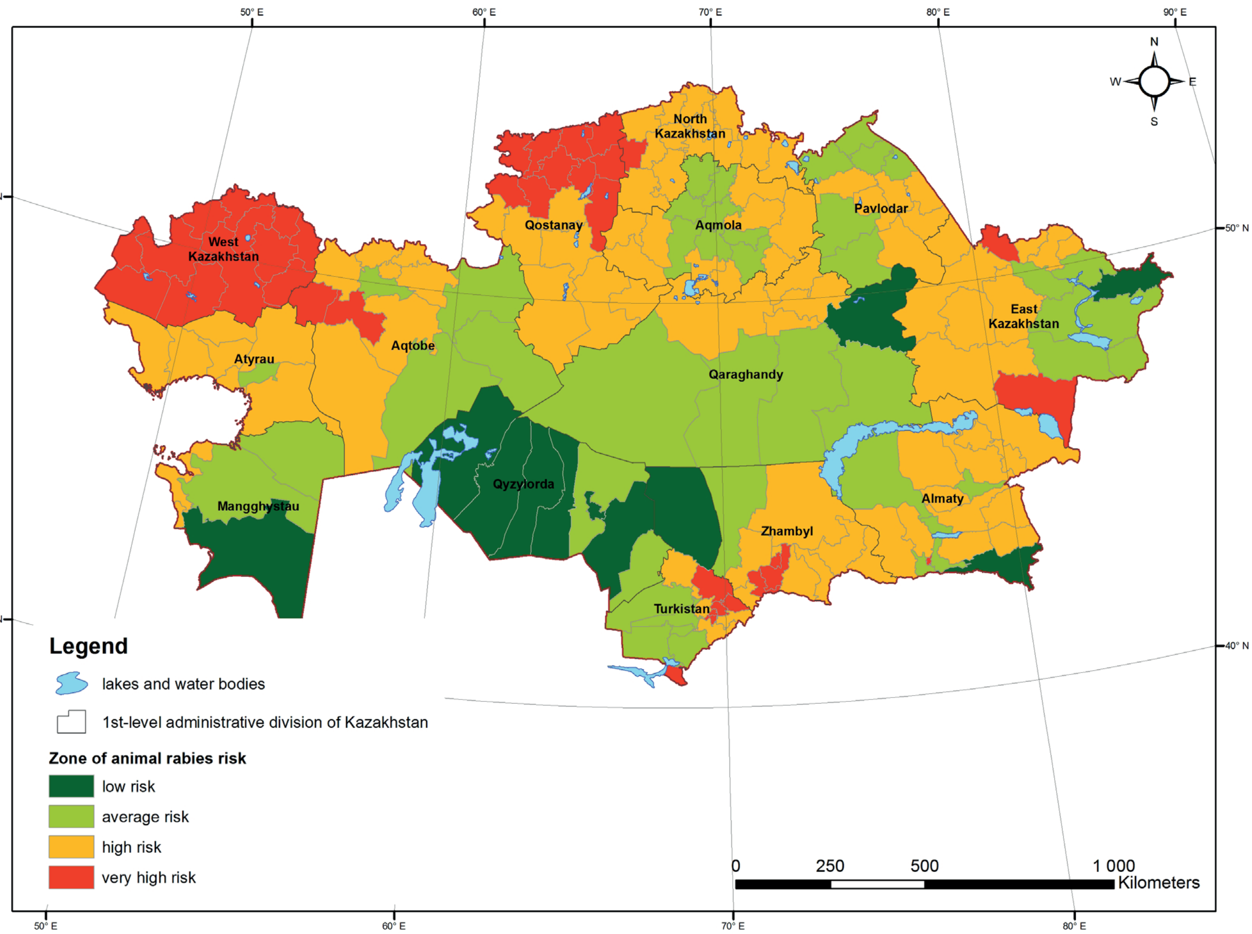


Fig. 4. Integral zoning map of the Republic of Kazakhstan as for risk of rabies in various species

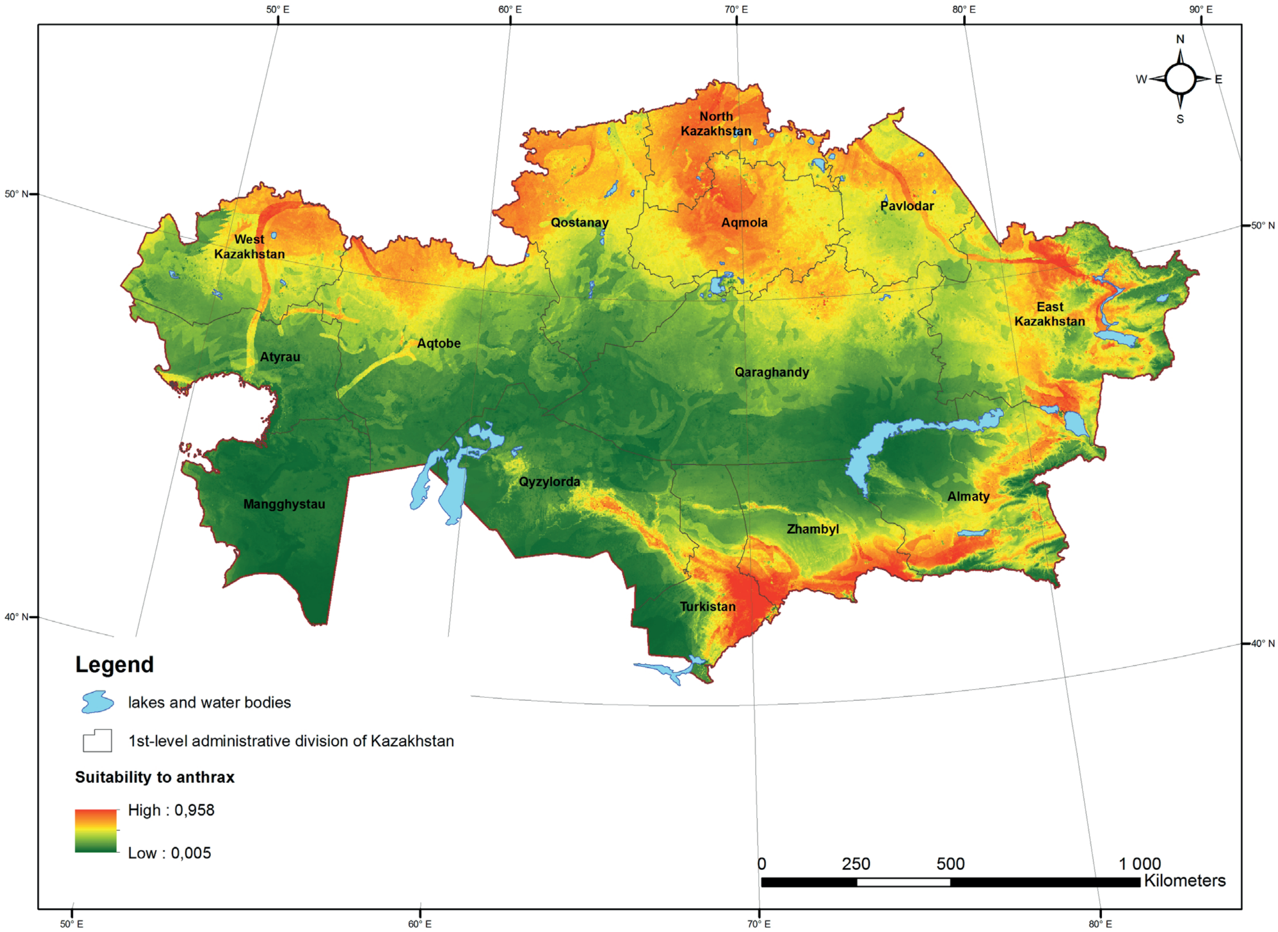


Fig. 5. Gridded distribution of suitability to anthrax in the Republic of Kazakhstan

However, in some studies, the maximum entropy method has been applied to simulate a risk area, in which the occurrence of some disease cases is possible, basing on the previously recorded cases (Stevens and Pfeiffer 2011; Abdrakhmanov et al. 2016).

In our paper, this method is also used to identify areas most at risk of outbreaks of particularly dangerous infectious diseases among animals. Places of registered disease cases are used as "presence locations".

This zoning method is based on the territorial location of the disease and implies the stability of its natural foci, as evidenced by the nature of the diseases, as well as numerous studies. A number of authors used Environmental Niche Modelling methods (Mullins et al. 2011, 2013; Kracalik et al. 2012), which gave similar results with respect to the probability distribution of anthrax outbreaks in the territory of the RK. The method of maximum entropy as used is essentially equivalent to the Poisson regression model, while providing a convenient and illustrative form of presenting the results and evaluating their statistical significance (Renner and Warton 2013).

Different to the studies of other authors, we applied an aggregation of a gridded pattern of the probability distribution by averaging the risk over the second-level administrative units. Such a generalization better fits the traditional practice of the country's veterinary service and gives more visible results in terms of their practical application. That is, the distribution of risk levels by administrative areas allows the country's veterinary service to use its administrative resources more efficiently within each territorial unit, and to make appropriate decisions depending on the current epizootic situation. In the meantime, such a generalization of the risk pattern leads to the loss of more detailed information about the local distribution of risk and specific locations with elevated values. Such information, however, remains available and can be easily displayed.

The key step in the application of the maximum entropy method is the selection of geospatial variables that presumably affect the probability of the disease outbreak being studied. As a rule, the factors under consideration are divided into several main groups depending on the epidemiology of the disease. We identified three main groups of factors: 1) landscape factors; 2) socio-economic factors; 3) climatic factors.

Since rabies and anthrax belong to natural focal diseases, we can assume a strong dependence of the disease cases on a combination of climatic and landscape factors. Moreover, due to the very rapid course of the infection process in infected animals, neither infection tends to spread over long distances but instead remains localized at the site of infection.

The obtained risk maps demonstrate that risks of both studied diseases distributed very similarly to the distribution of population in Kazakhstan. This could have been expected because of predominate tendency of disease diagnostics in close proximity to populated places or other easily accessible locations. In general, districts with medium to high risk can be identified along north, east and south-east borders of the RK. For rabies, very high risk is presented in the most districts of West Kazakhstan region, in Aqtobe and Qostanay regions. These areas do not demonstrate high population density, but provide suitable conditions for livestock breeding, specifically allowing transboundary migration of wild animals that facilitates contacts between livestock and wild species. For anthrax, very high risk is concentrated in high populated districts of Aqmola, North Kazakhstan and Turkistan regions as well as in some other districts providing suitable combination of soils, vegetation cover and amount of precipitation.

As a potential disadvantage of this method, we can mention a possible underestimation of the risk level in certain areas due of its significant variation, which occurs because of the data averaging. Thus, some areas with risk values close to

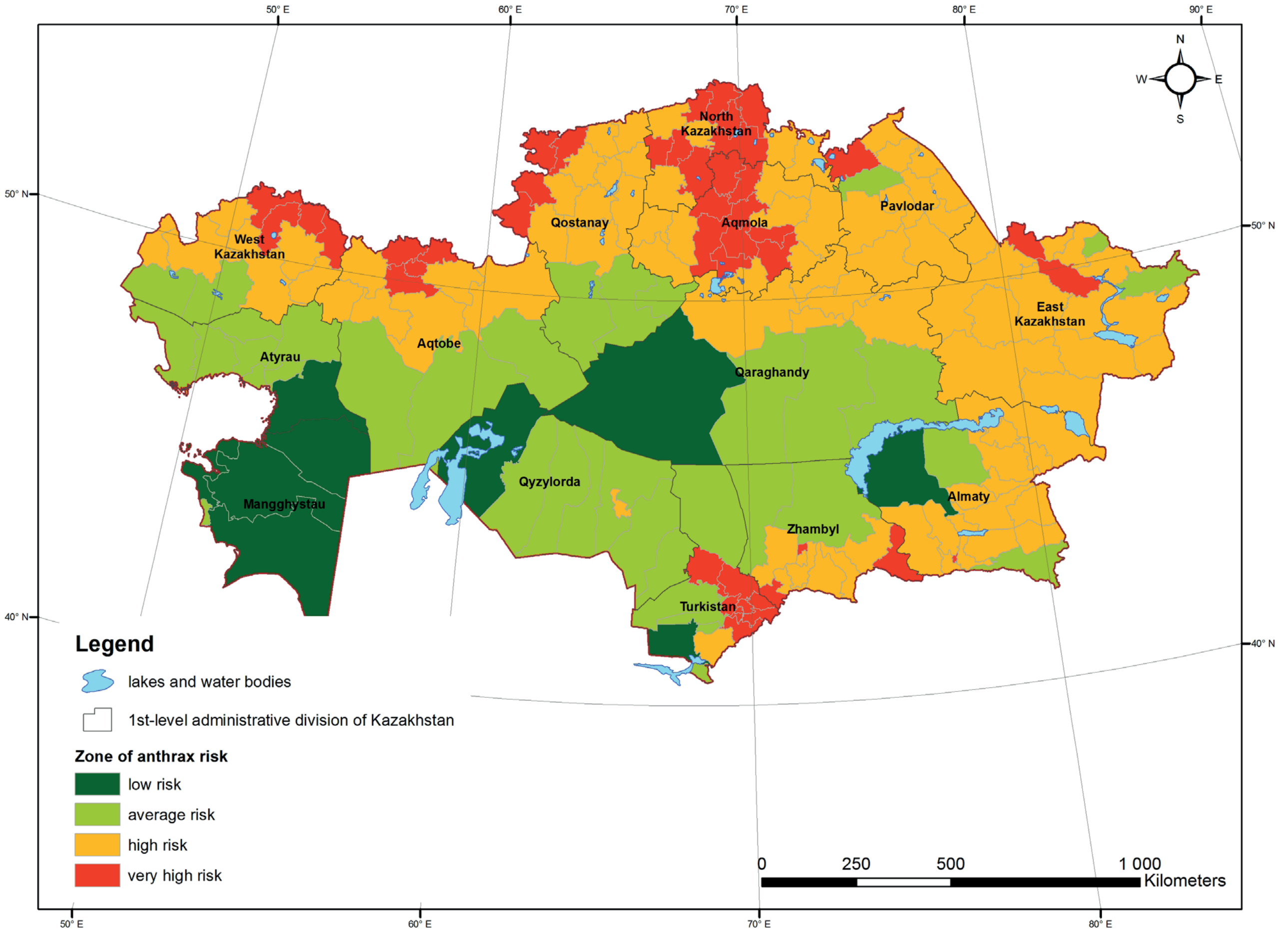


Fig. 6. Integral zoning map of the Republic of Kazakhstan as for risk of anthrax

categories separation may be assigned to a lower risk category. Therefore, for a more detailed epizootological analysis of the territory, it is required to assess the gridded distribution of risk within each administrative unit), also presented in our study.

CONCLUSIONS

The presented study summarizes an approach to veterinary zoning of the country as to the risk of two dangerous zoonotic diseases that is based on the application of Environmental Niche Modelling method with subsequent averaging risk values within administrative divisions. Such an approach allows not only considering presence or absence of historic disease cases throughout the study area but also accounts for potentially suitable areas, where the disease pathogen may survive based on the specific set of environmental factors.

REFERENCES

Abdrakhmanov S.K., Mukhanbetkaliyev Y.Y., Korennoy F.I. [et al.]. (2017). Maximum entropy modelling risk of anthrax in the Republic of Kazakhstan. Prevent. Vet. Med, Vol. 144, 149-157. DOI:10.1016/j.prevetmed.2017.06.003.

Abdrakhmanov S.K., Mukhanbetkaliyev E.E. and Kushubayev D.B. (2014). Visualization and cartographical analysis of the anthrax epizootic situation in Western and South-Western regions of Kazakhstan. Proceedings of KazNIVI "Towards theory and practice of today veterinary science, Vol. LX, 29-35.

Abdrakhmanov S.K., Sultanov A.A., Beisembayev K.K. [et al.]. (2016). Zoning the territory of the Republic of Kazakhstan as to the risk of rabies among various categories of animals. Geospat. Health,Vol. 11, 174-181. DOI:10.4081/gh.2016.429.

Abdrakhmanov S.K., Sytnic I.I. and Tursunkulov S.Zh. (2010). Visualization and analysis of veterinary and geographical rabies spread by using GIS technologies. Proceedings of Mat. V International scient.-pract. conf. (17-18 March 2010). Barnaul: Publishing house – AGAU. Book 3, 283-286 (in Russian);

Adamovich, V.L. and Nikonov, N.N. (1970). The importance of landscape-ecological factors in epizooto;ogy of anthrax. The journal of medical-epizootological researches, 8,113-117 (in Russian).

Aikembayev A.M., Lukhnova L., Temiraliyeva G., Meka-Mechenko T., Pazylov Y., Zakaryan S., Denissov G., Ryan Easterday W., Van Ert M.N., Keim P., Francesconi S.C., Blackburn J.K., Hugh-Jones M. and Hadfield T. (2010). Historical distribution and molecular diversity of bacillus anthracis, Kazakhstan. Emerg. Infect. Dis, 16, 789-796. DOI:10.3201/eid1605.091427

Anderson R.M., May R.M. (1992). Infectious diseases of humans. Dynamics and control. Oxford University Press, USA.

Antyuganov S.N., Ryazanova A.G. and Eremenko, E.I. (2012). Anthrax in the Russian Federation and abroad. Epidemiology and infectious diseases: actual questions, 4, 4-8 (in Russian)/

Araújo M.B., Pearson R.G. (2005). Equilibrium of species' distributions with climate. Ecography (Cop.). DOI:10.1111/j.2005.0906-7590.04253.x.

Bersagurov K.A. (2002). Epizootic and epidemiological situation of rabies in the West – Kazakhstan region and prevention measures. Official bulletin of the State Sanitary and Epidemiological Service of the Republic of Kazakhstan, P. 24-30 (in Russian).

Brown J.L. (2014). SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic, and species distribution model analyses. Methods in Ecology and Evolution. DOI: 10.1111/2041-210X.12200.

Broxton P.D., Zeng X., Scheftic W. and Troch P.A. (2014a).A MODIS-based global 1-km maximum green vegetation fraction dataset. J. Appl. Meteorol. Climatol. DOI:10.1175/JAMC-D-13-0356.1.

Broxton P.D., Zeng X., Sulla-Menashe D. and Troch P.A. (2014b). A global land cover climatology using MODIS data. J. Appl. Meteorol. Climatol. DOI:10.1175/JAMC-D-13-0270.1.

Cherkassky B.L. (1999). Territorial distribution and activity patterns of stationary unsafe anthrax sites. Epidemiology and infectious diseases, 2, 48-52.

Chubirko M.I., Chervanev V.A. and Efanova L.I. (2003). Epizootology of rabies in Voronezh region. In: Veterinary and medical aspects of zooanthroponosis. Pokrov, 102-107 (in Russian).

Digital Soil Map of the World. URL: http://www.fao.org/geonetwork/srv/en/metadata.show?id=14116 [Accessed 24.10.2019]

Domsky I.A. (2002). Natural foci of rabies and its principal hosts. In: Veterinary Pathology, 1, 119-122 (in Russian).

Dudnikov S.A. (2003). Foxes as a marker of risk in a rabies: epizootological aspects. Veterinary and medical aspects of zooanthroponosis. Pokrov, pp. 69-73 (in Russian);

FAO EMPRES-i. – URL: http://empres-i.fao.org/eipws3g/ [Accessed 18.12.2019].

Hijmans R.J., Cameron S.E., Parra J.L. [et al.] (2005). Very high resolution interpolated climate surfaces for global land areas. Intern. J. Climatol. DOI:10.1002/joc.1276.

Hugh-Jones M. and Blackburn J. (2009). The ecology of Bacillus anthracis. Mol. Aspects Med. DOI:10.1016/j.mam.2009.08.003.

Kracalik I.T., Blackburn J.K., Lukhnova L., Pazilov Y., Hugh- Jones M.E. and Aikimbayev A. (2012). Analysing the spatial patterns of livestock anthrax in Kazakhstan in relation to environmental factors: a comparison of local (Gi*) and morphology cluster statistics. Geospatial Health, 7(1), 111-126.

Lukhnova L.Yu., Pazylov E.K., Meka-Mechenko T.V., Sarmantaeva A.B., Izbanova U.A., Myrzabekov A.M., Tuleuov A.M. and Ospanova G.M. (2013). The analysis of epizootological and epidemiological anthrax processes in Kazakhstan, 2002–2012. In. "Life with no threats". Moscow. Vol VIII, 2, 66-72.

Makarov V.V., Sukhareva O.I., Gulyukin A.M. and Litvinov O.B. (2008). The trend of the rabies spread in Eastern Europe. Veterinary Medicine. 7, 20-22 (in Russian).

Maxent software for modelling species niches and distributions. URL: https://biodiversityinformatics.amnh.org/open_source/maxent/ [Accessed 01.11.2019].

Mullins J., Lukhnova L., Aikimbayev A., Pazilov Y., Van Ert M. and Blackburn J. (2011). Ecological Niche Modelling of the Bacillus anthracis A1.a sub-lineage in Kazakhstan. BMC Ecology, 11, 32. URL: http://www.biomedcentral.com/1472-6785/11/32.

Mullins J.C., Garofolo G., Van Ert M., Fasanella A., Lukhnova L. [et al.]. (2013). Ecological Niche Modelling of Bacillus anthracis on Three Continents: Evidence for Genetic-Ecological Divergence? PLoS ONE 8(8): e72451. DOI:10.1371/journal.pone. 0072451.

Norstrøm M. (2001). Geographical Information System (GIS) as a Tool in Surveillance and Monitoring of Animal Diseases. Acta vet. scand. 95, 79-85.

Nouvellet P, Donnelly CA, De Nardi M, Rhodes CJ, De Benedictis P, Citterio C, Obber F, Lorenzetto M, Pozza MD, Cauchemez S and Cattoli G. (2013). Rabies and canine distemper virus epidemics in the red fox population of northern Italy (2006–2010). PLoS One, 8(4).

Orłowska A., Smreczak M., Trębas P. and Żmudziński J.F. (2011). Rabies outbreak in Małopolska region in Poland in 2010. Bull. Vet. Inst. Pulawy, 55 (4), 555-809.

Phillips, S.J., Dudík M. and Schapire R.E. (2004). A maximum entropy approach to species distribution modelling. Proc. Twenty-First Intern. Conf. on Machine Learning, 655-662.

Renner I.W., Warton D.I. (2013). Equivalence of MAXENT and poisson point process models for species distribution modelling in ecology. Biometrics. DOI:10.1111/j.1541-0420.2012.01824.x.

Robardet E., Ilieva D., Iliev E., Gagnev E., Picard-Meyer E., Cliquet F. (2013). Epidemiology and molecular diversity of rabies viruses in Bulgaria. Epidemiol Infect, 5, 1-7.

Sanzybayev, E.B. (2003). Modern features of the natural type rabies. In: Proceedings of the International Scientific-Practical Conference, Almaty, 10, 201-203 (in Russian).

Smreczak M., Orłowska A. and Żmudziński J.F. (2009). Rabies situation in Poland in 2008. Bull. Vet. Inst. Pulawy, 53 (4), 557-843.

Smreczak M., Orłowska A., Trębas P. and Żmudziński J.F. (2012). Rabies epidemiological situation in Poland in 2009 and 2010, Bull. Vet. Inst. Pulawy. 56 (2). pp. 115–266.

Stevens K.B. and Pfeiffer D.U. (2011). Spatial modelling of disease using data- and knowledge-driven approaches. Spat Spatiotemporal Epidemiol, Vol. 2, 125-133. DOI:10.1016/j.sste.2011.07.007.

USGS EarthExplorer Home. (2019). URL: https://earthexplorer.usgs.gov/ [Accessed 24.10.2019].

Youla A.S., Traore F.A., Sako F.B., Feda R.M. and Emeric M.A. (2014). Canine and human rabies in Conakry: epidemiology and preventive aspects. Bull SocPatholExot, 7, 19-21.

Zavodskih A.I. and Sludov A.I. (2007). The behavior of raccoon dogs infected with the rabies. Veterinary Medicine, 2, 15-16 (in Russian).

Zholshorinov A.Z. and Sansyzbayev Y.B. (2004). Surveillance of rabies in a wild-type centers of domination: the methodical recommendations. Astana, (in Russian).

Received on January 1th, 2020 Accepted on January 21th, 2020

Dmitry V. Zhitin[1], Stanislav S. Lachininskii[1,2*], Anna A. Mikhaylova[2,3], Alexander V. Shendrik[1]
[1] Saint-Petersburg State University, St. Petersburg, Russia
[2] Immanuel Kant Baltic Federal University, Kaliningrad, Russia
[3] Saint Petersburg Electrotechnical University «LETI», St. Petersburg, Russia
***Corresponding author:** lachininsky@gmail.com

URBAN TRANSFORMATION OF A POST-SOVIET COASTAL CITY: THE CASE OF SAINT PETERSBURG

ABSTRACT. At the turn of the XX-XXI centuries, post-socialist cities of Europe experienced an active transformation of their socio-demographic and economic structure. A striking feature of post-Soviet cities was the preservation of the disproportionate weight of industry in the economy against the background of a long absence of the real estate market. This phenomenon highlighted the need to solve the problems of socio-economic inequality within the city and restructuring its economy. This is especially true for Russian cities experiencing the shifts in the territorial structure of the population under the influence of transition to market economy, the third industrial revolution and the change of economic- geographical location. This study focuses on identifying trends in the social segmentation of the urban space of St. Petersburg as the second largest city in Russia and a socio-economic center of national importance. The social stratification of the city was studied at the grassroots administrative and territorial level based on the assessment of spatial distribution and the formation of territorial groups of the population with certain qualitative characteristics. The object of the study was 111 municipalities of St. Petersburg. The dynamics of their five most important indicators of demographic, social and economic development in 1989–2018 was analyzed: real estate tax on individuals per capita; the proportion of entrepreneurs; own incomes of municipalities per inhabitant; the proportion of people with an academic degree; cost of housing. Using the rank method, a social welfare rating was compiled. Information for the study was taken from the materials of the general urban planes of St. Petersburg in 1966, 1987 and 2005, the All-Russian population censuses of 2002 and 2010, the databases of the Federal State Statistics Service and the Federal Tax Service, and from the real estate «CIAN» company. The increasing social segregation by income was revealed. The existing differentiation of municipalities in terms of welfare is shown. The poorest are the municipalities of the southern part of the city (Kolpinsky, Nevsky, Krasnoselsky districts and Kronshtadt), while the most prosperous are the municipalities of Petrograd and Central districts, as well as certain territories of the municipal district of Moskovskaya Zastava, the villages of Komarovo, Repino and Solnechnoe.

KEY WORDS: coastal agglomeration, urban zoning, population structure, resettlement determinants, city, St. Petersburg

CITATION: Dmitry V. Zhitin, Stanislav S. Lachininskii, Anna A. Mikhaylova, Alexander V. Shendrik (2020) Urban Transformation Of A Post-Soviet Coastal City: The Case Of Saint Petersburg. Geography, Environment, Sustainability, Vol.13, No 1, p. 145-158
DOI-10.24057/2071-9388-2019-77

INTRODUCTION

The post-socialist cities of Europe, including the largest cities of Russia, are a special phenomenon for research. Traditionally they differ from western cities in terms of their spatial features, morphology and social character, demonstrate a lower level of social segregation and marginalizaion, while also offering safer living environments (Hirt 2012). In 2007, Kirill Stanilov (2007) noted that these cities can develop as part of one of four types of urbanization: Western European, North American, East Asian, or the type typical of third world countries. According to him, the post-socialist cities combine some features of all four types of urbanization. In the 1990-2000s, business activity increased in the central districts of the cities under consideration, which is typical of the Western European type, and at the same time, there is uncontrolled growth of suburbanization and the all-encompassing nature of privatization processes in the urban economy, which is typical for the North American type. Over the past 20 years, several major studies have been published devoted to certain aspects of the transformation of the social, the demographic and economic structure of the population of post-socialist cities in Europe (Berki 2014; Szafranska 2015, 2019; Hirt 2012; Keivani et al. 2001; Kemper 1998; Kotus 2006; Musterd et al. 2017; Ouředníček 2009; Sailer-Fliege 1999; Stanilov 2007).

Transformation of the socio-demographic and economic structure of the post-socialist cities of Russia is defined by two main factors – the abandonment of the planned economy in favor of a market economy model and the transition to the third industrial revolution in the context of building up a new information economy. The former brought about changes in the social and economic structure of the population, the structure of employment and property in the city, and the latter – transformed all aspects of the life of the urban community from production and residential areas to employment and leisure.

Transformation of the territorial structure of cities, change of economic functions of urban space, social stratification and segregation, gentrification and suburbanization in the United States and Western Europe have been actively studied since the mid-1980s. (Humphery and Skvirskaja 2012; Kenneth 1985; Pacione 2005; Smith 1996). Today, special attention is paid to the issues of socioeconomic inequality in large agglomerations, including post-socialist cities of Europe, due to the growing concerns for social stability (Fedorov et al. 2018; Musterd et al., 2017). Studies suggest that socio-economic segregation has intensified, which is determined by social inequality, problems of globalization and economic restructuring, and the transformation of social and housing opportunities. Post-socialist cities of the former USSR are a particularly interesting subject for research in this sense. The territorial structure and the development pattern of these cities differ markedly from the cities of Western Europe and North America, which is due to a long period of absence of the real estate market (Bertaud 2013). Until the early 1990s, disproportionate weight of the industrial component in the urban economy remained in a number of cities, which has led to difficulties in the transition to a market economy model and the third industrial revolution (Ruble 1990).

In the 1990-2000s, several key studies were published that revealed the features of agglomeration, suburbanization and differentiation of the urban space of Moscow (Borodina 2017; Brade et al. 2013; Kirillov et al. 2009; Kurichev et al. 2018; Mahrova 2014, 2015; Mahrova et al. 2012; Mahrova et al. 2015; Pavljuk 2015). For instance, the study conducted by Mahrova and Kirillov (2015) concludes that the post-Soviet Russia developed a new model of urbanization, where the housing sector became an important indicator and factor of development. During this period, St. Petersburg also attracted the attention of researchers, although to a lesser extent. In particular, we will mention the study by Aksenov et al. (2006) on the transformation of the city's urban space in 1989–2002. Some scholars focused on the development of industry and tertiary sector in St. Petersburg addressing their impact on the urban environment (Aksenov et al. 2001, 2016, 2018; Bater 1980; Bater et al. 2000). They demonstrated that the transformations that took place in the post-Soviet period influenced not only the architecture and topography of St. Petersburg but also the socio-geographical makeup of the population, by means of – among everything else – altering the placement parameters of various social groups. In fact, uncontrollable de-industrialization and economic primitivization of the urban economy developed primarily against the background of the transformations of 1990–2000s mentioned above. Zubarevich (2017) emphasizes that in the 1990s, St. Petersburg lagged far behind the national capital in terms of population income. Convergence occurred only in the first half of the 2000s when St. Petersburg received target support from the federal authorities. As for the socio-demographic and spatial structure of the city and its agglomeration, we can recall a number of dedicated studies (Bugaev 2015; Degusarova et al. 2018; Reznikov 2017; Hodachek 2017; Krykova 2016), including those held by the authors of this article (Anokhin et al. 2017; Zhitin 2015). These studies analyze the change in population during the inter-census period, as well as the transformation of the social structure of the population.

This article sets itself the task to fill the existing gap in the complex socio-geographical studies devoted to the transformation of the socio-demographic and economic structure of the population of cities at the turn of the 21st century using an example of the St. Petersburg – the second largest city of Russia. The aim of the study is to identify the trends in the social segmentation of urban space. Particular attention is paid to the spatial divergence of the demographic development of the metropolis in the context of the post-socialist transition, which caused the growth of social inequality.

MATERIALS AND METHODS

Social inequality is an integral feature of the development of human society. The differences that exist in the employment, the level of income and consumption, the availability of spiritual and material goods of decent quality most fully and objectively characterize a society. With that the most important aspect of studying the social stratification of society is its spatial distribution, the formation of territorial groups of the population with certain qualitative characteristics. There are quite a few indicators of social stratification, all of which can be combined into several groups, reflecting the level income, education, professional occupation, compliance with certain standards of behavior. The last of the listed groups of indicators is typical for individual ethno-confessional and sociocultural communities and is most difficult to quantify. In the post-industrial society, there is no voluntary segregation on a professional basis – cohabitation within the framework of guild corporations was typical for the cities of the Middle Ages, but not for the cities of modernity.

Population income and education are most accessible and convenient indicators for studying social stratification of the urban population. Unfortunately, not all data reflecting population income and, as a result, their level of welfare, is available for analysis. In the framework of the 2010 all-Russian population census, there was no question asked about the volume of citizens' income. This information is also missing in the Municipal Education Indicators Database of the Russian Federation, published by Rosstat annually since 2006. The studies of the standard of living and population income conducted by Rosstat do not allow comparing these indicators at the local level (by municipal units of the constituent entities of the Russian Federation) due to the limited sample used.

Therefore, among the indicators of transformation of the socio-demographic and economic structure of the population of St. Petersburg available for analysis at the lower administrative-territorial level of municipalities and districts, five key demographic, social and economic indicators are selected, data on which can be obtained for all or most of the 111 municipalities:

Real estate tax of individuals, per capita.

The first indicator enables comparison of the average value of the real estate (housing, garage, non-residential premises) owned by citizens in the territory of the municipality. As a rule, this refers to residential houses, the main financial asset for the majority of Russians. The average value of the property tax paid per resident of a municipality allows one to determine the level of welfare of the population of a territory.

Proportion of individual entrepreneurs to total employed population.

The share of employers[1] in the total employed population in the economy makes it possible to make inferences about the distribution of a higher income level at a given point in time. It is, of course, possible for an employee of a company to have a bigger income than that of an entrepreneur, since a top manager of a large company will earn more than a self-employed individual, yet in the total aggregate of categories «employee» and «employer» the income level is higher for the latter.

Own income of municipalities, per capita.

In St. Petersburg, municipal budgets make up only about 2% of the consolidated city budget, and due to the limited powers of local governments have little effect on the socio-economic situation of their residents. However, the improvement of the territory is the prerogative of local authorities and the responsibility of municipalities. Thus,

[1] In the 2010 census form, this category is described as "self-employed (in their own enterprise or organization, in their own business) with the involvement of employees".

the size of budget revenue of municipalities is one of key indicators of the social well-being of a particular territory selected for the comparative analysis.

Proportion of scientific degree holders to total population.

Today, when higher education is becoming truly widespread, the number of college or university graduates can no longer be considered an important indicator of social stratification of society. Thus, in addition to the previously noted territorial distribution of the population with higher education, the concentration of persons with scientific degrees is considered[1] as an indication of a certain social status. The mere presence of a degree does not indicate that its holder belongs to an «elite group», but a higher proportion of citizens who have completed postgraduate education can be considered as an indicator of the social well-being of a particular residential area. Quality characteristics of urban environment are predominantly shaped by the people living in a territory. Like other information on the level of education, information on the number of scientific degree holders is collected in the Russian Federation only during population censuses.

Cost of real estate.

For most of the population, private accommodation has become the main financial asset and indicator of living standards. Therefore, market value of real estate was chosen as an important indicator of the social stratification of society. In the Soviet period, the division of «rich» and «poor» areas in Leningrad was rather conditional, although even then many city districts differed in the level of comfort and, accordingly, social composition of the population. The so-called «Stalinkas» – houses with improved layout built in the 1930-50s – had a much higher consumer value than «Khrushchevkas» and «Brezhnevkas» built in the 60-70s. In the 1990s, with the transition to market relations, territorial division of urban societies began to acquire the character of differentiation, primarily in terms of income.

The dynamics of indicators selected was analyzed for the period from 1989 to 2018. Territorial shifts in the distribution and structure of the population of St. Petersburg and its immediate suburbs were estimated broken down by the administrative-territorial level of municipal districts. A comparison of the five indicators made it possible to rank the municipalities of St. Petersburg in terms of social well-being. To do this, we ranked the municipalities for each of the indicators in descending order of value. For four indicators (real estate tax; share of individual entrepreneurs; own incomes of municipalities per capita; share of persons with scientific degrees), the rank value varied from 1 to 111, for one (real estate cost) – from 1 to 99, as for 12 municipalities of St. Petersburg data on the commercial value of one square meter of housing was not available. The ranks acquired were summed up and then the average rank value was calculated[2]. The minimum value of the obtained indicator (i.e. the average rank of social well-being – ARSW) is inherent in the most socially prosperous municipalities of the city. Conversely, the maximum value of the average value of the rank is typical for areas with the lowest standard of living. In theory, the ARSW of the municipalities of St. Petersburg can vary from 1 to 111.

The study relied on the materials of the master plans of St. Petersburg in 1966, 1987 and 2005, the data of the all-Russian population census of 2002 and 2010, the database of the Federal State Statistics Service, the database of the Federal Tax Service, as well as the largest real estate database in Russia and St. Petersburg, CIAN.

RESULTS AND DISCUSSION

Territorial changes in the settlement of the population of St. Petersburg

By the end of the Soviet period, the spatial structure of the city was composed of several large isolated residential areas. The central part of the city shaped by the end of the 19th century was surrounded by a belt of industrial enterprises and storage areas cut by transport corridors in the southern (Peterhof and Tsarskoye Selo (Moscow) avenues, Baltic, Warsaw, Tsarskoye Selo (Vitebsk) and Moscow railways) and the northern (Finnish railway) directions. By the end of the Soviet period, the spatial structure of St. Petersburg was set as a collection of 10-12 semi-isolated residential areas separated by wide strips of non-residential areas – industrial zones, radial transport corridors, green spaces (Fig. 1).

Fig. 1. View of St. Petersburg from space

[1] *Russian postgraduate degrees of candidate of science and doctor of science (habilitation).*
[2] *For 99 municipalities of St. Petersburg, the sum of the ranks is divided into five, and for 12 municipalities – by four.*

The master plan of 1966 for the development of Leningrad was dedicated to solving the housing problem and was perhaps the most successful in accomplishing this objective. The next master plan was adopted in 1987 and extended not only to Leningrad, but also to the Leningrad region, with the city and the surrounding region being considered as a single territory. It was planned to transfer part of the industrial production from the city to the region, securing a significant area of city land for housing. However, in the early 1990s, St. Petersburg, like the rest of the country, plunged into a deep crisis caused by the transformation of the political and socio-economic system, and most of the activities outlined in the master plan of 1987 were not implemented. The latest Master Plan for the Development of St. Petersburg was adopted in 2005. Unfortunately, in the new economic conditions, the possibility of pursuing a unified town-planning policy turned out to be quite difficult since it is necessary to take into account the interests of a large number of stakeholders. Over the past years, numerous changes have been repeatedly made to the master plan, distorting the original idea.

The St. Petersburg megalopolis, including the suburban areas (administratively included in St. Petersburg (Leningrad) since the late 1980s) and the areas of the Leningrad region adjacent to the city line is characterized by a high degree of unevenness in the placement of the population. The central part of the city that occupied less than 5% of its area in 1989[1] had almost 20% of total inhabitants. The population density in Vasileostrovsky, Admiralteysky and Central regions[2] was 14–19 thousand people per sq.km. This compact territory was surrounded by a belt of residential areas (Ozerki, Commandantsky, Grazhdanka, Rzhevka, Porokhovye, Kupchino, Rybatskoye, Sosnovaya Polyana, etc.), which emerged in the 1970s-80s. By the end of the Soviet era, 69% of the city's population lived in these areas[3]. The third belt of residential areas of St. Petersburg (Leningrad) consisted of suburban villages[4] and towns[5], administratively included in the borders of the «northern capital». Occupying more than 63% of St. Petersburg's area and only 11.5% of its population, the eastern and northeastern direction of the suburban belt extends beyond the administrative boundaries of the city and extends over part of Vsevolozhsk district of the Leningrad region (Fig. 2).

Coastal districts: 1. Kronshtadtsky – 19.53 sq.km; 2. Kurortny – 268.19 sq.km (marina Zelenogorsk); 3. Primorsky – 109.9 sq.km (marina of Lakhta Center, marina of Lisy Nos); 4. Petrogradsky – 19.54 sq.km; 5. Vasileostrovsky – 16.7 sq.km (Maritime station, Sea Passenger Port); 6. Kirovsky – 47.46 sq.km (Big port of St. Petersburg); 7. Krasnoselsky – 90.49 sq.km (marina Baltiets); 8. Petrodvortsovy – 107.08 sq.km (Bronka port, Port of Lomonosov, Peterhof and Strelna marinas).

Fig. 2. Administrative and structural division of St. Petersburg

[1] *In the administrative boundaries of Leningrad – St. Petersburg.*
[2] *Until 1994, there were two administrative districts (Leninsky and Oktyabrsky) in the territory of the Admiralteisky district, and three (Dzerzhinsky, Kuibyshevsky and Smolinsky) administrative districts in the Central district.*
[3] *Primorsky, Vyborgsky, Kalininsky, Krasnogvardeysky, Nevsky, Frunzensky, Moscovsky, Kirovsky, Krasnoselsky districts of St. Petersburg.*
[4] *Ushkovo, Serovo, Molodezhnoye, Smolyachkovo, Repino, Komarovo, Solnechnoye, Lisiy Nos, Pesochnoye, Pargolovo, Metallostroy, Ust-Izhora, Pontonny, Saperny, Tyarlevo, Aleksandrovskaya, Shushary, Strelna.*
[5] *Zelenogorsk, Sestroretsk, Kolpino, Pushkin, Pavlovsk, Krasnoe Selo, Petrodvorets (Peterhof), Lomonosov (Oraniembaum), Kronstadt.*

Over the following three decades (1989–2018) there was a significant redistribution of the population between these three zones of St. Petersburg. The resettlement of the overcrowded housing of the central part of the city led to the fact that today only 13% of its inhabitants live here. Thus, the population of the former Dzerzhinsky, Smolninsky and Oktyabrsky districts has declined over thirty years by a quarter, the former Leninsky district – by a third, the former Kuybyshevsky – by more than 2 times.

The number of residents of the Petrograd side – one of the most prestigious areas of the city, has decreased by a quarter. Because of the renovation, the «gray zone» of industrial territories is gradually disappearing, separating the historical center from the mass housing areas of the Soviet period. The territory of residential areas of the city have grown due to newly developed ones (Yuntolovo, Bogatyrsky, Kollomyagi – in the north-west, Baltiyskaya zhemchuzhina – in the south-west), as well as «spot buildings» in the existing microdistricts. As a result, the proportion of St. Petersburg residents living in the second zone increased to 71% of the city's population. In the third zone – the suburban area, changes in the population distribution were multidirectional. In some suburban villages and towns, the population has declined quite significantly. For 1989-2018, the population in a number of settlements in the Kurortny district (Ushkovo, Molodezhnoye, Smolyachkovo, Repino) and in the city of Pavlovsk (Pushkinsky district) decreased by 30-40%, in the villages of Komarovo (Kurortny district) and Pontonny (Kolpinsky district) by 16-17%. In others (the city of Kolpino, the city of Kronstadt, the city of Lomonosov, the settlement of Petro-Slavyanka, Solnechnoye, Aleksandrovskaya), the population practically did not change. In general, the proportion of the suburban area over thirty years saw a 1.4 times increase and in 2019 amounted to 15.7% of the total population of St. Petersburg.

Unfortunately, a relatively detailed analysis of the territorial changes in the distribution of the population of St. Petersburg is only possible for the period from the beginning of this century, after the municipal reform in the Russian Federation. Until this time, data on the population of St. Petersburg is only available in the context of 41 urban areas and individual villages. From 2002 on, statistics is available about the population of 111 intra-city municipalities of St. Petersburg[1]. At the same time, the size of the territory and the population of the municipalities of the city varies greatly. Thus, the population of the smallest of the municipality Serovo village (Kurortny district) is only 277 people, and the most populated – the municipal district Kolpino city (Kolpinsky district) – almost 148,000 inhabitants. The area of St. Petersburg municipalities can vary by tens of times (from 0.3 to more than 100 sq. km).

Since 2003, after 12 years of depopulation, the trend of population growth in St. Petersburg has been recovering. As seen in fig. 3, there is a center-peripheral effect in the dynamics of the number of residents of municipalities of the city over the past 16 years. A 15-35% population decline is registered in most of the Admiralteisky and Centralny districts. On the other hand, there has been more than 1.5 increase in the number of residents in the northern (Pargolovo, Yuntolovo, Kolomyagi, Lakhta) and southern (Shushary, Aleksandrovskaya, Zvezdnoye, Gorelovo) municipalities.

During this period, the St. Petersburg agglomeration has finally gone beyond the administrative boundaries of the «northern capital» proper and now includes rural and urban municipalities of several districts of the Leningrad region (Fig. 4).

Thus, over the past thirty years, there has been a redistribution of the population from the central and more established residential districts of St. Petersburg to the outskirts of the city and even beyond its administrative boundaries. The attractiveness of the «northern capital» for migrants from other regions of Russia and from abroad increases the demand for housing. To date, St. Petersburg has

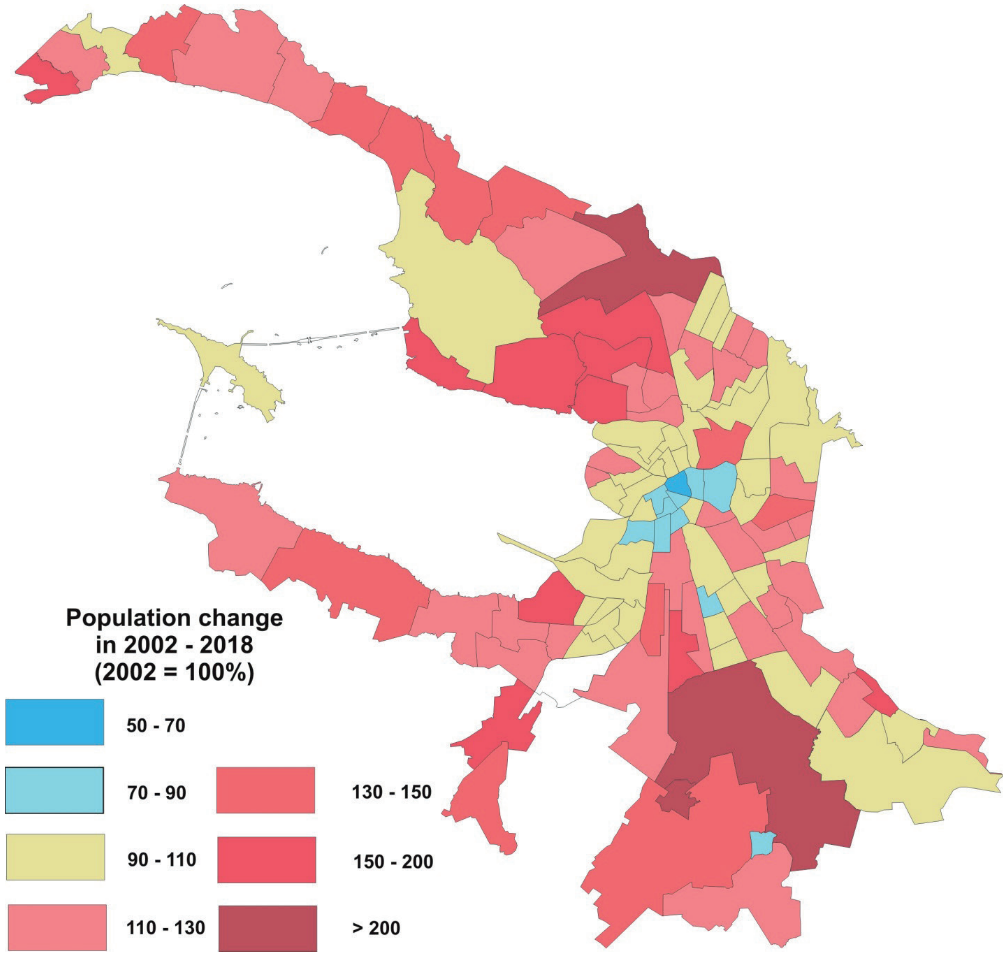


Fig. 3. Population dynamics by municipalities of St. Petersburg in 2002-2018

[1] *81 municipal districts, 9 cities and 21 villages.*

almost completely exhausted the possibilities of new housing construction in existing residential areas. Redevelopment of industrial areas, which has been actively pursued in St. Petersburg in recent years, can only partly address the shortage of land plots for housing development. Therefore, the city is expanding rapidly, de facto including in its composition the territory of the adjacent areas of the Leningrad region. In this regard, preservation and strengthening of the internal transport connectivity of the territory of the megalopolis will become the most acute task of the spatial development of St. Petersburg in the coming years.

«Rich» and «poor» districts of St. Petersburg: the beginning of segregation?

Considering such a large city as St. Petersburg, it is necessary to pay attention not only to the spatial differences of the demographic changes taking place but also to the social and economic heterogeneity of its territory. Let us consider the distribution of these indicators in the territory of St. Petersburg.

Real estate tax of individuals, per capita

According to the data for 2016, the average annual property tax was 434 rubles per capita. Among the municipalities, this indicator varied in almost 20-fold range from 50 rubles per capita in one of the most crowded and remote municipalities – Smolyachkovo village (Kurortny District) – to almost 2,000 rubles in Chkalovskoe municipality located in the center of the city (Petrogradsky district). However, it would be too simplistic to assume that the cost of real estate per one St. Petersburg resident depends only on the distance to the center.

Indeed, the amount of real estate tax paid by residents in most of the central municipalities of St. Petersburg is more than 1.5 times higher than the city average (Fig. 5). In such municipalities as Dvortsovyy district, No. 78, Smolninskoe (all in the Central District), Aptekarsky ostrov, Petrovsky, Chkalovskoe (all in the Petrogradsky District), this indicator is 2-5 times the average. High value of real estate is observed not only in the center of St. Petersburg, but also in a number of municipalities of Primorsky (municipality No. 65, Kolomyagi), Vyborgsky (Svetlanovsky municipality), Moscovsky (Zvezdnoye municipality), Krasnoselsky (Yuzhno-Primorsky municipality), Krasnogvardeisky (Pravoberezhny municipality), Pushkinsky (Aleksandrovskaya rural settlement) districts. All these territories (with the exception of the Svetlanovskoye municipality) have seen active development over the past decade with new residential complexes being constructed.

High volumes of the property tax are also characteristic of a number of municipalities in the Kurortny district of St. Petersburg, stretching over 50 km along the northern shore of the Gulf of Finland. In the several summer house («dacha») settlements (Komarovo, Repino, Solnechnoye, Ushkovo, Molodezhnoye) the tax burden per inhabitant is 2.0 – 3.5 times the city average. At the same time, according to this indicator, the greatest contrasts are observed in the territory of the Kurortny district; along with the richest municipality of St. Petersburg, the poorest are located here. Thus, in the village of Pesochny, property tax payments are almost two times less than the average values in the city, and in the above-mentioned village of Smolyachkovo, by more than 8 times.

Low property tax per capita is characteristic of municipalities located in the old residential («sleeping») districts of St. Petersburg, where the majority of construction was done in the 1960-80s. In most of Kalininsky, Kolpinsky, Krasnoselsky, Petrodvoretsky, Nevsky districts and Kronstadt municipal districts, this indicator is 1.5-3.0 times less than the average value in the city. In this regard, the municipalities of the Krasnoselsky district are especially distinguished – Sosnovaya polyana and Gorelovo, where tax payments for real estate are 3.5 – 4.5 times lower than the average value.

Proportion of individual entrepreneurs to total employed population

The distribution of business owners in St. Petersburg is also characterized by a high degree of unevenness. Unfortunately, complete information on this category of persons is collected only during population censuses and is now somewhat outdated since nine years have passed since the last census. According to the 2010 census, only 2.2% of the employed population of St. Petersburg were entrepreneurs who have employed someone. The largest share of employers was noted in the suburban settlements of the «northern capital»: Ust-Izhora (Kolpinsky district), Repino, Solnechnoye, Komarovo (all in the Kurortny district of the city). In these villages, the share of entrepreneurs is 2-3 times the average value in St. Petersburg (Fig. 6).

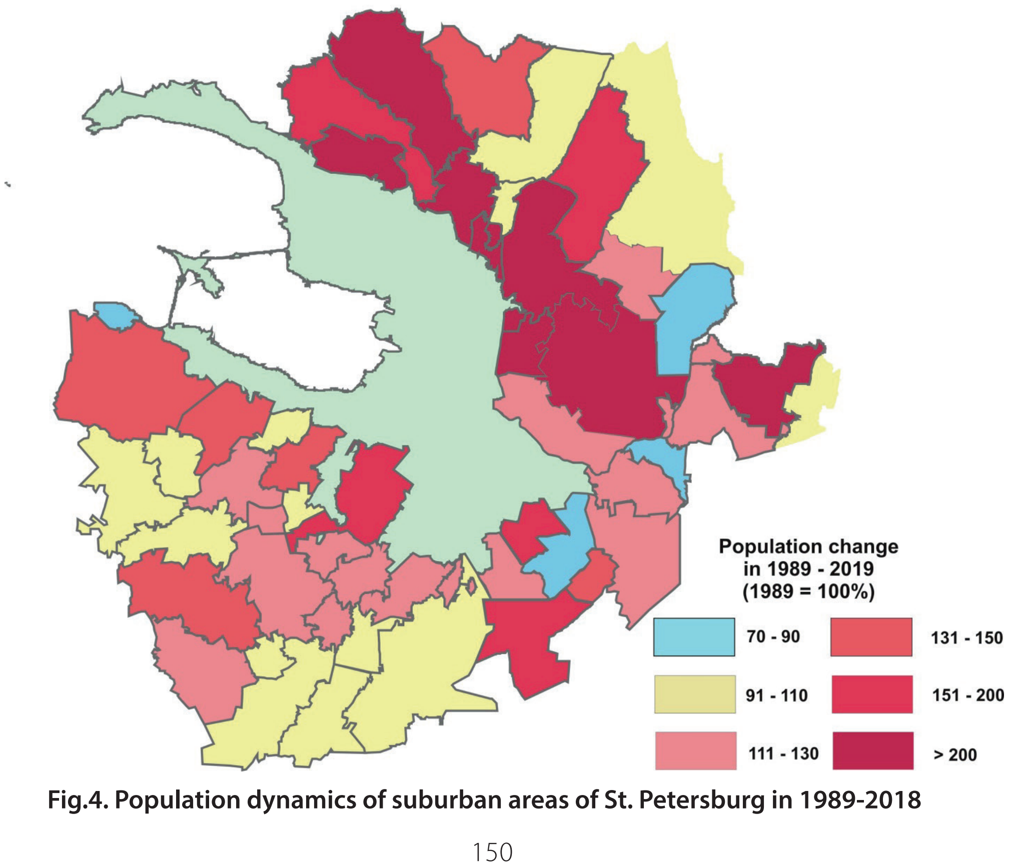


Fig.4. Population dynamics of suburban areas of St. Petersburg in 1989-2018

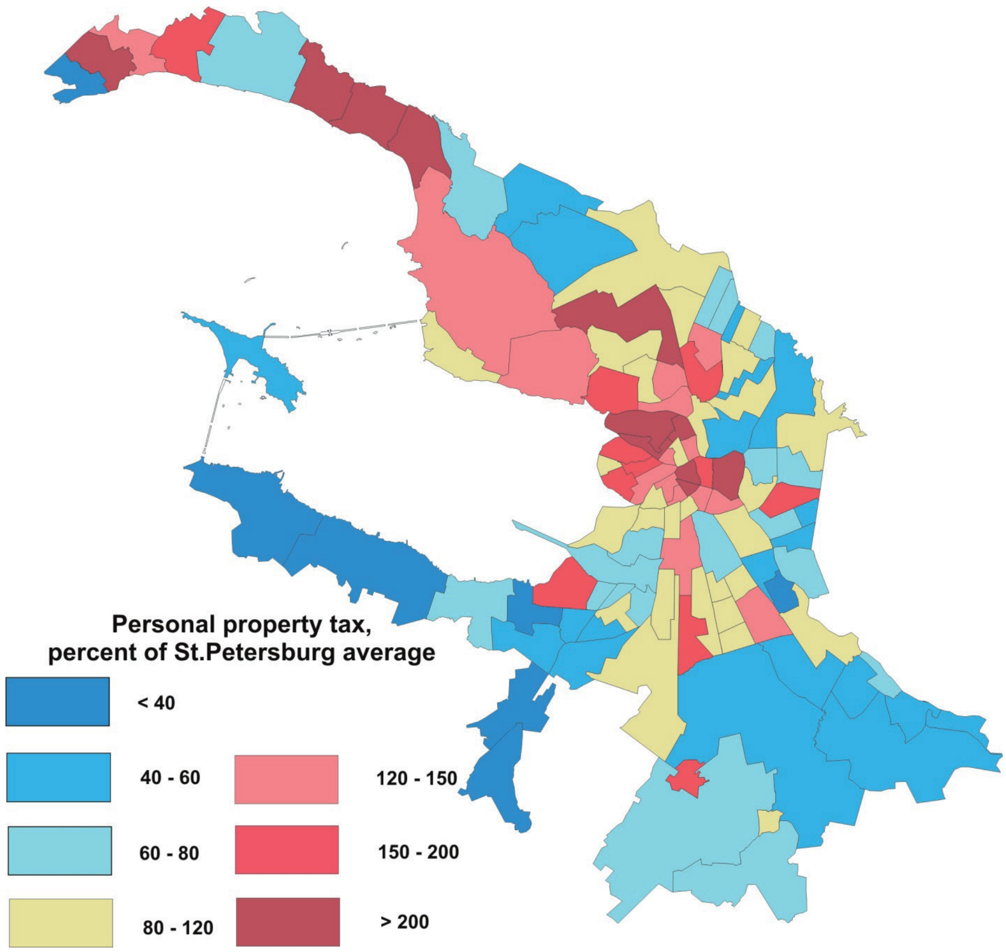


Fig. 5. Property tax in municipalities of St. Petersburg, 2016

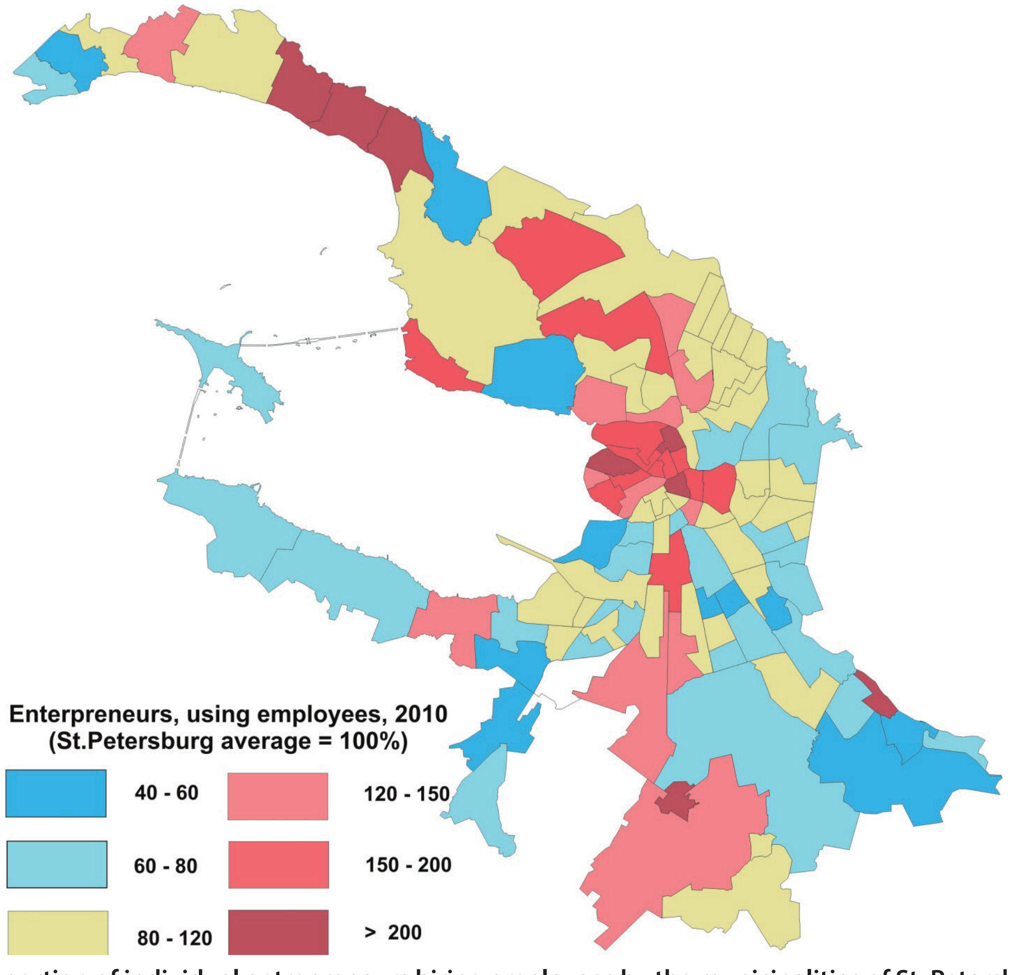


Fig. 6. The proportion of individual entrepreneurs hiring employees by the municipalities of St. Petersburg, 2010

The largest proportion of people hiring workers was registered for the Aleksandrovskaya village of the Pushkinsky district. Here, this category of persons includes almost 10% of the employed population. Among the districts of apartment buildings located directly in the city on the Neva River, more than twice the share of entrepreneurs is observed in the municipal districts of Petrogradsky (Aptekarsky ostrov municipality), Centralny (Petrogradksy municipality) and Vasileostrovsky (Ostrov Decabristov municipality) districts.

Cost of real estate

As has already been noted, for almost three decades of post-Soviet history in St. Petersburg there has been a marked stratification in terms of the quality of housing and its market value. At the same time, the market value of a building and its cadastral valuation, although related to each other, can vary considerably. According to CIAN[1], the cost of one square meter of residential housing at the beginning of 2019 varied across the municipalities of St. Petersburg[2] from 62.2 to 246.5 thousand rubles (Fig. 7).

Today the most expensive housing, expectedly, is located in the central part of the city – in the Petrogradsky, Centralny and Admiralteysky districts. The most fashionable district of St. Petersburg today is the Petrogradsky district (Central district), which occupies the territory of the historic part of the city between the Palace Embankment of the Neva and the beginning of Nevsky Avenue. The average cost of one square meter of housing today is almost 250 thousand rubles (about 3.8 thousand US dollars). In recent years, however, the cost of real estate in the Petrogradsky district, and especially on the territory of Krestovsky Ostrov, which until recently was almost exclusively reserved for recreational functions, has been growing rapidly. Today this territory, which is a part of the Chkalovskoye municipal district, is experiencing rapid development through the construction of luxury residential properties that dominate residential areas around the Nevsky Avenue, the city's main street.

Outside the central areas of the city, high cost of real estate is observed along the main transport artery of St. Petersburg – Moskovsky Avenue – stretching almost 10 km to the south. The axis of expensive housing continues to the north of the city center along the line connecting the Chernaya rechka, Pionerskaya and Udelnaya metro stations. The beginning of Engels Avenue – one of the main transport lines of the northern part of St. Petersburg, and the surrounding neighborhoods at the cost of real estate (on average, 120-130 thousand rubles per 1 sq.m.) being similar to residential areas adjacent to Moscovsky Avenue.

The cheapest housing is located in the industrial suburbs of St. Petersburg, mainly in its southern areas. In the villages of Pesochny (Kurortny district), Pontonny, Saperny (both in the Kolpinsky district), the Krasnoe Selo City (Krasnoselsky district), the cost of one square meter of housing at the beginning of the second quarter of 2019 did not exceed 70 thousand rubles (1.1 thousand US dollars). Figure 7 shows the center-peripheral nature of the distribution of the cost of

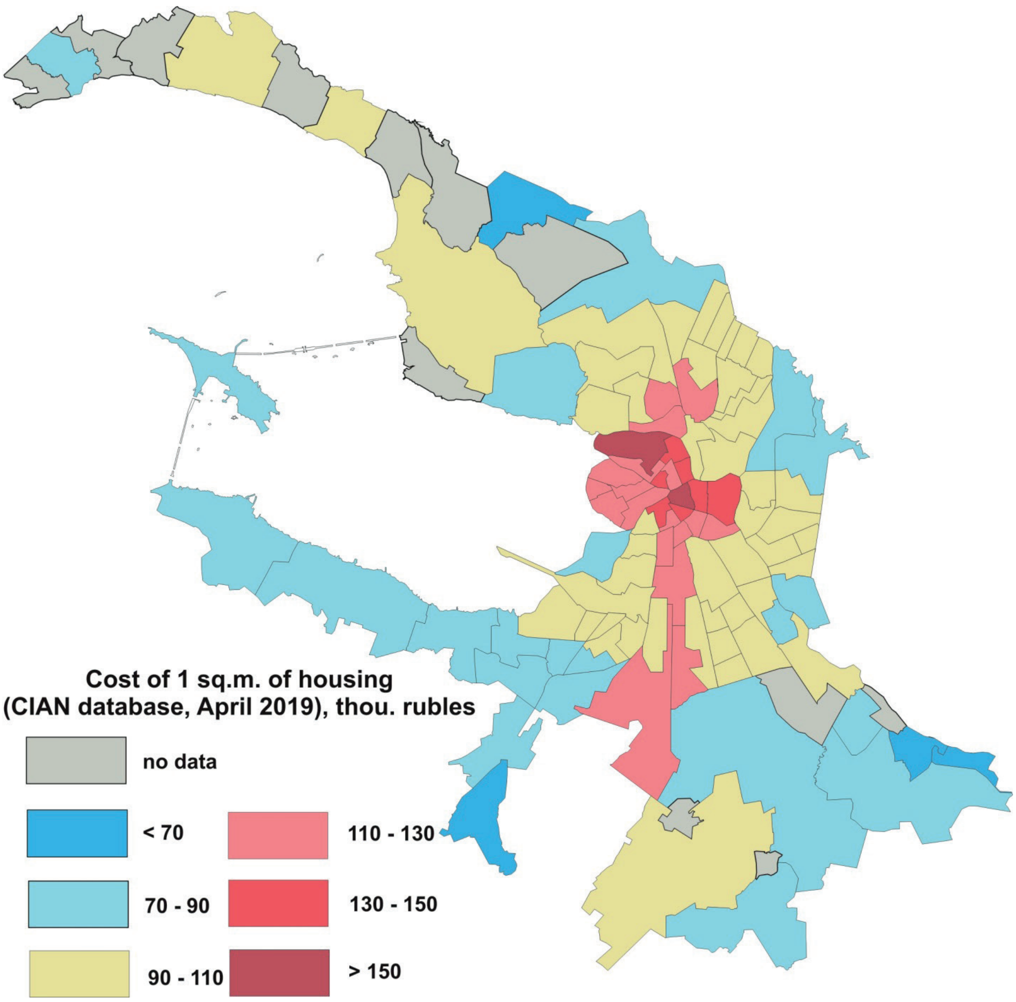


Fig. 7. The cost of housing in the municipalities of St. Petersburg, 2019

[1] A database of real estate – the largest in Russia, in general, and in St. Petersburg, in particular.

[2] There is no data on the cost of real estate in 12 municipalities of St. Petersburg – suburban settlements of Vyborgsky, Primorsky, Kurortny, Kolpinsky and Pushkinsky districts.

housing in St. Petersburg, with the separation of the axis of expensive housing in the direction from south to north.

Revenues of municipal budgets

By virtue of the system of budgeting of municipalities that has taken shape in the Russian Federation and the powers that are given to local governments, it is hard to talk about their financial viability. On average, more than a quarter of all budget revenues of municipalities of St. Petersburg are various transfers (subsidies, subventions, etc.) from higher-level budgets. In almost 20% of the municipalities of St. Petersburg, the share of such transfers exceeds 50% of the revenues of local budgets. At the same time, the amount and the proportion of subsidies to municipal budgets each year can vary in a very wide range. Therefore, to assess the level of well-being of municipalities and their financial independence, we consider the volume of own revenues of local budgets generated mainly from taxes on total income[1], property tax on individuals[2], income from the use of property owned by municipalities.

In 2017, the average size of budget revenues of municipalities of St. Petersburg amounted to 1,870 rubles per capita. In the city, this indicator varied in a very wide range from 767 rubles (Smolyachkovo rural settlement of the Kurortny district) to 18,382 rubles (Lakhta-Olgino municipality of the Primorsky district). Despite the 24-fold difference at the city level, the gap in the per capita revenues of local budget does not exceed 2.5 times for almost 75% of the municipalities of the «northern capital». With that, only some of the most financially secure municipalities are located in the central part of the city (Fig. 8).

The richest municipalities of St. Petersburg in terms of revenues of local budgets are the cities and villages of the suburban areas: Kurortny (Komarovo, Repino, Solnechnoye, Beloostrov, Serovo, Sestroretsk, and Zelenogorsk), Kolpinsky (Saperny, Petro-Slavyanka, and Ust-Izhora) and Pushkinsky (Tyarlevo, Aleksandrovskaya). Low-income municipalities are concentrated in the northern (Primorsky, Vyborg, Kalininsky) and southern (Krasnoselsky, Frunzensky) residential districts of the city.

Level of education of the population

According to the latest census (2010), 56.1 thousand candidates and 12.9 thousand doctors of sciences lived in St. Petersburg. Among the age group of the population over 25 years old[3], 18.4% on average held a scientific degree. The highest concentration of this category of persons was observed in the municipalities of the Petrogradsky and Centralny districts of the city – an average of 34-45 candidates and doctors of science per 1000 residents over 25 years old (Fig. 9).

In addition to the Centralny and Petrogradsky districts, a high (more than 50% of the average value in St. Petersburg) proportion of people with scientific degrees is observed in some municipal districts of Vasileostrovsky (Gavan, Morskoi municipalities), Admiralteysky district

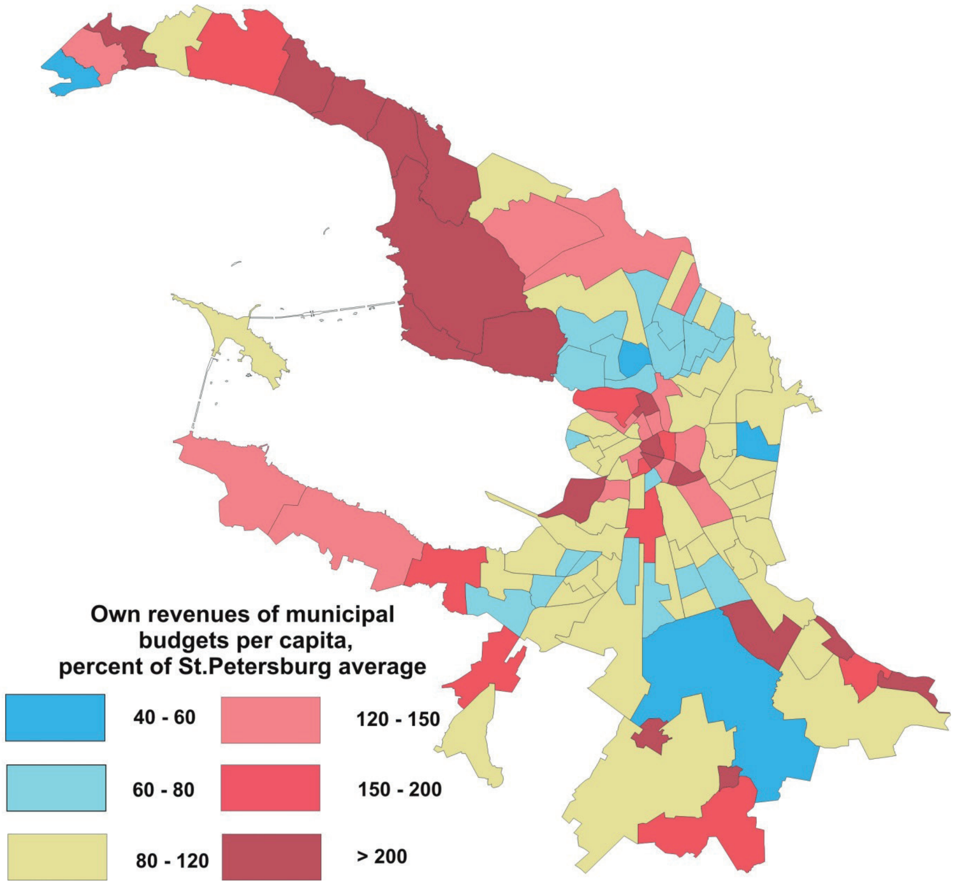


Fig. 8. Own revenues of municipalities of St. Petersburg per capita, 2017

[1] Taxes on gross income include tax levied with the use of a simplified system of taxation of enterprises and organizations; single tax on imputed income for certain types of activities; tax levied for the use of the patent system of taxation; single agricultural tax. Taxes on total income are only partially (the lesser part) transferred to the budgets of intracity municipalities of St. Petersburg. Most of this type of tax goes to the budget of St. Petersburg.

[2] Fully goes to the budgets of intracity municipalities of St. Petersburg.

[3] Given the length of the secondary and higher education levels in the Russian Federation, it is almost impossible to get a degree at the age of less than 25 years. According to the 2010 census, among residents of St. Petersburg younger than 25 years of age there are no persons with a degree of candidate or doctor of science.

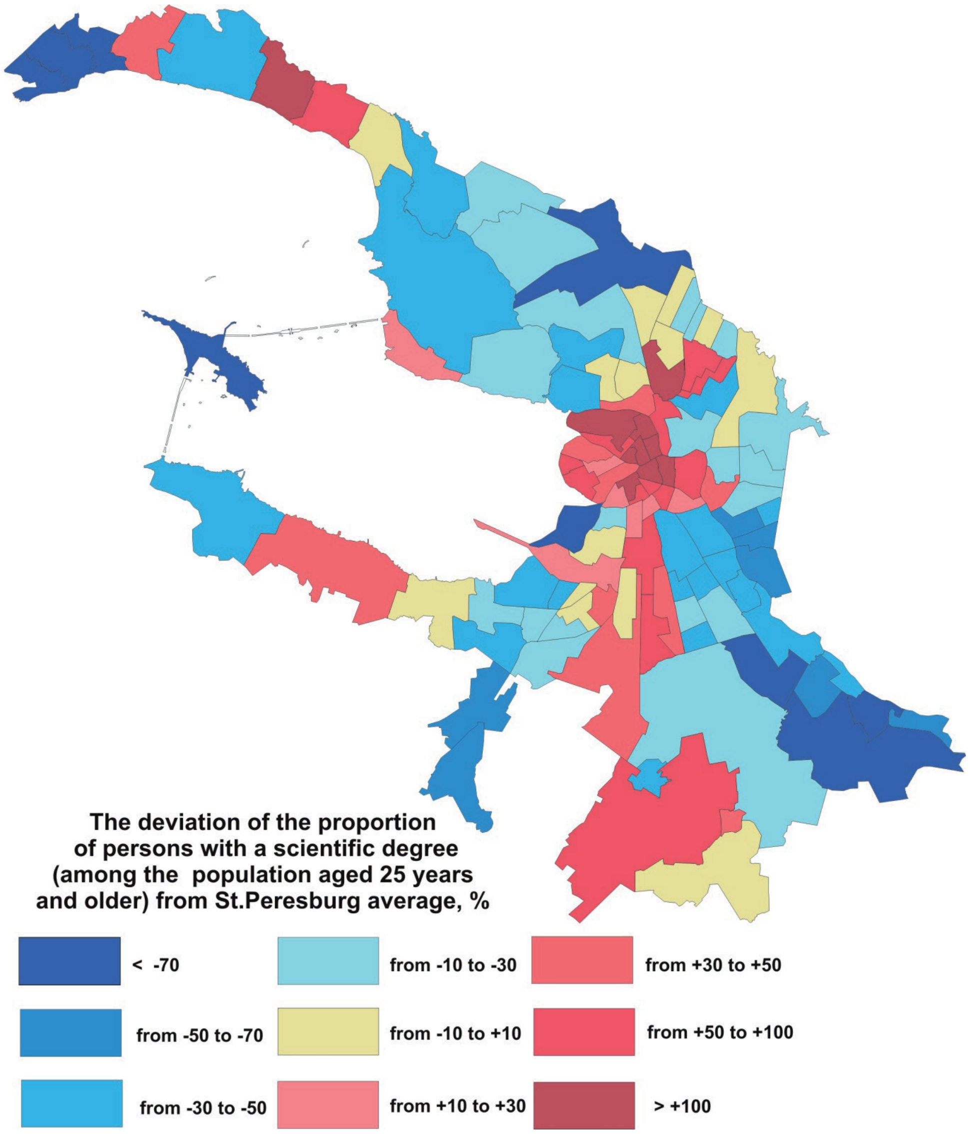


Fig. 9. Concentration of persons with a scientific degree in municipalities of St. Petersburg, 2010

(Sevvoy okrug, Admiralteysky district), Vyborgsky (Svetlanovskoe, Sampsonievskoe municipalities), Kalininsky (Akademicheskoe, Grazhdanka municipalities) and Moscovky (Moscovskaya Zastava, Zvezdnoye) districts. High concentration of academics is traditionally characteristic of such settlements in the Kurortny district of St. Petersburg, like Komarovo and Repino, as well as for the city of Pushkin[1]. At the same time, municipalities of the Kolpinsky, Krasnoselsky, Nevsky and Frunzensky districts are characterized by a very low proportion of residents with a scientific degree.

The calculated level of the average rank of social well-being (ARSW) ranged from 2.6 (Dvortsovyy okrug municipality of the Centralny district) to 107.3 (the Smolyachkovo Village community of the Kurortny District) (Fig. 10).

The most prosperous municipalities of St. Petersburg are located in the city center – in the Petrograd and Central districts. The most respectable territories also include the municipal district Moskovskaya Zastava (Moscow district) and the villages of Komarovo, Repino, Solnechnoye located in the Kurortny district. Municipalities of the southern part of the city – Kolpinsky, Nevsky, Krasnoselsky districts and Kronstadt have the worst indicators for the level of social well-being. Figure 10 shows that the social segregation that has been outlined so far in St. Petersburg has a clear spatial localization – the «poor» and «rich» municipalities are grouped in different parts of the city.

Coastal location and social effects of urban municipalities

The entire territory of St. Petersburg fits within the 50 km coastal zone boundaries. The areas located on the coast of the Gulf of Finland (the inner crescent coastal municipalities) have a direct coastal position: Kronstadtky district (Kotlin Island and the city of Kronstadt), Kurortny district (including the cities of Zelenogorsk and Sestroretsk, as well as a number of villages), Primorsky district, Petrogradsky district, Vasileostrovsky district, Kirovsky district, Krasnoselsky district, Petrodvoretsky district (including the city of Petrodvorets) and the city of Lomonosov.

The the inner crescent coastal municipalities experience various, often conflicting, social and economic effects, including population growth (Primorsky, Krasnoselsky, Petrodvortsovy and Kurortny); reduction or stabilization of the population (Kirovsky, Petrograd, Vasileostrovsky, Kronstadt); high property taxes in «rich areas» (Kurortny, Primorsky, Petrograd) and low property taxes in «poor areas» (Lomonosov, Petrodvoretsky district, Kronstadt district, Kirovsky district); the most socially prosperous (Petrograd) and the least socially prosperous (Kronstadt district, Petrodvorets district, Krasnoselsky district) areas; the richest municipalities by income per capita (Kurortny district) and the poorest municipalities (Primorsky district, Krasnoselsky district, Kronstadt district).

[1] *Pushkin Municipality*

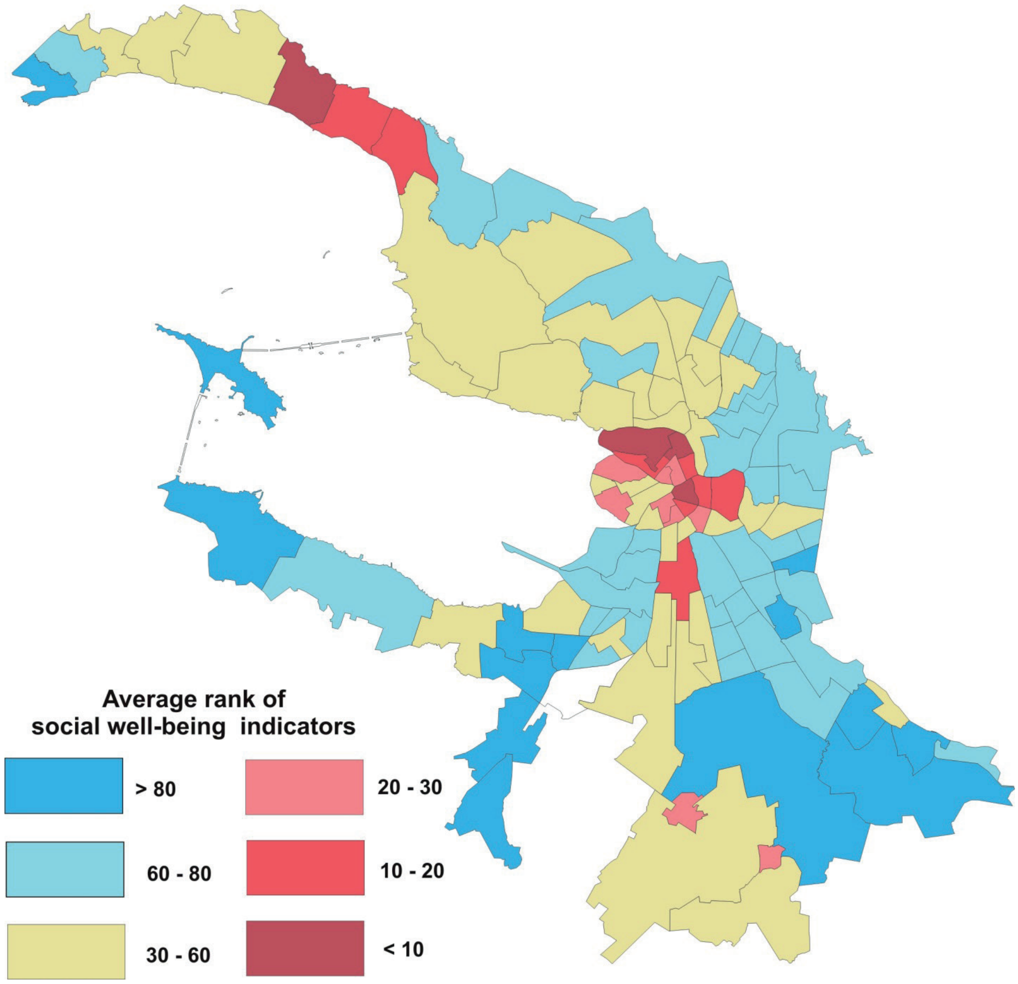


Fig. 10. Social well-being in St. Petersburg municipalities

The most ambitious construction projects of the 2000–2010s brought particular value to these territories. In the zone of coastal urban municipalities: «Western high-speed diameter» (2005–2016, connects the southwest and northwest territories), Gazprom Arena stadium (2007-2017, UEFA stadium of the 4th category); Lakhta Center (2012–2019, Gazprom headquarters); «Baltic Pearl» (a large residential neighborhood in the south-west of St. Petersburg has been under construction since 2005 with the participation of Chinese and Russian developers); «Marine Facade» (passenger sea port in St. Petersburg, located in the western part of Vasilievsky Island); State complex «Palace of Congresses» in Strelna (2000–2003) and others.

The highest real estate value is characteristic of the Petrograd and Vasileostrovsky districts, where the most desired residences are located, including in the coastal zone, as well as in the Kurortny district, with its summer cottages and new townhouses. At the same time, in the remote Kronstadt and Petrodvorets districts, as well as the city of Lomonosov, the cost of real estate is at the lowest level. The presence of a naval base, facilities and warehouse infrastructure negatively affects not only the value of real estate on the territory of the Kronstadt region and the city of Lomonosov and incomes of its population, but also contributes to the marginalization of these territories. In the Soviet period, the city of Kronstadt had the status of a closed city, and the lack of land connection with St. Petersburg led to the fact that until now this area has the strongest negative social effects (since 2011, a dam has been opened that connects Kotlin Island with the north and south coast of the Gulf of Finland and is part of the St. Petersburg ring road).

Experience in the development of European ports shows that today they are betting on the shift of port capacities from the central part of the city and the development of outports. Ports marginalize urban areas, and it seems that St. Petersburg is following this trend. Within the boundaries of St. Petersburg, on the territory of the Kirovsky and Admiralteysky districts, port capacities and infrastructure for 50 million tons are located. In recent years, the development of port facilities in the Leningrad region (the city of Primorsk, the village of Ust-Luga, the city of Vysotsk), as well as in the village of Bronka (the city of Lomonosov) has accelerated. Containers are then shipped inland using the «dry port» technology (Shushary, Yanino, 30 km from the coastline).

CONCLUSIONS

The analysis of a number of key demographic and economic indicators for the period from 1989 to 2018 shows the increasing importance of center-peripheral trends in the development of the city and its immediate environment. These trends suggest that St. Petersburg is no exception in the series of post-socialist cities of Eastern Europe, which are undergoing profound socio-demographic and socio-economic changes, both in general and at the level of the spatial structure of the city. It can be stated that in accordance with the approach of K. Stanilov, St. Petersburg is developing in line with the West European and North American types of urbanization. The central districts of the city are losing population, but are experiencing an increase in business activity – offices of Russian and foreign companies, banks, business centers, and hotels. At the same time, due to the introduction of information technologies, the quality of public services in the internal residential areas is increasing

Legend: Top left: Lakhta Center and Gazprom Arena; Top right: Marine Facade; Bottom left: Resort area; Bottom right: Western High-Speed Diameter and Gazprom Arena

Fig. 11. Key construction projects of the 2000-2010s

and large multifunctional shopping complexes are emerging, including leisure, food, trade, domestic services, and fitness. Signs of marked social segregation by income level appear, and the boundaries between such areas are becoming more visible.

After the collapse of the Soviet Union and the transition period of the 1990s, the territorial shifts in the distribution of the population of St. Petersburg and the transformation of its social structure at the lower administrative-territorial level suggest a significant increase in welfare differentiation and the development of natural segregation processes, social stratification, and a distinct formation of «rich» and «poor» territories.

A comparison of the five indicators of social well-being of the municipalities of St. Petersburg discussed above suggests that they are interrelated. With the help of correlation analysis, a high degree of mutual dependence between four of them was revealed. Thus, Pearson's correlation coefficient between the cost of housing and the proportion of people with a scientific degree, calculated for 111 municipalities of the city, was 0.813, between the cost of housing and the proportion of individual entrepreneurs with paid employees – 0.707, between the cost of housing and property tax paid per inhabitant – 0.739. The only indicator that has a weak relationship with others (the correlation coefficient is in the range from 0.222 to 0.347) is the size of municipal revenues per capita. This circumstance only confirms the «opacity» of the formation of the revenue part of the budgets of local governments in St. Petersburg and the weak dependence of the size of municipal revenues on the welfare of local societies.

Based on the research results we are confident about the increasing social segregation in terms of the level of education of the population, the cost of real estate, and the income of local budgets. The significant spatial contrasts that have emerged in St. Petersburg pose a serious challenge to sustainable development of the city in the coming decades.

ACKNOWLEDGEMENTS

The reported study was funded by RFBR according to the research project № 18-310-20016 «Coastal cities in innovation spaces of the European part of Russia». D. Zhitin and S. Lachininskii express gratitude to the 5-100 competitiveness program of the Immanuel Kant Baltic Federal University for supporting their research fellowship.

REFERENCES

Aksenov K., Brade I., Bondarchuk E. (2006). The transformation of urban space in post-Soviet Russia. New York: Routledge.

Aksenov K., Krupickaitė D., Morachevskaya K. and Zinovyev A. (2018). Retail sprawl in post-soviet urban residential communities: case studies of Saint-Petersburg and Vilnius. Moravian Geographical Reports, 26(3), 210-219.

Aksenov K.E. (2016). Evolution of the types of shopping and spatial organization of retail trade in the post-soviet metropolis. Regional Research of Russia, 6(4), 375-386.

Aksenov K.E., Bondarchuk E. and Brade I. (2001). Territorial shift in the tertiary sector in the economy of St.-Petersburg. Izvestiya Akademii Nauk, Seriya Geograficheskaya, 2, 58-70.

Anokhin A.A., Lachininskii S.S., Zhitin D.V., Shendrik A.V., Mezhevich N.M. and Krasnov A.I. (2017). Post-Soviet urban environment: the experience of St. Petersburg. Regional Research of Russia, 7(3), 249-258.

Bater J.H. (1980). The Soviet city: ideal and reality. Beverly Hills, Calif.: Sage Publications

Bater J.H. and Staples J.R. (2000) Planning for change in Central St. Petersburg. Post-Soviet Geography and Economics, 41(2), 77-97.

Berki M. (2014). Return to the road of capitalism: Recapitulating the post-socialist urban transition. Hungarian Geographical Bulletin, 63 (3), 319-334. DOI: 10.15201/hungeobull.63.3.6.

Bertaud A. (2013). The Spatial Structures of Central and Eastern European cities: more European than Socialist? International symposium on post-communist cities the Russian and East European Center (REEC) «Winds of Societal Change: Remaking Post-Communist Cities» [online] Available at: www.alainbertaud.com/wp-content/uploads/2013/08/AB_Central-European-Spatial-Structure_Figures_2.pdf [Accessed 9 March 2016].

Borodina T.L. (2017). Regional'nye osobennosti dinamiki naselenija Rossii v postsovetskij period. Izvestija RAN. Ser. Geograficheskaja, 1, 47-61. (In Russian).

Brade I., Mahrova A.G., Nefedova T.G. and Trejvish A.I. (2013). Osobennosti suburbanizacii v Moskovskoj aglomeracii v postsovetskij period. Izvestija Rossijskoj akademii nauk. Serija geograficheskaja, 2, 19-29. (In Russian).

Bugaev M.A. (2015). Majatnikovye migracii na rynke truda Sankt-Peterburga i Leningradskoj oblasti. Vestnik Sankt-Peterburgskogo universiteta. Serija 5. Ekonomika, 4, 86-116. (In Russian).

Degusarova V.S., Martynov V.L. and Sazonova I.G. (2018). Geodemography of the Saint Petersburg Suburbs. Baltic region, 10(3), 19-40.

Fedorov G.M. and Mikhaylov A.S. (2018). Regional divergence dynamics in the Baltic region: Towards polarisation or equalization? Geographia Polonica, 91(4), 399-411.

Hirt S. (2012). Iron Curtains. Gates, Suburbs and Privatization of Space in the Post-Socialist City. Chichester, West Sussex: John Wiley & Sons.

Hodachek A.M. (2017). O Sankt-Peterburgskoj aglomeracii na osnove koncepcii gradostroitel'nogo razvitija. Ekonomika Severo-Zapada: problemy i perspektivy razvitija, 4(1), 35-47. (In Russian).

Humphrey C. and Skvirskaja V. (2012). Post-Cosmopolitan Cities: Explorations of Urban Coexistence. New York and Oxford: Berghahn, Guide to Seattle Micro Urbanism. [online] www.microseattle.wordpress.com [Accessed 9 March 2019].

Keivani R., Parsa A., McGreal S. (2001). Globalization, institutional structures and real estate markets in Central European Cities. Urban Studies, 38(13), 245-2476.

Kemper F-J. (1998). Restructuring of housing and ethnic segregation: recent developments in Berlin. Urban Studies, 35(10), 1765-1789.

Kenneth J.T. (1985). Crabgrass Frontier: The Suburbanization of the United States. Oxford: Oxford University Press.

Kirillov P.L. and Mahrova A.G. (2009). Suburbanizacija v Moskovskom stolichnom regione: sovremennoe i perspektivnoe sostojanie. Regional'nye issledovanija, 4(5), 42-54. (In Russian).

Kotus J. (2006). Changes in the spatial structure of a large Polish city – The case of Poznan. Cities, 23(5), 364-381.

Kryukova O.V., Martynov V.L., Sazonova I.Y. and Polyakova S.D. (2016). Main spatial problems of St. Petersburg. European Journal of Geography, 7(2), 85-95.

Kurichev N.K. and Kuricheva E.K. (2018). Vzaimosvjaz' zhilishhnogo stroitel'stva v Moskovskoj aglomeracii i migracii v stolichnyj region. Izvestija RAN. Ser. Geograficheskaja, 1, 5-20. (In Russian).

Mahrova A. and Kirillov P. (2015). Rossijskaja urbanizacija i zhil'e gorozhan. Available at: www.demoscope.ru/weekly/2015/0645/tema01.php. Accessed: 21.01.2020.

Mahrova A.G. (2014). Osobennosti stadijnogo razvitija Moskovskoj aglomeracii. Vestnik Moskovskogo universiteta. Serija 5: Geografija, 4, 10-16. (In Russian).

Mahrova A.G. and Golubchikov O.Ju. (2012). Rossijskij gorod v uslovijah kapitalizma: social'naja transformacija vnutrigorodskogo prostranstva. Vestnik Moskovskogo universiteta. Serija 5: Geografija, 2, 26-31 (In Russian).

Musterd S., Marcińczak S., van Ham M. and Tammaru T. (2017). Socioeconomic segregation in European capital cities: Increasing separation between poor and rich. Urban Geography, 38(7), 1062-1083. DOI: 10.1080/02723638.2016.1228371.

Ouředníček M., Temelová J. (2009). Twenty years after socialism: the transformation of Prague's inner structure. Studia Universitatis Babeş-Bolyai, Sociologia, 54(1), 9-30.

Pacione M., (2005). Urban Geography: A Global Perspective. New York: Routledge.

Pavljuk S.G. (2015). Metodika differenciacii gorodskogo prostranstva (na primerah gorodov Rossii, Zapadnoj Evropy i SShA). Regional'nye issledovanija, 2(48), 26-36. (In Russian).

Reznikov I.L. (2017). Vyjavlenie granic Sankt-Peterburgskoj gorodskoj aglomeracii. Vestnik Sankt-Peterburgskogo universiteta, 62(1), 89-103. (In Russian).

Ruble B.A., (1990). Leningrad. Shaping a Soviet City. Berkeley, CA: University of California Press.

Sailer-Fliege U. (1999). Characteristics of post-socialist urban transformation in East Central Europe. GeoJournal, 49(1), 7-16.

Smith N. (1996). The New Urban Frontier: Gentrification and the Revanchist City. New York: Routledge.

Stanilov K. Ed. (2007). The Post-Socialist City. Urban Form and Space Transformations in Central and Eastern Europe after Socialism. Dordrecht: Springer.

Szafranska E. (2015). Transformations of large housing estates in Central and Eastern Europe after the collapse of communism // Geographia Polonica, 88 (4, 621-648.

Szafrańska E., Coudroy L. de Lille, Kazimierczak J. (2019). Urban shrinkage and housing in a post socialist city: relationship between the demographic evolution and housing development in Łódź, Poland. Journal of Housing and the Built Environment, 34(2), 441-464.

Vserossijskaja perepis' naselenija (2010). Itogi. Kaliningradstat. URL: www.kaliningrad.gks.ru/wps/wcm/connect/rosstat_ts/kaliningrad/ru/census_and_researching/census/national_census_2010/score_2010/ (In Russian).

Zhitin D.V. (2015). Prostranstvennaja obuslovlennost' social'nogo rassloenija naselenija Sankt-Peterburga. Social'no-jekonomicheskaja geografija. Vestnik Associacii rossijskih geografov-obshhestvovedov, 4, 85-102. (In Russian).

Zubarevich N.V. (2017). Koncentracija naselenija i jekonomiki v stolicah postsovetskih stran. Regional'nye issledovanija, 1(55), 4-15 (In Russian).

Received on July 1th, 2020

Accepted on January 21th, 2020

Diana Dushkova[1,2*], Maria Ignatieva[3]
[1] Humboldt University Berlin, Department of Geography, Berlin, Germany
[2] Lomonosov Moscow State University, Faculty of Geography, Moscow, Russia
[3] University of Western Australia, School of Design, Australia
***Corresponding author:** diana.dushkova@geo.hu-berlin.de

NEW TRENDS IN URBAN ENVIRONMENTAL HEALTH RESEARCH: FROM GEOGRAPHY OF DISEASES TO THERAPEUTIC LANDSCAPES AND HEALING GARDENS

ABSTRACT. Urban living style is associated with various negative impacts on human health, e.g. connected with the environmental problems. Thus, promoting health of urban population is nowadays one of the most challenging issues of the 21st century together with the growing needs for sustainable development and establishment of the biophilic or livable cities. It is increasing awareness among researchers and health practitioners of the potential benefits to the health from activities in natural settings and especially from regular contact with nature, which can be perceived as a preventive medical tool. This paper discusses the close relationship between the concepts of health-supporting landscapes and sustainability in modern cities based on literature review and case studies from EU, Russian and Australian projects. We first review the historical and modern paradigms (of the various disciplines) which determine the discourse in nature – human health and well-being research. This includes examination of Hippocrates «naturalistic history», Humboldt's concept of natural garden design; Oertel 's 'Terrain Kur'; «salutogenic approach» of Antonovsky; McHarg's Design with Nature; Ecopolis programme, Wilson's biophilia and some other approaches. Then there is a comparative analysis of structural similarities and differences in the past and current scientific schools devoted to understanding human – landscape interaction. One of the principal arguments is that nature also has another value for health, regardless of natural remedies. It includes, for example, the healing of space, outdoor training trails in parks, everyday use of urban green spaces and peri-urban recreation areas for sport and exercises. We provide an analysis of some examples based on the modern concepts of biophilic cities, therapeutic landscapes, healing gardens, green infrastructure and nature-based solutions. This article also discusses the main types of healing gardens and therapeutic landscapes and suggests the framework of design principles of healing and therapeutic landscapes. The analysis proved that healing gardens and therapeutic landscapes provide multiple benefits and can be regarded as nature-based solutions. These essential aspects of multifunctionality, multiculturality and social inclusion are well intertwined with the approach of biophilia.

KEY WORDS: urban environment, environmental health research, therapeutic landscape, healing garden, biophilic cities, nature-based solutions

CITATION: Diana Dushkova and Maria Ignatieva (2020) New Trends In Urban Environmental Health Research: From Geography Of Diseases To Therapeutic Landscapes And Healing Gardens. Geography, Environment, Sustainability, Vol.13, No 1, p. 159-171
DOI-10.24057/2071-9388-2019-99

INTRODUCTION

Modern cities worldwide face a number of problems relevant for all the cities of the 21st century – so called societal challenges, such as urbanization, climate change, ecological issues, environmental quality and sustainable development (Raymond et al. 2017). Urban living style is associated with various negative impacts on human health, e.g. connected with the environmental pollutants such as carbon and nitrogen dioxides, heavy metals, radionuclides, benzene, etc. There is a greater connection between urban pollution and health problems confirmed by the research results from highly polluted urban sites from around the world (Dushkova and Evseev 2012; Hankey and Marshall 2017; Janke et al. 2009; Landrigan and Fuller 2015; WHO-UNEP 2008). These pollutants are associated with environment connected diseases such as stroke, cardio-vascular problems, lung cancer, and both chronic and acute respiratory illnesses, headaches and dizziness, disruption of reproductive and immune systems, and premature death. Among the other adverse health consequences of urban living are high-fat diets, sedentary lifestyles, and increased levels of social and psychological stress (Beyer et al. 2014; Groenewegen et al. 2006; WHO-UNEP 2008). Thus, promoting health of urban population is nowadays one of the most challenging issues of the 21st century (EC 2015; Marcel et al. 2019; Souter-Brown 2014; Tzoulas et al. 2007) together with the growing needs for sustainable development and establishment of the biophilic or livable cities. The concepts of «healthy urban landscape» and «livable city» have acquired a new significance as an area of integration and balanced interaction of urban development processes, natural and cultural contexts and challenges of creating favorable environment for the urban citizens. Moreover, the concepts of biophilic cities, compact and sustainable cities, nature-based solutions and integrated blue-green infrastructure have the potential to become the theoretical foundation for health-supporting landscapes (EC 2015; Ignatieva 2018; Russo et al. 2017). There is a great number of research worldwide underlying the positive evidence / relationship of direct experience with natural environments to a wide range of health benefits (Beyer et al. 2014; Dushkova and Haase 2020; Hartig et al. 2014; Maas et al. 2009; Soga et al. 2016) and the key role of nature in achieving a healthy society (Brink et al. 2016; Groenewegen et al. 2006; Tzoulas et al. 2007). It is increasing awareness among researchers and health practitioners of the potential benefits to the health from activities in natural settings and especially from regular contact with nature, which can be perceived as a preventive medical tool (Frumkin et al. 2017; Groenewegen et al. 2006, Maller et al. 2005). A contact with nature in cities in forms of urban gardening, sport activities and community gardens increases life satisfaction, psychological wellbeing, social inclusion and social cohesion, sense of community, and cognitive function (Soga and Gaston 2016; Wood et al. 2016). Moreover, engagement with nature activities (for example, urban gardening, urban farming) was defined as not only a cost-effective health intervention and a type of nature-based solution (Dushkova and Haase 2020; Frumkin et al. 2017; Williams 2017) but also as a treatment for several physiological and mental health problems, so-called «therapeutic landscape and healing garden therapy» (Hartig et al. 2014; Söderback et al. 2004; Wood et al. 2016).The main goal of this article is to discuss the role of nature for creation of sustainable urban living spaces and to analyze, based on the literature review, the development of the concept of health-supporting landscapes for urban citizens. It suggested a historical discourse and analysis of different worldwide examples from antient time to nowadays. The starting hypothesis for our analysis is that researching historical precedents of landscape – health relationships can be beneficial to current and future planning of urban green spaces and making better the quality of life of urban population. We first review the development of the disciplines within the environment – human health interactions from theoretical and practical perspectives. This includes examination of the impact of changing the paradigms and re-thinking the problem. We then compare structural similarities and differences in the past and current scientific schools devoted to the human – landscape interaction, in order to provide a historical ground and a foundation for achieving more positive human – nature relationships. And finally, we provide some interesting examples based on the modern concepts of biophilic cities, therapeutic landscapes, healing gardens, green infrastructure and nature-based solutions which represent the results of collaborative systems thinking by involving different disciplines and professions in order to build healthy cities. These results were obtained from literature review and research projects based on case studies from Europe, Russian Federation and Australia. Compared to the recent publications on environmental health which primarily focused on potential threats from the natural and environmental surroundings (e.g. pollution, ecosystem disruption and different ecosystem disservices such as natural disasters, vector-borne pathogens and allergens), this paper focuses on how nature positively impacts on human health and wellbeing. One of the current societal challenges is to apply an interdisciplinary approach to researching resilient sustainable biophilic cities (with people and for people), which adopt an integrated conceptual framework based on human ecology, environmental health research, nature-based solution and therapeutic landscape concepts.

MATERIALS AND METHODS

In order to understand the evolution of the concepts in environment – health research, this article reviews and discusses the existing scientific evidence on the human – environmental relationships and health benefits from the contact with nature for individuals and at community scale, based on a number of different research methods and publications from different research disciplines. This review is primarily based on peer-reviewed literature. Furthermore, the article features numerous practical examples from around the world. We used the results of the project «Mathematical-cartographic assessment of medico-ecological situation in cities of European Russia for their integrated ecological characteristics» (2018–2020) under Grant number No 18-05-00236/18 supported by the Russian Foundation for Basic Research (RFBR); the Horizon 2020 Framework Programme of the European Union project «Connecting Nature» under Grant Agreement No 730222 and the research project «Perth as a biophilic resilient city model in the time of climate change» by the University of Western Australia (UWA) FABLE research grant (2018 – 2019). Some selected examples of good practice of urban green infrastructure, nature-based solutions and biophilic cities concepts were analysed in details. Another methodological approach of this article includes an explorative survey, which was carried out to assess to what extent the principles, and concepts that were revealed during the analysis of the literature and research within the above-mentioned projects, can be recognized and applied in urban planning and governance in cities. Thus, we identified different categories of innovative practices which incorporate those nature-based solutions and biophilic approaches that related to health and human well-being. We analyzed in detail to what extent these selected cases contribute to the creating and maintaining of resilient, sustainable biophilic cities and providing the health-supporting benefits for their citizens. In addition, a detailed on-site analysis was conducted, including a survey and a series of interviews with experts dealing with the issues of urban greening, landscape planning, health environmental research. This analysis provided additional insights into different aspects of nature-health relationships. The analysis was based on the methodical approaches proposed for assessment of therapeutic values of landscape (Belčáková et al. 2018; Williams 2017), healing gardens (Cooper Marcus and Sachs 2013; Frumkin et al. 2017; Hartig et al. 2013) and co-benefits from nature-based solutions, especially those related to human health and well-being (Marcel et al. 2019; Raymond et al. 2017).

The evidence brought together from a review of the literature and real-life applications has been discussed with a wider range of experts in the field of landscape ecology, health and social research working at local, regional, national or international levels at different project-related workshops (e.g. Connecting Nature partners workshop in Malaga on 30 September – 03 October 2019; IGU Moscow 2018 Special session: Health Geography – XII International Symposium for geospatial health on 5 June 2018; 25th IAPS Conference Transitions to sustainability, lifestyles changes and human wellbeing: cultural, environmental and political challenges, Rome, Italy, 8-13 July 2018; roundtables with environmental health experts at the Centre for Human Adaptation in the Arctic of the Kola Science Center of Russian Academy of Sciences on 12-13 August 2019 and 27 July 2018 and the seminar at Polar-Alpine Botanical Garden-Institute in Apatity, Russia on Perspectives and potentials of landscape therapy in the city on 26 July 2018; workshop at Humboldt university Berlin on Urban Biodiversity and Nature-Based Design: methodology and practical applications for interdisciplinary research on 27-29 November 2019, etc.). The discussions and insights from these workshops and round tables have been integrated in this article and have helped to formulate conclusions about using nature-based approach for solving health and social challenges across the cities worldwide. We also included the results from international summer schools «An interdisciplinary perspective on ecosystem services and human well-being» (2015–2019) and master's courses on «Environment and human health», «Landscape biophilic design», «Design with nature», «Green infrastructure as nature-based solutions» which were created and taught by authors at Humboldt university Berlin, Lomonosov Moscow State university and at the University of Western Australia.

RESULTS AND DISCUSSION

1. History of the development of the disciplines devoted to the environment – human health interactions from the theoretical perspective

The connection between nature and health is not new. Rather, nature has always had a high, often very well realized relevance to human health. Table 1 summarizes the evolution of disciplines dealing with environment – human health relationships. The

Table 1. The historical and modern paradigms (of the various disciplines) which determine the discourse in nature – human health and well-being research (created by authors)

Scientific approach / school, year(s)	Key concept and its definition	Theoretical background	Key publications
«Super-natural theory of disease» (prehistorical time)	Religious and beliefs often attributed diseases outbreaks or other misfortunes	Disease and human health were correlated with witchcraft, demons or the will of gods	see Lips-Castro (2015)
«Naturalistic theory» of classic world (from 460 B.C.)	Complex web of interconnections that influence health, well-being	One of the most frequent causes of diseases is the bad quality of environment	Hippocrates, ancient scientists philosophers
«Dark ages of medicine and health research» (5th-16th centuries)	n/a	Practical medicine turned back to the primitive one with domination of dogma, irrationality and superstition	See Valencius (2000)
Empirical research on environment and health, medical meteorology, 16th -17th centuries	Quantitative approach to the study of the environmental factors of illness	Era of empirical study of relationships between climate, topography, weather, geography, and disease was started and new science of medical meteorology emerged.	Thomas Sydenham (1676) – e.g. research on epidemics in London)
Foundation of classical medical geography as a discipline (18th century)	Theory of the natural tendency of zoonotic diseases to become localized in a specific habitat	Distribution of each of the diseases was related to the local environment of the place where they occurred. Further research on the interactions of health and environment (geographical pathology)	Finke (1792), Mühry (1856), Hirsch (1883–1886)
Medical geography, Humboldtian medicine and healing gardens (end of 18th – 19th centuries)	Medical geography as an exercise in mapping, cartography, and charting diseases and people. Healing gardens as example of design-with-nature concept.	Systematic study of the global variable of human diseases, making use of the concepts, terminology and representational forms of the new plant geography	A. von Humboldt and his followers
Terrain Cur and medical geography (especially, after cultural turn)	The concept of Terrain Kur – physical training in the form of walks over special routes with the positive influence of the nature's beauty for the treatment of several diseases.	Therapeutic landscape as places with natural or historic features for the maintenance of health and well-being, for achieving physical, mental and spiritual healing. Four dimensions: natural environment, built environment, symbolic environment and social environment.	Oertel (1886), Oblonsky (1901), Gesler (1992)
Ecology of Human Diseases	Description of the epidemiological constraints of various diseases	The main influenced environmental factors: inorganic, organic and socio-cultural were defined. Foci of infectious, zoonotic diseases.	Jacques M. May (1959), Pavlovsky (1966)
Ecological psychology and sociology (20th century)	Salutogenesis as theory of health and illness	Salutogenic environment and therapeutic landscape. Theories of environmental affordances; ecological psychology	Antonowsky (1976)
Environmental health, Health geography, Epidemiology (20-21th centuries)	Health and environmental pollution – health risks from natural disasters and anthropogenic health risks, use of spatial analysis	Health geography consists of two distinct elements: 1) geographies of disease and health involving descriptive research disease frequencies and distributions, and 2) analytical research to find what characteristics make an individual or population susceptible to disease	Curtis (2004), Landrigan and Fuller (2015), Kearn and Moon (2002), Malkhazova et al. (2019), Revich (2018)
Environmental psychology (a) (already in ancient China, Greek and Rom, with revival in 20-21th centuries)	Restorative environment	Attention-restoration theory, four features as restorative environment: being away, extent, fascination, and action and compatibility	Kaplan and Kaplan (1992)
Environmental psychology (b) (already in ancient China, Greek and Rom, with revival in 20-21th centuries)	Therapeutic landscapes and healing garden	Esthetic-affective theory, psycho-evolution theories, three features of healing gardens: relief from physical symptoms, illness or trauma; stress reduction for individuals dealing with emotionally and/or physically stressful experiences; and an improvement in the overall sense of well-being	Cooper Marcus and Sachs (2013), Williams (2017)
Horticultural Therapy (already in ancient China, Greek and Rom, with revival in 20-21th centuries)	Healing garden and therapeutic garden	Theory of «flow experience»; sensory stimulation theories. Gardens and landscape that facilitate health and well-being	Söderback et al. (2004), Jiang (2014), Souter-Brown (2014)
Biophilia, green-blue infrastructure and nature-based solution concepts (20th – 21th centuries)	To design with nature, to promote nature as a source of sustainable solutions as an answer to challenges associated with climate change	«Biophilic cities» inspired by and referred to human needs to have connections with nature. NBS inspired and supported by nature, which are cost-effective and provide multiple co-benefits (environmental, social and economic benefits and help build resilience)	McHarg (1969), Pötz and Bleuze (2012), Wilson (1986), Russo et al. (2017), EC (2015), Agavelov et al. (1985), Ignatieva (2000, 2018)

ancient scientists and philosophers such as Hippocrates (who wrote the book «On airs, waters and places») have already intuitively dealt with the complex web of interconnections that influence human health and well-being (Lawrence et al. 2017). According to Hippocrates, one of the most frequent causes of diseases is the bad quality of air. This approach was defined as a «naturalistic theory» and its development was intervened with cultural development of some civilizations (Table 1). However, in prehistorical time humans rather refer to «super-natural theory of disease» when disease was correlated with witchcraft, demons or the will of gods (Lips-Castro 2015). But with development of a hygiene and curative aspects of a disease as a powerful brunch of medicine, this super-natural approach has been abandoned.

The period of Middle Ages was rather defined as dark ages of medicine when the practical medicine turned back to the primitive one with domination of dogma, irrationality and superstition. Only in the late sixteenth – beginning of the seventieth century, the interest to the investigation in health relationship was arisen, thus, a number of studies appeared from Italian, British, French and German scholars in form of notes, letters, articles and monographs (Dzik 1997). Among them are works of Thomas Sydenham on epidemics in London who applied a quantitative approach to the study of the environmental factors of illness (Valencius 2000). Since this time the era of empirical study of relationships between climate, topography, weather, geography, and disease was started and the new science of medical meteorology had emerged. German researcher Leonhard Finke with his work «Versuch einer allgemeinen medizinisch-praktischen Geographie» (1792) was defined by many researchers as a founder of modern medical geography. In particular, Finke stressed that distribution of each of the diseases was related to the local environment where they occurred. In the late ninetieth century, other German scientists continued the research on the interactions of health and environment – Adolf Mühry (Die geographischen Verhältnisse der Krankheiten, oder Grundzüge der Nosogeographie, 1856) and August Hirsch «Handbook of Geographical and Historical Pathology» (1883–1886).

Famous German geographer and explorer Alexander von Humboldt (1769–1859) had also a great contribution to the medical geography, which is based on mapping, cartography and charting diseases and peoples. Humboldt had also developed the concept of natural garden design. In his «Ideas for a Physiognomy of Plants («Ideen zu einer Physiognomik der Gewächse) published in 1806, he referred to the «patriotic plants characters» (vaterländische Pflanzengestalten) and addressed the «decisive impact» provided by the native vegetation on the development of individual cultures. He also underlined the role of gardens in healing by proposing to design with nature and thus, the quality of existing environment. Further, the so-called «Humboldtian medicine» has expanded the scope of the Humboldtian approach and included a systematic study of the global variable of human diseases, concepts, terminology and representational forms of the new plant geography (Valencius 2000).

However, there was a down part of such disease-map-approach which later resulted in development of the nationalistic view declaring the specific pathway of nations and populations according to particular climate and region (Valencius 2000). This was a period of development of colonial geo-medicine and race distribution, used also for the purpose of nationalistic ideas.

There are several fundamental works in the field of medical geography – environmental health relationship from Russia and the Soviet Union. Russian physician Nikolay Oblonsky in the beginning of the twentieth century (1901) adopted and developed further on the concept of Terrain Kur suggested by the German physician Max Joseph Oertel (1886). «Terrain Cur» or Klima-Terrain-Kur (Terrain Curorte) means a scheduled physical exercise in the form of long walks over steep ascending routes (terrains) as a special positive treatment (Kur) by nature's beauty for several diseases (e.g. cardiovascular, mental etc.). Since this period, several routes were established in Russia (mostly in Caucasus Mountains) where the concept of landscape therapy was defined as one of the important drivers in the treatment of diseases. Another fundamental Russian concept developed by Evgeni Pavlovsky at early 1950's was the theory of the natural tendency of zoonotic diseases to become localized under the influence of a specific habitat. Pavlovsky identified the foci of infectious, zoonotic diseases by analyzing the associations of vegetation, animal and insect, soil and precipitation regime and other elements of the natural landscape (Pavlovsky 1966). Further research on such diseases caused by agents circulating in natural environments independently from humans was continued by Malkhazova et al. (2019). Even the concept of natural related diseases is not directly connected to therapeutic landscape, we intentionally include it in the list of research devoted to the analysis of the environment – human health relationships. This concept deals with the ecosystem disservices (risks and undesirable effects from direct contact with nature).

The Ecology of Human Diseases (1959) was introduced by the American geographer Jacques M. May. He described the epidemiological constraints of various diseases and defined the main influential environmental factors: inorganic, organic and socio-cultural.

The shift in research on human – environment relationship appeared in the 1970s with the «salutogenic approach» (health-illness continuum) developed by Aaron Antonovsky. The salutogenic perspective tries to explore how health is produced and what are the protective factors and resources designated for the good health (Antonovsky 1979). The concept concerned how specific personal dispositions serve to make individuals more resilient to the stressors of daily life and identified the characteristics, which claimed to help a person better cope with them (and remain healthy).

In the last decades, a great number of research in the field of medical to health geography appeared, focusing on the relationship between health and environmental pollution (Dzik 1997; Kearn et al. 2002; Landrigan et al. 2015; Revich 2018) and especially analyzing priority indicators of the life quality in cities, including urban environment indicators such as ambient air quality, level of noise pollution, temperature waves, population density, and the urban greening rate. The research of Curtis (2004), Malkhazova et al. (2019), Schweikart and Kistemann (2013) etc. presents a comprehensive analysis of how geographical perspectives can be used to understand the problems of health and its inequalities by explaining and demonstrating how different methodologies in the geography of health, both quantitative and qualitative, can be applied in research.

In this time, also the new concepts were established, which aim to create, design and planning healthier, sustainable and resilient urban environments. Among them is «green infrastructure» or «blue-green infrastructure», which in fact began in the 1870, when urban farming and allotment gardens were introduced as a special tool for providing healthy recreation activity and good air in polluted industrial cities (Pötz et al. 2012). «Design with nature» concept was introduced by I.McHarg in 1969 (McHarg 1969) with the direct message to respect component of the environment and natural process while designing and planning. Russian programme «Ecopolis» of the late 70s based on the concept of the «coherent development of nature and humans» and V.Vernadsky's concept of «Noosphere» aimed to create a new type of harmonies urban settlements. This new generation of human settlements is based on minimal ecological footprints (biophysical), which can maximise human potentials (human ecology) to repair, replenish and support human life (Agavelov et al. 1985; Ignatieva 2000, 2018).

The ideas of «Ecopolis» (Agavelov et al 1985) and «biophilic cities» inspired by Wilson (1986) referred to human needs to have everyday connections with nature. The recently emerged (in the 2000s) concept of nature-based solutions (NBS) was established to promote nature as a source of sustainable solutions and as an answer to challenges associated with climate change. The concept of NBS has been supported and broadened by the International Union for the Conservation of Nature and later by the European Commission. It defined NBS as solutions that are «inspired and supported by nature, which are cost-effective, simultaneously provide environmental, social and economic benefits and help build resilience (EC 2015). The concept has emerged as a way to operationalize and to promote the ecosystem services approach within spatial planning policies and practices (Russo et al. 2017). The opportunities of nature-based solutions for climate change adaptation are also particularly discussed in relation to nature conservation, public health, landscape architecture and urban planning (Marcel et al. 2019).

At the beginning of the 21st century, in addition to the global and regional changes in the natural resources supply and provision of the healthy living conditions, other aspects of the relationship between humans and nature and health are becoming increasingly important.

On the one hand, this is the direct, usually health-promoting (according to the concept of salutogenesis) effect of «nature» on humans and their health and well-being. Analysis of the publications in recent years from the field of natural medicine, has shown that the relationship of environment / nature and health is in the basic focus of the scientific and public debate. There are many direct links between nature and human health and well-being. Thus, connection with nature, in addition to satisfying elementary human needs (e.g. food and natural resources supply), heals or mitigates the most diseases and can be defined as a health resource (which keeps people healthy) (Groenewegen et al. 2006; Kabisch et al. 2018). The recreational and healing value of nature for physical health and mental well-being has long been discussed (Beyer et al. 2014; Hartig et al. 2014; Maas et al. 2009; Soga and Gaston 2016; Souter-Brown 2014). Thus, the effect of nature on humans was already treated in antiquity. Natural remedies (medicinal herbs, mineral waters, muds, extracts, etc.) are used through the thousands of years. Traditional medicine, holistic, complementary and alternative medicine – e.g. traditional Chinese medicine, Ayurveda, Hildegard's and Kneipp's teachings, homeopathy and many others have nature as the essential healing point of their approach. However, nature also has another value for health, regardless of natural remedies (though often not consciously perceived). For example, the healing of spas, outdoor training trails in parks, everyday use of urban green spaces and peri-urban recreation areas for sport and exercises (cycling, jogging and Nordic walking). These health aspects of outdoor nature are used for promotion healthy life-style, especially for children, through the active nature experience, since many children in urban spaces no longer have the opportunity to acquire nature in everyday life experience (Kabisch et al. 2018). Thus, as a source of healing, and source of inspiration, nature plays an important role in the identity of people and in the development of its own «sense of place» (Gesler 1992; Frumkin et al. 2017). The most important challenge for modern cities is to design urban landscapes accordingly with the idea of harmonic co-existence based on the concept of biophilia. Under the increasing urbanization and alienation of nature the creation of so-called health regions within the cities become the high priority.

2. Therapeutic landscape and healing gardens or where should you go when you are sick

In order to better explore the environment – human health relationship, we have to refer to the definition of a health which according to WHO/UNEP (2008) is «a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity». Such broad interpretation includes also the aspects of physical and psychological wellbeing. It means the high relevance of potential positive effects on health which can be caused by positive emotions (e.g. quality of life, life satisfaction, sense of community and happiness) and the minimization or absence of negative emotions (e.g. anger, loneliness, confusion) (Groenewegen et al. 2006; Söderback et al. 2004; Sullivan and Chang 2017). Being on nature promotes physical activities and supports healthy life style and minimize the risk of a number of modern urbanization-connected diseases (e.g. obesity, mental problems, stress etc.). These healthy activities connected to nature include not only outdoor sport, but also gardening when people grow, cultivate, and take care of plants (flowers and vegetables) for non-commercial use as well as activity in domestic gardens (including allotment and community gardens). In this case horticultural therapy often defined as a practice which helps to improve the physical, psychological, and social health (Jiang 2014).

Landscapes of healing were defined by Gesler (1992) as places, settings, situations, locales and milieus that encompass the physical, psychological and social environments associated with treatment or healing. According to them, there are four essential parts of such landscapes: a) natural, b) built, c) symbolic, d) social environments. All these parts of healing landscapes include a variety of elements, which characterize the relevance of natural and artificial (build) environments (for example, green elements, architectural style, scenic views, aesthetical features), their cultural / social / ritual value (such as social support, sense of place, sacrality) and political significance (especially, (inter)national connectivity, nation-building feeling etc.) (Figure 1).

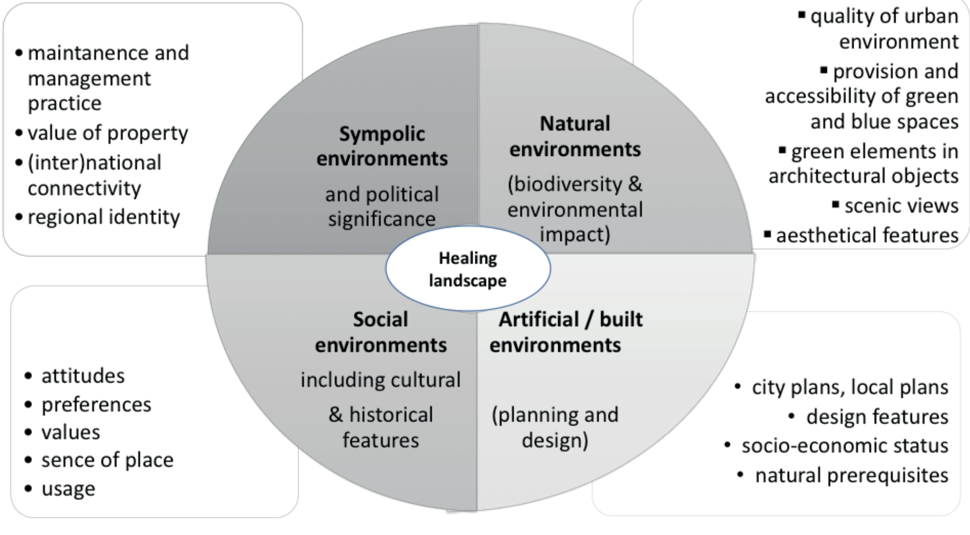


Fig. 1. Healing (therapeutic) landscape and its essential elements (source: authors)

The healing capability and potential of nature for health and well-being is discussed in the stress reduction theory (Cooper Marcus and Sachs 2013), which defines the main powerful and thus valuable factors of nature influencing the healing process, improving health and well-being (Table 2). This table illustrates the analysis of therapeutic impact of the selected nature-based solutions which were studied within authors' research projects and by applying current approaches (Cooper Marcus et al. 2013; Raymond et al. 2017; Dushkova et al. 2020).

Evidence from literature review revealed a broad spectrum of health effects from contact with nature (Table 3). There is a wide range of health benefits and other positive effects for humans from direct experience of natural environments and implemented nature-based solutions (Beyer et al. 2014; Dushkova et al. 2020; Hartig et al. 2014; Maas et al. 2009; Soga et al. 2016). Nature plays a key role in achieving a healthy society (Brink et al. 2016; Groenewegen et al. 2006; Tzoulas et al. 2007). A number of research (Frumkin et al. 2017; Groenewegen et al. 2006, Maller et al. 2005) highlighted benefits to the health from activities in natural settings and especially from regular contact with nature, which can be perceived as a preventive medical tool. A contact with nature via urban gardening, sport activities and community gardens increases life satisfaction, psychological wellbeing, social inclusion and social cohesion, sense of community, and cognitive function (Soga and Gaston 2016; Wood et al. 2016). Engagement with nature activities (for example, urban gardening and farming) was defined as not only a cost-effective health intervention and a type of nature-based solution (Dushkova and Haase 2020; Frumkin et al. 2017; Williams 2017) but also as a treatment for several physiological and mental health problems, so-called «therapeutic landscape and healing garden therapy» (Hartig et al. 2014; Söderback et al. 2004; Wood et al. 2016). According to Cooper Marcus and Sachs (2013), the garden is intended to be a healing garden: a garden that, in different ways, influences the visitor in a positive way.

The interest to the creation of gardens for healing purposes goes back to ancient Chinese, Greek and Roman cultures. Current research proposes the theories of healing effects of gardens. There are three main approaches (Stigsdotter et al. 2002):

1. *The Healing Garden School.* The health effects are derived above all from the experiences of the garden room as such, its design, and contents.
2. *The Horticultural Therapy School.* The health effects are derived primarily from the activities in the garden room.
3. *The Instorative or Cognitive School.* The health effects are derived from the experiences of the garden room as such, from the activities in the garden room, and the visitor's background and character: experiences of which give the visitor a feeling of belonging and an identity.

Based on the literature review, we can figure out the framework for the analysis of the potential of an urban environment to be designed as a healing or therapeutic

Table 2. Analysis of therapeutic impact of selected nature-based solutions based on the results from the Connecting Nature project

Factors of nature influencing the healing process, improving health and well-being	Type of NBS (selection from the database of EU Connecting Nature project)						
	Public parks	Community gardens	Educational facilities (sport-/ playgrounds)	Allotment gardens	Healing gardens	Terrain kur (walking in special routes)	Touching / smelling gardens
the feeling of control over own body and environment when being on nature							
the opportunity to feel the support from the fellows by spending time together							
the positive motivation to start physical exercise inspired by nature							
the feeling of calmness, serenity – positive impact produced by nature through reduction of stressful condition							
support the treatment							
Reduce depression							
decrease the effects of heart conditions							

Note: grey – contribute based on the scoring methods and analysis of project impact / benefit assessed using approaches of Raymond et al. 2017, Cooper Marcus and Sachs 2013; white – not relevant.

Table 3. The positive evidence / relationship of direct experience with natural environments – health benefits from nature-based solutions (healing gardens and therapeutic landscapes)

Type of NBS	Methodical approach	Health effects (evidence based)	References (study area)
Healing gardens	Evaluation of therapeutic designs, questionnaire survey	Support the treatment of patients, sociality, gives sense of control (person loses control in a hospital, led by the conditions and staff of medical facility), enable physical movement; access and bond to nature; general, various positive distractions	Belčáková et al. 2018 (Pezinok, Slovakia)
	Landscape evaluation using different criteria, questionnaire survey, scoring method	Well-designed healing hospital gardens form a social atmosphere through saving the patients from the monotony that the clinic environment has and positively affect the clinical results of the patients through reducing stress, and thus they enable patients to feel themselves good psychologically and physiologically	Duzenli et al. 2017 (Trabzon, Turkey)
	Epidemiological survey, psycholo-gical and physio-logical monitoring methods	Preventing stress and promoting mental health, it also demonstrates how environmental enrichment works (e.g. neuropsychological responses which improve behavioural and health outcomes) and how physical activity on nature reduces obesity and non-communicable diseases	van den Bosch and Bird 2018
	A quasi-experimental field study and a true experiment, multimethod assessments	Natural landscape areas of healing gardens reduced depression because nature takes people's interests and attention in an effective way and it takes negative thoughts away	Kaplan et al. 1992; Hartig et al. 2014; Rohde and Kendle 1994 (various)
	Remote sensing for NDVI measures, GIS, medical data	Natural spaces relieve people, create less anxiety and enhance sincerity	Gupta et al. 2012 (Delhi, India)
	Landscape design analysis, participatory approach, questionnaire survey	Feeling of self-confidence, self-respect, self-reliance and personality and self-development	Stigsdotter et al., 2002 (Alnarp, Sweden)
	Literature review	Healing gardens provide patients with healthy living environments through refreshing their memories, increasing motivation and physical activities which decreases the eeffects of heart conditions and other risky diseases	Elings 2006, Söderback et al. 2004 (various countries)
Therapeutic landscape	Landscape planning analysis, participatory design, questionnaire survey, medical data	Positive distraction (to help users get away, both physically and emotionally, from the stress and pain of interior environment), engagement with nature (biophilia), physical and emotional comfort, incl. safety, security and privacy	Cooper Marcus and Sachs 2013 (various cities of USA), Williams 2017 (concept)
	Literature review, biomedical studies; exposure science; epidemiology of health benefits	Reduced stress, better sleep, improved mental health (reduces depression and anxiety, greater happiness and life satisfaction), increased prosocial behavior and social connectedness	Beyer et al. 2014, Frumkin et al. 2017, Maas et al. 2009, Sullivan and Chang 2017
	Biomedical studies; exposure science; epidemiological survey, remote sensing, GIS	Lower blood pressure, improved postoperative recovery, improved congestive heart failure, reduced obesity, reduced diabetes, improved immune function, improved children health, including child development (cognitive and motor)	Kabisch and Haase 2018, Li et al. 2010, Van den Bosch and Bird 2018

landscape. This analysis should be done using an interdisciplinary approach based on the conceptual ideas of medical as well as environmental sciences, geography and sociology. For this purpose, different kind of data can be used:

1. Environmental data which encompasses the assessment of natural environment with specific elements such as green and blue spaces, fresh-water springs, scenic views which give a high aesthetic pleasure using different environmental indicators
2. Medical statistics data on public health
3. Results of sociological surveys including the participatory observation, questionnaires, semi-structured or short interview etc.
4.Geographical information which can be derived from the cultural landscape research – not only by reading the contemporary landscape but also looking at the history of its development (by exploring its palimpsest created throughout the time).

3. Design with nature to support the human health in cities: evidence from the cities worldwide

Here we provide some examples of successful application of biophilic design and nature-based solution concepts in urban settings which resulted in a wide spectrum of physical, mental and behavioral benefits.

The key idea of biophilic design is to create good and comfortable habitat for people as a biological organism in the modern built environment that advances people's health, fitness and wellbeing. Together with the main purpose of the nature-based solution approach the both concepts seek to prevailing paradigm of co-design and co-development with nature in creating sustainable, livable and resilient cities based on needs of their residents and in order to face the current urban societal challenges (e.g. climate change, food and water security or natural disasters). At the same time nature-based solutions provide wider benefits to human well-being and biodiversity. Nature-based solutions based on biophilic design help the urban societies to effectively solve environmental issues.

The most common nature-based solutions include parks and urban green areas which provide a range of natural benefits such as intercepting dust, toxins and noise, sheltering and cooling property, sinking carbon and buffering flooding, creating place for recreation, fostering well-being, and a host of other social benefits.

To the nature-based solutions also refer traditional healing or therapeutic gardens which can be found within or adjacent to indoor healthcare settings – mental health hospitals, schools and centers for the disabled, hospices and nursing homes. Along with the «green care», «farm care» and «farming for health» they represent a new social movement, which uses the benefits of horticultural therapy and dynamically develop throughout Europe, America, Australia and New Zealand. Different types of urban healing or therapeutic gardens, which in the same time can be described as nature-based solutions due to multiple co-benefits they provide, presented in the Table 4.

Thus, as the selected cases show, all types of healing gardens and therapeutic landscape provide multiple benefits (as nature-based solutions) and are developed according with the principles of biophilia. They mostly free assessable-

Table 4. Main types of healing gardens and therapeutic landscapes which characterize current socio-environmental health related movement in cities (Photos: D. Dushkova, M. Ignatieva)

Types	Selected case 1	Selected case 2	Main purpose and co-benefits
1. Touching and smelling gardens	 Friedenspark of Leipzig, Germany	 Botanical garden in Perth, Australia	Assessable-designed recreation for blind and visually impaired people in order to support learning about plants and flowers, social inclusion tool
2. Special green education facilities	 Participation in planting activities in Campus of Uppsala University, Sweden	 Palmgarden in Leipzig, Germany – place of recreation, environmental education, co-creation with nature	improved social interaction and integration, less hostility and aggression
3. Co-designed gardens (nature as healing tool for ill children)	 Summer Garden S.-Petersburg, place for co-creation (projects of children with hearing, sight, physical and mental problems)	 Green project of Association for children with mental and physical health problems, Duisburg, Germany	Work in garden provides health benefits to its users, improving both their physical and mental condition. Important element of such green projects – to create opportunities for patients to wrench away from health problems, arouse involuntary attention and fascination

4. Community gardens as urban farming and recreation	 Annalinde community garden, Leipzig	 McDugall Farm Community Garden, Perth, Australia	Urban farming as a recreation space, cultural hotspot and educational source for residents (to enjoy tasting organic fresh foods, and learn environmental stewardship)
5. Sensory gardens and nature playground	 Playground in King's Park, Perth (Australia)	 Healthy track in kinder garden Moscow region, explore nature	Outdoor environments that have a series of curved, raised and by different types of concreat-covered surfaces' tracks enclose a number of «activity stations» and provide space for sensory exploring and plantings. It not only promotes children's health but also allows children to explore different textures, learning and experiencing nature around
6. Terrain cur within the city	 Terrain in park Khimki /Moscow (Russia)	 Terrain cur in park of Bad Düben (Germany)	Walking in such specially developed routes has not only therapeutic effect, but also directs attention from difficult internal experiences to an attractive, friendly outside world. It gives possibility of performing various types of recreation – active rest, relaxation and rehabilitation in a natural environment, by introducing a sense of peace, order and harmony

designed and open to the larger community what gives people sense of social inclusivity and social cohesion, enable the needs of disabled users to accommodate and experience the nature. This explain why they became a perfect area for recreation and rehabilitation and more and more attract people as welcoming space that could be used for exercise, gardening and an escape from the normal nursing home routine. These essential aspects of multifunctionality, multiculturality and social inclusion are well intertwined with the approach of biophilia (Ignatieva 2018; Dushkova et al. 2019; Dushkova et al. 2020):

1) They incorporate the elements of natural ecosystems in form of interconnected plants, animals, water objects, rocks, and geological forms.

2) *Natural colors* and materials as well as naturalistic shaped forms which are drawn from design principles and characteristics of the natural world – they can stimulate and reflect the dynamic properties of organic matter in adaptive response to the stresses and challenges of the everyday life (Figures 2a, 2b).

3) *Biomimicry* – using of forms and functions found in nature, especially among other species, whose properties have been adopted or suggest solutions to human needs and problems. Examples include the lawn presented by the native species (Figure 3 a), sustainable urban-drainage systems (Figure 3 b) etc.

4) *Place attachment* – culturally relevant designs can promote a connection to place and the sense that a setting has a distinct human (and also regional) identity, and provide emotional attachment to an area, particularly an awareness of local landscapes, indigenous flora and fauna, and finally motivate people to protect and sustain the environment of their living.

CONCLUSIONS

One of the tasks of this article was critical review of the historical and modern paradigms that determine the discourse in nature – human health and well-being research. Several concepts from various disciplines were overviewed, for example, Hippocrates «naturalistic history», Humboldt's concept of natural garden design; Oertel 's 'Terrain Kur'; «salutogenic approach» of Antonovsky; McHarg's Design with Nature; Ecopolis programme and Wilson's biophilia.

In the beginning of the 21st century the most important concepts are biophilic cities, therapeutic landscapes, healing gardens, green infrastructure and nature-based solutions.

Analysis of the publications presented in the paper shows, that there are a great number of research which presents the empirical findings by applying a set of different indicators / variables to measure varying exposures to natural elements within a landscape: character and coverage of vegetation within a neighborhood, proximity to parks, participation in outdoors activities in urban green and blue spaces, proximity to water, other contacts with nature (urban gardening etc.). There are also publications which refer to the quality of the nature experience by presenting evidence from the impact of different landscape elements and features on mental, physical, and social health and also their relation to / or value for quality of life, wellbeing, mood states and children's

Fig. 2. a) Playground of Nachbarschaftsgarten – one of the oldest community gardens in Leipzig, Germany, (use of natural materials, forms and colors as alternative to conventional playgrounds as well as planting native vegetation attractive to wildlife and edible for humans). (b) Public park of Perth, Australia, with native vegetation (Photos: D. Dushkova, M. Ignatieva).

Fig. 3. (a) Interpretation of nature-based solution in SLU (Swedish University of Agricutural Sciences) Ultuna Campus-grass free (tapestry) lawn (alternative to conventional lawn biodiverse community attractive for human and wildlife); (b) Low Impact Design practice example (rain garden) in the center of Oslo, Norway (Photos: M. Ignatieva).

health. Another big number of related research deals with the assessment of physical health parameters and analysis of meaning of contact with nature for fixing health problems (cardiovascular diseases, brain functioning, birth outcomes, asthma, heat-related accidents, etc.). Some research revealed that exposure to neighborhood street trees, small parks, or views of nature from a window all have salutary impacts on health. However, it is still not clear – what is the minimum threshold of contact with nature for urban citizens (e.g. the frequency and duration of exposure to nature to provide health benefits as well as density of nature to which people are exposed on health outcomes). The vast majority of the evidence demonstrates that more frequent contact with nature predicts better health outcomes.

The analysis of therapeutic impact of the selected nature-based solutions which were studied within authors' research projects and by applying current approaches revealed the factors of nature influencing the healing process, improving health and well-being. Among them are the opportunity to feel the support from the fellows by spending time together, the feeling of control over own body and environment when being on nature, the positive motivation to start physical exercise inspired by nature, reduction of stressful condition, support the treatment, reduce depression, decrease the effects of heart conditions. It was shown that the particular benefit is connected with the type of nature-based solutions.

By analyzing a variety of nature-based solutions studies within the above-mentioned research projects, we also revealed and characterized the main types of healing gardens and therapeutic landscapes, their main features and different benefits provided by them. We also discussed the multiple benefits of the selected healing gardens and therapeutic landscape (as nature-based solutions) and highlighted the essential principles of biophilia which were used by their creation and development. They include the incorporating of the elements of natural ecosystems, using of natural colors, materials and naturalistic shaped forms, biomimicry (using of forms and functions found in nature), and idea of place attachment (culturally relevant designs which promotes a connection to place and the sense of belonging in order to motivate people to care of and protect it).

Nevertheless, there are several important questions to address in future research: a) we still need to understand about the dose – response (dose – effect) relationship between exposure to natural elements in the landscape and health outcomes; b) what kind of contacts with nature (visual / tactile, direct / through a window / on a screen) can more effectively promote health?

Different types of urban healing gardens and therapeutic landscape, which in the same time can be described as nature-based solutions due to multiple co-benefits they provide, discussed in the paper, provides the evidence demonstrating that exposure to landscapes with natural elements has pervasive, positive, prolonged impact on human health.

New generation of such research is needed to go deeper and explore the process of human – nature relationship and especially their potentials and real benefits in different parts of the world. This is the acute issue and one of the main challenges for the modern cities and their citizens: How we can co-design and co-work with nature in order to have the healthy environment and opportunity to have a long and happy life? As Henry D. Thoreau said „Nature is but another name for health» (Thoreau 1965: 364-365, cited in Da Rocha 2009). It highlights the importance of being in touch with the natural world, posing ourselves as part of nature and thus always have the opportunity of the healing value of nature. And continuing the words of Thoreau that «Health resides in nature» (cited in Da Rocha 2009), the urgent task of urban societies today is to reside and to rediscover the nature in the cities that it can be further source of life, living environment, educational tool, sacral place, inspiration and healing landscape.

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research(RFBR) project «Mathematical-cartographic assessment of medico-ecological situation in cities of European Russia for their integrated ecological characteristics» (2018–2020) under Grant number No 18-05-00236/18 as well as by the Horizon 2020 Framework Programme of the European Union project «Connecting Nature» under Grant Agreement No 730222. and the research project «Perth as a biophilic resilient city model in the time of climate change» by the University of Western Australia (UWA) FABLE research grant (2018 – 2019).

Disclosure statement

No potential conflict of interest was reported by the authors.

REFERENCES

Agavelov V., Brudny A., Bozhukova E., Kavtaradze D. et al. (1985). Ecopolis Programme. Tsukuba: The International Exposition.

Antonovsky A. (1979). Health, Stress and Coping. San Francisco: Jossey-Bass publishers.

Belčáková I., Galbavá P., Majorošová M. (2018) Healing and therapeutic landscape design – examples and experience of medical facilities. Archnet-IJAR, 12 (3), 128-151. DOI: 10.26687/archnet-ijar.v12i3.1637.

Beyer K.M., Kaltenbach A., Szabo A., Bogar S., Nieto F.J. and Malecji K.M. (2014). Exposure to neighborhood green space and mental health: evidence from the survey of the health of Wisconsin. International Journal of Environmental Research and Public Health, 11, 3453-3472.

Brink P., Mutafoglu K., Schweitzer J-P., Kettunen M. et al. (2016) The Health and Social Benefits of Nature and Biodiversity Protection. A report for the European Commission (ENV.B.3/ETU/2014/0039). London/Brussels: Institute for European Environmental Policy.

Cooper Marcus C. and Sachs N. (2013). Therapeutic Landscapes: An Evidence-Based Approach to Designing Healing Gardens and Restorative Outdoor Spaces. Hoboken, NJ: John Wiley & Sons.

Curtis S. (2004). Health and Inequality: Geographical Perspectives. London: SAGE Publications.

Da Rocha A.C. (2009). The value of health in the writings of H. D. Thoreau. Environmental values, 18, 201-215.

Dushkova D. and Evseev A. (2012). The Russian North: Environment and human health risk assessment. In: H. Kremers and A. Susini, eds., RISK Models and Applications. LNIS 6 – Lecture Notes in Information Sciences. Berlin: CODATA Germany, 89-102.

Dushkova D., Ignatieva M. and Melnishuk I. (2020). Urban Greening as a Response to Societal Challenges. Towards Biophilic Megacities (Case Studies of Saint Petersburg and Moscow, Russia). In: J. Breuste, M. Artmann, C. Ioja and S. Qureshi, eds., Making green cities – concepts, challenges and practice. «Cities and Nature» Series. New York, NY: Springer (in press).

Dushkova D. and Haase D. (2020). Not Simply Green: Nature-Based Solutions as a Concept and Practical Approach for Sustainability Studies and Planning Agendas in Cities. Land, 9 (1), 19. DOI: 10.3390/land9010019.

Duzenli T., Tarakci E., Yılmaz S. (2017) A study on healing effects of hospital gardens. Fresenius Environmental Bulletin, 26, 12, 7342-7352.

Dzik A.J. (1997). Looking for Dangerous Places: Some Aspects of Medical Geography and Disease Mapping. West Virginia Medical Journal, 93, 250-253.

EC – European Commission (2015). Towards an EU Research and Innovation policy agenda for nature-based solutions & re-naturing cities. Final Report of the Horizon2020 Expert Group on Nature-Based Solutions and Re-Naturing Cities. Brussels: European Commission.

Elings M. (2006). People-plant interaction: The physiological, psychological and sociological effects of plants on people, farming for health. In: Hassink J., van Dijk M. (eds). Springer agriculture book. Dodrecht: Springer, 43-55.

Frumkin H., Bratman G.N., Breslow S.J., Cochran B. et al. (2017). Nature Contact and Human Health: A Research Agenda. Environ Health Perspectives, 125 (7), 075001. DOI: 10.1289/EHP1663.

Gesler W.M. (1992). Therapeutic landscapes: Medical issues in the light of the new cultural geography. Social Science and Medicine, 34 (7), 735-746.

Groenewegen P.P., van den Berg A.E., de Vries S. and Verheij R.A. (2006). Vitamin G: effects of green space on health, well-being, and social safety. BMC Public Health, 6, 149. DOI: 10.1186/1471-2458-6-149.

Gupta K.P., Kumar P., Pathan S.K., Sharma K.P. (2012). Urban neighborhood green index – a measure of green spaces in urban areas. Landscape and Urban Planning 105 (3), 325-335.

Hankey S. and Marshall J.D. (2017). Urban Form, Air Pollution, and Health. Current Environmental Health Reports, 4, 491-503. DOI: 10.1007/s40572-017-0167-7.

Hartig T., Mitchell R., de Vries S. and Frumkin H. (2014). Nature and health. Annual Review of Public Health, 35, 207-228.

Hartig T., Mang M., Evans G.W. (1991). Restorative effects of natural environment ex- periences. Environ Behav. 23(1), 3-26.

Humboldt A. (1806). Ideen zu einer Physiognomik der Gewächse. Tübingen: Cotta Verlag (in German).

Ignatieva M. (2000). Ecopolis – towards the total holistic city: lessons in integration from throughout the world. In: G. Stewart and M. Ignatieva, eds., Urban biodiversity and ecology as a basis for holistic planning and design. Proceedings of a Workshop at Lincoln University, Lincoln University International Centre for Nature conservation, Publication № 1. Christchurch: Wickliffe Press, 102-105.

Ignatieva M. (2018). Biodiversity-friendly designs in cities and towns: Towards a global biodiversinesque style. In: A. Ossola and J. Niemelä, eds., Urban Biodiversity: From Research to Practice. Oxon: Routledge (Routledge Studies in Urban Ecology), 216-235.

Janke K., Propper C. and Henderson J. (2009). Do Current Levels of Air Pollution Kill? The Impact of Air Pollution on Population Mortality in England. Health Economics, 18(9), 1031–55.

Jiang S. (2014). Therapeutic landscapes and healing gardens: A review of Chinese literature in relation to the studies in western countries. Frontiers of Architectural Research, 3, 141–153.

Kabisch N. and Haase D. (2018). Urban nature benefits – Opportunities for improvement of health and well-being in times of global change. Newsletter on Housing and Health, 29, 1-11.

Kaplan S. (1992). The restorative environment: nature and human experience. Role of horticulture in human well-being and social development: a national symposium, 19-21 April 1990, Arlington, Virginia. Portland: Timber Press, 134-142.

Kearn R.A. and Moon G. (2002). From medical to health geography: theory, novelty and place in a decade of change. Progress in Human Geography, 26, 587-607Landrigan P.J. and Fuller R. (2015). Global health and environmental pollution. International Journal of Public Health, 60 (7), 761-762. DOI: 10.1007/s00038-015-0706-7.

Lawrence R., Capon A. and Siri H. (2017). Lessons from Hippocrates for contemporary urban health challenges. Cities & Health (1) 1, 72-82. DOI: 10.1080/23748834.2017.1372967

Lips-Castro W. (2015). A brief history of the natural causes of human disease. Gaceta Medica de Mexico, 151, 749-761.

Li X, Meng Q, Li W, Zhang C, Jancso T, Mavromatis S. 2014. An explorative study on the proximity of buildings to green spaces in urban areas using remotely sensed imagery. Ann GIS 20 (3), 193–203. doi:10.1080/19475683.2014.945482.

Maas J., Verheij R.A., de Vries S., Spreeuwenberg P., Schellevis F.G. and Groenewegen P.P. (2009). Morbidity is related to a green living environment. Journal of Epidemiology and Community Health, 63, 967-973.

Malkhazova S., Mironova V., Shartova N. and Orlov D. (2019). Mapping Russia's Natural Focal Diseases: History and Contemporary Approaches. Basel: Springer Nature. DOI. 10.1007/978-3-319-89605-2.

Maller C., Townsend M., Pryor A., Brown P., Leger L.S. (2006) Healthy nature – healthy people: 'contact with nature' as an upstream health promotion intervention for populations. Health Promotion International, 21 (1), 45-54. DOI: 10.1093/heapro/dai032.

Marcel M.R., Stadler J., Korn H., Irvine K. and Bonn A. eds. (2019). Biodiversity and Health in the Face of Climate Change. Cham: Springer International Publishing. DOI: 10.1007/978-3-030-02318-8.

May J. M. (1959). The Ecology of Human Disease. New York, NY: MD Publications Inc.

McHarg I.L. (1969). Design with nature. New York, NY: Garden City.

Oertel M.J. (1886). Über Terrain-Curorte zur Behandlung von Kranken mit Kreislaufstörungen, Kraftabnahme des Herzmuskels, Lungenkreislauf etc. Leipzig: Vogelverlag (in German).

Pavlovsky E.N. (1966). Natural nidality of transmissible diseases. Urbana, IL: University of Illinois Press.

Pötz H. and Bleuze P. (2012). Urban green-blue grids for sustainable and dynamic cities. Delft: Coop for Life.

Revich B. (2018). Priority factors in urban environments that affect the quality of life for metropolitan populations. Studies on Russian Economic Development, 29(3), 267-273. DOI: 10.1134/S1075700718030115.

Raymond C.M., Franzeskaki N., Kabisch N. et al. (2017). A framework for assessing and implementing the co-benefits of nature-based solutions in urban areas. Environmental Science & Policy, 77, 15-24. DOI: 10.1016/j.envsci.2017.07.008

Rohde C.L., Kendle A.D. (1994). Human well-being, natural landscapes and wildlife in urban areas. English Nature Science Publications, Royal Jubilee Hospital Patient Care Centre Project. Pacific Green.

Russo A., Ignatieva M., Cirella G. T. et al. (2017) Biophilia: Nature-based solutions for sustainable cities. In: M. Ignatieva et. al., eds., Three pillars of landscape architecture: design, planning and management. S.-Petersburg: S.-Petersburg State Polytechnical University, 105-112.

Schweikart R. and Kistemann T. (2013). Mapping health and health care. Kartographische Nachrichten, 63(1), 3-11.

Söderback I., Söderström M. and Schälander E. (2004). Horticultural therapy: the «healing garden» and gardening in rehabilitation measures at Danderyd hospital rehabilitation clinic, Sweden. Developmental Neurorehabilitation, 7 (4), 245-260.

Soga M. and Gaston K.J. (2016). Extinction of experience: the loss of human-nature interactions. Frontiers in Ecology and the Environment, 14, 94-101.

Souter-Brown G. (2014). Landscape and Urban Design for Health and Well-Being: Using Healing, Sensory and Therapeutic Gardens. 1st ed. New York, NY: Routledge.

Stigsdotter U.A., Grahn P., Stigsdotter U.A., Grahn, P. (2002). What makes a garden a healing garden? Journal of Therapeutical Horticulture 13, 60-69.

Sullivan W.S. and Chang Ch.-Y. (2017). Landscapes and Human Health. International Journal of Environmental Research and Public Health, 14(10), 1212. DOI: 10.3390/ijerph14101212.

Tzoulas K., Korpela K., Venn S. et al. (2007). Promoting ecosystem and human health in urban areas using green infrastructure: a literature review. Landscape and Urban Planning, 81, 167-178.

Valencius C.B. (2000). Histories of medical geography. Medical History. Suppliment, 20, 3-28.

van den Bosch M. and Bird W. (2018) Oxford Textbook of Nature and Public Health: The Role of Nature in Improving the Health of a Population. Oxford: Oxford University Press.

WHO-UNEP Health and Environment Linkages Initiative (2008). Health environment: managing the linkages for sustainable development: a toolkit for decision-makers: synthesis report. World Health Organization. Available at: www.apps.who.int/iris/handle/10665/43946 [Accessed 1 June 2019].

Williams A. (2017). Therapeutic Landscapes. London: Routledge. DOI: 10.4324/9781315551166.

Wilson E.O. (1986). Biophilia: the Human Bond with Other Species. Cambridge: Harvard University Press.

Wood C.J., Pretty J. and Griffin M. (2016). A case-control study of the health and well-being benefits of allotment gardening. Journal of Public Health, 38, e336-e344.

Received on August 3th, 2019 Accepted on January 21th, 2020

Yurij Vasil'chuk[1*], Julia Chizhova[2], Natalia Frolova[1], Nadine Budantseva[1], Maria Kireeva[1], Alexander Oleynikov[1], Igor Tokarev[3], Ekaterina Rets[1], Alla Vasil'chuk[1]
[1] Lomonosov Moscow State University, Faculty of Geography, Moscow, Russia
[2] Institute of Geology of Ore Deposits, Petrography, Mineralogy, and Geochemistry (IGEM RAS), Moscow, Russia
[3] The Centre for X-ray Diffraction Studies, Research park at St. Petersburg University, St. Petersburg, Russia
***Corresponding author:** vasilch_geo@mail.ru

A VARIATION OF STABLE ISOTOPE COMPOSITION OF SNOW WITH ALTITUDE ON THE ELBRUS MOUNTAIN, CENTRAL CAUCASUS

ABSTRACT. This study aims to analyze the stable isotope composition of the snow cover of the Elbrus Mountain – the highest mountain in Europe. Snow sampled in the middle accumulation period in January 2017, February 2016, January 2001 and during snowmelt in July 1998 and August 2009. Snow sampled at the south slope of Mt. Elbrus at different elevations, and the total altitude range is approximately 1700 m. A significant altitude effect in fresh snow precipitation was determined in February 2001 with gradient –1.3‰ $\delta^{18}O$/100 m (–11.1‰ δ^2H /100 m) at 3100-3900 m a.s.l. and inverse altitude effect in February 2016 with gradient +1.04‰ $\delta^{18}O$ /100 m (+8.7‰ δ^2H /100 m) at 3064-3836 m a.s.l. There is no obvious altitude effect of the δ^2H and $\delta^{18}O$ values in snow at the Elbrus slope in 2017, except for the height range 2256-3716 m a.s.l., where altitudinal effect of $\delta^{18}O$ values was roughly -0.32‰/100m. The $\delta^{18}O$ values in the winter snowpack in some cases decrease with increasing altitude, but sometimes are not indicating a temperature-altitude effect. Post-depositional processes cause isotopic changes, which can result from drifting, evaporation, sublimation, and ablation. The study of altitude effect in snow is important for understanding the processes of snow-ice and snow-meltwater transformation and the snow/ice potential to provide paleo-environmental data.

KEY WORDS: stable isotopes, spatial variability, snow cover on glaciers, high altitude, Caucasus

CITATION: Yurij Vasil'chuk, Julia Chizhova, Natalia Frolova, Nadine Budantseva, Maria Kireeva, Alexander Oleynikov, Igor Tokarev, Ekaterina Rets, Alla Vasil'chuk (2020) A variation of stable isotope composition of snow with altitude on the Elbrus Mountain, Central Caucasus. Geography, Environment, Sustainability, Vol.13, No 1, p. 172-182
DOI-10.24057/2071-9388-2018-22

INTRODUCTION

For all studies that require information about the isotopic composition of the snowpack in a catchment, detailed information on the spatial and temporal variability of the isotopic content of snow is valuable. The water input in snow-dominated watershed for residence time analysis, end member mixing analysis or the detection of source water contribution requires a detailed understanding of the effects of factors that modify the isotopic composition of snow cover.

During orographic lift of air mass, the heavier water molecules condense at first, i.e. the precipitation is isotopically enriched, and the cloud moisture is subsequently depleted due to continuous precipitation under equilibrium fractionation.

Depletion of the isotopic composition of precipitation with elevation is the "altitude isotope effect" – the altitude isotope effect in precipitation well-known since the Dansgaard (1964) research. The altitude effect is temperature-related because the condensation is caused by the temperature drop due to the increasing altitude. Due to the decreasing pressure with increasing altitude, a larger temperature decrease is required to reach the saturated water vapor pressure than for isobaric condensation. Moser and Stichler (1970) observed altitudinal isotope effect in fresh snow at the Kitzsteinhorn in Austrian Alps. An elevation gradient for $\delta^{18}O$ values between is –0.6 and –1.0‰ per 100 in high mountains in the snow was determined (Niewodniczanski et al. 1981). However, there was in a wide range of isotope values with small-scale inverse gradient and thus only partly attributable to a linear elevation gradient. These variations explained by conditions during snowfall and after snow deposition, such as wind drift and fractionation by melting processes, as well as orographic and climatic features of the studied areas (Niewodniczanski et al. 1981). For the fresh snow in the Canadian Rocky Mountains, elevation gradients range from –0.3 to +1.8‰ per 100 m, depending on snowfall and accumulation conditions (Moran et al. 2007).

In contrast to the altitude isotope effect in fresh snow, the isotopic composition of an entire snowpack is more complex (Moser and Stichler 1974). The snowpack is altered by sublimation, evaporation, metamorphism of snow crystals, percolating of meltwater and isotopically enriched precipitation, especially in temperate climatic conditions (Judy et al. 1970; Ambach et al. 1972; Stichler et al. 2001; Sokratov and Golubev 2009). These processes may conceal the altitudinal effect in fresh snow and result in inverse gradients (Moser and Stichler 1970). In some cases, there was no significant relation between the isotopic signature of the entire snowpack and elevation (Raben and Theakstone 1994; Kang et al. 2002; Königer et al. 2008).

Dietermann and Weiler (2013) observed only a limited altitude isotope effect in Swiss Alps. The altitudinal effect for δ^2H values at the south-facing slopes ranged from –6.2 to +2.6‰/100 m with a wide variability for the individual samples. These results confirm the influence of melting processes altering the mean isotopic composition of snow cover. The north-facing slope of this catchment is a steep avalanche-prone slope popular for mountain skiing. These factors likely lead to snow mixing and disturbance of the altitude isotopic effect (Dietermann and Weiler 2013).

We have found the altitudinal effect on $\delta^{18}O$ and δ^2H values of fresh snow on the southern slope of Elbrus in January 2001, decreasing with increasing altitude (Vasil'chuk and Chizhova 2010). The of snow becomes more and more depleted in ^{18}O and ^2H content at higher elevations. In the range of 3100–3400 m a.s.l., the gradients are –2.4 ‰ / 100 m for $\Delta\delta^{18}O$ and –20 ‰ for $\Delta\delta^2H$ values, and at altitudes of 3400–3900 m they are –0.6 / 100 m for $\Delta\delta^{18}O$ and –6 ‰ for $\Delta\delta^2H$ values. We found a decreasing d-excess values in snow at 3100 to 3400 m a.s.l. We associated high gradients at altitudes from 3100 to 3400 m with intensive washing out of the air mass and isotopic depletion during precipitation. However, if we assume, based on the isotope data, the decrease of $\delta^{18}O$ values from –17 to –29‰ in snow with increasing altitude from 3100 to 3400 m due to progressive precipitation, the decreasing of d-excess in this snow could not be explained.

The progressive rainout process based on the Rayleigh fractionation/condensation model predicts increasing d-excess values in the latest stages of precipitation. Also, the equilibrium Reyleigh condensation model including the isotopic kinetic effect (Jouzel and Merlivat 1984) as well as isotopic model including mixed cloud processes (Ciais and Jouzel 1994) predict relatively high values of d-excess.

In many cases, the d-excess is found to increase with altitude on the mountain slopes, possibly for a variety of reasons. This issue has not been finally resolved. Hereby, a decreasing d-excess with decreasing $\delta^{18}O$ values in fresh snow in January 2001 may indicate a very ambiguous formation of the isotopic composition of snow cover on the southern slope of Elbrus.

Recent studies of glaciers of Elbrus are focused on obtaining information about the environmental conditions of ice accumulation, including sources of air masses, atmospheric conditions, and the transformation of snow to ice. The stable isotopes and chemical composition are indicators of the processes involved in atmospheric precipitation.

Observations of recent retreat of the Elbrus glacier system (Vasil'chuk et al. 2006, 2010; Zolotarev and Kharkovets 2010; Holobâcă 2016; Tielidze and Wheate 2017) show significant changes of the glaciers volume and their 'tongues' retreats during the past 100 years, however, there appears to be no signature in the isotopic composition of the glacier ice (Vasil'chuk et al. 2006).

Although most of the incoming moisture to the Elbrus is of Atlantic origin, some air masses drift from the southern deserts. The dust that originates from the foothills of the Djebel Akhdar in eastern Libya and transported to the Caucasus along the eastern Mediterranean coast, Syria and Turkey (Shahgedanova et al. 2013) was found in snowpack of Garabaschi glacier.

Elbrus glacier's ice is paleo-archive, especially at altitudes above 4900 m, where isotope record is undisturbed by the meltwaters due to the absence of melting at this elevation. The isotope records of low-latitude and high-mountain glaciers cores have the potential to provide detailed paleo-environmental proxy record and to prove extremely valuable in producing continuous records of atmospheric chemistry and climate (Thompson et al. 1998, 2006a, 2006b; Tian et al. 2003; Yao et al. 2006). The low-latitude Tibetan cores records, especially Dunde and Dasuopu, are consistent with the local temperature records (Yao et al. 2006), the more northern sites, similar to Dunde, are thought to be more temperature-dominated (e.g., Tian et al. 2003).

In recent years, deep drilling of Elbrus glaciers allowed to obtain $\delta^{18}O$ and δ^2H records (Mikhalenko et al. 2015; Kozachek et al. 2017). There was no significant correlation between ice core $\delta^{18}O$ records from the western Elbrus plateau (height 5115 m a.s.l) with regional temperature, neither with the reanalysis data nor with the data of meteostation (Mikhalenko et al. 2015; Kozachek et al., 2017). At the western Elbrus plateau, the snow accumulation rate is high and moreover, pronounced seasonal variations of $\delta^{18}O$ and δ^2H values were noted in the core. In spite of the presence of $\delta^{18}O$ amplitude in ice core, which could indicate the existence of a $\delta^{18}O$-temperature relationship, conditions of individual snowfall play an important role. Such conditions include snow mixing by wind and different air masses with different source characteristics affect the precipitation at the base and crest of a mountain.

In this case, the study of the formation of an isotope altitude effect in snow is important for understanding this exclusion of temperature effect in ice. Here, the results of isotope analysis of the snow samples are presented to provide a better understanding of the spatial and temporal features of snow accumulation on Elbrus.

STUDY SITE

The Caucasus Mountains are located between the Black and the Caspian Seas and generally oriented from east to southeast, with the Greater Caucasus range often considered as the divide between Europe and Asia. The total area of glaciers in the Caucasus is about 1121±30 km^2 (Mikhalenko et al. 2015). Glaciers on the Elbrus Mountain are located in the altitudinal range 2800-5642 m.

The coldest conditions occur above 5200 m a.s.l., where the mean summer air temperature does not exceed 0°C, while the Elbrus glaciers between 4700 and 4900 m a.s.l. are prone to surface melting. Snow accumulation measurements from 1985 to 1988 showed total snow accumulation of 400–600 mm w.e. a^{-1} with considerable wind-driven snow erosion at the col of Elbrus (5300 m a.s.l.).

The summer atmospheric circulation pattern in the Caucasus is dominated by the subtropical high pressure to the west and the Asian depression to the east. In winter, circulation is affected by the western extension of the Siberian High (Volodicheva 2002). The Caucasus is located in the southern part of the vast Russian Plain and therefore buffeted by the unobstructed passage of cold air masses from the north. High mountain ridges in the southern Caucasus deflect air flowing from the west and southwest. The influence of the free atmosphere on the Elbrus glacier regime is greater than local orographic effects as the glacier accumulation area lies above main ridges.

Most part of the annual precipitation falls in the western and southern parts of the Caucasus. For the southern slope of the Caucasus, the amount of precipitation ranges from 3000–3200 mm a^{-1} in the west to 1000 mm a^{-1} in the east. The proportion of winter precipitation (October–April) also decreases eastward from more than 50 to 35–40% for the northern Greater Caucasus and from 60–70 to 50–55% for the southern slope (Rototaeva et al. 2006). The proportion of solid precipitation increases with altitude and reaches 100% above 4000–4200 m.

Our research is focused on the southern slope of Elbrus, from the Azau station (2330 m a.s.l.), along the Garabashi Glacier (43°20´ N, 42°26´ E) to the summit (Fig. 1). The paper discusses data obtained in 2017, 2016, and, also the data we have obtained in previous years (Vasil'chuk et al., 2006, 2010; Vasil'chuk and Chizhova 2010).

In terms of temperature, the 2015/16 season continued a unique series of warm winters, which began in 2009/10. The temperature anomaly was formed due to warm months at

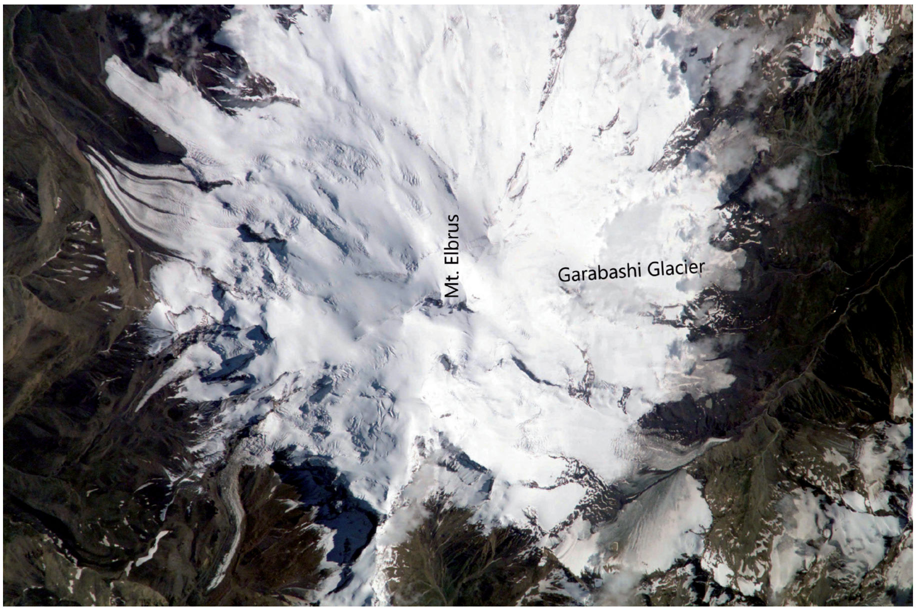


Fig. 1. Space image of Elbrus from SPOT 7 satellite, August 20, 2016

the beginning (November) and the end of winter (February, March). While the traditionally cold months (December, January) were slightly different from the long-year norm.

According to the amount of precipitation, the 2015/16 season was 18% below normal, and the winter maximum precipitation was recorded in January. During the January snowfall, 101.9 mm of precipitation fell, which was 36% of the precipitation of the entire cold period (XI-III). The thickness of the snow at the bottom of the valley during the second decade of January increased from 30 cm to 72 cm (weather station Terskol, 2141 m a.s.l.) and from 59 cm to 103 cm (Azau weather station).

Winter 2016/17 was characterized by extremely low precipitation and long periods without precipitation, for example, until January 26, 2017, snow fell only on 26 November. The level of temperature drop with altitude (lapse rate) for the southern slope of Elbrus is concerned to be 0.6 °C / 100 m. This lapse rate was determined by comparison of automatic weather station (installed on the western Elbrus plateau at 5115 m a.s.l.) record with measurements from the nearest meteorological station (Mikhalenko et al. 2015).

METHODS

Snow was sampled on the south slope of Mt. Elbrus in the middle of the snow accumulation period in January 2017, February 2016, January 2001 and during snowmelt season in July 1998 and August 2009. In 2017, fresh snow was sampled. In 2016, the surface snow (1 Feb) and fresh snow (3 Feb) were sampled, in 2001 and 2009 fresh snow was sampled. During the ablation season of 1998, surface snow was sampled. The sampling was performed at an altitude of about 1700 m (Fig. 2).

Samples of surface snow (from a depth of 0-15 cm) were collected on 1 February 2016 (according to the Terskol weather station, precipitation events were on 28 and 29 January). Samples of fresh snow collected on 3 February 2016 (from a depth of 0-15 cm) represent snow fell on 2 February from morning till evening. According to the Terskol weather station on 2 February, 9 mm of precipitation fell.

In January 2001 and August 2009, samples of just deposited snow also have been collected within three hours after snowfall at 0-10 cm depth of snow cover. Samples of melted snow were collected at the Garabaschi glacier in July 1998.

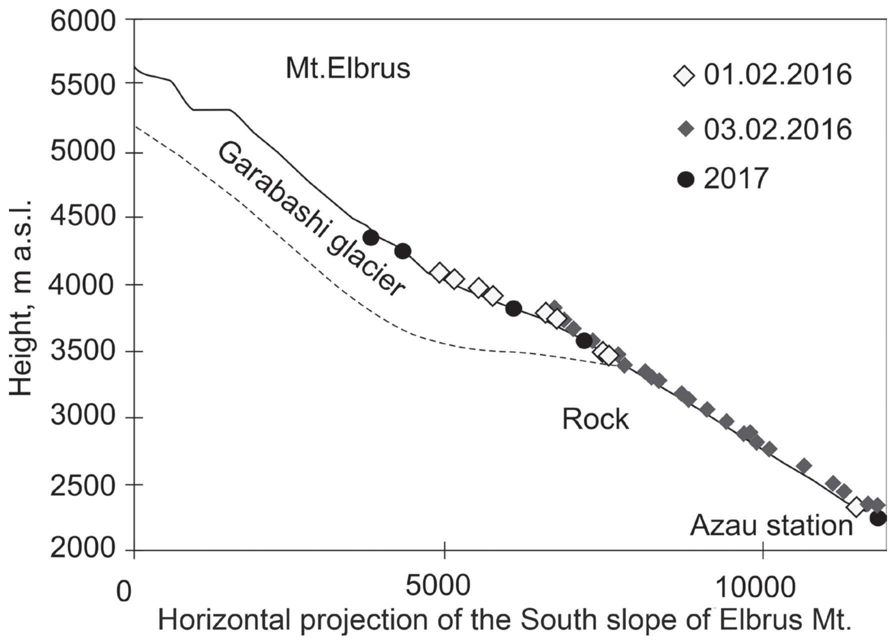


Fig. 2. Sampling profile on the southern slope of Elbrus

In January 2017, samples of fresh snow were collected within the range of 2300-3800 m a.s.l. On January 27 and 28, in the valley and on the slope of Elbrus, the snow fallen on 26 January was sampled. On January 29, snow fell in the afternoon and newly deposited snow was sampled on 30 January. All snow samples were taken from 0-10 cm depth of snow cover. In the valley (2332 m a.s.l.) and at the slope (3345 m a.s.l.) snow pits had been excavated at 10 cm intervals.

Isotope ratios in snow of 2016, 2017 and 2009 were measured by a Finnigan Delta-V continuous flow mass spectrometer in Stable Isotope Laboratory of Geographical Department of Lomonosov Moscow State University. Concurrently, isotope composition of snow sampled in 2016, was determined in Saint Petersburg State University Resource center for Geo-Environmental Research and Modeling (GEOMODEL) by Picarro L-2120i. The differences in measured $\delta^{18}O$ values for the same samples in two laboratories does not exceed ±0.3‰. Isotope composition of snow sampled in 2001 and 1998 was measured by W.Papesch in Research center "Arsenal" in Seibersdorf, Austria.

Isotope data are expressed conventionally as δ-notion (‰), representing a deviation in parts per thousand, relating to the isotopic composition of V-SMOW (Vienna Standard Mean Ocean Water). International standards V-SMOW2, GISP, SLAP2 were used for the calibration. The measurement precision for $\delta^{18}O$ values is ±0.1‰.

RESULTS

The $\delta^{18}O$ - δ^2H ratios for all snow samples (accumulation and melt) are shown in Fig. 3. Most of them are very close to the global meteoric water line.

Altitude isotope effect in fresh winter snow is clearly visible in 2016 at elevation up to 3000 m a.s.l and in 2001 at elevation from 3000 to 4000 m a.s.l (Fig. 4, a). Inverse altitude effect was observed in fresh snow sampled in 2016 above 3000 m a.s.l and in August 2009 (Fig. 4, b).

The values of d-excess (d_{exc}) in fresh snow have seasonal variations increasing in summer (Table 1). The d_{exc} values depend on the relative humidity of the air masses at their oceanic origin (Merlivat and Jouzel 1979). The lower d_{exc} values of precipitation in the northern hemisphere during the summer months correspond with the higher relative air humidity which relates to the SST in the oceanic source regions of the air masses concerned. In the Chinese Tien Shan, high d_{exc} values were noted during the winter months when pre-

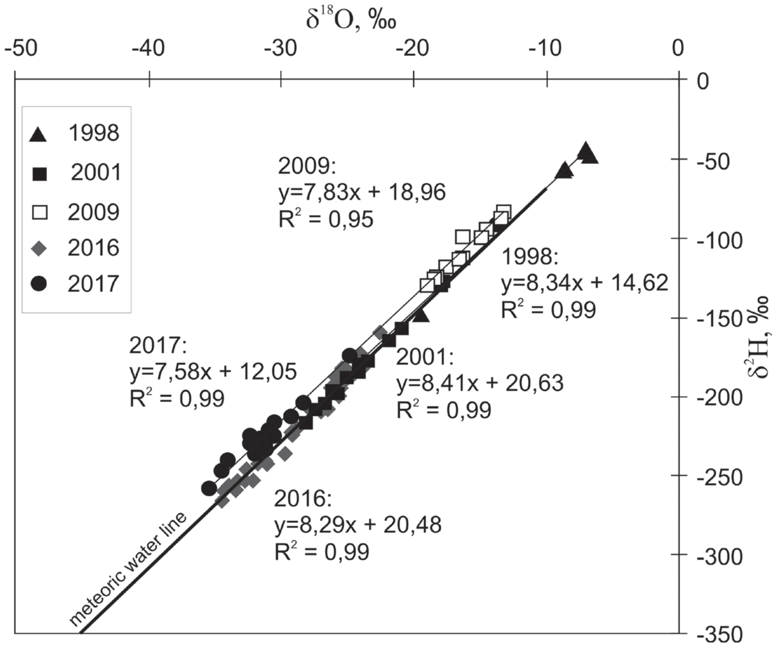


Fig. 3. The $\delta^{18}O$ - δ^2H plot for all snow samples: 1998, 2001, 2009, 2016, 2017

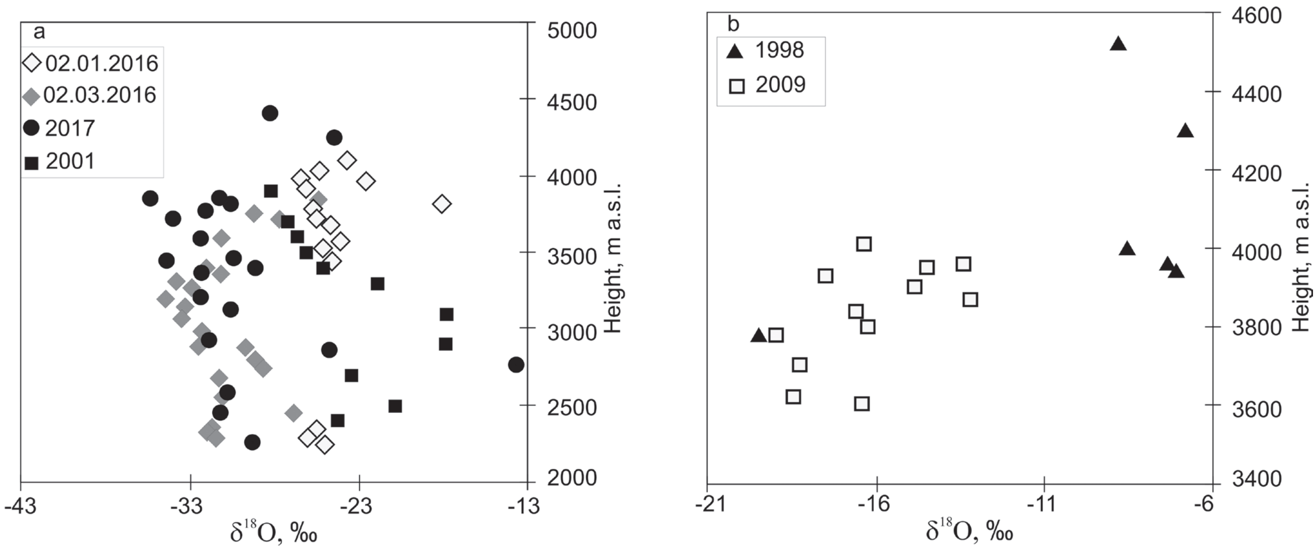


Fig. 4. The $\delta^{18}O$ values in snow cover of Elbrus Mountain in winter (a) and summer season (b)

Table 1. Deuterium excess of snow cover of Elbrus Mountain

Year	Type of snow	d_{exc}, ‰		
		mean	max	min
1998	Summer melted	11	15	7
2001	Winter fresh	11	14	9
2009	Summer fresh	22	32	18
2016	Winter fresh	12	22	2
2017	Winter fresh	24	33	17

cipitation fell at low temperatures and low relative humidity (Pang et al. 2011). On the contrary, the higher d_{exc} values during the winter months are caused by the lower relative humidity at the oceanic source regions. The inverse trend for d_{exc} values in summer precipitation may occur in regions where the atmospheric water vapor dominates due to moisture evaporation from continental basins (Schotterer et al. 1993; Schotterer et al. 1997).

We consider separately the surface snow (which lay after falling out for some time and was subjected to various processes: re-deposition, sublimation, melting, drifting, etc.) and fresh newly deposited snow, which was sampled either immediately after snowfall or the next day after.

Newly deposited (fresh) snow

In fresh snow sampled on 27 and 28 January 2017, the $\delta^{18}O$ values vary from −24.84 to −34.46‰, the δ^2H from −173.2 to −247.9‰ (Table 2). Regardless of the date of selection, all samples are in the altitude range of 2256-3850 m a.s.l. Practically, there is no clear relationship between the $\delta^{18}O$ and δ^2H values and altitude (Fig. 5). However, some weak trend to decreasing $\delta^{18}O$ values with altitude can be identified by calculating the difference between isotope content at 2256 m a.s.l. and 3850 m a.s.l. It gives the gradient to be −0.32‰ $\delta^{18}O$/100 m.

On 3 February 2016, the $\delta^{18}O$ values increased from −34.5‰ to −25.5‰ in fresh snow between 2287 and 3836 m a.s.l., the lowest $\delta^{18}O$ values were obtained near 3000 m a.s.l. (Fig. 6). It was found that there is a clear inverse altitudinal isotope effect between 4000 and 3000 a.s.l. with a gradient of $\delta^{18}O$ = +1.04‰/100 m (Table 2). Below 3000 m a.s.l., the $\delta^{18}O$ and δ^2H values distributed randomly which could be attributed to the lower boundary of air mass or turbulent mixing inside of it.

In August 2009, fresh snow showed weak positive isotope trend with altitude with a low statistical significance (Fig. 7).

In 2001, the altitudinal isotope effect has been observed above 3000 m in fresh snow with gradient of $\delta^{18}O$ −1.3‰/100 m and δ^2H −11.1‰/100 m (at 3100 m $\delta^{18}O$ = −17.81‰, δ^2H = −128.1‰, at 3900 m $\delta^{18}O$ = −28.24‰, δ^2H = −217.1‰, see Fig. 6).

Simultaneous temperature measurements on the southern slope of Elbrus in the altitude range of 2355−3853 m showed temperature drop with altitude for different types of weather in the 2016/17 season (Table 3).

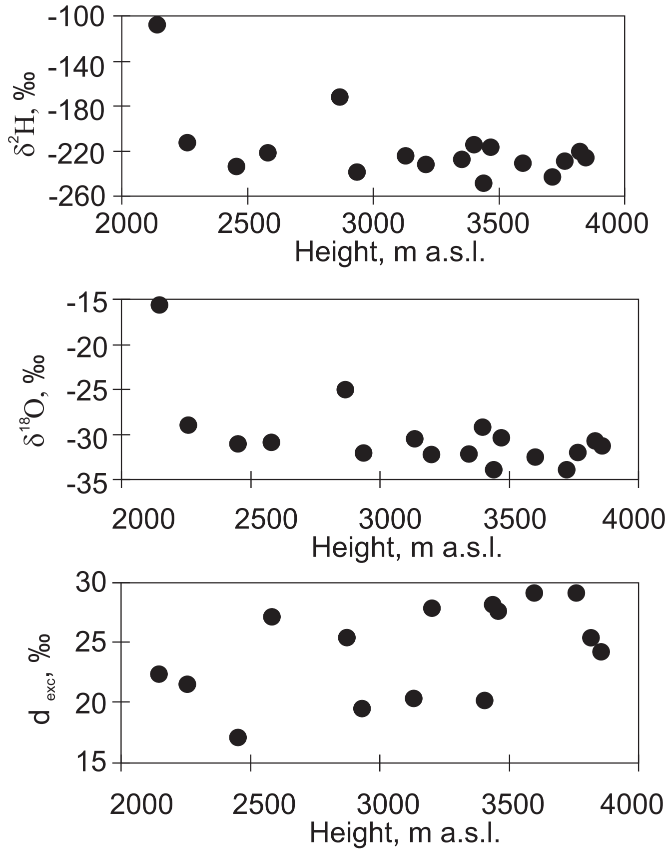


Fig. 5. The altitudinal effect on $\delta^{18}O$, δ^2H and d_{exc} for fresh snow in 27 January 2017

Table 2. The altitudinal distribution of δ18O and δ2H values in fresh snow on Elbrus in 2016 and 2017

Sample ID	Height, m	$\delta^{18}O$, ‰	Δ $\delta^{18}O$, ‰/100 m	$\delta^{2}H$, ‰	d_{exc}, ‰
03.02.2016					
E19 c	3836	–25.5	+1.04	–191	12.68
E20 c	3747	–29.2		–224	9.63
E21 c	3705	–27.7		–209	12.04
E22 c	3588	–31.1		–243	6.45
E23 c	3457	–34.4		–260	15.59
E24 c	3403	–32.1		–242	14.87
E26 c	3351	–31.3		–237	12.96
E27 c	3307	–33.9		–257	14.51
E28 c	3255	–32.8		–253	10.02
E29 c	3197	–34.5		–266	10.13
E30	3142	–33.3		–253	13.43
E31	3064	–33.5		–259	8.98
E32	2978	–32.2	Not pronounced	–253	5.18
E35	2353	–31.8		–242	12.11
E34	2321	–32.1		–242	14.71
E33	2287	–31.6		–241	11.32
04.02.2016					
E36 c	2908	–32.2	Not pronounced	–252	6.12
E37 c	2462	–27.0		–209	6.51
E38 c	2884	–32.6		–247	14.26
E39 c	2548	–31.2		–236	13.57
E40 c	2872	–29.7		–236	1.80
E41 c	2796	–29.2		–222	11.61
E42 c	2744	–28.7		–219	10.83
E43 c	2665	–31.3		–241	9.12
27.01.2017					
E1-2017	3850	–31.28	Not pronounced	–225,9	24.3
E7-2017	3820	–30.65		–220	25.2
E2-2017	3764	–32.12		–228	28.96
E9-2017	3716	–33.97	–0.32	–241.6	30.16
E3-2017	3598	–32.47		–230.8	28.96
E4-2017	3468	–30.47		–216.5	27.26
E15-2017	3443	–34.46		–247.9	27.78
E5-2017	3400	–29.15		–213	20.2
E8-2017	3351	–32.31		–225.2	33.28
E6-2017	3205	–32.31		–230.8	27.68
E16-2017	3130	–30.61		–224.5	20.38
E10-2017	2931	–32.02		–236.8	19.36
E11-2017	2865	–24.84		–173.2	25.52
E13-2017	2579	–30.94		–220.3	27.22
E14-2017	2454	–31.24		–232.9	17.02
E18-2017	2256	–29.28		–212.7	21.54
E19-2017	2145	–15.63		–102.7	22.34

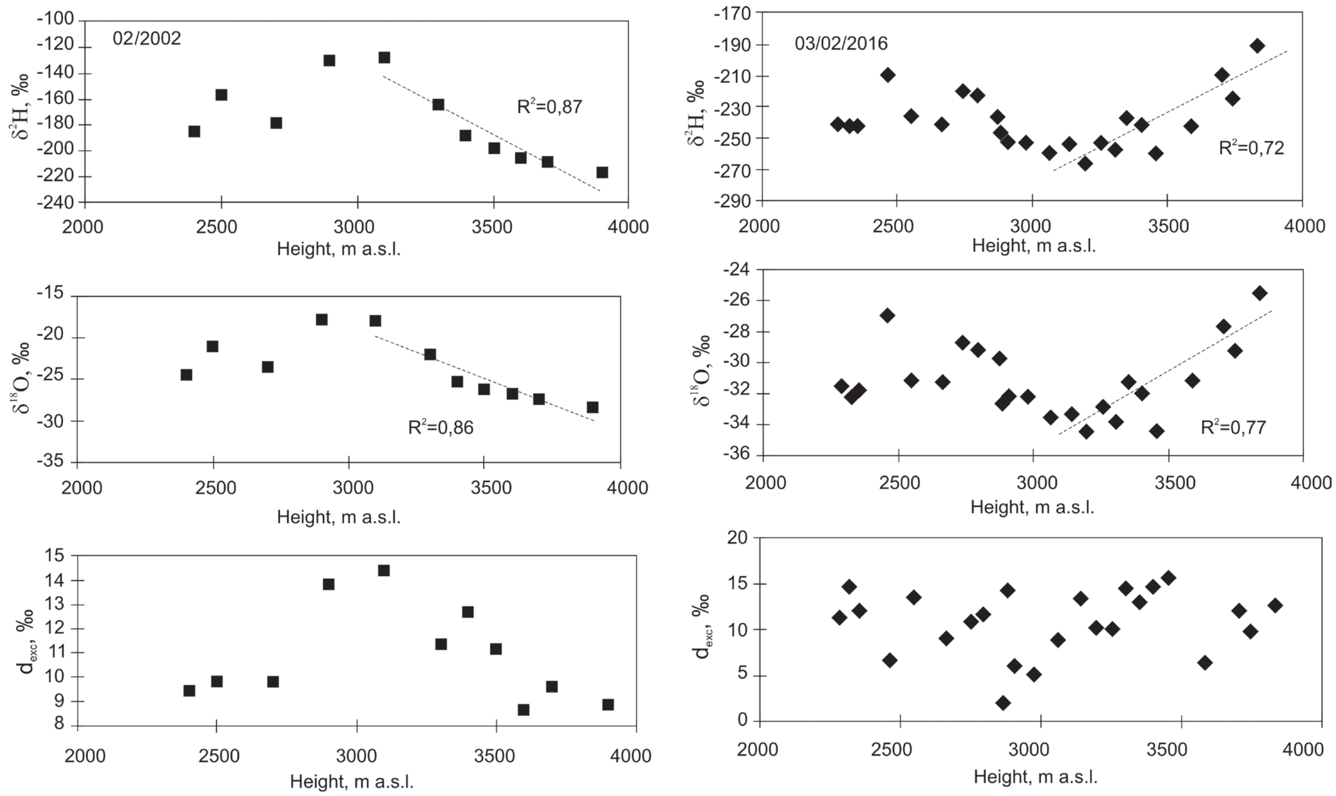


Fig. 6. The altitudinal effect on $\delta^{18}O$, δ^2H and d_{exc} for fresh snow in 8 February 2001 and 3 February 2016

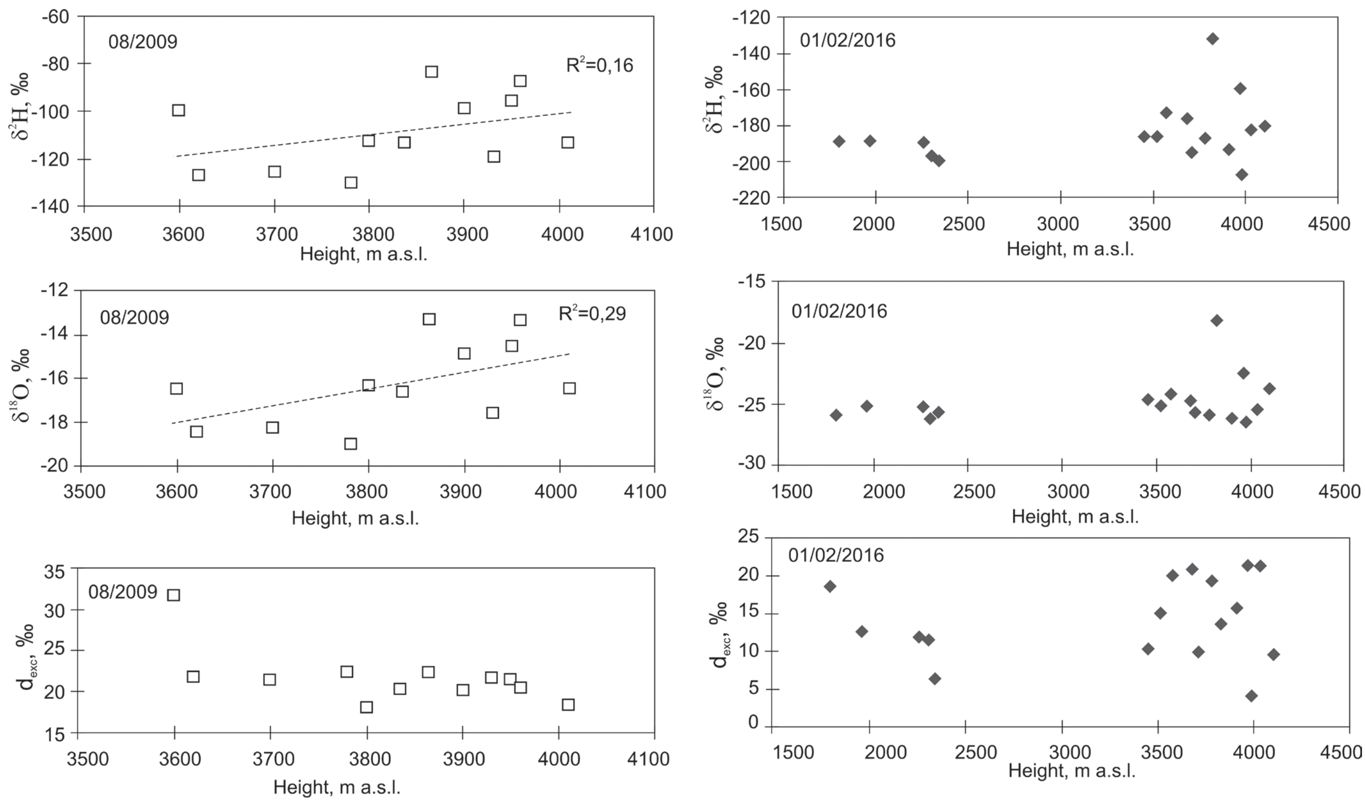


Fig. 7. The altitudinal effect on $\delta^{18}O$, δ^2H and d_{exc} for fresh snow in August, 2009 and for surface snow 1 February, 2016

These values mean that for the snow samples of 2001, in which the altitude isotope effect is pronounced (from 2900 to 3900 m a.s.l.), the relationship coefficient $\delta^{18}O$ with T is in the range from 0.55 ‰/˚C to 0.76 ‰/˚C (based on Table 3 data for different types of weather). This corresponds to the Rayleigh model of equilibrium isotope fractionation.

The decrease of d_{exc} values with altitude was revealed (see Fig 6). Such decreasing d_{exc} values during snowfall contradicts other model calculations (Jouzel and Merlivat 1984; Ciais and Jouzel 1991) and field observations (Vasil'chuk et al. 2005).

Surface snow

Snow sampled on February 1, 2016, in the range of 1900 m – 4100 m a.s.l. is characterized by insignificant isotope variations (see Fig. 7). The possible reasons are: 1) the formation of isotope composition of snow at single condensation level from extensive cloud; 2) the initial isotope signal of the snow may be modified by processes of drifting or wind erosion. In the ablation season of 1998, the residual surface melted snow had the highest δ18O values from –6.82 to –8.79‰ and δ2H from –41.9 to –57.0‰ (see Fig. 4, b). That probably was a result of spring-summer snow accumulation modified by sublimation and partial melting. The lower value at 3780 m a.s.l. indicates partial melting of surface snow and exposure of winter snow.

The absence of altitudinal isotope effect can be explained by the fact that snow-bearing air masses undergo no small-scale orographic uplift and secondly that the source and the trajectory of air masses are essential to the average isotopic content (Moran et al. 2007).

Snow pits

The mean values of δ18O for two snow pits in January 2017 at 2332 m and 3345 m a.s.l. were –26.4 ‰, –24.8 ‰ and the values for δ2H were –189.7 ‰, –174.8 ‰, respectively. While the mean δ18O and δ2H values in snow pit at a lower altitude are more negative than the values at a higher altitude (Fig. 8). In the pit at 2332 m, the upper horizon is formed by snow with low values of δ18O (–29.4 ‰) and δ2H (–215.4 ‰), this is clearly freshly fallen January snow, in other snow horizons the δ18O and δ2H values are close to a uniform.

Similar values of δ18O and δ2H were obtained in the middle snow horizons at 3345 m a.s.l., while the lower horizon here is characterized by relatively high values (see Fig. 8), indicating the accumulation and preservation of snow, which was fallen most probably in autumn.

Interpretation

In fresh snow in February 2016, the most negative values of δ18O from –31.3 to –34.4‰ in the range of 3064–3457 m a.s.l. (Table 2) are extreme for Elbrus, especially for elevation below 4000 m a.s.l. Snow pit and firn core isotope records obtained at 5115 m a.s.l. (Kutuzov et al. 2013) show a clear season variation from –27 ‰ to –5.5 ‰ for δ18O values. Extremely negative isotope values in fresh snow of 2016 may be explained by drying of the air mass. The evolution of the isotope composition of water vapor during condensation and rainout from an idealized air mass is commonly modeled as a Rayleigh distillation process. The late stages of rainout are associated with rapid decreases of δ18O values in precipitation and in vapor. The rate of decrease of δ18O values also increases exponentially as the air mass dries out, and is greater at lower temperatures (Moran et al. 2007).

Table 3. Measured air temperature on the southern slope of Elbrus in 2017

Date and type of the weather Height a.s.l.	30/01/2017 Cloudy with clearings, overhead fog	31/01/20017 Clear, little cloudy	4/02/2017 Clear, in the morning cloudy	5/02/2017 Cloudy with clearings, the bottom of the sun, above the fog
2355	–15	–18	–4	–2
2934	–22	–21	–7	–8
3465	–27	–28	–13	–12
3853	–23	–26	–15	–14
calculated lapse rate ˚C/100 m	–0.53	–0.53	–0.734	–0.80

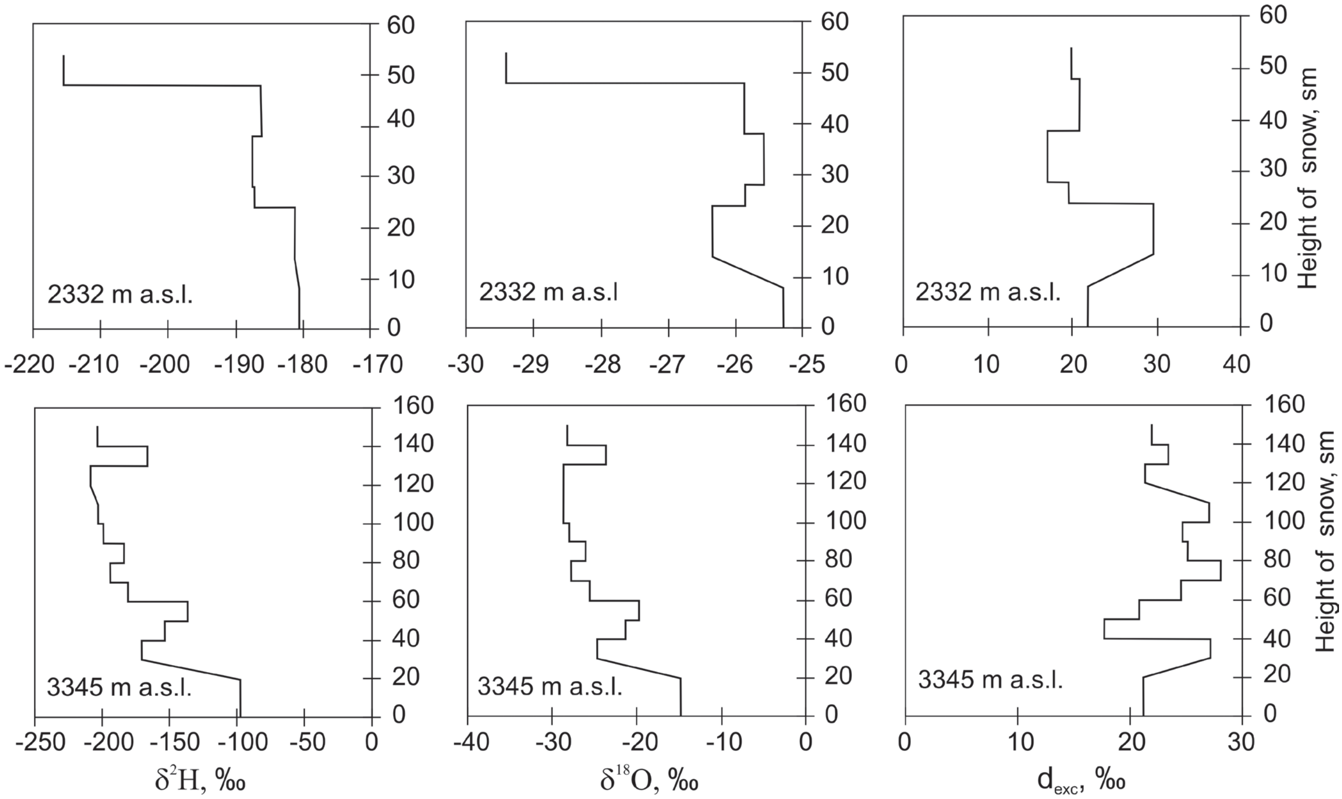


Fig. 8. The values of δ18O, δ2H and d_{exc} in snow pits at 2332 and 3345 m a.s.l.

In any case, we can not ignore the empirical data obtained for snow on 3 February, even if the isotope values distribution was not described by any model.

The formation of an altitude isotopic effect is not always associated with local conditions like a windward / leeward slope, temperature, etc. The absence of altitude effect is often explained by the fact that air masses do not follow the orographic uplift and, secondly, the source and trajectories of air masses, that could change pretty fast, are both important for the formation of isotope composition.

Moore et al. (2016) investigated the importance of non-local processes through the analysis of the synoptic scale circulation during a snowfall event at the summit of Mount Wrangell in south-central Alaska. During this event, there was over a 1-day period in which the local temperature was approximately constant, a change in $\delta^{18}O$ values that exceeded half that normally seen to occur between summer and winter in the region. It may be suggested that a change in the source region for the snow that fell on Mount Wrangell during the event from the subtropical eastern Pacific to northeastern Asia.

In order to explain isotope signal in the sow collected on the 3rd of February, we suppose a simultaneous coming of one air mass to the mountain slope, but by two ways. Backward air masses trajectories to Mt. Elbrus are provided by NOAA using the HYSPLIT model (Draxler and Rolph 2011), calculated for the 3rd of February at 3000 and 5000 m showed one source and one path of moisture from the north. The most negative $\delta^{18}O$ and $\delta^{2}H$ values corresponded to 3064–3457 m a.s.l. range. One of the reasons for this distribution of $\delta^{18}O$ and $\delta^{2}H$ values in snow is the removal of a part of the snow from the summit zone to a height of 3000 m. In this case, snow with low $\delta^{18}O$ and $\delta^{2}H$ values, deposited on summit is blown downward, forming a reverse altitude isotope effect. Another reason is that the altitude of 3000 m corresponds to the zone of maximum accumulation. Progressive precipitation leads to strong isotopic depletion of the remaining vapor and the last precipitation. It is obvious that the inverse high-altitude isotope effect is associated with these very negative values at an altitude of about 3000 m.

In fresh snow sampled in 2017 with a weakly decreasing of $\delta^{18}O$ and $\delta^{2}H$ values with altitude, there is also a very slight increase of d_{exc}. The main feature of the isotope signature of snow in 2017 is high d_{exc} values reaching 33‰ (see Table 1, Table 2). Backward HYSPLIT trajectories to Mt. Elbrus for the 26th of January 2017 at 3000 and 5000 m showed the source of moisture was the Mediterranean area. It is known that this region during the winter months due to low relative humidity over the sea is a source of precipitation with high deuterium excess (Gat and Carmi 1970).

In fresh snow sampled in 2001, there was a "classical" altitudinal isotope effect in precipitation due to orographic uplift of air masses and the related decrease in the condensation temperature (see Fig. 6).

When precipitation falls from air masses as they traverse topographic barriers, continued Rayleigh distillation on the lee slope should indeed produce an inverse relationship with altitude-lighter isotopic ratios with decreasing altitude. This would suggest a systematic altitudinal relationship that is the opposite of that which is observed on windward slopes, but an inverse relationship of this type is not well established or widely reported.

Poage and Chamberlain (2001) provide a compilation of observed $\delta^{18}O$-elevation gradients from 68 different studies worldwide, with only two of these studies reporting $\delta^{18}O$ depletion of precipitation with altitude. Ambiguous or inverse $\delta^{18}O$-elevation relationships have been reported from eastern (lee) slopes in Sierra Nevada (Friedman and Smith 1970) and the Canadian Rockies (Grasby and Lepitski 2002). Complex altitudinal relationships are also evident in high alpine snow samples (Niewodnizański et al. 1981). This study indicates the necessity to further study the isotope variability of the snow cover to predict the isotopic composition of snowmelt water and to better understand the accumulation processes and the sources of snow in high mountains.

CONCLUSIONS

The results suggest that $\delta^{18}O$-elevation gradients in fresh snow on south slope of Elbrus Mountain have similar values but opposite trends in different years and seasons. Above 3000 m in 2001, the $\delta^{18}O$ values decreased with altitude by –1.3 ‰/100 m (–11,1‰ $\delta^{2}H$ /100 m), in 2016, the $\delta^{18}O$ values increased with altitude by +1.04 ‰ /100 m (+8.76 ‰ $\delta^{2}H$ /100 m).

In 2017, the relationship between the values of $\delta^{18}O$ and $\delta^{2}H$ with the altitude of the terrain is not clearly pronounced with a weakly decreasing $\delta^{18}O$ values in an altitude range of 2256–3716 m a.s.l., there is also a very slight increase of d_{exc}. Such an uneven distribution of the isotope composition of snow with altitude in different seasons most likely is explained by various mechanisms of snow deposition – orographic uplift of the air mass along the slope or over-climbing through the main Caucasian ridge and considerable drift of dry snow on the slope.

Below 3000 m, the disruption of $\delta^{18}O$-elevation gradients has been attributed to post-depositional altering, wind movement, turbulent mixing of air masses or simultaneous coming of an air mass to the slope.

ACKNOWLEDGEMENTS

Researches are financially supported by Russian Foundation for Basic Research (grants RFBR № 16-05-00977 field research, 18-05-60272, isotope analyses and 18-05-60021, data compilation).

REFERENCES

Ambach W., Eisner H. and Pessl K. (1972). Isotopic oxygen composition of firn, old snow and precipitation in alpine regions. Z. Gletscherk. Glazialgeol., 8, pp. 125–135.

Ciais P. and Jouzel J. (1994). Deuterium and oxygen-18 in precipitation: Isotopic model, including mixed cloud processes. J. Geoph. Res., 99(D8), pp. 16793-16803.

Dansgaard W. (1964). Stable isotopes in precipitation. Tellus, 16, pp. 436–468.

Dietermann N. and Weiler M. (2013). Spatial distribution of stable water isotopes in alpine snow cover. Hydrol. Earth Syst. Sci., 17, pp. 26057–2668.

Draxler R.R. and Rolph G.D. (2011). HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) Model access via NOAA ARL READY Website (http://ready.arl.noaa.gov/HYSPLIT.php). NOAA Air Resources Laboratory, Silver Spring, MD.

Gat J.R. and Carmi I. (1970). Evolution of the isotopic composition of atmospheric waters in the Mediterranean Sea Area. Journal of Geophysical Research, 75, pp. 3039–3048.

Grasby S.E. and Lepitzki D.A.W. (2002). Physical and chemical properties of the Sulphur Mountain thermal springs, Banff National Park, and implications for endangered snails. Canadian Journal of Earth Sciences, 39, pp. 1349–1361.

He Y., Pang H., Theakstone W.H., Zhang D., Lu A., Song B., Yuan L. and Ning B. (2006). Spatial and temporal variation of oxygen isotopes in snowpacks and glacial runoff in different types of glacial area in western China. Annals of Glaciology, 43, pp. 269–274.

Holdworth G., Fogarasi S. and Krouse H.R. (1991). Variation of the stable isotopes of water with altitude in the Saint Elias Mountains of Canada. J. Geophysical Res., 96(4), pp. 7483–7494.

Holobâcă J-H. (2016). Recent retreat of the Elbrus glacier system. Journal of Glaciology, 62(231), pp. 94–102.

Friedman I. and Smith G. (1970). Deuterium content of snow cores from Sierra Nevada area. Science, 169, pp. 467–470.

Jouzel J. and Merlivat L. (1984). Deuterium and oxygen-18 in precipitation: modeling of the isotopic effects during snow formation. Journal Geophysical Res., 89(7), pp. 11749–11757.

Judy C., Meiman J.R. and Friedman I. (1970). Deuterium variations in an annual snowpack. Water Resour. Res., 6, pp. 125–129.

Kang S., Kreutz K.J., Mayewski P.A., Qin D. and Yao T. (2002). Stable-isotopic composition of precipitation over the northern slope of the central Himalaya. Journal of Glaciology, 48(163), pp. 519–526.

Kozachek A., Mikhalenko V., Masson-Delmotte V., Ekaykin A., Ginot P., Kutuzov S., Legrand M., Lipenkov V. and Preunkert S. (2017). Large-scale drivers of Caucasus climate variability in meteorological records and Mt Elbrus ice cores. Climate of the Past, 13, pp. 473–489.

Königer P., Hubbart J.A., Link T. and Marshall J.D. (2008). Isotopic variation of snow cover and streamflow in response to changes in canopy structure in a snow-dominated mountain catchment. Hydrol. Process., 22, pp. 557–566.

Kutuzov S., Shahgedanova M., Mikhalenko V., Ginot P., Lavrentiev I. and Kemp S. (2013). High-resolution provenance of desert dust deposited on Mt. Elbrus, Caucasus in 2009–2012 using snow pit and firn core records. The Cryosphere, 7, pp. 1481–1498.

Merlivat L. and Jouzel J. (1979). Global climatic interpretation of the deuterium oxygen 18 relationship for precipitation. J. Geophys. Res., 84, pp. 5029–5033.

Mikhalenko V., Sokratov S., Kutuzov S., Ginot P., Legrand M., Preunkert S., Lavrentiev I., Kozachek A., Ekaykin A., Faïn X., Lim S., Schotterer U., Lipenkov V. and Toropov P. (2015). Investigation of a deep ice core from the Elbrus western plateau, the Caucasus, Russia. The Cryosphere, 9, pp. 2253–2270.

Moore G.W.K., Field R. and Benson C.S. (2016). Impact of Source Region on the δ^{18}O Signal in Snow: A case study from Mount Wrangell, Alaska. Journal of Hydrometeorology, 17, pp. 139–151.

Moran T.A., Marshall S.J., Evans E.C. and Sinclair K.E. (2007). Altitudinal gradients of stable isotopes in lee-slope precipitation in the Canadian Rocky Mountains. Arct. Antarct. Alpine Res., 39(3), pp. 455–467.

Moser H. and Stichler W. (1970). Deuterium measurements on snow samples from the Alps. In Isotope Hydrology (Proceedings of a Symposium on Use of Isotopes in Hydrology), IAEA, Vienna, 43–57.

Moser H. and Stichler W. (1974). Deuterium and oxygen-18 contents as an index of the properties of snow covers. In Snow Mechanics (Proceedings of the Grindelwald Symposium, Switzerland, April 1974), IAHS Publ., 114: 122–135.

Niewodnizański J., Grabczak J., Barański L. and Rzepka J. (1981). The altitude effect on the isotopic composition of snow in high mountains. Journal of Glaciology, 27(95), pp. 99–111.

Pang Z., Kong Y., Froehlich K., Huang T., Yuan L., Li Z. and Wang F. (2011). Processes affecting isotopes in precipitation of an arid region. Tellus, 63(3), pp. 352–359.

Poage M.A. and Chamberlain C.P. (2001). Empirical relationships between elevation and the stable isotope composition of precipitation and surface waters: considerations for studies of paleoelevation change. American Journal of Science, 31, pp. 1–15.

Raben P. and Theakstone W.H. (1994). Isotopic and ionic changes in a snow cover at different altitudes: observations at Austre Okstindbreen in 1991. Ann. Glaciol., 19, pp. 85–91.

Risi C., Bony S., Vimeux F., Chong M. and Descroix L. (2010). Evolution of the stable water isotopic composition of the rain sampled along Sahelian squall lines. Q. J. R. Meteorol. Soc., 136(s1), pp. 227–242.

Rototaeva O.V., Nosenko G.A., Tarasova L.N. and Khmelevskoy I.F. (2006). General characteristics of glacierization of the north slope of the Gteater Caucasus). In: V.M. Kotlyakov, ed. Glaciation in North and Central Eurasia at Present Time. Moscow: Nauka Publ. House, pp. 141–144. (in Russian).

Rozanski K. and Araguas-Araguas L. (1995). Spatial and temporal variability of stable isotope composition over the South American continent. Bull. Inst. Fr. Etud. Andin., 24, pp. 379–390.

Schotterer U., Fröhlich K., Stichler W. and Trimbom P. (1993). Temporal variations of oxygen-18 and deuterium excess in alpine regions of Switzerland. In: Isotope Techniques in the Study of Past and Current Environmental Changes in the Hydrosphere and the Atmosphere (Proc. Symp. Int. Atomic Energy Agency, Vienna), pp. 53–64.

Schotterer U., Gäggeler H.W., Fröhlich K., Sandjordj S. and Stichler W. (1997). Isotope records from Mongolian and alpine ice cores as climate indicator. Climatic Change, 36, pp. 519–530.

Shahgedanova M., Kutuzov S., White K.H., and Nosenko G. (2013). Using the significant dust deposition event on the glaciers of Mt. Elbrus, Caucasus Mountains, Russia on 5 May 2009 to develop a method for dating and "provenancing" of desert dust events recorded in snow pack. Atmos. Chem. Phys., 13, pp. 1797–1808.

Siegenthaler U. and Oeschger H. (1980). Correlation of O-18 in precipitation with temperature and altitude. Nature, 285, pp. 314–317.

Sokratov S.A. and Golubev V.N. (2009). Snow isotopic content change by sublimation. Journal of Glaciology, 55, pp. 823–828.

Stichler W., Schotterer U., Fröhlich K., Ginot P., Kull C., Gäggeler H. and Pouyaud B. (2001). Influence of sublimation on stable isotope records recovered from high-altitude glaciers in the tropical Andes. J. Geophys. Res., 19, pp. 22613–22620.

Tielidze L.G. and Wheate R.D. (2017). The Greater Caucasus Glacier Inventory (Russia/Georgia/Azerbaijan). The Cryosphere Discuss., doi:10.5194/tc-2017-48. 27 p.

Vasil'chuk Y.K. and Chizhova J.N. (2010). Altitude gradient of $\delta^{18}O$ and δD in precipitation and snow cover in high mountain. Earth Cryosphere, 14(1), pp. 13–21 (in Russian).

Vasil'chuk Y.K., Chizhova J.N., Budantseva N.A. and Mukhina J.S. (2010). Rapid reduction of the glacier Big Azau, Elbrus on stable climatic conditions and risks appearing in this case. Georisk, 2, pp. 16–29 (in Russian).

Vasil'chuk Y.K., Chizhova J.N. and Budantseva N.A. (2006). Isotope composition of Bolchoy Azau glacier tongue, Elbrus. Earth Cryosphere, 1(10), pp. 56–68 (in Russian).

Vasil'chuk Yu.K., Chizhova Ju.N. and Papesch W. (2005). Trend of isotope composition of a separate winter snowfall in northeastern Europe. Earth Cryosphere, 9(3), pp. 81-87. (in Russian).

Volodicheva N.A. (2002). The Caucasus. In: M. Shahgedanova, ed. The Physical Geography of Northern Eurasia, Oxford: Oxford University Press, pp. 350–376.

Yamanaka T. and Yamada Y. (2017). Regional Assessment of Recharge Elevation of Tap Water Sources Using the Isoscape Approach. Mountain Research and Development, 37(2), pp. 198–205.

Zolotarev Y.A. and Kharkovets Y.G. (2010). Evolution of the Elbrus glaciation since the mid XIX century under changing climate. Key findings of the glacio-cartographical monitoring. Geogr. Environ. Sustainability, 2, pp. 13–32.

Received on May 31th, 2018 Accepted on January 21th, 2020

Alexey Yu. Sidorchuk[1*] and Tatiana A. Matveeva[1]
[1] Lomonosov Moscow State University, Moscow, Russia
***Corresponding author:** fluvial05@gmail.com

PERIGLACIAL GULLY EROSION ON THE EAST EUROPEAN PLAIN AND ITS RECENT ANALOG AT THE YAMAL PENINSULA

ABSTRACT. The net of dry valleys, gullies and shallow hollows is typical for the East European Plain. Dense vegetation usually covers their bottoms and slopes, so the modern erosion there is negligible in the pristine conditions. This erosion landscape formed in periglacial conditions during the terminations of the last two glaciations. The same kind of the erosion landscape is typical for the Arctic regions, especially for the Yamal, Gydan, and Tazovsky peninsulas. The size and the density of such valleys and gullies are quite similar to those existing on the East European Plain, but these erosion features are active there, especially in the conditions of natural or anthropogenic deterioration of the vegetation cover. As the density of dry valley network is an indicator of hydrological conditions in the river basin, the landscapes of the Arctic regions can be used as the modern analogs of the territories with the past periglacial erosion.

The recent hydrological characteristics of the west-central Yamal Peninsula were used to estimate the parameters of erosion network at the Khoper River basin, formed in periglacial conditions. For these purposes gully erosion and thermoerosion model GULTEM was verified and calibrated based on the observation of the modern processes on the Yamal Peninsula. The meteorological characteristics were taken from ERA-Interim Reanalysis grid. To calculate the flow characteristics a synthetic hydrological model was used. These verified and calibrated models were used to find the most suitable characteristics of climate and vegetation cover, which can explain the structure and density of the Perepolye dry valley in the Khoper River basin. This dry valley with the main trunk length of 6400 m was formed at the end of the Late Valdai Glaciation (MIS 2). The conditions required for the formation of a periglacial gully of such length were estimated with the GULTEM model. The critical velocity of erosion initiation was within the range 0.8-0.9 m/s, and the surface runoff depth was close to the recent one on the Yamal Peninsula (330 mm). The system of shallow hollows in the Perepolye catchment (the gullies formed at the end of the Moscow Glaciation, MIS 6) is denser and longer than the dry valley system, and the modelling estimates showed that the surface runoff during that period was almost 3.3 times more than the recent one on the Yamal Peninsula.

KEY WORDS: The Yamal Peninsula, active gullies, East European Plain, periglacial erosion, gully erosion and thermoerosion modelling, estimates of past surface runoff

CITATION: Alexey Yu. Sidorchuk and Tatiana A. Matveeva (2019) Periglacial Gully Erosion On The East European Plain And Its Recent Analog At The Yamal Peninsula. Geography, Environment, Sustainability, Vol.13, No 1, p. 183-194
DOI-10.24057/2071-9388-2019-01

INTRODUCTION

One of the most interesting relief features of the East European Plain is the net of dry valleys (Fig. 1A). These dry valleys typically have lengths in the range of 1-10 km. At present, these systems drain water mostly during snow thaw period or after heavy rains. Dense vegetation usually covers their bottoms and slopes, so that in the pristine conditions the modern erosion there is negligible.

The main part of these dry valleys (the Russian term is «balka») was formed during erosional events of the Late Quaternary (Eremenko et al. 2011; Matoshko 2012) in the periglacial conditions with the permafrost. This type of ancient erosion features, some buried, some renewed by the modern erosion, is widespread in Europe (Larsen et al. 2016; Hošek et al. 2019), Northern America (Bettis et al. 1985), and Australia (Prosser et al. 1994). The reconstruction of the environmental conditions of Quaternary gullies initiation and evolution are critical for understanding their morphology and history.

The erosion landscapes in the Arctic regions are quite similar to those on the East European Plain in their morphology, but these erosion features are currently active in the Arctic (Fortier et al. 2007), especially in the conditions of natural or man-induced deterioration of vegetation cover (Sidorchuk 2015). The size and the density of such valleys and gullies, especially at Yamal, Gydan and Tazovsky peninsulas in Russia (Fig. 1B) are nearly the same as those of the net of «balkas» in the lowland regions of the East European Plain, such as the upper Dnepr and Don River basins. As the erosion network density depends on the hydrological conditions on the watershed (Carlston 1963; Gregory 1966; Gregory et al. 1968), the erosion processes at the gully catchments in the Arctic regions can be used as the modern analogs of the former periglacial erosion.

The main goal of this paper is to determine the most suitable characteristics of surface runoff and vegetation, which can explain the density of the ancient dry valleys systems in the Don River basin on the East European Plain. For this purpose, the modern erosion features on the Yamal Peninsula were used as the closest morphological analog. Calculations of erosion net density were performed with the gully erosion and thermoerosion model (GULTEM). The model was verified and calibrated using observations on the modern hydrological and erosion processes in the Arctic. The climatic and meteorological characteristics for such verification are available through reanalysis.

Fig. 1. The net of dry valleys and gullies on the East European Plain, Kursk district (A) and in the west-central Yamal Peninsula (B). Space images from Google Earth

MATERIALS AND METHODS

The dry valley systems in the Don River basin

The Don River basin on the East European Plain was not glaciated at least for the last 250 ka. Therefore, mostly erosional processes determined the landscape there, and dry valley systems are of the main importance for the long-term palaeohydrological reconstructions. There were at least two major hydrological and erosional events, which governed the morphology of erosional landscape: the catastrophic surface runoff at the end of the Moscow Glaciation (MIS 6), about 130 ka ago, and the one at the end of the Late Valdai (MIS 2), 16-19 ka ago. The dry valleys formed during these events were described by Eremenko and Panin (2011). We used the dry valley Perepolye as a typical example for palaeohydrological reconstructions (Sidorchuk et al. 2018).

The Perepolye dry valley with a catchment area of 41.7 km^2 is situated on the left bank of the Khoper River, the left tributary of the Don River, near Povorino settlement (Fig. 2). The modern dry valley with the main trunk length of 6400 m is characterized by steep slopes (10-15°) and a concave longitudinal profile; the valley depth is 6-8 m in the middle part and 8-10 m in the lower part. Together with the tributaries, the dry valley forms a tree-like network with a total length of 11.3 km, and the density of the modern network is 0.27 km / km^2.

In the upper part of the catchment, the modern dry valley and its tributaries are continued by shallow hollows, clearly visible on the plowed fields by their brighter tone on the space images. The heads of such hollows almost reach the watershed (see Fig. 2). The length of the entire dry valley-hollow network is 40.5 km and its density of 0.97 km / km^2; its longest element (from the mouth to the top of the main dry valley) reaches 8400 m.

To study the sediments within this dry valley, the coring along nine transverse profiles was performed. The sedimentary structures uncovered by drilling indicate the presence of two major buried erosion features (Fig. 3). The first (most ancient) erosion feature is embedded into the sands of the third terrace of the Khoper River. Its bed is incised 5-6 m be-

Fig. 2. The Perepolye dry valley with the system of hollows in the Khoper River basin near Povorino

low the bottom of the modern dry valley in its lower and middle parts and 2-3 m below it in the upper part. This buried gully has an average longitudinal slope of about 2.5 m/km. On its bed buried soil layers with well-defined humus and illuvial horizons, with numerous filled animal burrows («krotovinas») exist. This buried soil (Fig. 4) probably formed during the Mikulino Interglacial (=Eemian, MIS 5e), since its structure corresponds to the Mezin soil complex (Velichko et al. 2007). Therefore, the incision event most probably took place at the end the Moscow Glaciation (MIS 6).

This first erosion feature was filled later almost to the top by the loess-like yellow-brown silt deposits with the maximum thickness over 13 m. These deposits are dated by OSL method to about 65 ka ago (Panin et al. 2011). The hollows in the upper part of the catchment were also formed during the late Moscow (MIS 6) erosion event and subsequently filled with similar loess-like yellow-brown sediments.

The second, younger erosion feature has a steeper longitudinal profile and is narrower than the first one. The average slope of its bottom is 3.7 m/km; the width is 20-30 m at the head of the dry valley and up to 100 m near its mouth. In the lower part of the dry valley, the loess-like deposits filling the older incision were completely eroded along the thalweg; they preserved only at the sides of the younger incision. Here the depth of this incision is about 20 m.

The second incision is filled with gray and black loams and clays with sandy interlayers. The top of these deposits forms the main part of the bottom of the modern dry valley. The AMS radiocarbon dating showed that the age of black heavy loams and clays which overlie the sands of the third terrace is 11441 +/- 60 14C yr BP (AA104015). This agrees with the date 11900 +/- 120 (Ki-5305), obtained for the deposits filling a similar incision in the nearby dry valley Khoprets (Sidorchuk and Borisova 2000).

The second erosional feature can be attributed to the powerful surface runoff event at the beginning of the degradation of the last glaciation, 16-19 ky ago, which left numerous traces in the relief of the periglacial zone of the Northern Hemisphere (Panin et al. 2006). The modern dry valley network was formed during this erosion event. The network

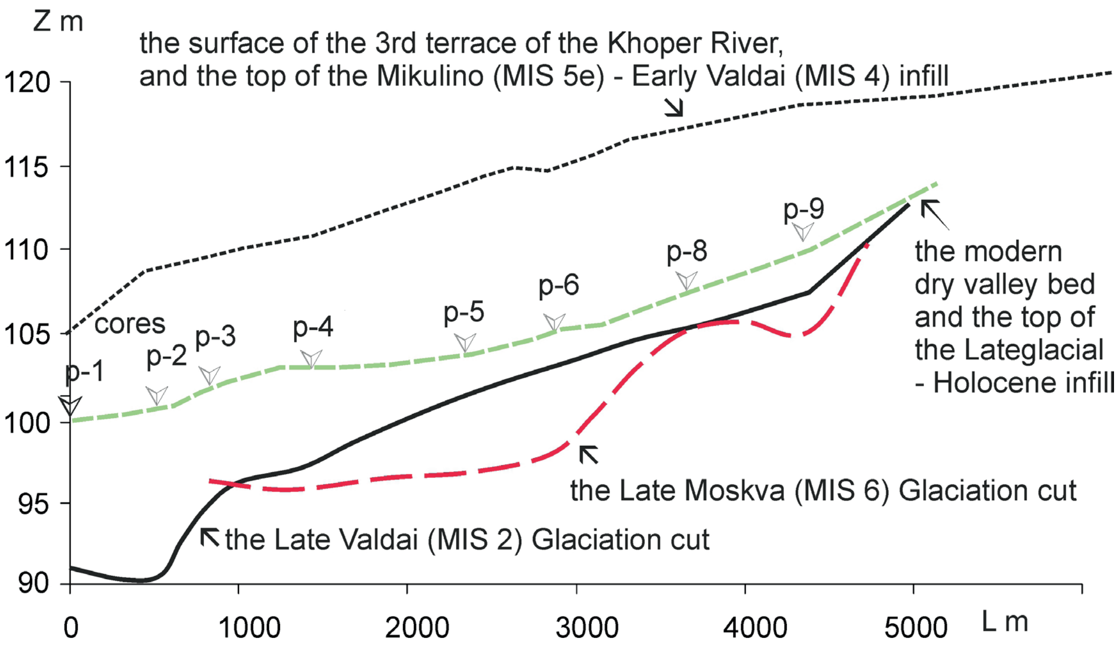


Fig. 3. Longitudinal profiles of the Perepolye dry valley and two major erosion features, buried by the deposits in the dry valley

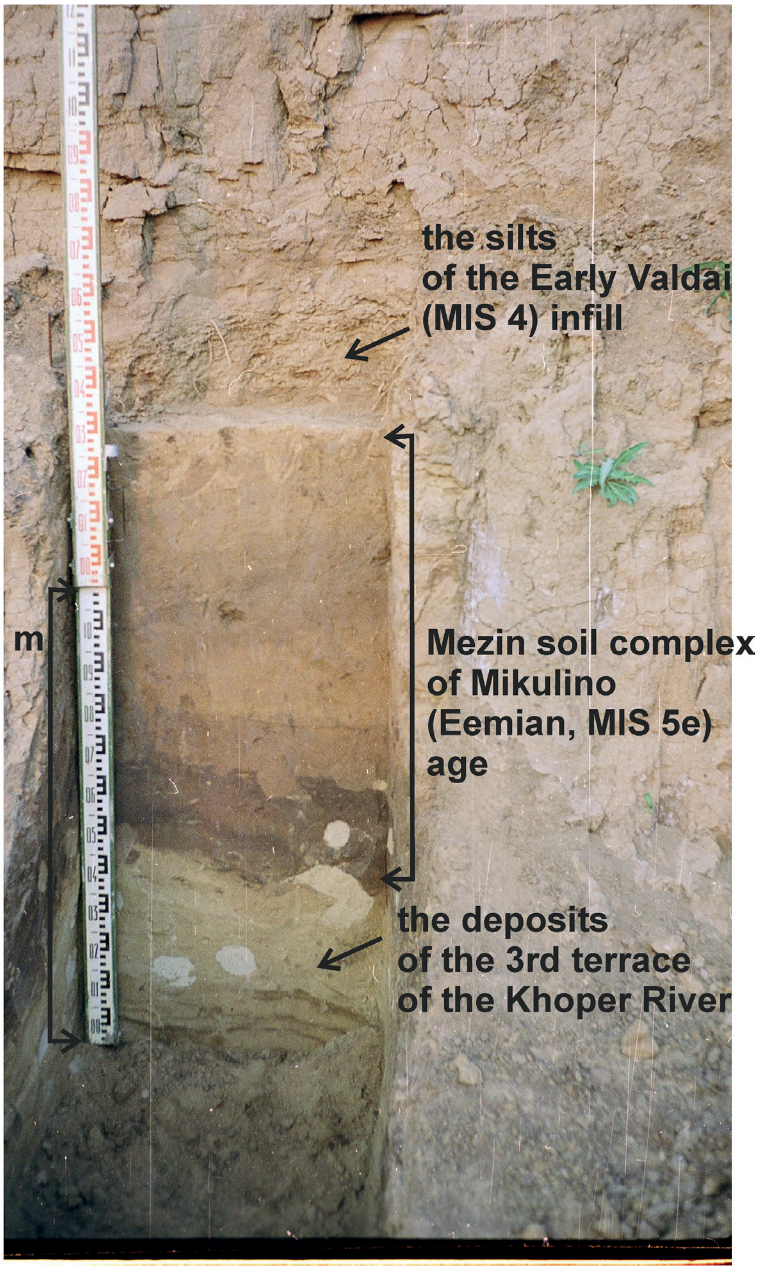


Fig. 4. Mezin-type buried soil complex on the erosional bed of the Perepolye dry valley

of hollows, which was formed during the previous erosion event, was not affected by erosion at this stage.

The modern dry valleys and hollows functioned as active deep gullies during the periods of catastrophic surface runoff. They were mainly in general inactive in the climatic and landscape conditions of the Holocene. The morphology of these systems including the maximum lengths of their main trunks can be regarded as the 'memory' of the fluvial system (Knighton 1984). This memory can be used to find the most probable characteristics of the hydrological regimen for the period of the dry valley systems formation. To resolve such inverse problem of geomorphology, an optimisation of the input characteristics of gully erosion model, which leads to the given output morphology of erosion landscape, is required.

The vegetation on the East European Plain during the periods of catastrophic surface runoff represented a so-called 'tundra-steppe' or 'periglacial forest-steppe'. The closest modern analogs of such past vegetation (and climate) on the East European Plain are found in the Altai and Sayan mountains (Borisova et al. 2006; Sidorchuk et al. 2011). As the erosion processes in the mountain regions cannot be applied as the analogs for erosion on lowlands, different indicators have to be used to find such analogs. At present, the lowland territories with the periglacial climate are situated in the Arctic and Subarctic regions of European Russia and Siberia. Therefore, the first approximation of input hydrological regimen and vegetation cover was taken from the west-central Yamal Peninsula. The main reason to use this analog is the similarity of erosion relief of the watersheds in the Don River basin and on the Yamal Peninsula, especially the similarity in drainage density, which is the main indicator of hydrological conditions (Gregory 1966; Sidorchuk et al. 2018).

The gully erosion in the west-central Yamal

Natural processes of the gully erosion are typical for the arctic tundra of the Yamal Peninsula (Sidorchuk 1996, 2015). The main natural factors of erosion in this region are: 1) sufficient annual precipitation (up to 350-400 mm); 2) low soil permeability due to the permafrost and therefore high runoff coefficients (up to 0.9-1.0); 3) high erodibility of frozen and thawing soils; 4) natural triggers of gully erosion, such as thermokarst processes and shallow landslides, which often cause vegetation cover deterioration.

This natural high erosion risk has been greatly increased recently by human impact. In the areas of gas production and transportation facilities the erosion potential increases due to: 1) deterioration of the vegetation cover due to industrial development; 2) increased snow depth on the slopes due to excessive snow accumulation near buildings and roads; 3) an increase of the runoff coefficient on impermeable surfaces of the urbanized territories and roads; 4) local industrial and urban sources of warm water; 5) exploitation of sand-pits, gas and oil fields, and construction of pipe lines and ditches. The combination of high natural erosion potential and human interference causes extremely intensive gully, rill and sheet erosion (Fig. 5). Erosion follows the old drainage lines, mostly formed by the former shallow slumps, and new gullies cut into previously gently sloping elongate depressions. On the marine terrace in the west-central Yamal Peninsula, composed from frozen loams and sands, it takes anthropogenic gullies 4-10 years to reach their potential (possible) length.

The investigations of 1990–97 within the Scientific Program «Yamal» of RSC «GAZPROM» gave the opportunity to collect the information about hydraulic characteristics of the flows and morphological changes in the gullies, both natural and anthropogenic. The main feature is the alternating processes of quick intensive incision and rapid sidewall slumping in the gullies. Repeated instrumental measurements of the gully longitudinal profiles showed relatively low resulting mean rates of deepening. For the period of 1990-1995, the increase of the mean depth of gully K_1 was 0.9 m, and that of gully K_2 was 0.6 m. Note that during one rainfall on 8-9 Aug. 1990 the bed of the gully K_2 was incised by 0.58 m (Sidorchuk 2015). The discharge from the catchment with area 64300 m2 reached 21.3 l/s; the maximum sediment concentration was 74 kg/m3. This incision was already partly filled by sidewall slumping only 2-3 days after the flood. The mean rate of gully deepening is about 10 times less than the peak rate of incision during snowmelt and rainfall runoff, due to backfilling by material from slumping walls.

The gully increase in length is also a complicated process. The main head of gully K_1 remains stable since 1970, but in the 1990s a long and deep rill was formed in its middle part, and the second active head developed in the convex upper part of the slope. The length of gully K_2 was 165 m in 1988, 190 m in 1989, 210 m in 1990, 230 m in 1991, and 280 m in 1995. The growth rate of the main gully decreased in time, but in 1990–95, a new additional gully head was formed 120 m upstream from the old head, so that the gully became discontinuous. Nevertheless, both the photos taken in 2007 and space images of 2016 showed a general stabilisation of these gullies.

The modelling of gully erosion

The most typical for the gully erosion is to develop in two stages. At the first stage, gully channel formation is very intensive, and morphological characteristics of the gully (its length, depth, width, area, and volume) are changing rapidly. This stage takes about 5% of the whole gully lifetime, but

Fig. 5. Dynamics of the gullies K_1 and K_2 on the west-central Yamal Peninsula, shown on the Google Earth image of 2016

about 90% of its length is formed during this stage (Kosov et al. 1978). To describe this first stage of gully evolution, we developed the gully thermoerosion-erosion model GULTEM. Water flow is calculated with the appropriate hydrological model. Meteorological data are taken from the reanalysis, as the station net is sparse in the Arctic.

Meteorological data from the reanalysis

Reanalysis is a grid data of surface and upper-air meteorological parameters. It spans a period over 30 years and can be used for climate analysis. The reanalysis combines observational data assimilation system and a forecast model. The atmospheric forecast model uses various hydrodynamic equations solved using finite-difference methods of forecast. Therefore, grid data of reanalysis are formed step-by-step using the assimilation data system and forecast model.

ERA-Interim Reanalysis is one of the best global atmospheric reanalyses (Lindsay et al. 2014). It was produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The atmospheric model in ERA-Interim has a horizontal resolution approximately 79 km (a grid of 0.75°×0.75°). The vertical resolution is inconstant and includes 60 layers with the highest level at 0.1 hPa. ERA-Interim covers the period from 1979 to the present. Time resolution is 3 hours for surface parameters and 6 hours for upper-air parameters. Atmospheric model of ERA-Interim assimilates many observation data: various types of satellite data (Dee et al. 2009), surface observations from land stations, ships, drifting buoys, reports from radiosondes and pilot balloons launched from land stations and ships, drop sondes, aircraft reports, wind profiler, and METAR airport weather reports.

The ERA-Interim reanalysis quite successfully reproduces most of the atmospheric parameters, including air temperature and precipitation (Dee et al. 2011; Donat et al. 2014; Sun et al. 2018). Over the Yamal Peninsula, observational data of weather stations Marresale and Novy Port are assimilated in the ERA-Interim. In this study we used daily 2 meter temperature and daily precipitation to compare station observations with reanalysis data. As observational data, we used the Marresale weather station (69.72°N, 66.80°E) records from RIHMI-WDC archive (Bulygina et al. 2008). The data sampling of this station is representative because missing data amounts to not more than 10%. Selected reanalysis grid point for comparison has coordinates 69.75°N, 66.75°E.

There is a consistent overestimation of the daily 2 meter temperature in winter, spring and autumn and an underestimation in the summer by the ERA-Interim. The grid cell, which contains the Marresale weather station, includes both land and sea. Therefore, the temperature in winter, spring and autumn according to reanalysis is higher than according to station observations, and in summer, it is lower. The standard deviation of the difference between daily 2-meter temperature data series from ERA-Interim and from the station is 3.2°C. This is the effect of the sea area included in this reanalysis grid cell.

At the same time, ERA-Interim reanalysis underestimates the daily precipitation data compared to Marresale station. A decrease of precipitation compared with observational data is noted in most reanalyses and grid archives for the entire globe. Also, the reanalysis significantly underestimates the values of extreme precipitation, one of the reasons for that being the local character of the extreme values (Donat et al. 2014; Matveeva et al. 2015), while reanalysis data are averaged over the cell. The standard deviation of a series of daily precipitation is 1.8 mm.

Comparison of hydrological parameters (including runoff) from ERA-Interim reanalysis data with observational data showed significant inconsistency (Betts et al. 2009). To increase the accuracy of gully erosion calculation, we apply our own hydrological model.

Hydrological model

Any erosion modelling begins with water flow calculations. In the conditions of the Arctic climate with deep permafrost and high density of linear erosion features hydrological modelling requires specific characteristics described further. Two main sources of water flow are typical for this environment: snow thaw during the spring and rainfall, mostly during the summer. The runoff depth *Hrunoff* for the period of snow thaw was calculated with the models of Komarov (Komarov et al. 1969), Vinogradov (1988), Vinogradov et al. (2014), and Gelfan (1999), using the maximum depth of snow fall *Hsnow* (water equivalent), precipitation *P*, air temperature *T*, and evaporation *E* from reanalysis. Snow thaw rate was calculated with the air temperature and snow thaw coefficient k_{sm}. Runoff occurs when the content of water in snow exceeds the maximum potential water content, which was calculated with current snow density. Snow thaw period ends when *Hsnow*=0.

The initial snow pack depth on the catchment is irregular. In the natural conditions, snow driven by wind during the winter accumulates in depressions (river valleys and gullies), at the base of steep slopes and in dense vegetation. In the conditions of human impact, snow accumulation around the buildings and roads becomes significant. The measurements on the gullied territory of the west-central Yamal Peninsula (Bobrovitskaya et al. 1999) show the typical distribution of snow depth/mean snow depth ratio *(Nratio)* at the beginning of snow thaw period. The distribution is bimodal, the first mode describing the snow pack variability in the areas near the water divide and on gentle slopes, and the second one in the gullies. The relative weight of these two modes depends on the areas proportion of gentle slopes and gully cuts in the area. To take into account the uneven thickness of the snow cover, the calculations of the runoff depth *Hrunoff* for each time period were performed with different values of *Nratio*, weighted with their probabilities and then integrated to get the resulting value. The runoff depth for the period of summer rains was assumed to be equal to rainfall depth *P(t)* from reanalysis. Evaporation *E* for the periods with rainfall was assumed to be 0, and the losses were calculated as the amount of water required to fill depressions on the catchment and water evaporated from them during the periods between rainfalls using the formula of Popov (1956).

The sensitivity analysis of the hydrological model shows that the most important are absolute values of input climatic characteristics, type of snow pack depth distribution, and snow thaw coefficient k_{sm}. The climatic characteristics were taken from reanalysis. They are quite valid for air temperature, but differ at random from measured values by up to -50% to +100% for snow pack depth at the beginning of snow thaw period and for precipitation depth. The coefficient of snow thaw k_{sm} was calibrated using the measurements over the spring period of 1993, when both snow pack depth distribution and runoff were measured (Bobrovitskaya et al. 1999). For k_{sm}=0.75 10^{-4} the measured and calculated runoff depth are well correlated (Fig. 6A). Runoff sequence during the spring 1992, calculated with this calibrated value of k_{sm}, have the pattern similar to measurements (Fig. 6B). The measured snow depth (in water equivalent) at the beginning of snow thaw period in 1992 was nearly two-fold that from reanalysis, and the measured and calculated runoff depths are well correlated only when snow depths from reanalysis were corrected according to measurements. Without such correction, the calculated maximum runoff depth is considerably underestimated (65 instead of 80 mm), and the date of this calculated maximum was one day earlier than the real one.

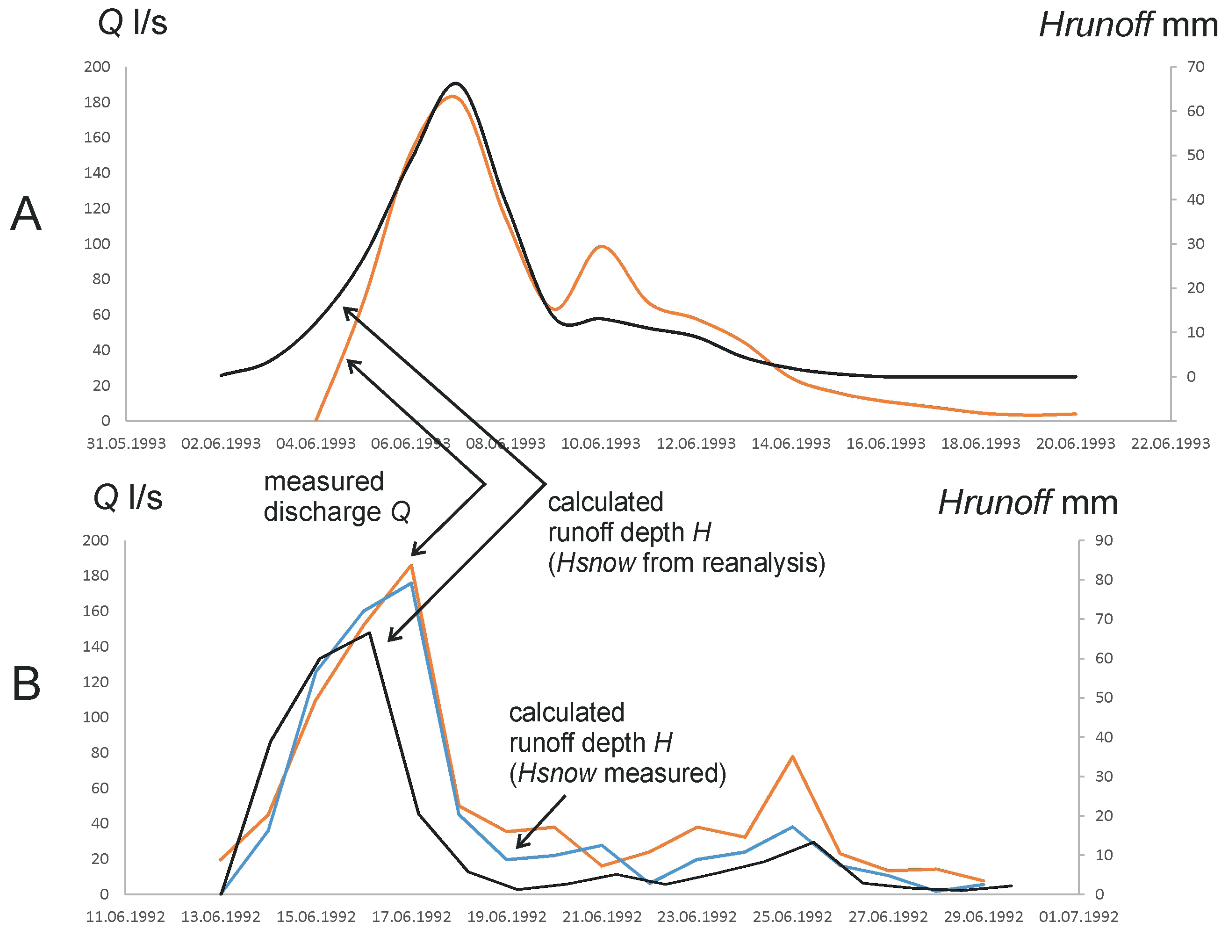


Fig. 6. Measured discharges and calculated runoff depths during the spring of 1993 (A) and 1992 (B)

Erosion model

Gully thermoerosion and erosion model GULTEM was previously described in detail in (Sidorchuk 2015). Therefore, only the major formulae are listed below. According to the model, two main processes occur along each flowline, as in nature: 1) formation of erosion cut in the topsoil or at the gully bottom by flowing water during snowmelt or a rainstorm event; 2) transformation of gully walls by shallow landslides during the period between adjacent water flow events.

In the first process, the rate of erosion is linearly correlated with the product of bed shear stress $\tau=g\rho DS$ and the mean flow velocity U:

$$E \approx U\tau = k_E qS \qquad (1)$$

Here S is gully bottom mean slope, q is specific discharge, i.e. the ratio of discharge and flow width Q/W, g is acceleration due to gravity, ρ is water density and k_E is erosion coefficient, which depends on soil properties. If flow velocity is less than critical value for erosion initiation U_{cr}, then erosion is negligible.

In the case of gully erosion in the soil with permafrost, another type of bed lowering develops – the combined action of thawing of frozen soil by running water with positive temperature and mechanical removal of thawed layer – so-called thermoerosion. Thermoerosion rate is equal to the rate of soil thawing and linearly correlated with water temperature T^0C:

$$E_T = k_{TE} T \qquad (2)$$

The linear relationship of thermoerosion rate with water temperature, as well as k_{TE} values for different types of frozen soil were obtained in the course of laboratory and field experiments (see details in Sidorchuk 2015). The same type of relationship follows from solution of Stefan's problem for the case of thermal erosion (Yershov et al. 1982).

The erosion rate of gully banks can be estimated only in a first approximation as a function of the rate of gully bed erosion:

$$E_b = \frac{V}{U} E = k_b E \qquad (3)$$

Here V is the lateral flow velocity. For k_b estimates see Eq. 42 and 43 in Sidorchuk 2015. Together, erosion and thermoerosion produce gully volume V_0.

The second process, rapid sliding and slumping following the incision, leads to formation of stable straight slopes of the gully with an inclination equal to the angle of repose φ for a given soil type. This process transforms the erosion cut of volume V_0 into trapezium with bottom width W_b, depth

$$D_t = \left(\sqrt{W_b^2 + \frac{4V_0}{\tan(\varphi)}} - W_b \right) \frac{\tan(\varphi)}{2} \qquad (4)$$

and top width

$$W_t = W_b + 2.0 D_t / [\tan(\varphi)] \qquad (5)$$

The altitudes Z of gully longitudinal profile are transformed in time t along flowline length x according to the following partial differential equation for the process of erosion

$$\frac{\partial Z}{\partial t} = k_E q \frac{\partial Z}{\partial x} \qquad (6)$$

and according to ordinary differential equation for the case of thermoerosion

$$\frac{\partial Z}{\partial t} + k_{TE} T = 0 \qquad (7)$$

If the rate of thermoerosion is less than erosion rate, the thaw layer is removed, and Eq. 7 is used for calculating the gully bed lowering. If the rate of thermoerosion (i.e., the rate of the thermal front shifting in soil) is greater than erosion rate, a thaw layer is formed, and calculation with Eq. 6 is used in the model.

GULTEM calibration procedure

The input morphological and hydrological characteristics in the model are: longtitudinal profile *Z(x)* of initial slope, discharge *Q(x,t)*, flow width *W(x,t)*, and water temperature *T(x,t)*. As the gully basin is relatively small, discharge can be calculated as product of runoff depth and contributing catchment area *Hrunoff(t)F(x)*. The flow width, as well as other flow hydraulic characteristics, are calculated using the empirical relationships with the discharge (see Eq. 55 and 56 in Sidorchuk 2015).

The coefficients and parameters used in the model are k_E, k_{TE}, k_b, U_{cr}, and φ. Some of these values are used for the model calibrations, and others must be known. Usually k_E and/or U_{cr} are used for the calibration. Field and laboratory experiments were used for k_{TE} estimations (see Table 6 in Sidorchuk 2015) and k_{TE}=0.000025 m/(°C s) was taken for the loam deposits of the west-central Yamal. The measurements of bank erosion in the gullies show that the coefficient of bank erosion depends on the ratio between gully flow width *W* and gully bottom width W_b, as well as on the flow depth in the gully *D* (see Eq. 42 and 43 in Sidorchuk 2015). The angle of repose φ is estimated from the stable gully sections or calculated with the straight slope stability model (see Eq. 46 and 47 in Sidorchuk 2015). The angle of repose φ for the gullies investigated in 1990-97 was about 30° (the mean value for 11 stable cross-sections).

Calibration procedure was two-staged. At the first stage, the critical velocity of erosion initiation was calibrated according to the information about gully length growth, as this morphological characteristic mostly depends on critical velocity in given hydrological conditions. Two main scenarios were used: 1) initial catchment topography, which existed prior to the gully incision, with undisturbed natural vegetation cover; 2) catchment topography in the year 1990 with vegetation cover completely disturbed due to human impact. Critical velocity U_{cr} was 0.8 and 0.12 m/s, respectively. At the second stage, the optimal erodibility coefficient k_E was found according to the information about gully mean depth change during the calibration period 1990-95 (Fig. 7); optimal erodibility coefficient k_E value was 0.0025 1/s. Finally, with the values of all coefficients and parameters k_E=0.0025 1/m, k_{TE}=0.000025 m/(°C s), U_{cr}=0.12 m/s, and φ=30° the gullies evolution calculated for the period 1986-2016 was compared with the surface photos of 2007 and space images of 2016 (see Fig. 5).

RESULTS AND DISCUSSION

Before discussing the results it is useful to consider two questions: 1) the interconnection of thermokarst and erosion processes in formation of the gullies in the regions with permafrost, and 2) the use of morphological analogs as hydrological ones.

The first question was raised by Kosov and Konstantinova (1970) during their investigation of erosion processes in the Arctic. Among other types they distinguished so-called thermokarst - erosion gullies, which were formed mostly due to in situ melting of polygonal ice wedges without significant influence of running water from the catchment. These features were rather small and narrow, and their pattern resembles the relief of ground hummocks of conical shape (so-called «*baydjarakhs*»). This type of gullies was only briefly mentioned in the subsequent works of the same authors (see Yershov et al., 1982); instead, after more extensive investigations they described typical thermo-erosion gullies formed by running water (see, for example, Konstantinova 1973). Special works devoted to gully erosion processes in the Arctic (Poznanin 2012), as well as general works in geocryology (e.g., Yershov and Williams 2004) did not discuss the formation of such gullies without running water, and this hypothesis was nearly forgotten.

Interest in the formation of such ravine-like features formed in the absence of water has been renewed due to planetary studies. Numerous attempts to find terrestrial analogs for the gullies on Mars (see, for example, review by Conway et al. 2019) showed that water-lacking processes (including ground ice melting) can produce only small gully-like features, which cannot be compared with large gully nets formed by water erosion. A general degradation of permafrost on the East European Plain with former periglacial conditions left numerous remnants of thermokarst features in recent relief morphology (Velichko 1965). However, that is the relief of small flat hummocks and shallow pits, but not the linear erosion pattern. In the modern conditions in the

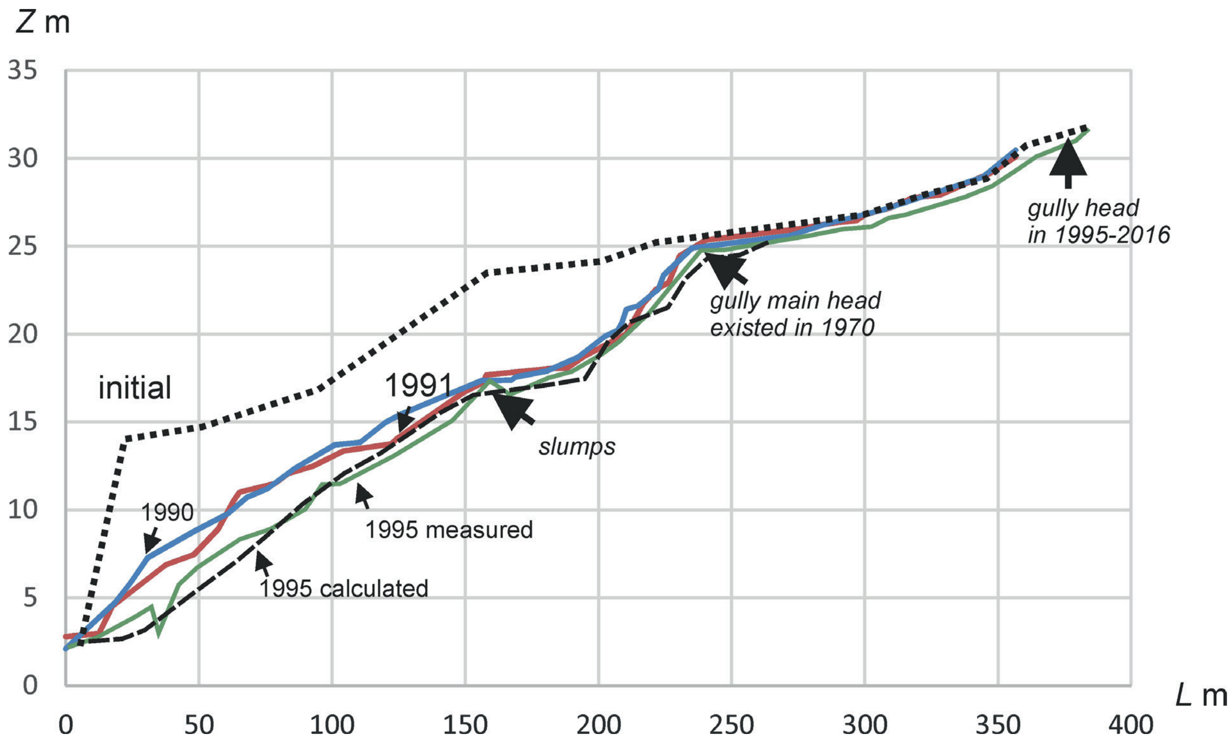


Fig. 7. Measured and calculated longitudinal profile of gully K_1 for calibration period 1990-95

Arctic, some small tributaries not more than 10-20 m long can be formed as the result of ice wedges melting, but only running water prevents them from filling by thaw ground slumps from the slopes. The general shape of gully network can follow the system of cracks in frozen ground. Often in this case, not thermokarst initiates erosion, but erosion and thermoerosion by running water leads to ice wedges melting in the gully cuts (Godin et al. 2012). In some cases, an opposite process was described (Levy et al. 2008): filling of the former gullies by ice wedges. Certainly, the processes of ground ice melting influence the morphology of gullies in permafrost conditions, but this influence remains secondary to the main process of erosion and thermoerosion by water from the catchment even in the Antarctic (Dickson et al. 2019).

The second question arose after Horton (1945) formulated the basic laws of channel network morphology. Gregory (1966) came to conclusion that for hydrological indication it is preferable to use a channel network, which includes both permanent and temporary watercourses. He was one of the first researchers to appreciate the structure of erosion networks as a source of paleohydrological information. Gregory pointed out that the presence of dry valleys not occupied by permanent watercourses may reflect a significant decrease in water flow and pointed out that paleohydrological estimates based on changes in the shape of river channels and cross sections can be supplemented by studying the patterns of fluvial systems.

Gartsman (1968) found a linear relationship between the total length of watercourses ΣL in a system of order K and the mean output discharge Q of this system. He introduced a hydro-morphological coefficient γ_Q (HMC) as an indicator of the length of the river network necessary for the formation of a unit discharge

$$\gamma_Q = \frac{\Sigma L}{Q} \quad (8)$$

The hydro-morphological coefficient represents a way of using morphological analogs as hydrological ones, since the networks of channels of the same density are hydrologically similar if certain geomorphological and lithological characteristics coincide. This opens a possibility to use a morpho-paleohydrological analogy, that is, correspondence between the structure of modern channel networks and their hydrological characteristics can be applied to estimate the hydrological characteristics of the ancient channel systems by their structure.

The relationship between the annual runoff depth (mm) and the averaged density of the dry valley-river network (km/km^2) was established for nine large lowland regions with tundra and sparse northern taiga vegetation and permafrost (Sidorchuk et al. 2018). This relationship is linear, with the coefficient of determination R^2=0.85 (Fig. 8). A relatively high network density of dry valleys and rivers (1 km/km^2) is formed by the annual runoff depth of ~270 mm, which indicates a high erosion efficiency of water flow in such landscapes. The density of the network of dry valleys and rivers K in the Khoper River basin (a tributary of the Don River) is 0.73 km/km^2. This erosion landscape was formed during the powerful surface runoff event at the beginning of the degradation of the last glaciation 16-19 ka ago. The closest modern morphological and hydrological analogs of this landscape are the Northern Siberian Lowland (K=0.7 km/km^2) and Yamal Peninsula (K=0.77 km/km^2), where drainage networks were formed in the Holocene. For further calculations, we used the hydrological data on the Yamal Peninsula, as the region is much better investigated than the Northern Siberian Lowland.

Estimations of the past periglacial erosion rate

After successful calibration of GULTEM for the gullies in the west-central Yamal, we used this model for palaeohydrological reconstructions. Calculations were carried out for the development of the erosion network within the catchment area of the Perepolye dry valley in the Khoper River basin. The

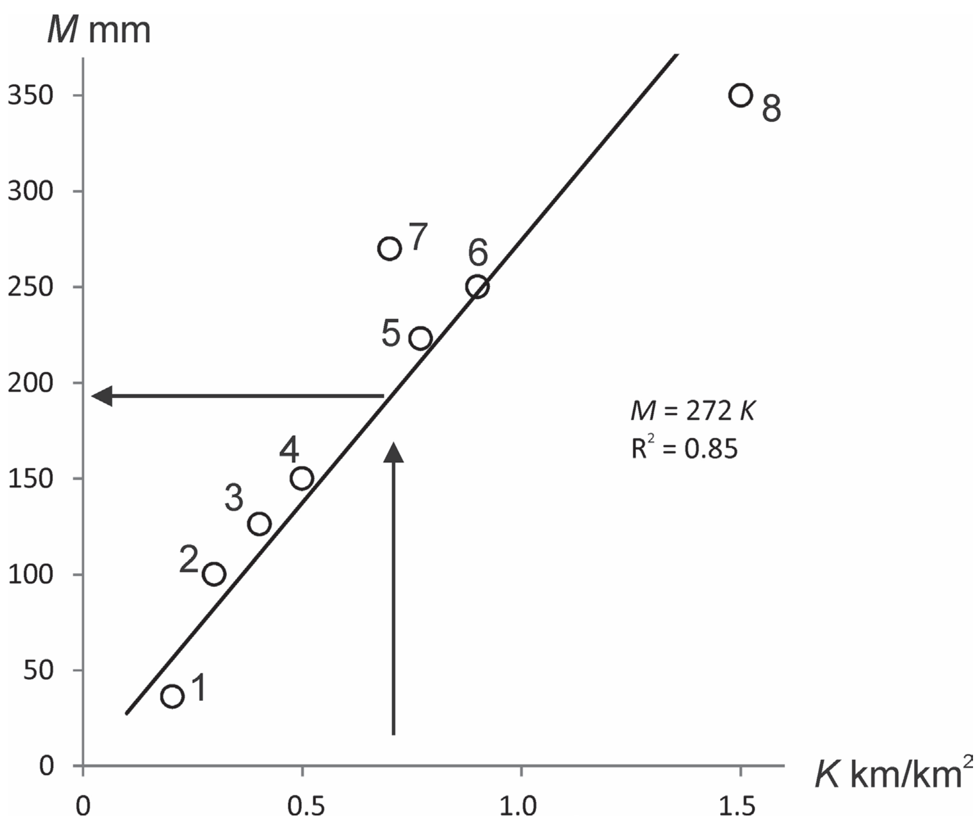


Fig. 8. The relationships between the annual runoff depth *M* (mm) and the averaged density of the dry valley-river network *K* (km/km^2) for the large regions with tundra and sparse northern taiga vegetation and permafrost.

The numbers indicate the lowland regions: 1 – Central Yakutian Lowland; 2 – Yana-Indigirka L.; 3 – Kolyma L.; 4 – Olenek L.; 5 – the Yamal Peninsula; 6 – Northern Siberian L.; 7 – the Gydan Peninsula; 8 – Bolshezemelskaya Tundra initial relief of the catchment was reconstructed by complete "filling" the existing erosion forms, while the modern structure of the drainage network was preserved in the calculations. The main task was to estimate the total length of the gully system and the maximum length of the main trunk for two main periods of catastrophic surface runoff – the Late Moscow time (MIS 6) and the Late Valdai time (MIS 2). As up to 90% of a gully length forms during the first 5% of its lifetime, a 30-year simulation is sufficient for the purpose.

The length of the main trunk of the active paleogully was calculated with critical velocity of erosion initiation U_{cr} varying within the range 0.2-1.6 m/s and the ratio M/M_{Yamal} from 0.2 to 4.1 due to increased or decreased surface runoff depth M (Fig. 9). The recent mean annual runoff depth M_{Yamal} at the reanalysis grid point 70.25°N, 68.25°E on the Yamal Peninsula is 340 mm, the maximum runoff depth for six-hour period is 50 mm for the snow thaw and 24 mm for rainfall period. The main trunk length of the modern Perepolye dry valley system inherited from the Late Valdai time is 6400 m. The critical velocity of erosion initiation U_{cr} corresponding to this value is 0.87 m/c, which is more than the value obtained for the Yamal Peninsula with natural (undisturbed) vegetation. Since the hydrological conditions in the Yamal Peninsula (M_{Yamal}) are not completely analogues to the conditions in the Perepolye catchment in the Late Valdai, an uncertainty of the M/M_{Yamal} ratio of ±20% was included in the calculations. This uncertainty interval leads to variability in critical velocity of erosion initiation U_{cr} within the range 0.8-0.9 m/s. If considering the difficulties of estimating this characteristic for the natural objects, such a range is not especially broad.

The value of U_{cr} =0.87 m/s was then used to reconstruct the hydrological and erosion conditions in the late Moscow time (MIS 6). The system of shallow hollows (that were gullies at the end of the Moscow Glaciation) is denser and longer than the system of dry valleys. The length of the main trunk of an active paleogully was then 8400 m. The calculations with GULTEM showed a high annual and maximum depth of the surface runoff (see Fig. 9). The calculated value of *M* (1090 mm), which corresponds to the network of hollows in the case of U_{cr}=0.87 m/s, is almost 3.3 times greater than M_{Yamal}. The uncertainty of this estimate of M/MYamal is very high and is in the range 2.3-4.1. Such high errors are explained by the very weak relationship between the surface runoff and the growth of the gully when its top is close to the watershed.

CONCLUSION

The same type of erosion landscape is characteristic of the Arctic regions, especially for the Yamal, Gydan and Tazovsky peninsulas with active gullies, and for various regions of the East European Plain, where ancient active gullies were later transformed into stable dry valleys. In the both cases, this type of landscape develops due to erosion in periglacial conditions. Thus, the Arctic regions can be used as modern analogs to reconstruct the conditions for the development of periglacial erosion features in regions with a modern temperate climate. For these purposes, gully erosion and thermoerosion model GULTEM was verified and calibrated

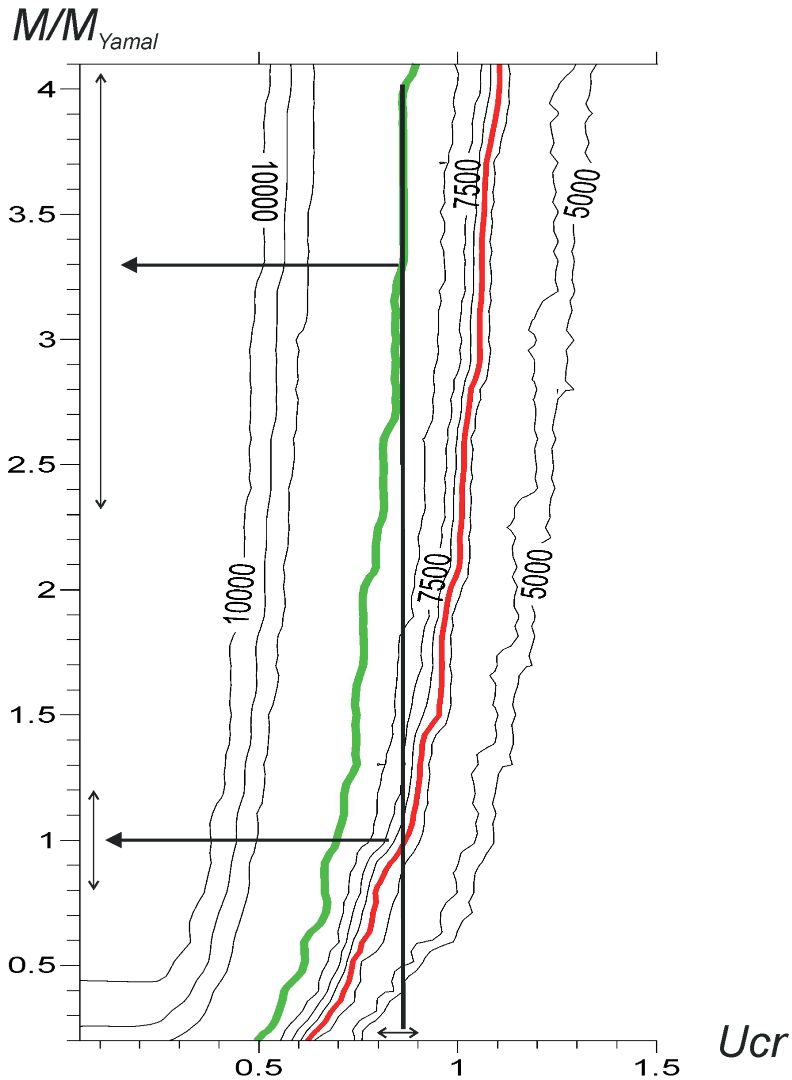


Fig. 9. The diagram for estimating the length of the main trunk of an active paleogully at the Perepolye dry valley catchment using the critical velocity of erosion initiation U_{cr} and the ratio of runoff depths M/M_{Yamal}

based on the observations of the modern processes of gully erosion in the Yamal Peninsula. The most suitable characteristics of climate and vegetation cover, which can explain the structure and density of dry valley systems in the East European Plain were then estimated with this model. The main input characteristics, which control calculation results are critical velocity of erosion initiation U_{cr} and surface runoff M. These characteristics were varied in numerical experiments to obtain the best similarity with the morphology of existing dry valleys. The topography of the modern dry valley and hollow system in the basin of the Perepolye dry valley (the Khoper River basin) was used to calculate the length of the active paleogullies. The length of the main trunk of this erosion system was 8400 m in the end of the Moscow Glaciation (MIS 6) and 6400 m in the end of the Late Valdai Glaciation (MIS 2). The critical velocity of erosion initiation U_{cr} corresponding to these values is within the range of 0.8-0.9 m/s for both time intervals. For the first period, the surface runoff, which fits the network of hollows, is almost 3.3 times greater than the recent one in the Yamal Peninsula, or more than six-fold the modern value in the Khoper basin. For the second period, the mean annual surface runoff, which could produce these large gullies, was close to the present annual surface runoff in the Yamal Peninsula, approximately 330 mm. This is about twice the modern runoff on this catchment. The uncertainty of these estimates is relatively low for the middle parts of the catchments and increases significantly near the watershed.

ACKNOWLEDGEMENTS

This study was carried out with support of the Russian Science Foundation project 18-05-60147 "Extreme hydro-meteorological phenomena in the Kara Sea and the Arctic coast".

REFERENCES

Bettis III E.A. and Thompson D.M. (1985). Gully Erosion. Rangelands, 7(2), 70-72.

Betts A.K., Köhler M. and Zhang Y. (2009). Comparison of river basin hydrometeorology in ERA–Interim and ERA–40 reanalyses with observations. Journal of Geophysical Research: Atmospheres, 114, D02101. DOI: 10.1029/2008JD010761.

Bobrovitskaya N.N., Baranov A.V., Vasilenko N.N. and Zubkova K.M. (1999). Hydrological conditions. In: A. Sidorchuk and A. Baranov, eds., Erosion processes of central Yamal. St. Petersburg: RNII KPN, 90-105 (in Russian).

Borisova O.K., Sidorchuk A.Yu. and Panin A.V. (2006). Palaeohydrology of the Seim River basin, Mid-Russian Upland, based on palaeochannel morphology and palynological data. Catena, 66, 53-78. DOI: 10.1016/j.catena.2005.07.010.

Bulygina O.N., Razuvayev V.N., and Aleksandrova T.M. (2008). Description of the data set of daily air temperature and precipitation at meteorological stations in Russia and the former USSR (TTTR). Certificate of State Registration of the Database № 2014620942 (in Russian).

Carlston C.W. (1963). Drainage density and streamflow. Geological Survey Professional Paper, 422.

Conway S., de Haas T., and Harrison T.N. (2019). Martian gullies: a comprehensive review of observations, mechanisms and insights from Earth analogues. Geological Society London Special Publications, 467, 7-66. DOI: 10.1144/SP467.14.

Dee D.P. and Uppala S.M. (2009). Variational bias correction of satellite radiance data in the ERA–Interim reanalysis. Quarterly Journal of the Royal Meteorological Society, 135, 1830-1841. DOI: 10.1002/qj.493.

Dee D.P., Uppala S.M., Simmons A.J., Berrisford P., et. al. (2011). The ERA–Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553-597. DOI: 10.1002/qj.828.

Dickson J.L., Head J.W., Levy J.S., Morgan G.A. and Marchant D.R. (2019). Gully formation in the McMurdo Dry Valleys, Antarctica: multiple sources of water, temporal sequence and relative importance in gully erosion and deposition processes. Geological Society London Special Publications, 467, 289-314. DOI: 10.1144/SP467.4.

Donat M.G., Sillmann J., Wild S., Alexander L.V., Lippmann T. and Zwiers F.W. (2014). Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets. Journal of Climate, 27(13), 5019-5035. DOI: 10.1175/JCLI-D-13-00405.1.

Eremenko E.A. and Panin A.V. (2011). An origin of the net of hollows in the central and southern regions of the East European Plain. Vestnik Moskovskogo Universiteta, Seriya 5, Geografiya, 3, 59-66 (in Russian).

Fortier D., Allard M. and Shur Y. (2007). Observation of rapid drainage system development by thermal erosion of ice wedges on Bylot island, Canadian Arctic Archipelago. Permafrost and Periglacial Processes, 18(3), 229-243. DOI: 10.1002/ppp.595.

Gartsman I.N. (1968). River network and basin outflow in the southern Far East. Trudy DVNIGMI, 27, 15-22 (in Russian).

Gelfan A.N. (1999). Model of formation of river flow during snowmelt and rain. In: A. Sidorchuk and A. Baranov, eds., Erosion processes of central Yamal. St. Petersburg: RNII KPN, 205-225 (in Russian).

Godin E. and Fortier D. (2012). Geomorphology of a thermo-erosion gully, Bylot Island, Nunavut, Canada. Canadian Journal of Earth Sciences, 49(8), 979-986. DOI: 10.1139/e2012-015.

Gregory K.J. (1966). Dry valleys and the composition of the drainage net. Journal of Hydrology, 4, 327-340. DOI: 10.1016/0022-1694(66)90096-5.

Gregory K.J. and Walling D.E. (1968). The variation of drainage density within a catchment. International Association of Scientific Hydrology Bulletin, 13(2), 61-68. DOI: 10.1080/02626666809493583.

Horton R.E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Bulletin of the Geological Society of America, 56, 275-370. DOI: 10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2

Hošek J., Prach J., Křížek M., Šída P., Moska P. and Pokorný P. (2019). Buried Late Weichselian thermokarst landscape discovered in the Czech Republic, central Europe. Boreas, 48(4), 988-1005. DOI: 10.1111/bor.12404.

Knighton D. (1984). Fluvial Forms and Processes. London: Edward Arnold.

Komarov V.D., Makarova T.T and Sinegub E.S. (1969). Calculation of the hydrograph of floods of small lowland rivers based on thaw intensity data. Proceedings of the Hydrometeorological Center of the USSR, 37, 3-30 (in Russian).

Konstantinova G.S. (1973). The peculiarity of the erosion relief of the Yamal interfluve. In: Soil erosion and channel processes, 3. Moscow: Izdatelstvo MGU, 105-115 (in Russian).

Kosov B.F. and Konstantinova G.S. (1970). Features of ravine erosion in the tundra. In: Soil erosion and channel processes, 1. Moscow: Izdatelstvo MGU, 157-161 (in Russian).

Kosov B.F., Nikol'skaya I.I. and Zorina Ye.F. (1978). Experimental studies of ravine formation. In: N.I. Makkaveev, ed., Experimental geomorphology, 3. Moscow: Izdatelstvo MGU, 113-140 (in Russian).

Larsen A., Heckmann T., Larsen J.R. and Bork H.-R. (2016). Gully catchments as a sediment sink, not just a source: Results from a long-term (~12,500 year) sediment budget. Earth Surface Processes and Landforms, 41(4), 486-498. DOI: 10.1002/esp.3839.

Levy J., Head J., and Marchant D. (2008). The role of thermal contraction crack polygons in cold-desert fluvial systems. Antarctic Science, 20(6), 565-579. DOI: 10.1017/S0954102008001375.

Lindsay R., Wensnahan M., Schweiger A., and Zhang J. (2014). Evaluation of seven different atmospheric reanalysis products in the Arctic. Journal of Climate, 27(7), 2588-2606. DOI: 10.1175/JCLI-D-13-00014.1.

Matoshko A.V. (2012). Balkas – a new look at a common landform of the East European Plain from a Quaternary perspective. Earth Surface Processes and Landforms, 37, 1489-1500, DOI: 10.1002/esp.3255.

Matveeva T.A., Gushchina D.Y. and Zolina O. (2015). Large-scale indicators of extreme precipitation in coastal natural-economic zones of the European part of Russia. Russian Meteorology and Hydrology, 40(11), 722-730. DOI: 10.3103/S1068373915110023.

Panin A.V. and Sidorchuk A.Yu. (2006). Macrobends («large meanders»): The problems of origin and interpretation. Vestnik Moskovskogo universiteta, Seriya 5, Geografiya, 6, 14-22 (in Russian).

Panin A.V., Matlahova E.Y., Belyaev Y.R., Bulyart J.-P., Dubis L.F., Murray A., Pakhomova O.M., Selezneva E.V. and Filippov V.V. (2011). Sedimentation and formation of terraces in river valleys of the central Russian Plain during the second half of the Late Pleistocene. Bulletin of the Commission for Study of the Quaternary, 71, 47-74 (in Russian).

Popov Ye.G. (1956). Analysis of river flow formation. Leningrad: Gidrometeoizdat (in Russian).

Poznanin V.L. (2012). Erosion processes in cryolitzone. Space and Time, 1(7), 127-132 (in Russian).

Prosser J.P., Chappell J.M.A. and Gillespie R. (1994). Holocene valley aggradation and gully erosion in headwater catchments, South-Eastern Highlands of Australia. Earth Surface Processes and Landforms, 19, 465-480.DOI: 10.1002/esp.3290190507

Sidorchuk A. (1996). Gully erosion and thermoerosion on the Yamal Peninsula. In: O. Slaymaker, ed., Geomorphic Hazards. New York: Wiley, 153-168.

Sidorchuk A. (2015). Gully erosion in the cold environment: Risks and hazards. Advances in Environmental Research, 44, 139-192.

Sidorchuk A.Yu. and Borisova O.K. (2000). Method of paleogeographical analogues in paleohydrological reconstructions. Quaternary International, 72, 95-106. DOI: 10.1016/S1040-6182(00)00025-2.

Sidorchuk A., Panin A., and Borisova O. (2011). Surface runoff to the Black sea from the East European plain during the Last Glaciation maximum – Late Glacial time. In: Geology and Geoarchaeology of the Black Sea Region: Beyond the Flood Hypothesis. Geological Society of America Special Paper, 473, 1-25. DOI: 10.1130/2011.2473(01).

Sidorchuk A.Yu., Panin A.V., Borisova O.K. and Eremenko E.A. (2018). Geomorphological approach to the river flow evaluation in the geological past. Paper 3. Drainage net structure analysis. Geomorfologiya (Geomorphology RAS), 1, 18-32 (in Russian). DOI: 10.7868/S0435428118010029.

Sun Q., Miao C., Duan Q., Ashouri H., Sorooshian S. and Hsu K.L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. Reviews of Geophysics, 56(1), 79-107. DOI: 10.1002/2017RG000574.

Velichko A.A. (1965). The relict permafrost relief of the periglacial zone (cryolithozone) of the Russian Plain. In: Quaternary and its history. Moscow: Nauka, 104-120 (in Russian).

Velichko A.A., Morozova T.D. and Panin P.G. (2007). Polygenetic soil complexes as a system phenomenon of the Pleistocene macrocycles). Izvestiya Rossiiskoi Akademii Nauk, Seriya Geograficheskaya, 2, 44-54 (in Russian).

Vinogradov Yu.B. (1988). Mathematical modeling of flow formation processes. Leningrad: Gidrometeoizdat (in Russian).

Vinogradov Yu.B., Vinogradova T.A., Zhuravlev S.A. and Zhuravleva A.D. (2014). Mathematical modeling of hydrographs from the unstudied river basins on the Yamal Peninsula. Bulletin of St. Petersburg State University, Seriya 7, 3, 71-81 (in Russian).

Yershov E.D. and Williams P.J. (2004). General Geocryology (Studies in Polar Research). Cambridge: Cambridge University Press.

Yershov E.D., Malinovskiy D.V., and Kuchukov E.Z. (1982). Thermoerosion of dispersed sediments. Moscow: Izdatelstvo MGU (in Russian).

Received on January 7th, 2019 Accepted on January 21th, 2020

Vyacheslav V. Dolotov[1*], Yuri N. Goryachkin[1], Andrey V. Dolotov[1]
[1] Federal Research Center «Marine Hydrophysical Institute of the RAS», Russia
***Corresponding author:** vyadol@gmail.com

ANALYSIS OF THE CURRENT STATE OF UNSTABLE GEOMORPHOLOGICAL STRUCTURES WITH MODERN METHODS

ABSTRACT. The paper gives results of the digitization of the status and spatial position of a cliff in the Western Crimea coastal zone. The modern equipment and methods accelerate the survey from the time perspective and improve the quality of the outcomes; namely a high precision GNSS receiver in RTK mode and PHANTOM-3 PRO copter. The digital terrain model was generated with used the Agisoft Photoscan software. The paper shows that the precision of the mathematical model of the relief constructed by aerial photographs provides more detailed data in comparison to those obtained in the field observations. Furthermore, aerial photography makes it possible to calculate the number of spatial characteristics of hazardous for surveying and latent natural objects out of reach for an on-location investigation. As a result, the very detailed data about current condition of dangerous cliff were obtained. The paper also evaluates the linear and volumetric characteristics of cleavages that are prone to collapse.

KEY WORDS: Crimea, coastal zone, field observations, relief modeling, photogrammetry

CITATION: Vyacheslav V. Dolotov, Yuri N. Goryachkin, Andrey V. Dolotov (2019) Analysis Of The Current State Of Unstable Geomorphological Structures With Modern Methods. Geography, Environment, Sustainability, Vol.13, No 1, p. 195-199
DOI-10.24057/2071-9388-2019-20

INTRODUCTION

The information on the sea coastal zone relief with its variability is always of particular interest in the safety of economic activities and the sustainability of nature.

The leading Russian scientists Zenkovich V. (Zenkovich 1960, Zenkovich 1962) and Longinov V. (Longinov 1963) focus on these issues in particular; a specialized terminological guidebook on the marine geomorphology (Geomorphology 1980) was published in 1980 that marked the end of the first stage of the seacoast study. Further, it served as a tool for the coastal zone dynamics studying (Shuisky 1984).

One of the most important parts of the Russian coastal zone is the Black seashores that have always been attracting people with their amazing nature. That caused the recognized Man vs. Nature Conflict that was rooted in the coastal areas reconstruction that broke the fundamental natural laws (Agarkova 1999, Ivanov, etc. 2006). At the same time, tourism and recreation as an important part of medical and social spheres are inevitably facing challenges.

At the beginning of the 21st century, the situation started worsening on the Black Sea coasts of the Caucasus and Crimea, which initiated in-depth research both of the Caucasus (Peshkov 2003), and of Crimea (Shuysky 1974, Romanuk 1992, Goryachkin 2009). At the same time, the most severe damage was noticed in the beach areas of the coasts; and earlier attempts to strengthen the coasts with concrete structures have resulted in more dramatic damage only (Peshkov 2003, Shuysky 2005, Goryachkin 2011).

The adoption of the Crimea Supreme Council Resolution (The main… 2004) by the municipal authorities of Crimea, as well as the scientific methods development of cadastral evaluation of the beaches recreational resources (Dolotov 2010), proved they are ready to recognize the importance of the up-to-date knowledge and methods to ensure the stability of coastal zones, and the beaches preservation in particular. The unique technique covers the exhaustive indicators of beaches and their recreational areas, taking all geomorphologic characteristics. The particular attention should be given to the state of dynamic landforms, especially unstable ones such as cliffs. The interdisciplinary study, carried by Crimea economists and rehabilitation specialists (Efremov 2003), developed the method of monetary assessment of the beaches' recreational resources, based on the evaluation of the individual recreational factors on health. This paper indicates the importance of permanent monitoring of the dynamics in coastal zones.

The great importance of the permanent monitoring of high dynamic landforms in the coastal zone was noted in the past, e.g. (Goryachkin et al. 2009; Dolotov et al. 2017-a; Udovik 2009). The natural cliff foundering, investing in the formation and recharge of beaches, causes the gradual extension of the coastal zone inland challenging the economic activity within the seashores. The significant changes in the spatial position of the coastline are mainly the result of the storms affecting extended sites of the coast.

There was a requirement to upgrade the methods and measuring equipment, and to increase the spatial positioning accuracy for any natural and artificial object as a priority, as proved in the paper.

MATERIALS AND METHODS

The former methods of the coast studying were quite primitive, though up-to-date once, with limited capacity in measuring the width of beaches and the distance between the cliff edge and some fixed ground reference points. They were quite time-consuming and therefore used to fail to pinpoint the contours of extended curved objects.

The aerial survey technologies and the gradual rise of the field survey accuracy (Kosyan 2012, Manuel 2012) ensured the significant progress in those studies. However, the current compliance of the accuracy with the uncertainty of a few meters, taking into account the spatial dimensions of the measured objects, requires georeferencing with the use of highly accurate GNSS devices (Pulling… 2010). The more accurate photographs are too expensive (Goryachkin et al. 2009). Thus, the monitoring of the coastal zone is only possible periodically, mainly along with approved expeditions. Given this, the analysis of a particular storm impact on the coastal landforms faces challenges, which affects the accuracy of the assessment and prognosis.

The introduction of drones as aerial survey equipment rocketed the quality of the coastal instruments. The drone advantages in the coastal zones studying are obvious. These include first of all the simplicity and cheapness in obtaining high-resolution images (Tao 2001). Drones along with the up-to-date software allow taking photographs in the photogrammetry mode in a short time, which allows you to elaborate the significantly accurate, and, what is of the most importance, - the detailed digital terrain models. The estimation of the exactness of these models (Dolotov 2017-b, Pikelj 2017, Krylenko 2018) for that case meets the requirements of the tasks. It is to be stressed that to ensure the methods assessment completeness, a significant model accuracy in terms of the spatial location of the horizontal and vertical points one can only get when georeferencing the very particular relief element in the images thus ground control points (GCP) to its real location.

However, according to the recommendations of the software developers (Unmanned… 2018) the number of these GCPs can be relatively small. Those GCPs can be assigned as temporary marks, particularly at the breakpoints of relief assessment that lacks constants, or as fixed marks at stationary elements of the terrain. Corners of retaining walls and berths and nearly any low capital constructions can serve as the stationary elements to pinpoint the location. As a result of the long-term monitoring, the staff of the Marine Hydrophysical Institute of RAS constructed an exhaustive database of landmarks with the use of the GNSS-receiver in RTK mode (Geodesy… 2018) on the sites that use by the institute, and that extends from Saky, the town in the North of Crimea, to the Tolstoy Cape in the South.

This database makes it possible to survey the area promptly with the use of a Phantom3 Pro drone equipped with a high-resolution camera. All the landmarks registered in the base go along with the images and survey points location map as well as with the fixed coordinates matching the elevation level.

RESULTS

The study focuses on the possibilities of digitization of natural objects on the base of the prone to slumping coastal area located to the North from Nikolaevka village in Western Crimea (Fig. 1).

The study site is a beach 4-20 meters wide and 1.1 km long, located at the foot of the nearly vertical cliff with a height of around 15 meters. Following the basic concepts of the dynamic of this geomorphological structure, the cliff would slump periodically (traces of these process are in Fig. 1, right), and at the same time, there are several natural cleavages (box in Fig. 1) at its upper part, what implies the further contribution to the beach by this source.

While traditional measuring, the upper edge coordinates of the cliff are surveyed like extra points to the beach parameters, but the measurements are to be conducted every 2-5 m along the edge, i.e. from 200 to 500 measurements, to ensure the pinpointing of its contour. There were 48 measurements only taken, however, throughout the real survey of the cliff, i.e. about 5-10 times less than assumed.

There were two drone flights within the survey based on the above-mentioned approach from the center to both sides, traversing twice, taking 87 images with overlapping 60% along the direction of the flight, as well as across (Fig. 2). The flights were performed at 80 m altitude from the cliff level at a speed of 3.5 m/s. By the speed increase up to 4 m/s, there were occasional omissions seen in a set of images, and therefore the speed of 3.0 – 3.5 m/s was considered optimal. The flight routes were developed by the Basic version of the Drone Harmony program (Professional… 2018).

The images were processed by the trial version of Agisoft Photoscan (Unmanned… 2018). After the images uploading, they were spatially aligned with 16 GCP, six out of them located on the plane of the beach, eight at the height of the upper limit of the cliff and two on the concrete staircase leading to the sea in the southern part of the beach. The model with the following characteristics was elaborated as a result of the survey (Fig. 3):

- images: 87;
- GCP: 16;
- tie points: 94625;
- dense cloud of 53654822 points;
- model of 10604753 polygons.

Fig. 1. The coast near the Nikolaevka village; one of the sites with a typical cleavage (box)

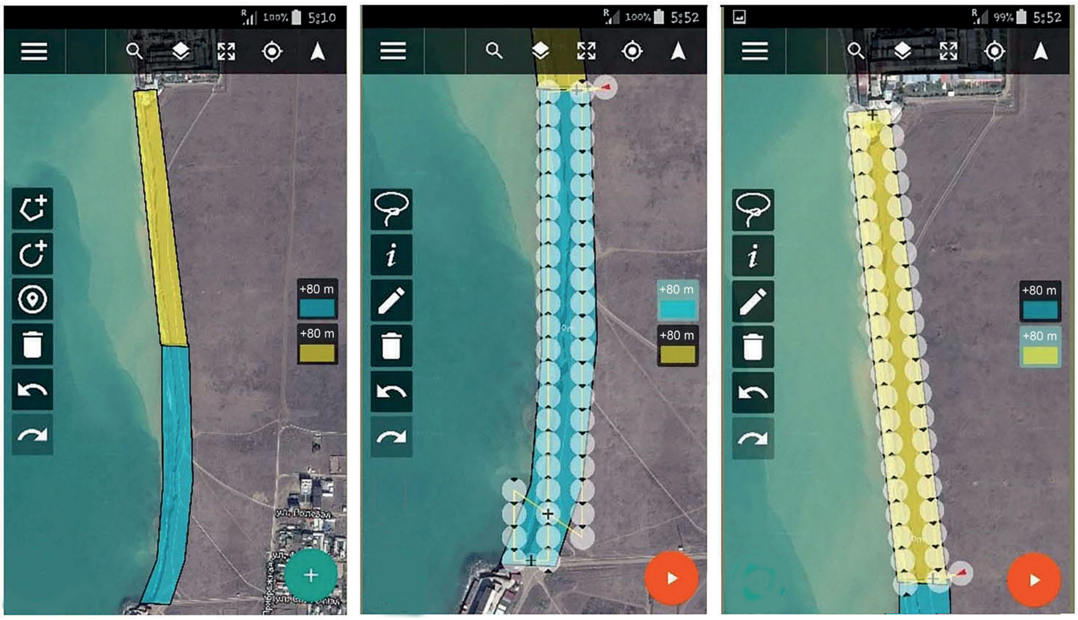


Fig. 2. Flight plan (left) and image locations taken by each drone (center, right)

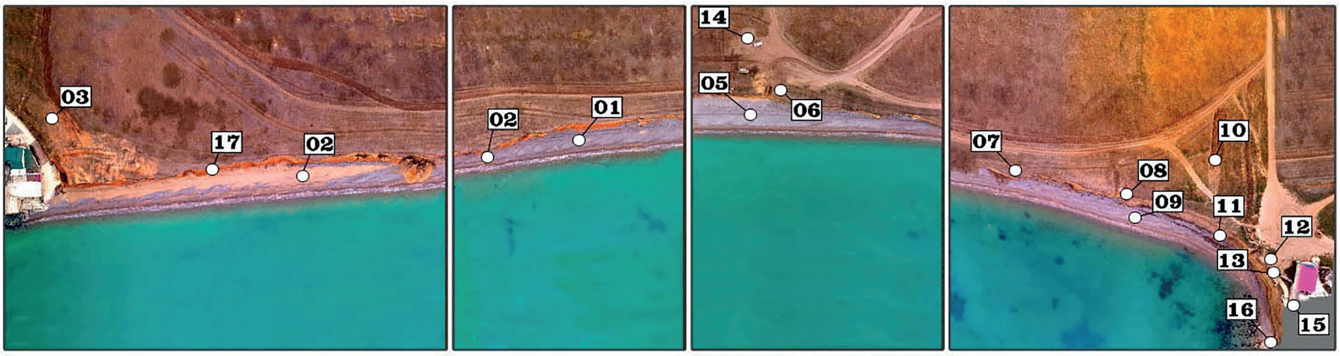


Fig. 3. Visualization of the model with the use of an image texture

The height map of size 6508 x 35504 pix with a resolution of 4.88 cm/pix and the orthophotomap of size 8412 x 49350 pix with a resolution of 2.44 cm/pix were elaborated based upon the received data.

The digitizing of the cliff edge was also carried out with the use of the traditional drawing procedure of the polyline from the top look. For the case when the point fell to the flat of a beach, its position was corrected shoreward and controlled in the vertical direction by switching to the front look. There were 468 points along the edge of the cliff digitized totally with a mean distance of 2.3 m. Forty-seven points extra were obtained by a GNSS-receiver (Fig. 4), and the corresponding points coincided.

It is clear that for the case of a crooked survey area the distance between the points tends to be reduced to 1 m or less (Fig. 5) and if necessary, the total quantity of digitized points can be significantly increased.

The extra evaluation of some linear and volumetric characteristics of cleavages, which did not lie in the horizontal flat in contrast to the edge of the cliff, seemed interesting throughout the survey. These characteristics cannot be measured within on-location measurements. Eight explored cleavages were digitized in the same way, i.e. by drawing a line along the contour of the segment that split from the landmass (Fig. 6).

Thus, on the presented model, the length of the line within horizontal points was 7.4 m long, and the total volume of the landmass within the extreme points was 18.4 m^3. It is interesting to note that this model allows building and displaying the vertical section constructed by digitized points (box in Fig. 6) as well, which shows that many of digitized points initially hid in the cleavage. This result confirms that to perform accurate measurements of the geometric dimensions of objects it is necessary to perform pre-alignment of digitized points in the vertical flat.

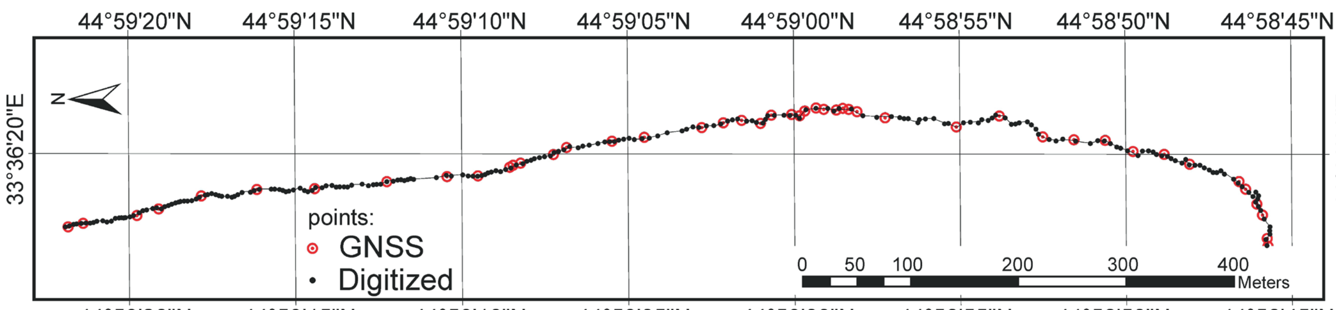


Fig. 4. Spatial position of points throughout the survey

Fig. 5. The site of the cliff edge digitized by means of Photoscan

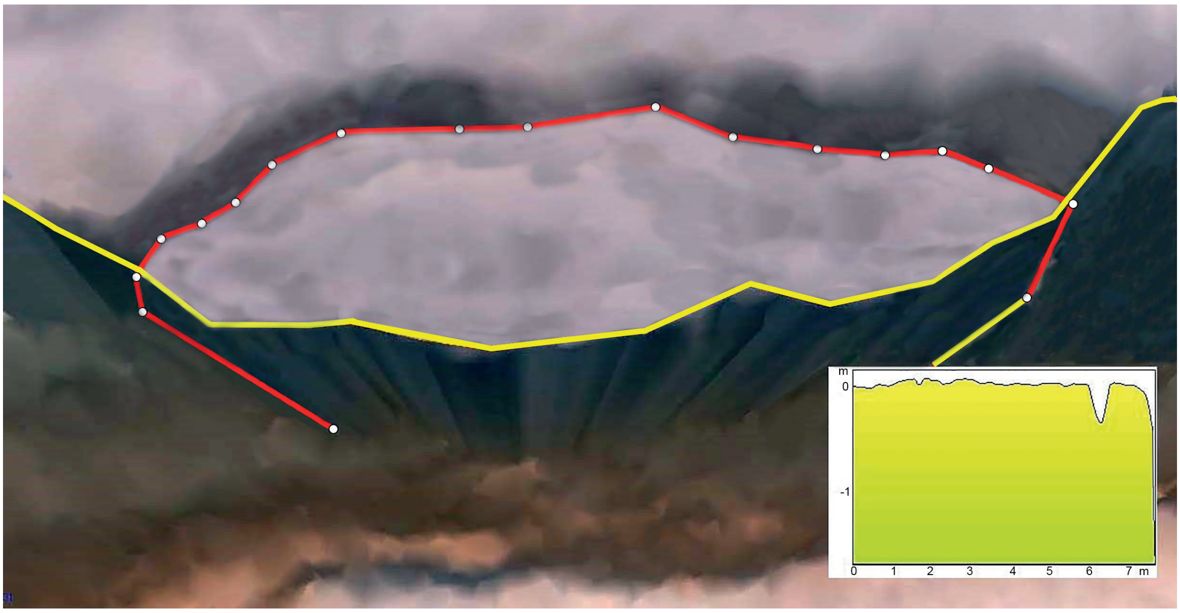


Fig. 6. The digitization of cleavage with a polyline

DISCUSSION

The results show that the modern survey technologies provide significant progress in the determining of the spatial position of individual geomorphological structures, including those inaccessible to on-location measurement. The main advantages of the method are the speed of taking photographs resulting in the high-precision digitization and measurements of the spatial position of individual natural objects, including those that cannot be measured in on-location surveys. The results of the omitting in this paper video-filming also allow us to estimate the thickness of the upper gravel layer, which together with the results of measuring the volume of existing natural cleavages allows quantifying the amount of landmass expected to recharge the beach in the near future.

CONCLUSIONS

No question that the most significant results of the survey are expected to be later in further monitoring of the coastal zone in Western Crimea. At present, according to the results of long-term observations of two sites accomplished by the authors, the beach extension from 2004 to 2014 was estimated to be 1.6 m with a mean rate of 0.14 m/yr. In this regard, the obtained results make it possible to assess the coast retreat in some areas with high accuracy. This allows early identification of both dangerous and dynamic areas of the coast to make proper decisions on its protection. At present, the results are the first detailed data array on the spatial position of the cliff edge facilitating the identification and estimation of the potential slump in the study site of the coast.

ACKNOWLEDGEMENTS

The present study was undertaken within «Complex Interdisciplinary Studies of Oceanological Processes that Determine the Functioning and Evolution of the Black and Azov Seas Ecosystems Based on Modern Methods of Monitoring the Marine Environment and Grid Technologies» State Task №0827-2014-0010, and RFFI project №18-45-920021 (2018).

REFERENCES

Agarkova I.V. (1999). Influence of economic activity on the dynamics of the Saki coast. Scientific notes of Vernadsky Taurida National University, 12, Simferopol, 15-19 (in Russian).

Dolotov V.V, Ivanov V.A. (2010). Cadastral Assessment of Crimean Beaches as an Instrument for Sustainable Coastal Development. GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY, 3(2), 98-119.

Dolotov V.V., Goryachkin Yu.N., Dolotov A.V. (2017-a). Statistical analysis of changes in the shoreline of the village Lyubimovka beach. Environmental safety of coastal and shelf zones and integrated use of shelf resources, 1, 40-47. (in Russian with English summary).

Dolotov V.V., Udovik V.F. (2017-b). Use of innovative technologies for monitoring the coastal zone of Crimea, Environmental safety of coastal and shelf zones and integrated use of shelf resources, 4, 67-77. (in Russian with English summary).

Efremov A.V. (2003). Considerations for the beache's economic evaluation. Crimea Economics, 6, 57-60 (in Russian).

Geodesy with pleasure. (2018). Effective technology. [online] Available at: www.eftgroup.ru [Accessed 28 Marth 2019].

Goryachkin Yu.N., Kharitonova L.V., Dolotov V.V. (2009). North-Western Crimea coastline variability. Environmental safety of coastal and shelf zones and integrated use of shelf resources, 20, 18-26 (in Russian).

Goryachkin Yu.N., Dolotov V.V. (2011). Changes in the shoreline of the accumulative shores of Western Crimea. Environmental safety of coastal and shelf zones and integrated use of shelf resources, 25, 8-18 (in Russian).

Goryachkin Yu.N. (2015). The current state of the coastal zone of Crimea. Sevastopol: ECOSI-Gidrophysics (in Russian).

Ivanov V.A. Yastreb V.P., Goryachkin Yu.N., Prusov A.V., Zima V.V., Fomin V.V. (2006). Nature management on the black sea coast of Western Crimea: current state and development prospects, Sevastopol: MHI (in Russian).

Kosyan R.D., Goryachkin Yu.N., Krylenko V.V., Dolotov V.V., Krylenko M.V., Godin E.A. (2012). Crimea and Caucasus accumulative coasts dynamics estimation with using satellite pictures. Turkish Journal of Fisheries and Aquatic Sciences. 12, 385-390.

Geomorphology of sea shores (1980). Ed. Zenkovich V.P. Moscow: Mysl (in Russian).

Krylenko V.V., Krylenko M.V. (2018). High-precision topography bakalskaya spit. Environmental safety of coastal and shelf zones, 4, 65-72 (in Russian with English summary).

Longinov V.V. (1963). Dynamics of the non-tidal seas coastal zone. Moscow: Academy of Science (in Russian).

Manuel A. Aguilar, Fernando J. Aguilar, María del Mar Saldaña and Ismael Fernández (2012). Geopositioning accuracy assessment of GeoEye-1 panchromatic and multispectral imagery. Photogrammetric Engineering & Remote Sensing, 78, 3, 247-257.

Peshkov V.M. (2003). Coastal zone of the sea. Krasnodar: Lacont (in Russian).

Pikelj K. (2017). Ocean and Coastal Management. DOI: 10.1016/j.ocecoaman.2017.11.019 [Accessed 16 May 2018].

Professional flight planner for DJI Drones (2018). DH DRONE HARMONY. [online] Available at: www.droneharmony.com [Accessed 16 May 2018].

Pulling in All Signals. (2010). PPP with GPS and GLONASS: The New G2. GPS World, 3, 28-36.

Romanuk O.S., Lushik A.V., Morozov V.I. (1992). Conditions of formation and dynamics of the sea coast in the area of Saki resort zone, Simferopol: IMR, 12-21 (in Russian).

Shuisky Yu.D. (1974). Processes and rates of abrasion on the Ukrainian shores of the Black and Azov seas. News of USSR Academy of Sciences. ser. Geography, 6, Moscow, 107-117 (in Russian).

Shuisky Yu.D., Vykhovanetz G.V., Karasev L.M., Makeev I.A. (1984). Morphology and dynamics of the coast of the South-Western part of the Crimean Peninsula. Physical Geography and Geomorphology, 31, 83-88 (in Russian).

Shuisky Yu.D. (2005). The main regularities of morphology and dynamics of the West coast of the Crimean Peninsula. Environmental safety of coastal and shelf zones and integrated use of shelf resources, 13, 62-72 (in Russian).

Tao, C.V. and Y. Hu (2001). A comprehensive study of the rational function model for photogrammetric processing. Photogrammetric Engineering & Remote Sensing, 67(12), 1347-1357.

The main directions of development of sanatorium-resort and tourist (recreational) complex of the Crimea until 2010. (2004). Resolution of the Supreme Council of Crimea, March 17(849-3/04).

Udovik V.F., Dolotov V.V. (2009). Modern trends in the dynamics of the coastal zone in the area of the beach village Lyubimovka. Environmental safety of coastal and shelf zones and integrated use of shelf resources, 20, 92-99. (in Russian).

Unmanned technologies for professionals (2018). GEOSCAN'S Official Website. [online] Available at: www.geoscan.aero/ [Accessed 16 May 2018].

Zenkovich V.P. (1960). Morphology and dynamics of the Soviet black sea coast. Moscow: Academy of Science. (in Russian).

Zenkovich V.P. (1962). Fundamentals of the development of sea shores doctrine. Moscow: Academy of Science (in Russian).

Received on February 16th, 2019 Accepted on January 21th, 2020

Vladislav V. Sysuev[1]
[1] Faculty of Geography, Lomonosov Moscow State University, Moscow, Russia
***Corresponding author:** v.v.syss@mail.ru

GEOPHYSICAL ANALYSIS OF LANDSCAPE POLYSTRUCTURES

ABSTRACT. The objective identification of landscape cover units is very important for sustainable environmental management planning. The article proposes a method-algorithm for describing the formation of landscape structures, which is based on the classic landscape analysis and applies the parameters of geophysical fields. The main driving forces of all structure-forming processes are the gradients of gravitational and insolation fields, parameters of which were calculated using the digital elevation models and the GIS-technologies. A minimum number of principal parameters are selected for typological and functional classification of landscapes. The number and importance of parameters were identified basing on the results of numerical experiments. Landscape classifications elaborated on the basis of standard numerical methods take a fundamental geophysical value. In this case, a concept of polystructural landscape organization is logical: by selecting different structure-forming processes and physical parameters, different classifications of landscapes could be elaborated. The models of geosystem functioning are closely related to their structure through boundary conditions and relations between parameters. All models of processes and structures are verified by field experimental data obtained under diverse environmental conditions.

KEY WORDS: Geosystem, Modeling, Geophysical field, Morphometric parameters, Flow

CITATION: Vladislav V. Sysuev (2019) Geophysical Analysis Of Landscape Polystructures. Geography, Environment, Sustainability, Vol.13, No 1, p. 200-213
DOI-10.24057/2071-9388-2019-17

INTRODUCTION

The identification of multi-scale polystructural geosystems and the boundaries between them is among the principal problems of landscape research. The fundamentals of non-equilibrium thermodynamics show the principles on which the classifications of natural-territorial complexes (NTC) should be based. In accordance with the Onsager bilinear equation, classifications should account for: 1) the system-forming flows; 2) the force fields and their gradients; 3) the phenomenological coefficients of generalized conductivity. The fields of gravity and insolation are the most common for any geosystem. Selection (classification, integration) of geographical objects by the parameters describing the geophysical fields and their gradients leads to the identification of geosystems according to the flows of matter and energy. This is a functional approach to the identification and investigation of geosystems, which is being developed in the works by Armand (1988), Reteum (1975), and other. In such a case the boundaries of geosystems are determined by the magnitude and sign of flow divergence. For example, if we consider the behavior of elementary water volumes in a geopotential field, we obtain a hierarchy of catchment geosystems (river basins) which corresponds to the formalized Horton-Strahler-Tokunaga schemes. The drastic change in the phenomenological coefficients is the basis for the classification of NTC according to the principle of homogeneity (typological approach in accordance with N. A. Solntsev's theory (Solntsev 1948)). Considering the spatial distribution of plants and animals in both geopotential and other physical fields - insolation, chemical, thermodynamic etc. allows obtaining the hierarchy of ecosystems (biogeocenotic systems) and their spatial distribution. Such approaches to classification of geosystems are mutually complementary, and should not be contrasted. For example, V.N. Solntsev (Solntsev 1997, 2006) considers three mechanisms of landscape structuring - geostationary, geocirculatory and biocirculatory, which could operate individually or simultaneously.

The objective identification of landscape cover units is essential for the planning of sustainable nature management. The development of territorial planning assumes the need to use different methods for selection of spatial units, depending on the objectives of environmental management. For example, the maps showing the structure of biocentricity are necessary to embed nature reserves; those representing positional-dynamic and morphological structures are essential for industrial facilities (Pozachenyuk 2006); for agroforestry purposes the catenary differentiation should be considered, and different types of units, such as urochishche, mestnost etc., and catchments should be used (Rulev 2008). In landscape-agroecological planning the preference is given to genetics-morphological approach, the principal units being the urochishche and a group of urochishches (Orlova 2014). An example of the use of landscape planning is water protection zoning (Landscape Planning… 2002). Complicated environmental situations and a variety of conflicts between land and water users are characteristic of the water protection zones. On the other hand, the most complex landscape-hydrological systems (LHS) are presented in the areas adjacent to water bodies. The combination of landscape and hydrological indicators at the basin level is used to refine the calculation of hydrological parameters or to assess the distribution of LHS (Antipov and Fedorov 2000). The concept of LHS as applied to landscape hydrology is a set of natural-territorial complexes (NTC) similar in runoff formation conditions. The NTC only indirectly characterizes the catchment area with similar capacitive features; therefore the experimental observations in each small river basin are necessary for an accurate estimate of the LHS area. However, in most cases LHS are formed from a set of NTC. The calculation accuracy (*ceteris paribus*) could be improved by increasing the number of taxonomic units. Moreover, the landscape typological approach is the most appropriate for the large-scale work (Tkachev and Bulatov 2002).

Modern landscape ecology is based on the patch mosaic paradigm, in which landscapes are conceptualized and analyzed as mosaics of discrete patches (Forman 2006; Turner et al. 2005).The strength of the patch mosaic model lies in its conceptual simplicity and appeal to human intuition. In addition, the patch mosaic model is consistent with well-developed and widely understood quantitative statistic techniques designed for discrete data (e.g., analysis of variance). Developing this rather limited approach to environmental considerations McGarigal and Cushman (McGarigal and Cushman 2005) introduced the «landscape gradient» model, as a general conceptual model of landscape structure based on continuous spatial heterogeneity. Based on the continuous characteristics of the earth's surface, obtained from DEM (slope, topographic wetness index, topographic position index, normalized difference vegetation index NDVI etc.), the «landscape gradient» model is constructed using the statistical characteristics of patch mosaic (patch density, largest patch index, edge density, mean patch area, area-weighted mean patch area, coefficient of variation in patch area, mean patch shape index, etc.). Statistical characteristics of patches are called landscape metrics (McGarigal et al. 2009). However, these are metrics of just the sizes and shapes of patches in mosaics, and not of complex geosystems, such as landscapes of any dimension and hierarchy. The structure (pattern) is understood, first of all, as a combination of interacting spatial elements with their area, configuration, orientation, neighborhood, connectedness or fragmentation (Turner and Gardner 2015), i.e. close to the concept of "landscape pattern" (Viktorov 1986). The structure is interpreted as an indicator (on the one hand) and a condition (on the other hand) of radial and lateral processes. This interpretation turned out to be especially productive for regions highly transformed by anthropogenic activity, where the zonal landscape was preserved in the form of a few "islands". From the point of view of modeling the landscape structure based on structure-forming processes these methodological approaches are not entirely correct. Amount of empirical data, composite indices with a fuzzy (intuitive) physical meaning cannot be used directly as parameters of the equations of mathematical physics. As a result, a gap arises between landscape-ecological and physical-mathematical models, and the empirical and semi-empirical parameters need to be introduced into rigorous descriptions of the processes of transfer of matter and energy to overcome it. For example, GIS SAGA software (Olaya 2004) is supplemented by the TOPMODEL hydrological unit (Beven 2012), which describes the migration of moisture based on the Darcy equation, with a significant number of empirical parameters that are difficult to determine, so parameterization, approximation, and similar fitting methods are required.

Planning decisions based on the landscape-ecological analysis could become more reasonable if the following problems are solved (Landscape Planning… 2002):

• Identification of quantitative indicators describing both the structure and the functioning and development of a landscape. We need objective indicators which are relatively easy to calculate.
• Development of classifications of landscapes according to their sustainability, vulnerability, suitability and capacity for particular types of environmental management.
• Identification of quantitative characteristics of the landscape self-organization, or at least qualitative description of the regional developments of this process.
• Search for relationships between the spatial structures of natural and socio-economic systems.
• Determination of minimum natural ranges within which ecological stabilization of cultural landscapes could be implemented.
• Development of regional norms or recommendations for planning spatial relationships between the main landscape elements (by area and configuration).

These tasks are currently relevant. In our work, possible ways of implementing some of these tasks are given.

The aim of the work is to justify the choice of the least number of objective parameters characterizing the landscape structure, which is interpreted in the classical definition of the Moscow State University School of Landscape Sciences. In fact, this is a synergistic task of determining the main parameters of structure-forming processes. The method proposed in the article allows application of numerical modeling to describe the landscape polystructure by the parameters of major continuous geophysical fields.

COMMON PRINCIPLES OF THE LANDSCAPE STRUCTURE MODELS

The elaboration of any physical-mathematical model begins with basic axioms and postulates. The principal point is the identification of elementary material objects (particles, points) forming the system and the assignment of independent variables and functions of the system's states. Further, it is necessary to adopt a number of binding postulates so that it is possible to apply particular physical laws. It is essential that physical laws and their parameters be applied relevant to the description of the structure-forming landscape processes (Sysuev 2014).

To describe geosystems, it is first of all necessary to substantiate the potentials of the main geophysical fields that determine the structure-forming processes, and then to formalize the description of elementary geosystems and hierarchical invariants of geosystems. The quantitative values of spatially distributed physical parameters of the state of landscapes could be obtained: 1) from digital elevation models (DEM) – morphometric parameters describing the gradients of the gravity and insolation fields; 2) from digital remote sensing data – parameters of the Earth's surface cover; 3) from field and laboratory measurements, and 4) during special experiments.

The space of geographical coordinates is provided by the construction of a digital elevation model (DEM). Pixels of the 3D DEM are elementary material points (similar to the material points in theoretical mechanics), from which the NTC structure is synthesized using the formalized procedures. DEMs are constructed to achieve the maximum resolution of a particular hierarchical level of geosystems. For example, if a regular-grid DEM based on the contours from a detailed topographic map M 1:10 000 is constructed, the pixel size could be 10x10 m. However, the pixel size depends on the resolution of aerial photo or satellite image as well. So, the resolution of Landsat images (30x30 m) allows us to identify the NTC of just urochishche level.

Differentiation of geographical space could be realized using various mathematical methods (cluster analysis, neural networks, etc.). However, we need numerical parameters of the state of elementary material objects (pixels) to distinguish uniform areas. Theoretical description of the geostructure, i.e. stationary (for a certain time interval) state of a dynamic geosystem, begins with identifying morphometric parameters (MP) which describe the force fields that determine the main structure-forming processes. Morphometric formalization of the Earth's surface in the gravitational field was systematized in (Shary 1995). Logically meaningful association of morphometric parameters includes three groups characterizing: 1) the distribution of solar energy – the dose of direct solar radiation (daily, annual), aspect and illumination of slopes; 2) the distribution and accumulation of water under the influence of gravity – the specific catchment area and the specific dispersive area, depth of *B*-depressions and the height of *B*-hills, the slope gradient; 3) the mechanisms of matter redistribution under the influence of gravity – horizontal, vertical and mean curvature, slope gradient, height (Sysuev 2014). It should be noted that the minimum number of simple (non-composite) state parameters is selected that independently describe the gradients of geophysical fields generating the main structure-forming processes. Thus, the selection (classification) of geosystems is not carried out according to relief elements, or elementary locations, or patches, or other spatial units, but directly by the parameters of geophysical fields and structure-forming processes.

The physical meaning of the morphometric parameters is quite clear. For example, the dose of radiation characterizes the potential input of direct insolation. Slope aspect and gradient are the components of the geopotential vector gradient. Horizontal curvature is responsible for the divergence of flow lines. Vertical curvature is the derivative of the steepness factor and characterizes the slope convexity/concavity. These parameters are directly included in the equations of mathematical physics. Specific catchment area shows the area from which suspended and dissolved substances could be transported to a surface element. It is a substance balance parameter (included in conservation laws), and a component of a number of related indices (water flow capacity index, erosion index, etc.).

The state of the Earth's surface (vegetation cover, snow cover, soil cover, etc.) is detected from the digital data of space spectral image and from the related indices (e.g., the normalized difference vegetation index, snow index, humidity index – NDVI, NDSI, NDWI, etc.). The most important sources of data are field studies, which also allow verifying the interpretation of the state of the covers. In addition to the traditional complex methods of landscape science, it is necessary to use automated complexes for recording geophysical parameters of the lowest atmospheric layers, natural waters and soils. Methods of applied geophysics, based on measuring the spatial distribution of gravitational, electromagnetic, and other parameters, are very promising as well.

The parameters for describing the structure are chosen in accordance with the classical approaches of landscape studies. All formal algorithms for selecting the smallest and the higher order units of relief surface based on the parameters of the gravity and insolation fields acquires a fundamental geophysical meaning. In this case, the concept of polystructural landscape organization is absolutely logical: by selecting different structure-forming processes and physical parameters, different classifications of landscapes could be elaborated. Let us demonstrate the approach through particular cases below.

RESULTS AND DISCUSSION

Typological model of the landscape structure

Typological approach allows obtaining a hierarchy of classical NTC (*facies - urochishche - mestnost - landscape*) in accordance with N.A. Solntsev's theory (Solntsev 1948). The parameters are chosen in accordance with the widely-known definitions. For example, "An elementary NTC - *facies* … is confined to one element of mesorelief; this territory is homogeneous in terms of its three principal characteristics: the lithological composition of the rocks, the slope aspect and gradient. In this case, the total solar radiation and atmospheric precipitation coming to the surface are the same within any part of it. Therefore, one microclimate and one water regime are formed, … one biogeocenosis, one soil unit and a uniform complex of soil mesofauna" (Dyakonov and Puzachenko 2004). As it follows from the definition, elementary NTCs could be selected by the parameters of solar radiation and water distribution over the surface; more precisely, by the distribution of the gradients of insolation and gravity fields. Thus, the classical definition already requires the description of the NTCs differentiation using the theory of field and the morphometry of the Earth's surface. The classification results essentially depend on weight values and number of parameters. By changing the latter, it is possible to optimize the classification of landform elements according to a known landscape structure. On the other hand, the change in the set of parameters and their numerical values makes it possible to model landscape structure changes under the influence of climate change, neotectonic events, etc. A rigorous landscape approach is needed for such modeling that allows identification of the main factors of differentiation and exclusion of derivative or dependent variables. Automatically obtained classes of landscape cover require identification and verification of their physical content.

The investigated territory of the Valdai National Park is located in the central part of the Valdai Upland, which belongs to the end moraine belt of the last Valdai (Würm) Ice Age in the northwestern East-European Plain. The loamy moraine deposits with residual carbonates reach a thickness of 25 m in the ridges of the Crestets end moraine belt and overlie the glacio-fluvial sands of preceding stages. Locally, the moraine is covered by kame silty-sandy loam sediments. The fluvioglacial plains are covered with peat bog sediments, the massifs of which are separated by sandy eskers. Peatland systems are connected by streams and the Loninka and the Chernushka rivers. Such a variety of landforms and sediments causes the high degree of biological and landscape diversity within the study area.

The digital elevation model (DEM) was constructed on horizontally a detailed topographic map with a scale of 1:10 000 using the regular grid method with 28 × 28 m pixel size attached to the Landsat-7 DTM+. The pixel size makes it possible to reliably distinguish NTCs of the locality (mestnost) rank in the study area with dimensions of about 10x10 km. Based on the values of the height and size of pixels (vertical and horizontal steps), the main geomorphometric parameters were calculated, i.e. slope, dose of direct solar radiation, aspect and illumination of slopes, specific watershed area and dispersive area, B-depression depth, B-hill height, horizontal and vertical curvature. GIS ECO (P. Shary), GIS FractDim (G. Aleschenko, Yu. Puzachenko), GIS DiGem (O. Conrad), and GIS SAGA (V. Olaya) were used for calculations. The best result of building a smoothed DEM and morphometric parameters of the studied territory was shown by GIS ECO.

Next, a matrix (database) was built: the rows correspond to the relief surface elements (DEM pixels), and the columns to the parameters (MP) describing the state of an element (height, geomorphometric parameters, as well as digital values of the brightness of the Landsat-7 DTM+ channels and NDVI). The parameters describing the same surface element have different physical meanings and are not comparable in dimensions and size. They are therefore normalized and reduced to a standard form. The resulting matrix is ready for any algebraic operations.

The row vectors of the data matrix characterize a multitude of DEM pixels. Geometrically, the closer two such vectors in the parameter space, the less different the parameter values for both objects. This suggests that the closer two vectors in the parameter space, the more "similar" and less distinguishable the corresponding objects in many of their other properties, not only for those included in the data matrix. Therefore, if it is possible to distinguish a geometrically sufficiently isolated "group" of vectors close to each other within the set of all object vectors, then a class of objects with similar internal properties could be identified. The Euclidean metric of the distance between the corresponding object vectors was used as a measure of the proximity of two objects. If the proximity measure function is selected and calculated, the matrix of communication between objects is thus constructed, and the task of automatic classification of objects is reduced to the problem of diagonalizing such a communication matrix. The automatic classification can be understood as a geometric task of distinguishing "dense" concentration of points in a certain space. Such geometric approach allows elaborating the methods of solving the task of automated classification of spatially distributed objects, such as elements of relief surface in the DEM, remote sensing data, etc. The number of objects in the geographic data matrix is very large and could reach dozens and hundreds of thousands. Efficient recurrence algorithms could be applied in these cases which provide that the calculations are performed with only one successive object (or with one row of the corresponding communication matrix) at each step. In our work, the FractDim software was used to classify the relief according to the MP matrix.

The classes of vegetation cover automatically decoded from space spectral image were verified on the basis of field data (Akbari et al. 2006). Some results are shown in Fig. 1. The investigated transect (integrating a series of analogous transects) was about 5 km long; leveling was performed at 5 m interval; integrated descriptions were made at sample plots 20 m apart. Along with complex descriptions, a complete forest inventory was carried out at

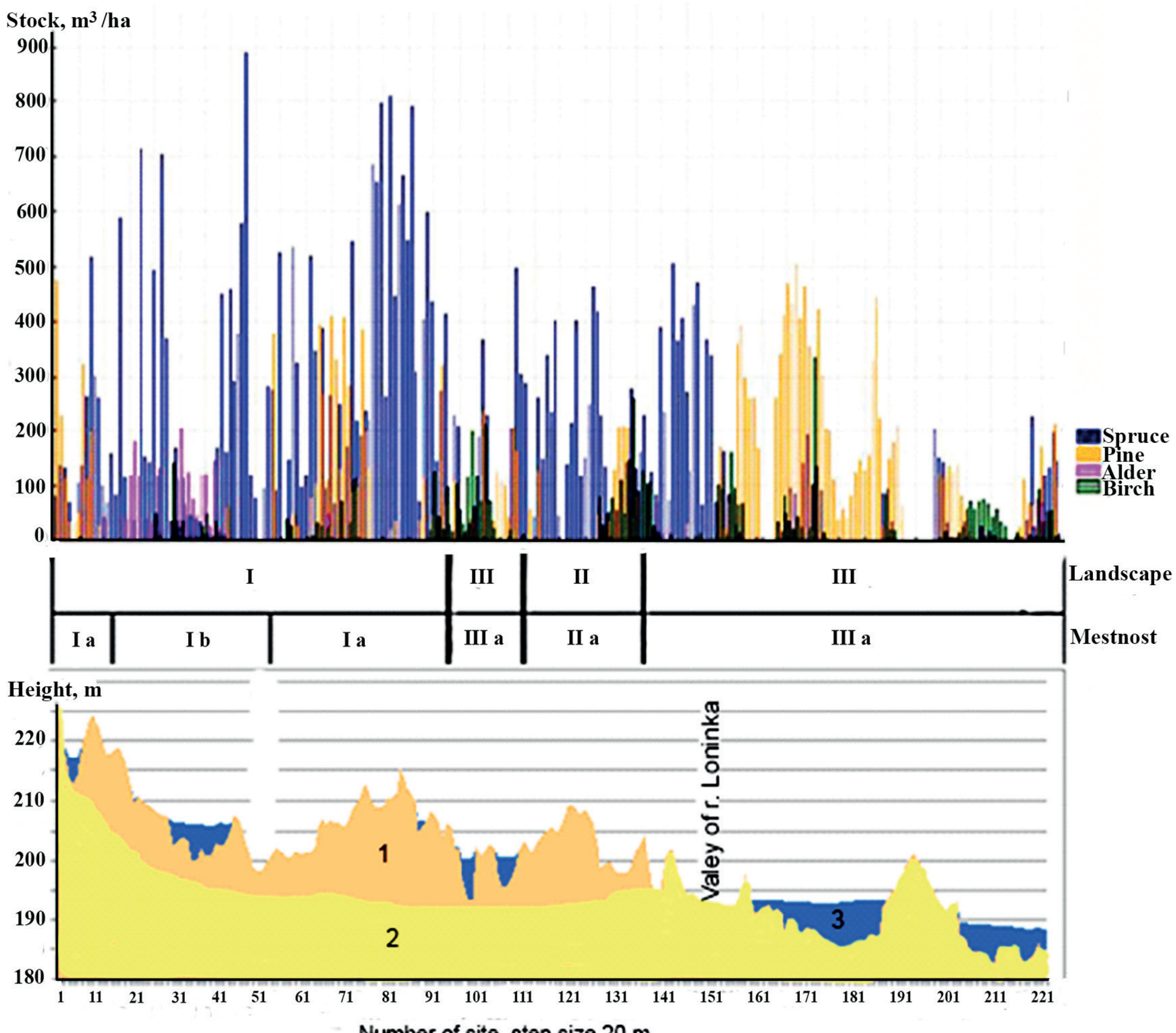


Fig. 1. Distribution of the stand volume (m³/ha) along the transects crossing typical landscapes of the Valdai National Park: I - landscape of the ridge-hilly moraine-kame plain on carbonate moraine loams often covered with silt-sandy loam deposits of low thickness; II - landscape of ridge-hilly kames-eskers plain on sandy-loam sediments; III - landscape of outwash glacio-fluvial plain with sand ridges. The alphabetic indices characterize the localities ("mestnost"). In the lower part – elevations a.s.l. (m) and a schematic lithological section: 1 - moraine deposits, locally overlaid by kames, 2 – glacio-fluvial sands, 3 - peat deposits

239 plots of 20x20 m area. The forest inventory included strip enumeration of trees with measurement of standard parameters of each tree and the sample plot (composition, stratum, height, diameter, age, crown attachment height, stocking, canopy density, underwood, advance growth, grass-shrub cover, type of soil, type of station, etc.)

The map of automatically decoded classes (Fig. 2, A) compiled from a priori geophysical data (annual dose of direct solar radiation, slope gradient, numerical data of Landsat 7 spectral channels and the NDVI) was refined on the basis of field data geo-referenced to space image. The map of vegetation cover (Fig.2, B) shows classes according to the field-based verification. Interpolation of continuous forest inventory data on the studied territory was performed using the discriminatory methods.

Correlation of the compiled stand map and the calculated type of site conditions with scale 1:10000 forest compartment map for this territory showed that the simulation results are significantly more detailed compared to the standard forest inventory data.

The scheme of the geosystems structure (NTC at the *urochishe* level) is shown in Fig. 3. We used successive dichotomous grouping of landform elements (DEM pixels) based on the parameters of geophysical fields and the state of landscape cover. Independent morphometric parameters were chosen based on preliminary analysis (digital modelling): the annual dose of direct solar radiation, elevation, slope gradient, horizontal and vertical curvature, specific catchment area, as well as numerical data of Landsat 7 spectral channels and the NDVI index.

Classification results depend significantly on the number of parameters and their weight values. By changing them, it is possible to optimize the classification of relief elements according to the known (assumed) landscape structure. The selection of parameters was carried out in accordance with the classical landscape science defini-

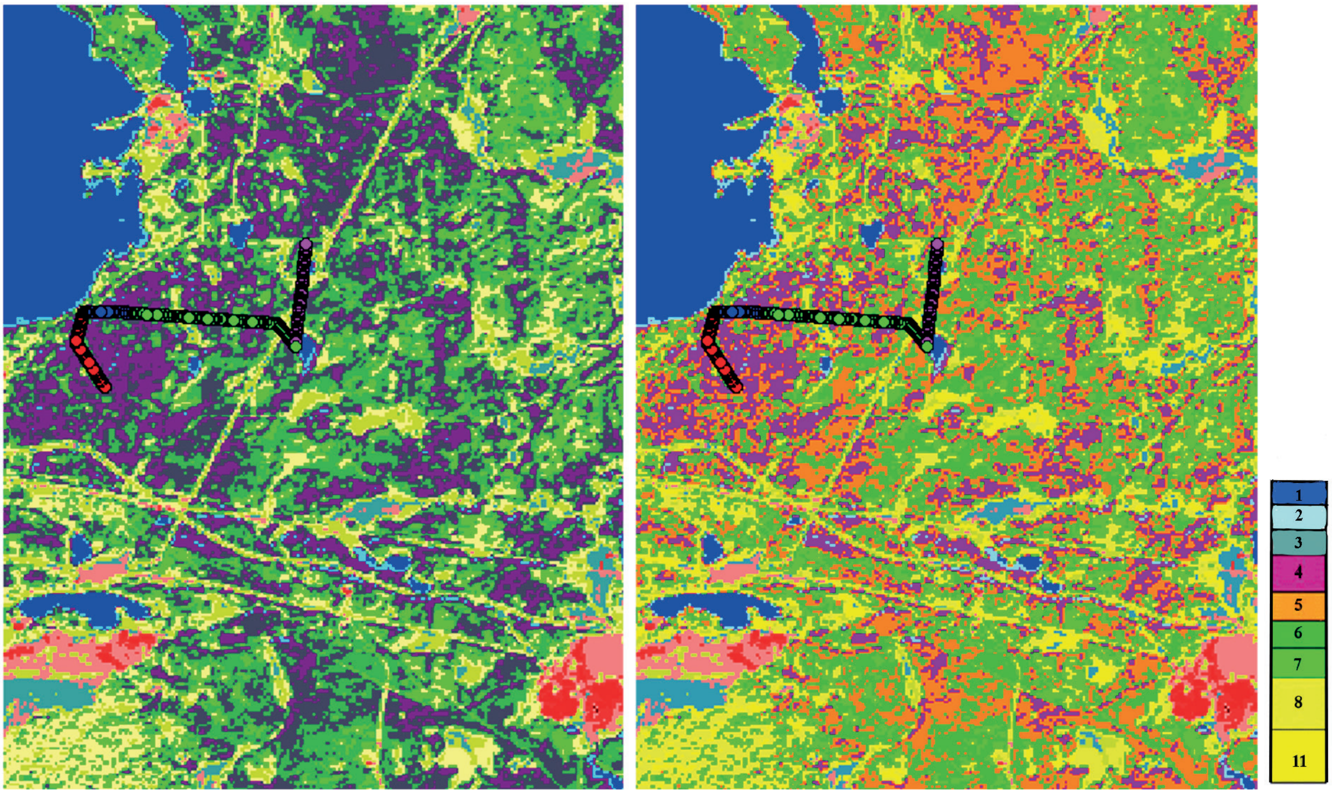


Fig. 2. Identification of the physical content of landscape cover classes in the Valdai National Park according to a priori information (A) and the vegetation cover identified by continuous forest inventory along the transect (B). Legend see in Table 1. The dots show the transect location

Table 1. Identification of decoded classes basing on the results of the continuous forest inventory along the transect

Class (Fig. 2, B)	Number of sites	Stand
1	14	Open water
2	40	Water plants
3	31	Meadow swamp
4	66	Closed spruce forests
5	59	Spruce-pine and pine forests
6	50	Mixed and spruce-pine forests
7	25	Boggy pine forests
8	22	Small-leaved and boggy pine forests
11	9	Bogs with stunted pine trees

tions and a substantial number of preliminary numerical experiments (Sysuev 2003, 2014). For example, if the heat (energy) supply of the territory is modeled with the same weight coefficients for all insolation parameters, the obtained groups do not satisfy the landscape structure revealed in the field studies. The problem is that at first and subsequent levels of the dichotomic classification the classes of relief surface are distinguished, first, against the exposure of slopes, which is not true for a significant part of the territory, which is a swampy landscape of the fluvioglacial outwash plain. On the other hand, the cooling properties of swampy landscape are not taken into account, and even vast massifs of upland bogs with lakes are not identified. To obtain more correct classification, the weight coefficient of the slope parameter was increased, which improved the classification of the relief surface. At each stage, the verification of the classification of relief elements with the selected values of weight coefficients was checked by the method of discriminative analysis. At all levels, according to the values of F-criterion, the distinguished classes differ statistically reliable. Thus, the leading role of waterlogging of the territory, obvious to landscape scholars, could be numerically expressed by the value of weight coefficients of slope, i.e. the parameter of gravity field gradient. Similarly, the significance of other geomorphometric parameters was substantiated.

Because of numerical experiments, the level of required numerical classification was also objectively revealed. The 5-level dichotomy adopted at the beginning with identification of 32 classes turned out to be excessive. The size of urochishche (simple and complex) corresponding to the main mesostructures of the relief of the studied territory is the most adequately displayed at the 4th level of classification. Moreover, as can be seen in Fig. 3, a good number of identified classes are automatically combined into larger groups according to close-valued colors of a palette.

The need for an objective justification of the significance of the state parameters and the level of numerical classification dichotomy is emphasized by the crucial role of the landscape approach, which makes it possible to single out the role of individual factors (structure-forming processes) of NTC differentiation in specific geographical conditions.

Finally the distinguished classes were characterized using the parameters obtained during field study of experimental landscape transects. The characteristics of the grass-shrub and soil cover, and the lithological structure were extrapolated by discriminative methods.

In addition, an independent team of researchers created a landscape map basing on the classical method. The comparison showed that the landscape map resulting from numerical experiments is sufficiently accurate in reproducing the boundaries of the NTC independently obtained by the classical field methods (Sysuev and Solntsev 2006).

Revealing spatial structures in such a way is a process of synthesis, since the material points (pixels) are integrated into elementary natural-territorial complexes (at a given hierarchical level) according to the selected parameters of main geophysical fields and the state of covers.

Functional model of geosystem structure

The functioning of low-order geosystems is largely determined by water flows. Hence, the classification is aimed at the construction of a hierarchy of catchment geosystems according to the morphometric parameters describing the

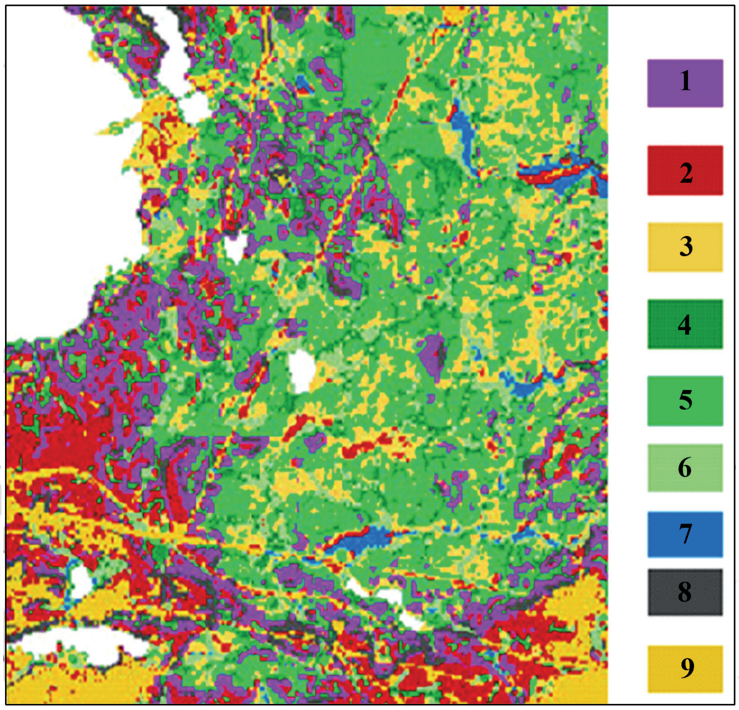


Fig. 3. The structure of the NTC based on the relief classification using the geophysical field gradients and the space image data Landsat-7.

1 - moraine ridges and kame hills with loamy Umbric Albeluvisols (sod-podzolic soils) under Piceetum oxalidosum forests; 2 - summits of kame hills and ridges with sandy Umbric Podzols under Pinetum cladinosum, Pinetum cladinoso-hylocomiosum and Pinetum herbosum forests; 3 - foot slopes of the hills and flat concave hollows with Umbri-Endogleyic Albeluvisols and Umbric Albeluvisols under mixed forests; 4 - river and lake terraces with Endogleic Umbrisols and Distri-Fibric Histosols under spruce forests and mixed forests; 5 - dune ridges and sandy hills with Umbric Albeluvisols under pine forests; 6 - flat and convex upland bogs with deep Histosols under sparse pine forests; 7 - river floodplains with Endogleic Umbrisols under flooded meadows; 8 - steep slopes of hills with Umbrisols under coniferous forests; 9 –anthropogenically modified and anthropogenic landscapes (roads, power lines, quarries, farmland, nurseries and residential areas)

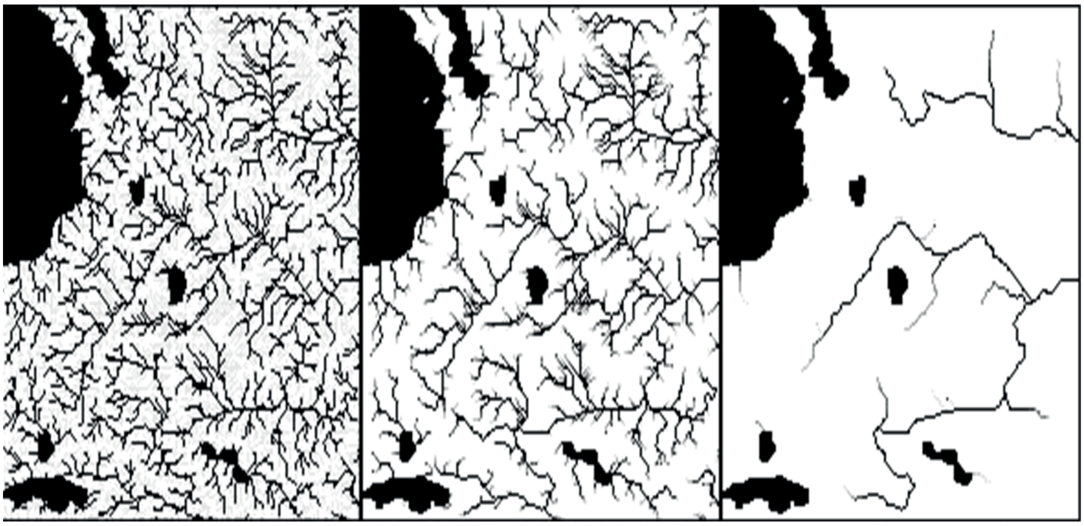


Fig. 4. Automated procedure for cutting off the drainage channels shorter than the critical length (catchment basin of the Loninka River, the Valdai National Park) in MapWindow GIS using the TauDem module (Tarboton et al. 1991)

redistribution of water in the gravity field. These parameters, namely slope gradient, specific catchment area, horizontal and vertical curvatures, determine the boundaries of the divergence zones and the convergence of streamlines. The hierarchy of catchment geosystems is determined in accordance with the Horton-Strahler-Tokunaga scheme (Tarboton et al. 1991; Dodds and Rothman 1999).

The automated algorithm for identifying drainage channels on the basis of the GIS raster layers involves three main steps. First, we use the digital map of an above-mentioned parameter to select cells with values exceeding a predetermined threshold that are considered to be potential source points. At the second step, channels from the given sources are drawn, and the sources which have transit flow from higher elevations are removed. At the third step, channels smaller than a certain minimum length are cut off (Fig. 4).

The process can be easily adjusted by changing the limit values of the catchment area and the minimum length of the drainage channels. The resulting array of morphometric characteristics with geographic reference to watersheds of various orders is a characteristic of landscape-hydrological systems (LGS), or catchment geosystems (Sysuev 2014).

Application of typological approach to obtain landscape characteristics of catchment geosystems could be demonstrated by the example of the Upper Mezha River basin (Central Forest Reserve, Tver Region). The southern taiga spruce forests with broad-leaved species, shrubs and nemoral herbs dominate this southwestern part of the Valdai Upland (the East European Plain). The territory is a combination of flat ridges and inter-ridge depressions, a weakly dissected plain composed of glacial and fluvioglacial deposits. Modern drainage streams are mostly temporary, with poorly developed alluvial relief and poorly pronounced valley forms. They occupy the drained inter-ridge depressions, which are ancient valleys of the glacial meltwater runoff. The inter-ridges depressions lacking the active runoff are occupied by upland and transitional bogs.

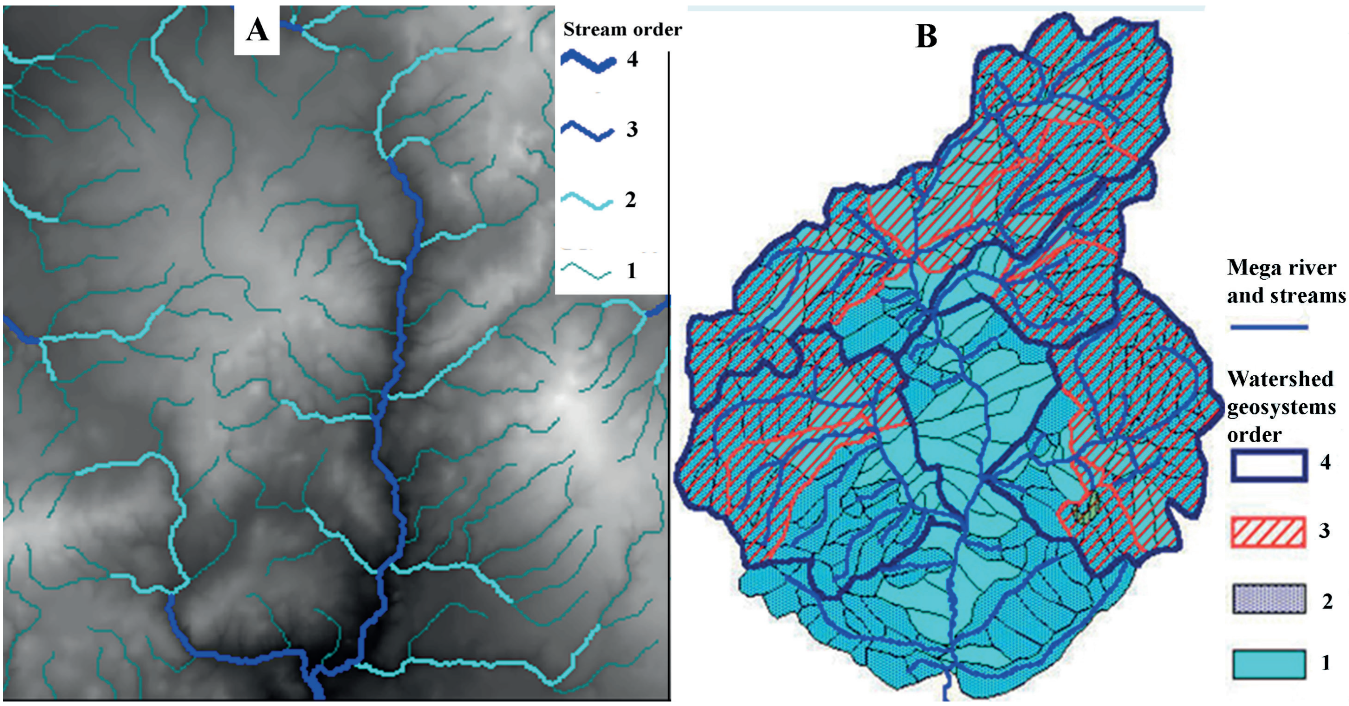


Fig. 5. The structure of watercourses on a digital elevation model (A) and catchment geosystems (B) of the Upper Mezha River (Central Forest Reserve, Tver Region). The higher is the territory, the lighter is the DEM tone; a specific catchment area (SCA) was selected as a critical value for calculating the order of watercourses (A). Watersheds as delineated by zero SCA values (B)

The upper reaches of the Mezha River are located in the study area. The river channel is winding within the valley with a wide floodplain (50-100 m) winding channel.

The methods for the DEM construction and calculation of parameters are similar to those described above. The algorithm for constructing a drainage network in SAGA numbers the selected segments of watercourses in the order of R. Shreve. To obtain data on the orders of catchments according to the Straler classification, the channels of the first, second, and subsequent orders were sequentially cut off, until maps of watercourses with a breakdown in order were received (Fig. 5 A). An analysis of the use of the method of the specific catchment area critical values in different landscape showed that as the age of the relief increases (in the series: secondary glacial plain → periglacial plain of the marginal zone of the Valdai IceAge → end moraine zone of the Valdai Ice Age) the average area of watercourse formation noticeably reduced.

The geosystem order is the same that the order of a watercourse (Fig. 5B). A special category - zero order geosystems – was introduced for lower rank complexes that do not have a pronounced drainage watercourse. About 400 such geosystems have been identified within the territory under study. The maximum fourth order geosystem is the catchment of the Mezha River in the vicinity of the Fedorovskoye village (Table 2).

The dependence of the average values of a catchment area Y on its order *X* is described by the equation $Y=b0*(X+1)^{b1}$, where the parameter values are b_0=0.419665, b_1=2.526742, and the model reliability is R = 0.99977. Areas of zero order geosystems have the approximately lognormal distribution.

The qualitative and quantitative characteristics of runoff for any watercourse depend on the physical and geographical features of its catchment. To identify these dependencies, a map of the landscape structure of catchment geosystems within the river basin was compiled (Fig. 6).

The methodology for compiling the map of the landscape structure of geosystems is as follows. Maps of relief structure and vegetation cover were created for the study area by means of classification analysis of the digital relief model and satellite imagery of Landsat 7. The relief structure map was created using "with training" classification according to the digital relief model. Using the K-means method, eleven classes were distinguished, which reflect the differentiation of the territory into flat sections and slopes with convex (spurs) and concave (hollows) sections. A map of the structure of vegetation cover was created through the interpretation of the Landsat 7 satellite imagery. Eleven classes were also identified, and two of them were then combined to reflect anthropogenically modified territories (village, roads, fields and hayfields, deposits etc.). At the next stage the number of pixels corresponding to a particular class of vegetation and topography was calculated for each geosystem. As a result, the areas occupied by the main classes of relief and vegetation within each geosystem were estimated. These data were summarized in a single table. Then, the percentage of the area occupied by each class was calculated for each geosystem. To simplify the map compilation, classes with the area exceeding 10% of a geosystem total area were left for further consideration. A generalized matrix was then obtained and the map of the landscape structure of catchment geosystems of various orders was compiled (Fig. 6). Thus, the landscape structure is described within physically determined watersheds boundaries.

The technique simplifies the application of landscape characteristics and differs from methods used for identifying the landscape-hydrological systems (LGS). According to Antipov and Fedorov (2000), the area of LGS varies from year to year, from season to season, and from day to day. Therefore, the selection of NTCs that have similar state in relation to runoff in particular time period (for example, floods) is not simple.

In mathematical modeling of surface runoff, there is a clearer concept of an "active runoff area" that changes during the process (Troendle 1985). We suggested calculating the active runoff area by the value of the territory elevation excess above the mouths of the particular order watercourses for a certain period of time (Fig. 7).

The relationship of functioning and the structure of geosystems could be analyzed through the gradual decline in flow hydrographs (Fig. 7, Table 3). During spring floods, and after heavy prolonged rains, practically all first order catchments function, i.e. a characteristic surface runoff along the hollow-like depressions is observed. These depressions are usually not well pronounced in the relief and their depth could be only dozens of centimeters. However they are well marked by moist grasses and humus-gley soils. So for the beginning of June (the final stage of 2001 spring flood), the active runoff area for the first-order geosystems with an excess above the watercourses mouths of <1.0 m was 0.2 to 1.7 ha. As a rule, such areas accounted for about 60-70% of the total area of a geosystem (Fig. 7). For some geosystems, particularly those with flat surface, the active areas could not estimated. During the low water period associated with the low drainage from soil and ground, the flow continued only in the largest streams and the Mezha River itself.

Analysis of the catchments parameters and hydrological measurements showed a close relationship between the structure and functioning of geosystems. This provides opportunity to calculate the water flow discharge basing exclusively on the geosystems structure and precipitation data.

Hydrological functioning and water protection zoning of geosystems

The calculation of surface runoff from a priori topographic DEM data can be performed in different GIS supporting hydrological procedures. For example, in order to calculate the water flow in SAGA (Olaya 2004), a sufficiently

Table 2. The main statistical parameters of catchment geosystems

Order of catchment	Number of catchments	Average area, km^2	Minimum area, km^2	Maximum area, km^2	Dispersion	Average discharge, June, l/s
0	393	0.093	0.001	0.363	0.005	0.1
1	52	0.478	0.120	1.583	0.123	6
2	9	2.391	0.795	4.679	2.046	25
3	4	6.743	2.720	10.940	11.423	60
4	1	36.645	-	-	-	120

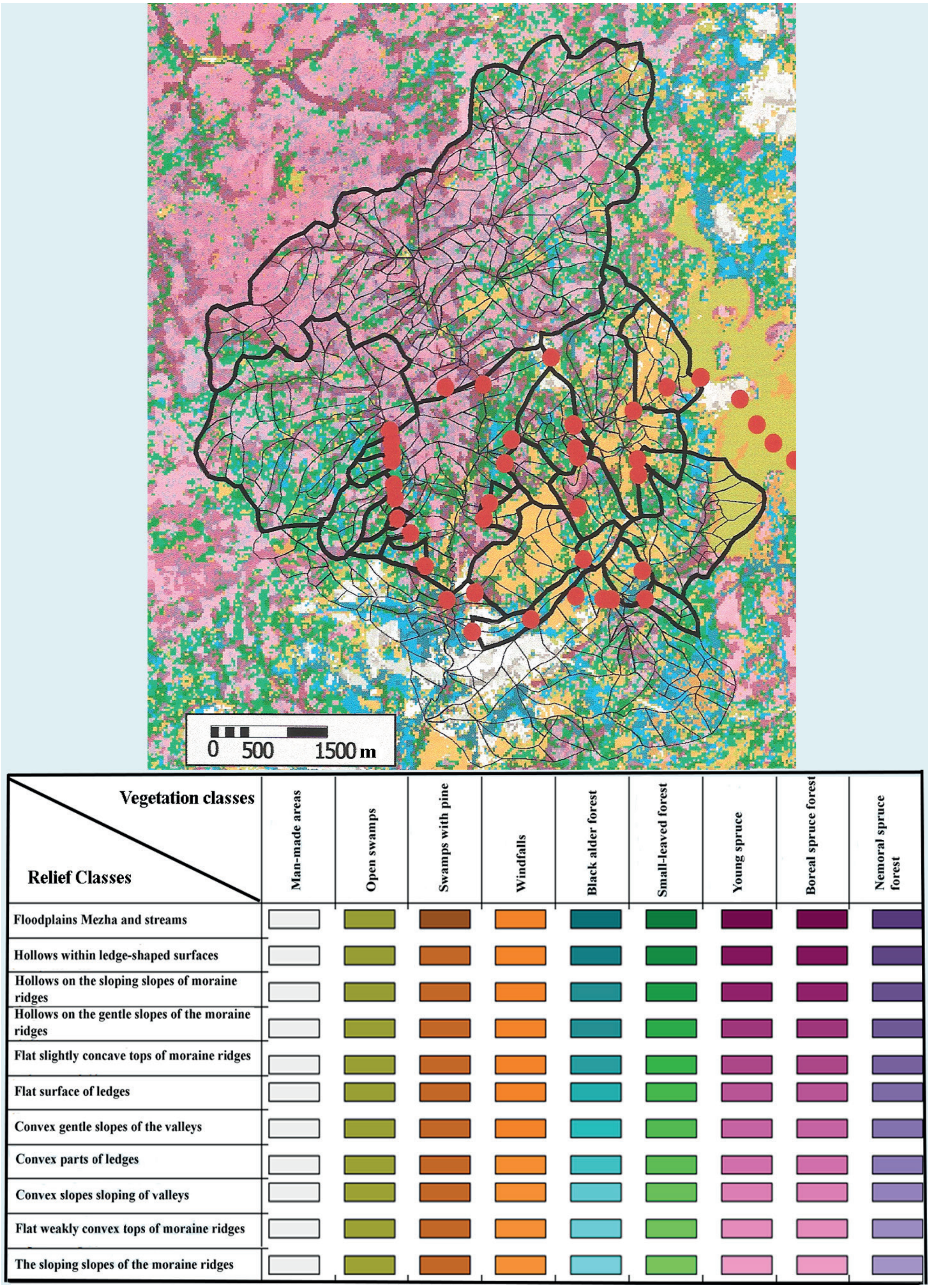


Fig. 6. Map of the landscape structure of catchment geosystems. The table legend is below the map. The color indicates the type of vegetation cover, the intensity of color shows the moisture gradient distribution. Dots indicate the sites of hydrological and hydrochemical testing

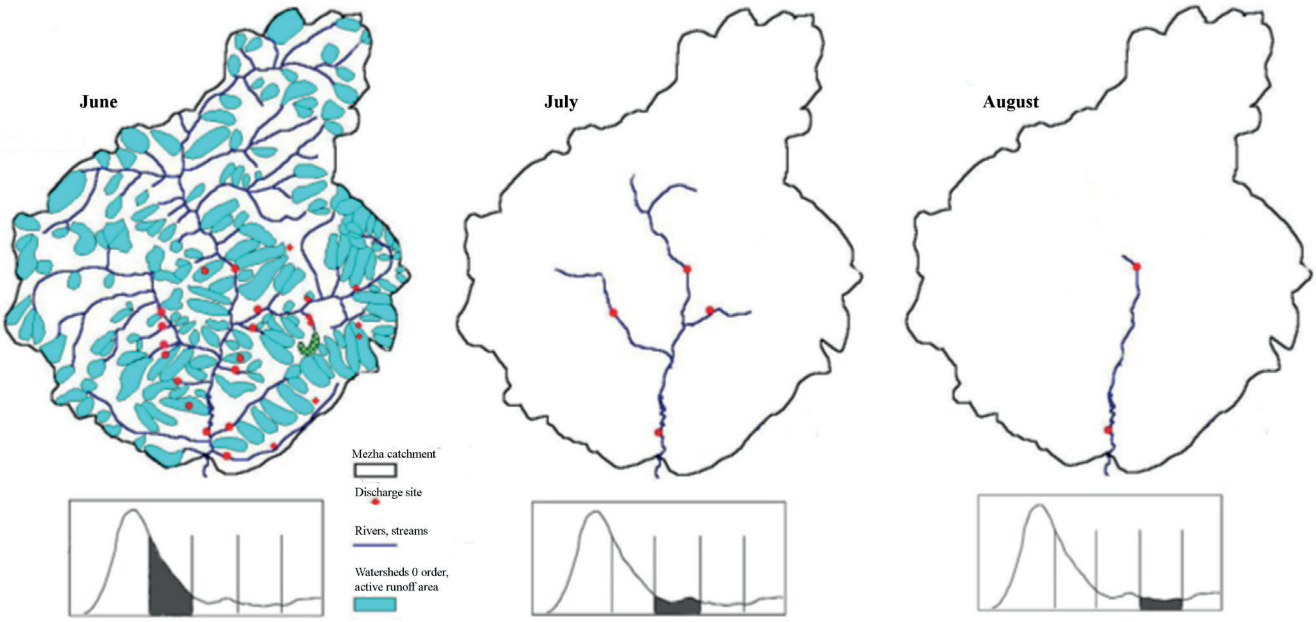


Fig. 7. The discharge dynamics for the measuring sites (red points) in the Upper Mezha River basin (the Tver region) during 2001 summer, and the decrease in active runoff area as a result of the gradual depletion of spring flood water. Schematic hydrographs of runoff are shown below in small graphs; the black area is the period of discharge averaging for Table 3

Table 3. Average water discharges and mean values of some hydrochemical indicators for the sections of different order catchments in the Mezha River basin (the Tver region)

Order catchment	Number of measuring sites	Discharge, l/s	pH	Electrical conductivity, μC/cm.
June				
0	5,9,12,13,16,17	0.2	5.71	35.4
1	7,11,15,23,24,25,26,27,28,36,37	5.8	6.61	38.9
2	10	56	5.79	24.3
3	18,19,22	33	6.58	36.5
4	29	117.5	6.59	41.7
July				
0	No runoff	0	-	-
1	No runoff	0	-	-
2	No runoff	0	-	-
3	18,19,22	5.3	7.54	42.6
4	29	15	7.7	42.9
August				
0	No runoff	0	-	-
1	No runoff	0	-	-
2	No runoff	0	-	-
3	19	3.5	7.15	42.2
4	29	5	7.01	42.7

large number of individual catchment parameters, such as the Shezi-Manning coefficient of surface roughness ("Manning's n" - MN) and the coefficient of soil influence on the intensity of surface runoff ("Curve number" - CN), are required in addition to general parameters (elevation, slope gradient, specific catchment area, etc.).

The values of parameters must be assigned to each pixel of the model, which is objectively possible only with reliance upon the information on the landscape structure. In Fig.8, the numerical values of MN, taken from the standard Chow tables (Chow 1959), are depicted in accordance with the typological structure of landscape (see Fig. 3). These data are the basis for setting the spatially distributed parameters of the hydrological model aimed at calculating the runoff rates within the Loninka River basin.

The average precipitation intensity was assumed 0.0, 0.66, 10.0, or 100.0 mm/hour in numerical experiments for the calculation of runoff. In addition, Channel Site Slope (CSS) parameters, runoff characteristics (slope surface, Mixed Flow Threshold – MFT, and Channel Definition

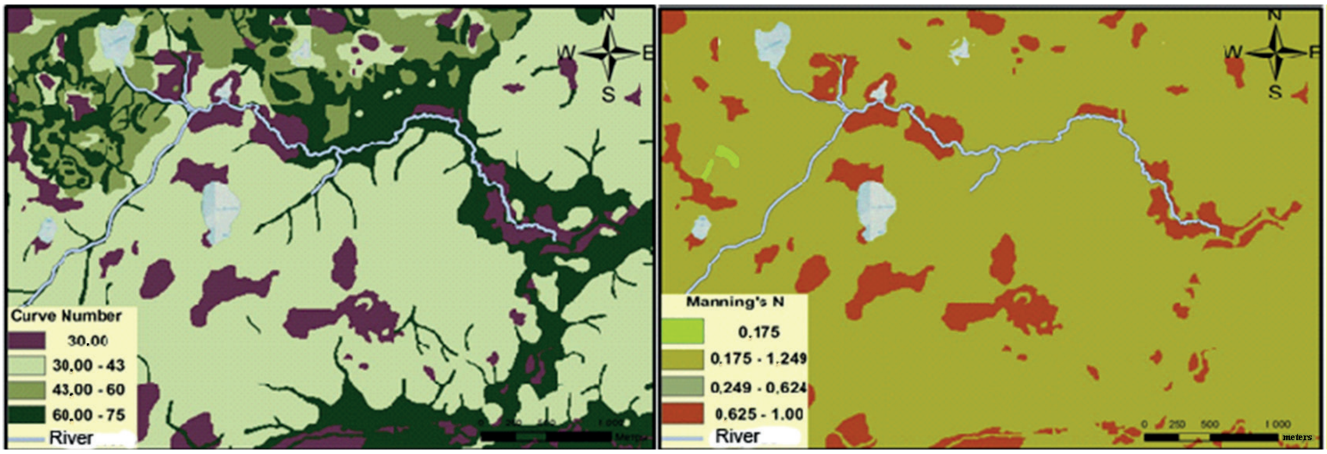


Fig. 8. Distribution of hydrophysical parameters in the basin of the Loninka river for modeling water flow on the basis of the landscape structure (SAGA GIS). A - Curve number (CN); B - Manning's n (MN)

Threshold – CDT) and some other parameters were subject to changes in calculations. Numerical modeling has shown that even tabular values of MN, CN, CSS, MFT and CDT, not adapted to taiga wetlands, reveal significant features in the distribution of surface water runoff in various geosystems (Fig. 9). Extremely low runoff values (<0.01 m/s), were observed for most of the basin. Higher rates are characteristic only for the channels of streams and rivers (0.025-0.2 m/s) and the runoff increases up to 2 m/s within certain sections of the Loninka River. The pattern of runoff rates distribution is quite realistic, since the catchment of the Loninka River is a flat swamped hummocky sandy plain, cut by rare channels with water flow.

The results of runoff simulation using various parameters of Average Rain Intensity (ARI) and Channel Site Slope (CSS) revealed some regularities. In all cases, the increase in ARI caused higher runoff, for example, at the source point located near the drainage pipe under the railway embankment (the source of the Loninka River). This site is highly modified by human activities, and, consequently, it has low MN values (0.025) contributing to the surface runoff, and high CN (98) impeding water infiltration into the soil. Thus, the changing intensity of precipitation successively leads to the change in surface runoff rates. That is, in this section of the river the sensitivity of runoff to the precipitation intensity is great, although the flow rates are not very high. Most other observation sites are located in natural forest and mire landscapes. These sites are characterized by high MN values (0.5-0.9) and low CN (30-40). High values of MN prevent surface runoff, while low CN favor active infiltration. As such, the runoff rates decrease substantially at these sites and weakly respond to changes in the precipitation intensity. Thus, the different location of the observation sites in terms of landscape structure results in significant differences in runoff characteristics. Low regular runoff rates at zero precipitation intensity confirm the high capacity of flat over-moisturized catchment to accumulate water and regulate runoff in geosystems.

Verification of model calculations was carried out in the field. Experimental measurements of flow rates and discharge of the Loninka River at the gauging stations suggest the following. In all cases, the predicted rates differ

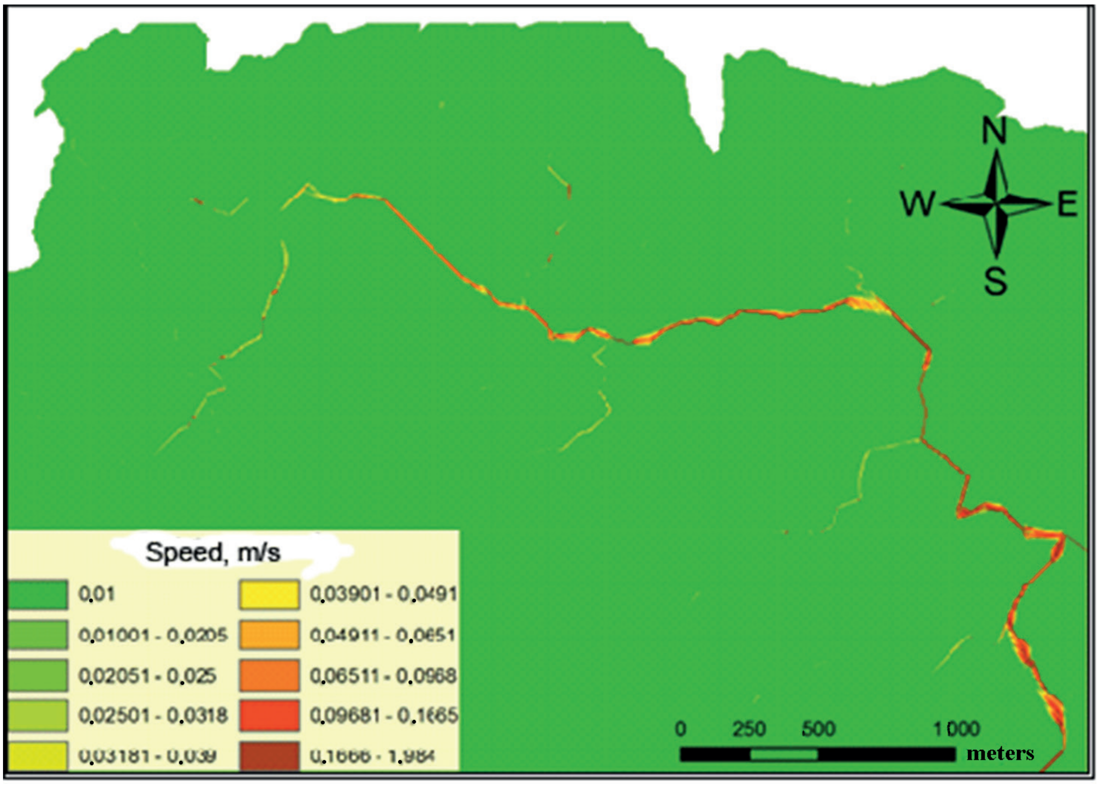


Fig. 9. Calculation of surface runoff rates in the Loninka River basin by means of the SAGA GIS. The precipitation intensity is 10 mm/h, CSS = 1, MFT = 180, CDT = 360

from the measured ones, but the calculated values were not so far from the real ones as it was expected (Fig.10). The closest results were obtained for the precipitation intensity of 10 mm/h, although lower-intense precipitation is more probable.

More accurate simulation results could be obtained by adjusting the values of model coefficients (tabular values for non-waterlogged rivers were used for calculations). A more detailed DEM could also be useful. It turned out that the channel width of 1.0-1.5 m, and the 30 m pixel size does not allow delimiting valleys, as well as the microrelief which is very important for runoff from the flat plains. On the other hand, errors in the measurement of flow rates are quite possible for flat, boggy meandering channels, often blocked by forest debris and beaver dams. Nevertheless, under the lack of information, the values obtained during GIS modeling could become a basis for predicting runoff values in areas where the direct measurements are labor-consuming or otherwise impossible.

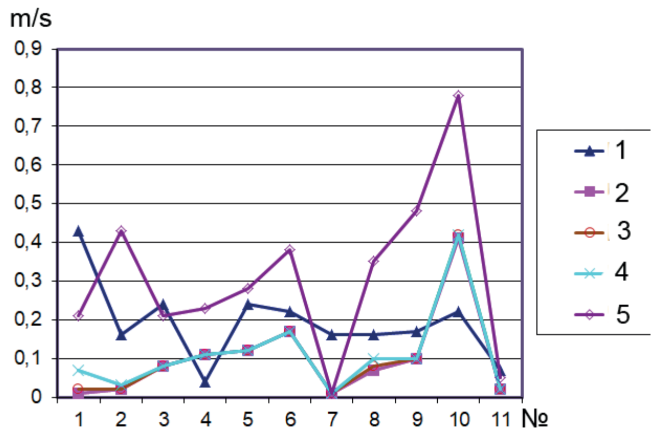


Fig. 10. Comparison of the measured and calculated flow rates in the Loninka River at the gauging stations. 1- field measurements; 2 – model calculation, precipitation intensity 0.0 mm/h; 3 - model calculation, precipitation intensity 0.66 mm/h; 4 - model calculation, precipitation intensity 10.0 mm/h; 5 - model calculation, intensity of precipitation 100.0 mm/h. The abscissa shows the numbers of gauging stations downstream the river sources

Let us demonstrate the possibility of water protection zoning of geosystems based on modeling the structure of catchment basins and the runoff from their areas using a priori data. An important environmental characteristic of the processes in the catchment basin is the delay time of water flowing to the river or control stations. The isochrones of flow delay time were calculated using the SAGA GIS (Fig. 11, A). However, this method of calculating is difficult to use to predict the time of pollutants arrival from side streams, which is important for taking measures to localize pollution before it gets into the river channel. We developed a modified cascade algorithm (Sysuev et al. 2011) to calculate the running time to first-order channel for each first-order catchment, combine them into second-order catchments, and then calculate the running time for each second-order catchment before merging together all second-order catchments, and so forth (Fig. 11, B).

CONCLUSION

The formation of landscape structures is described in traditional empirical concepts using the geomorphometric parameters of geophysical fields, i.e. gravity and insolation. The concept of landscape polystructure becomes physically defined: by choosing the main structure-forming processes and their principal parameters different classifications of landscapes could be elaborated. Formal mathematical algorithms of selecting surface relief units acquire fundamental geophysical meaning if combined with the state parameters. Implementation of the typological approach makes it possible to obtain a hierarchy of natural-territorial complexes (facies - urochishche - mestnost - landscape); implementation of the approach of the hydrological functioning of the landscape results in the hierarchy of catchment geosystems; implementation of the classification approach for parameters and normalized coefficients of remote sensing data produces the structure of vegetation cover.

Geomorphometric values describing the gradients of gravity (height, slope, horizontal, vertical and average curvature, specific collection area and specific dispersive area; B-depression depth) and insolation (direct solar radiation dose; aspect and illumination of slopes) fields are considered to be the parameters of physical state of individual units of relief surface, i.e. the DEM pixels, which form the

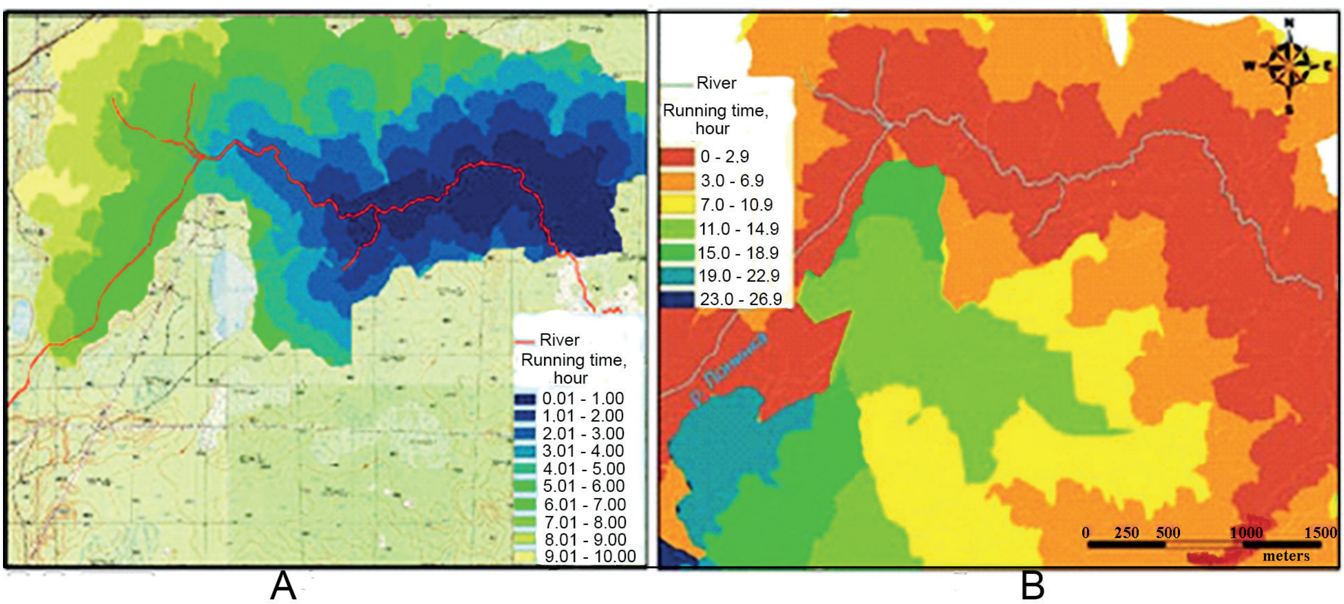


Fig. 11. Running time of surface water runoff, hours. A - to the control section of the Loninka River, according to the SAGA GIS algorithm; B - to the river for each particular point of the channel, according to the cascade algorithm

geosystems. The state parameters were preferred due to their simple form and direct description of physical fields For example, slope is the modulus of the geopotential gradients; horizontal/ planar curvature is the divergence of streamlines; vertical curvature is a derivative of the steepness factor along the streamline; the dose of direct solar radiation is the relative amount of incoming energy, etc. The state parameters are also independently included into the description of structure-forming processes. Digital remote sensing data are also physical parameters of the state of individual units of relief surface and geosystems.

Parameters of the typological model of the landscape structure are selected in accordance with classical definitions and preliminary numerical experiments. The need for professionally correct and justified selection of the physical state parameters, i.e. principal structure-forming processes, as well as their weights suggests the crucial role of the landscape approach.

The functional model of the landscape structure is based on morphometric parameters describing the redistribution of water over the surface in the gravitational field (slopes, specific catchment area, horizontal and vertical curvature). Such classification makes it possible to identify the contours of various-order catchment geosystems in accordance with the Horton-Strahler-Tokunaga scheme.

The structural parameters obtained from the typological description of landscapes allow simulating the hydrological functioning of catchment geosystems with satisfactory accuracy for particular types of water protection zoning.

REFERENCES

Akbari Kh.Kh., Bondar Yu.N., and Sysuev V.V. (2006). Indicative properties of the stand in the landscapes of the edge zone of the Valdai glaciation. Proceedings of Moscow University, series 5 Geography (6), 59-66. (in Russian with English summary)

Antipov A.N. and Fedorov V.N. (2000). Landscape and hydrological organization of the territory. Novosibirsk: Nauka. (in Russian with English summary)

Armand A.D. (1988). Self-organization and self-regulation of geographical systems. Moscow: Nauka. (in Russian)

Beven K. J. (2012). Rainfall-runoff modelling: the primer. Chichester: Wiley-Blackwell.

Chow V.T., Maidment D.R., and Mays L.W. (1988). Applied Hydrology. New York: McGraw-Hill.

Chow V.T. (1959). Open-channel hydraulics. New York: McGraw-Hill.

Dodds P.S. and Rothman D.H. (1999). Unified view of scaling laws for river networks. Phys. Rev. E, 59(5), 4865–4877.

Dyakonov K.N. and PuzachenkoYu.G. (2004). Theoretical positions and directions of modern landscape studies. In: K.N. Dyakonov, ed. Geography, society and environment. V. II. Functioning and present-day state of landscapes. Moscow: Gorodets, 21–36. (in Russian).

Dyakonov K.N. (2008). Basic concepts and intention of landscape studies. In: N.S. Kasimov, ed. Geographical scientific schools of Moscow University. Moscow: Gorodets, 348-381. (in Russian)

Forman R.T.T. (2006). Land mosaics: the ecology of landscapes and regions. Cambridge University Press, Cambridge.

Kleidon A. (2010). Life, hierarchy, and the thermodynamic machinery of planet Earth. Physics of Life Reviews, 7(4), 424–460.

Khoroshev A.V. (2012). Geographical concept of landscape planning. Proceedings of the RAS. Geographical series (4), 103-112. (in Russian with English summary)

Khoroshev A.V. (2016). The large-scale organization of the geographical landscape. Moscow: Fellowship of scientific publications KMK.

Landscape Planning: Principles, Methods, European and Russian Experience. (2002). In: A.N. Antipov and A.V. Drozdov, ed. Irkutsk: Publishing House of the Institute of Geography SB RAS. (in Russian with English summary)

Likens G.E. and Bormann F.H. (1995). Biogeochemistry of a forested ecosystem. New York: Springer-Verlag.

McGarigal K., and Cushman S.A, (2005). The gradient concept of landscape structure. In: J. Wiens and M. Moss, ed. Issues and perspectives in landscape ecology. Cambridge: Cambridge University Press, 112–119

McGarigal K., Tagil S., Cushman S.A. (2009). Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landscape Ecol (24), 433–450. DOI 10.1007/s10980-009-9327-y

Olaya V. (2004). A gentle introduction to SAGA GIS. Edition 1.2.

Orlova I.V. (2014). Landscape-agroecological planning of the territory of the municipal district. Novosibirsk: Publishing House of the SB RAS. (in Russian)

Pozachenyuk E.A. (2006). Territorial planning. Simferopol: Publ. House of the TNU. (in Russian)

Rulev A.S. (2008). Landscape planning of agroforestry land improvement. Environmental planning and management, 1(6), 25-28. (in Russian with English summary)

PuzachenkoYu. G. (2014).Organization of a landscape. In: K.N. Dyakonov, V.M. Kotlyakov, T.I. Kharitonova, ed. Issues in Geography. Vol. 138. Horizons of landscape studies. Moscow: Kodeks, 35-64. (in Russian)

Reteyum A. Yu. (1975). Physical-geographical zoning and the allocation of geosystems. Issues in Geography, 98, 5-27.(in Russian)

Shary P.A. (1995). Land surface in gravity points classification by a complete system of curvatures. Mathematical Geology, 27(3), 373–390.

Solntsev N.A. (1948).The natural geographic landscape and some of its general rules. Proceedings of the Second All-Union Geographical Congress. Moscow: State Publishing House for Geographic Literature, 258-269. (in Russian). See also in: J.A. Wiens, M.R. Moss, M.G. Turner, D.J. Mladenoff, ed., (2006). Foundation papers in landscape ecology. New York: Columbia University Press, 19-27.

Solntsev V.N. (1997). Structural landscape studies. Moscow: Faculty of Geography MSU. (in Russian)

Sysuev V.V. (2003). Morphometric analysis of geophysical differentiation of landscapes. News of the Academy of Sciences. Series geographical, (4), 36-50 (in Russian with English summary)

Sysuev V.V. (2014). Basic concepts of the physical and mathematical theory of geosystems. In: K.N. Dyakonov and V.M. Kotlyakov, ed. Issues in Geography. Vol. 138. Horizons of landscape studies. Moscow: Kodeks, 65-100. (in Russian with English summary)

Sysuev V.V., and Solnetsev V.N. (2006). Landscapes of the edge zone of the Valdai glaciation: classical and morphometric analysis. In: K.N. Dyakonov, ed. Landscape science: theory, methods, regional studies, practice. Proceedings of the XI International landscape conference. Moscow: MSU Publishing House, 249-252. (in Russian)

Sysuev V.V., Sadkov S.A., Erofeyev A.A. (2011). Basin principle of functional zoning: modeling of the structure and drainage of catchment geosystems based on a priori data. In: K.N. Dyakonov, ed. Landscape science: theory, methods, regional studies, practice. Proceedings of the XI International landscape conference. Moscow: MSU Publishing House, 101-105. (in Russian)

Tarboton D.G., Bras, R.L., Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. Hydrological Processes, 5(1), 81-100.

Tkachev B.P., and Bulatov V.I. (2002). Small rivers: state-of-the act and ecological problems: Analit. review. Novosibirsk: Publishing House of the SB RAS.

Troendle C.A. (1985). A variable source area model for stormflow prediction of forested watersheds. In: M.G. Anderson and T.P. Burt, ed. Hydrological forecasting. Chichester: John Wiley & Sons, 431-495.

Turner MG (2005). Landscape ecology: what is the state of the science? Annu. Rev. Ecol. Evol. Syst. (36), 319–344

Turner M., Gardner R.H. (2015). Landscape Ecology in Theory and Practice. Pattern and Process. New York: Springer.

Viktorov A.S. (1986). Landscape pattern. Moscow: Mysl. (in Russian with English summary).

Received on February 12th, 2019 Accepted on January 21th, 2020

Natalia N. Matinian[*], Anastasia L. Gusareva, Kseniia A. Bakhmatova, Anastasia A. Sheshukova
* Saint Petersburg University, Universitetskaya nab., 7-9, Saint Petersburg, 199034, Russia
***Corresponding author:** geosoil@mail.ru

MICROBIOLOGICAL INDICATORS AND HEAVY METAL CONCENTRATION IN ECOLOGICAL ASSESSMENT OF URBAN SOILS OF SAINT PETERSBURG, RUSSIA

ABSTRACT. This paper aimed to characterize urbostratozems (Urbic Technosol, WRB) of Saint Petersburg located in industrial ("Electrodepo" railway station) and residential (region Polish Garden) zones. These soils were also compared with background (natural) soddy podzol soil (Umbric Albic Gleic Podzol, WRB) sampled in recreational zone (suburban park "Oranienbaum"). Soil samples were collected from soil horizons for chemical analysis and from top of soils for microbialogical analysis in June of 2012. Chemical properties (pH, total organic carbon, mobile forms of K and P) and content of heavy metals (Pb, Cu, Zn, Ni) in soils were determined. Culturable forms of microorganisms (bacteria and fungi) were studied. Assessment of the enzymatic activity of the soil was carried out by culturing of microorganisms-producers of protease, amylase, cellulase and lipase on special media. Biotesting using cress (*Lepidium sativum* L.) seeds had been carried out for assessment of soil phytotoxicity. It was found that chemical properties of urban and natural soils differ greatly. Heavy metal pollution was evident in both urban soils, but maximum concentrations of heavy metals were found in the soil of the industrial zone. Phytotoxocity had been also most pronouncend in the soil of the industrial zone. The natural soil exhibited significantly higher respiration activity than urbostratozems. The greatest difference in the structure of the bacterial and fungal communities was observed between the natural soil of the recreational zone and the urbostratozem of the industrial zone. Algae had been present in the urban soils of the residential zone that was not observed in the natural podzol. The minimum number of producers of all enzymes, except for cellulase, was observed in the soddy podzol in the recreational zone. The maximum number of protease and amylase producers was found in the soil of the industrial zone. Lipolytic activity was almost the same in all samples. It was found that more sensitive biological methods are needed for environmental assessment of urban soils. The results of the article can be used by soil scientists and environmental engineers for a comprehensive environmental assessment of the condition of urban soils and for creating new urban green spaces.

KEY WORDS: urban soil, heavy metal, microbiological indicator, enzyme activity, soil phytotoxocity

CITATION: Natalia N. Matinian, Anastasia L. Gusareva, Kseniia A. Bakhmatova, Anastasia A. Sheshukova (2019). Microbiological Indicators And Heavy Metal Concentration In Ecological Assessment Of Urban Soils Of Saint Petersburg, Russia. Geography, Environment, Sustainability, Vol.13, No 1, p. 214-223 DOI-10.24057/2071-9388-2019-63

INTRODUCTION

Urbanization significantly influenced the environment and results in irreversible changes of relief, hydrological conditions, vegetation cover and soils. Soils in urban environment are characterized by the presence of many artefacts, alkaline pH, high content of organic carbon and nutrients, as well as high content of technogenic contaminants (Lehmann and Stahr2007; Burghardt et al. 2015). Also typical for urban soils is spatial heterogeneity of structure and properties (Greinert 2015). According WRB (IUSS Working group WRB 2006, 2014), the most of urban soils belong to the Technosols group, with different qualifiers, used to represent their diversity (Charzynski et al. 2013). The morphological and chemical features of these soils reflect the history of land use of each particular urban area (Prokofieva and Poputnikov 2010). The direction of pedogenesis in urban environment also depends on characteristics of soil forming material and climate (Aleksandrovskiy et al. 2012). Biological properties of urban soils can serve as indicators of their ecosystem function (Piotrowska-Dlugosz and Charzyńsky 2015), with the microbial diversity and activity being among the most meaningful biological characteristic of soils (Schindelbeck et al. 2008; Terekhova et al. 2014; Rozanova et al. 2016). Microbial properties (including substrate-induced and basal respiration) are influenced by specific chemical properties in different functional zones of city (Ivashchenko et al. 2019). The main groups of soil microorganisms include bacteria, fungi, actinomycetes and algae. Content of organic matter, texture and redox potential determine the profile distribution of microorganisms in natural soils (Marfenina 1987, 2005). Despite the wide using molecular techniques in modern studies (Bouchez et al. 2016; Huot et al. 2017), cultivation of microorganisms is considered a useful tool for investigation of urban soil communities (Braun et al. 2006; Marfenina et al. 2017).

Heavy metals (HM) are one of the most dangerous pollutants of urban soils, reducing their biological activity and productivity (Yungen et al. 2001; Papa et al. 2010; Muhlbachova et al. 2015). Elevated concentaration of HM reported to reduce microbial diversity (Jieng et al. 2010). The urgency of this problem is underlined by the impact of HM on soil microorganisms, which play a key role in soil processes, required for viable ecosystem (Nannipiery et al. 2003). The scope of toxic effects due to HM is determined by the nature of the HM, the form of the compounds in question, soil pH, and the the specific sensitivity of different microorganisms. Parameters of biodiversity and physiological activity of soil microbiota (Yuangen et al. 2006; Niemeyer et al. 2012), fungal diversity of soil (Marfenina et al. 2017), as well as the species composition and amount of soil algae can be used as biological indicators of soil contamination with HM (Zenova et al. 1995).

The sharp decrease in the activity of enzymes: amylase, dehydrogenase, urease, invertase, catalase was previously observed in soil with the influence of elevated HM concentrations (Zhigareva et al. 2006). Restructuring of the microbiocenosis had been found in the soil containing heavy metals in concentrations of 1-2 orders of magnitude higher than the background. A decrease in richness and diversity of soil micromycetes species was also observed, with few, down to a single one species dominating. The number of colored micromycetes increased and in some cases, toxin-forming microorganisms and resistant species dominated in such contaminated soils (Zvyagintsev et al. 2005). The effect of HM on enzyme activities of soil depends on soil organic matter content (Stefanowicz et al. 2010; Calvarro et al. 2014; Muhlbachova et al. 2015). It was shown that in soils that poor in nutrients, micromycetes are more susceptible to the negative effects of HM. It was reported that HM inhibit the metabolism of microorganisms, change their growth kinetics and morphology. A catastrophic decrease of soil microbiological activity or total death of microorganisms was found at concentrations of heavy metals of 4 and more orders of magnitude above the background (Rabotnova and Pozmogova 1979). However, Hagmann et al. (2015) showed that the highest activities of several enzymes (alkaline phosphatase, celobiohydrolase, L-leucine-aminopeptidase) were positively correlated with the highest HM contamination of soil.

Ammonifying bacteria, along with some spore bacteria, cellulolytic bacteria and actinomycetes are most sensitive to soil pollution. However, some authors reported increase of bacterial numbers in the most anthropogenically transformed soils along with decrease of their enzymatic activity (Naylo et al. 2019). An increase in the total number of microflora has been found in the soils with increased content of HM (Castaldi et al. 2004). Active development of resistant forms had been observed possibly due to the absence of competition as more sensitive microorganisms did not survive in the environment. Marfenina (2005) reported the increase of soil microfungae diversity in urban environment in comparison with zonal conditions. On the other hand, several heavy metals (e.g., Cu, Zn, Mo) are known to be micronutrients which can stimulate the microbial growth.

The aims of this paper were to study morphological and chemical properties of urban soils and to find easily identifiable biological indicators for environmental assessment of soils of different types of urban land use. Such indicators are in demand by environmental practitioners, limited in material resources and specialized laboratory equipment. It is known that modern molecular genetic methods for studying soil biota require high material costs and the availability of high-tech analytical equipment.

MATERIALS AND METHODS

Saint Petersburg is located in the northwestern part of Russia (59°57´N, 30°19´E) on the banks of the river Neva and the east coast of the Finnish gulf. The climate of humid continental type, the average annual temperature of 5.8°C and the average annual precipitation of about 660 mm. There are several main functional areas on the territory of the city: residential, industrial, agricultural, recreation. The working hypothesis of the study was that there is a direct relationship between the type of land use, the level of pollution and biological indicators.

The soils of the following zones were selected for the study:
1. Recreation zone (the State Art and Architectural Palace and Park Museum-Reserve "Oranienbaum"; area - 165 ha) - profile 1. Research site vegetation – coniferous forest with ground cover of bilberry and green moss.
2. Residential zone (central part of the city, Polish Garden is located inside residential areas with dense buildings, between the Fontanka embankment, Izmailovsky avenue, Derzhvinsky lane and 1st Krasnoarmeisky street; area -2.3 ha) - profile 2. The inner garden of the former G.R. Derzhavin manor was choosen to characterize the soil of residential zone, because it is difficult to find green space in the historical center of Saint Petersburg. Research site vegetation–grassy lawn.
3. Industrial zone (railway station Elektrodepo, near electromechanical workshops) – profile 3. Research site vegetation –ruderal herbaceous plants.

Soil names and indexes of natural soil horizons were given according to Classification System of Russian Soils (2004) and WRB (IUSS Working group WRB 2014). Diagnostics of urban soils was made according to Prokof´eva et al. (2014).

Samples for chemical analysis were taken from each horizon of soil profiles, then air-dried and sieved (1mm). Traditional analytical methods were used to obtain physical and chemical characteristics of soils (Vorobieva 2006). Total organic carbon (TOC) content was determined titrimetrically by dichromate oxidation, pH was measured in suspension at a ratio of soil to water = 1 : 2.5; the content of plant available potassium (K) and phosphorus (P) compounds in soils with pH>7 – using a 1% solution of $(NH_4)_2CO_3$ (ratio soil to solution = 1:20), in soils with pH<7 – using 0,2N solution of HCl (ratio soil to solution = 1:5). P was measured photometrically (FEK-3) in an extract tinged with ammonium-heptamolybdate and ascorbic acid reagent. Potassium was determined by photoelectric flame photometry. The content of mobile forms of phosphorus and potassium was estimated using the benchmarks designed for agriculture plants. Heavy metal (HM) total concentrations, namely of Cu, Pb, Zn, Ni were determined by atomic emission spectroscopy (AES). Digestion of samples was carried out by evaporation from the carbon electrode channel (GOST 41-08-265-04 2004; Vorobieva 2006). The analysis was done at the Central laboratory of A.P. Karpinsky Russian Geological research institute (Saint Petersburg) (included in the International Association of Geoanalysts – IAG).

Soil samples (1 mixed sample consisted of 5 subsamples from each site) for microbiological analysis were collected in the upper 10 cm of soil profile and analysed immediately after sampling. The following methods were used to determine the biological properties of the soil (Zvyagintsev 1991; Labutova 2008; Labutova and Bankina 2013):
1. The biological activity of soils in respiration (actual, potential) was determined by the modified Golovko method, based on producing of CO_2 under laboratory conditions. To determine the amount of actual soil respiration, a sample of soil in a small Petri dish was placed in a hermetically sealed container next to a cup with 20 ml of 0.02 N NaOH. The sample was left for a day and then the amount of CO_2 released per day and bound by NaOH was determined titrimetrically (0.02 N H_2SO_4, phenolphthalein). A nutrient substrate (10% glucose solution) was preliminarily added to the soil sample to determine potential respiration.
2. The number and diversity of bacteria was determined by spreading on meat-peptone agar (MPA) media. Desorption of microorganisms was carried out with 0.1% solution of $Na_2P_2O_7$ (Szegi 1983), then soil suspension was shaken for 10 minutes on laboratory shaker and settled. A series of 10-fold dilutions of soil sample was prepared. Determination of fungi was carried out on the Czapek medium. Inhibition of bacterial growth during the isolation of fungi was carried out by adding streptomycin sulfate (50 mg /l) (Microbiology laboratory manual 2005). Plating was performed in 5 replications per variant. The numbers of colonies formed

on the medium were counted after the end of the incubation period: 3 days for bacteria, and 5-7 days for fungi. Cultivation have been done at a temperature of 28°C. Variety of soil bacteria and fungi was evaluated using MCT, a number of Morpho-Cultural Types, representing a number of species of soil fungi and bacteria in the sample which correlated to numbers of colonies with the same characteristics formed by a single type of microorganism. The frequency of occurrence of MCT (%) was counted as n·100/N, where n – the number of plates where this MCT was found and N – the total number of plates. The abundance of MCT (%) was conted as a·100/A, where a – the number of colonies of this MCT, A –the total number of colonies (Labutova 2009). According to values of abundance and frequency MCT were assessed as dominant (25-30% and more), common (15-25%), rare (5-15%), random (less than 5%).

3. The number of soil algae was determined by sowing on Gromov's agar medium (№1). Soil particles were placed on the surface of agar medium and cultivated at a temperature of 24°C, illuminated with fluorescent lamps, for 30 days.

4. Assessment of the enzymatic activity of the soil was carried out by culturing of microorganisms-producers of protease, amylase, cellulase and lipase on special media. This method shows potential biological activity since during the cultivation of microorganisms in laboratory conditions, the share of them which were active in real soil remains unknown. However, this potential activity differs across different soil types (Zvyagintsev et al. 2005). The soils a series of 10-fold dilutions of soil sample was prepared, with 3 replications per variant, and incubation time of 3-5 days:

1) Milk agar was used to determine the protease activity of the soil. The method is based on the ability of protease producers to break up milk protein (casein). 0.5 ml aliquots of soil suspension were placed on the medium. The milk agar has cleared around the colonies of microorganisms which were considered active protease producers. Microorganisms that formed colonies, but did not produce cleared zones, were defined as having a weak protease activity.

2) Starch-ammonium agar was used to determine the amylolytic activity. Then the surface of the agar was covered with Lugol's solution and the total number of colonies was counted. The activity and the number of producers of amylase were determined depending on the stained areas around the colonies. The activity of amylase producers (very active and active producers) was evaluated by the comparison of the diameters of the hydrolysis zones from the edge of the colonies. Microorganisms that formed colonies, but did not produce the hydrolysis zones, were defined as having a weak amylase activity.

3) To determine the cellulolytic activity of the soil, Getchinson's agar medium was used. Filter paper was placed on the surface of the medium and 1 ml of soil suspension was applied evenly. The number of colonies exhibited cellulose decomposed around them was determined after 2-week incubating at 28°C. The method is based on the ability of the producers of cellulases to use cellulose of filter paper as the sole carbon source.

4) Lipolytic activity of the soil was determined using agar medium with castor oil. Small soil agregates (20 pieces) were placed in each dish and incubated for 3-5 days. The numbers of agregates exhibited opaque zones around them were counted. The method is based on the fact that fat decomposition occurs in the presence of lipases and free fatty acids accumulate in the medium, causing its opaqness.

5. Cress (*Lepidium sativum* L.) seeds were used for assessment of the soil phytotoxicity.

Phytotoxicity of the soil was measured by the difference in the number of germinated seeds and the length of seedlings and roots in the studied soil and in the control (wet filter paper) in triplicate. The duration of the growth was 5-7 days.

Statistical data processing was carried out using STATISTICA 10.0.

RESULTS AND DISCUSSION

Gleyed soddy podzol (Umbric Albic Gleic Podzol, WRB) on glacio-lacustrine sands (profile 1) has well defined diagnostic horizons: AY-AYe-E-BF-BCg-Cg. Water appeared from the depth of 82 cm. The soil meets the characteristics of the zonal type. Profile 2 exhibited urbostratozem (Urbic Technosol, WRB) on buried gray-humus gley soil (Umbric Gleysol, WRB) consisted of several filled-in horizons and underlyed by gray-humus alluvial soil (UR1ay-UR2-UR3rt-UR4-UR5rt-UR6-UR7-[AY]-CG). The garden planning had been changed several times; the filling and leveling of the surface were carried out. The thickness of stratified humus-accumulative horizons was 30 cm. The soil profile was characterized by a series of infilled layers of different color, thickness, and composition. Anthropogenic material consisted mainly of construction and household waste (fragments of bricks, ceramics, pieces of asphalt, broken glass, coal, bones, pieces of tin, wire, nails, wood, etc.). It was mixed with sandy or loamy-sandy fine earth material. Some anthropogenic layers contained organic matter, which allows suggest that the filling of humus-enriched material (e.g. compost) and the terrain levelling were done repeatedly. The total thickness of the cultural layer was 155 cm. At a depth of 155 cm, a dark gray homogeneous humus horizon of buried natural soil was found. Profile 3 – Technogenic urbostratozem on technogenic material (Urbic Technosol, WRB) (profile 3) consisted of infilled layers (thickness 60 cm). The soil characterized by loamy-sandy and sandy texture and the significant amount of artefacts (granite rubble, broken glass, pieces of metal, polyethylene, textile fragments etc.). At a depth of 60 cm was a concrete block.

Soddy podzol (profile 1) had low pH value (3.5-4.1) (Table 1). Urbostratozem (profile 2) characterized by slightly acid reaction of upper horizons (6.0-6.2), which increases with depth up to pH = 8.4. The soil alkalinization was likely caused by the release of calcium from various construction debris (e.g. cement, bricks, etc.). Lime layer was found at a depth of 112-124 cm. Technogenic urbostratozem (profile 3) has slightly acid and near neutral pH (6.2-6.8).

Organic carbon content in soddy podzol is low and decreased with depth (from 1.8 to 0.32 %), demonstrating ordinary distribution for this soil type. The carbon content in the soil of the Polish Garden (profile 2) varied greatly throughout the profile, due to the abundance of anthropogenic inclusions and differences in the composition and origin of anthropogenic layers. In the soil of the industrial zone (profile 3), the carbon content varies from 1.4 to 4.7%, which also depended on the origin of the anthropogenic material.

The studied soils had a low content of exchangeable potassium, which can be explained by their light texture. The amount of mobile forms of phosphorus in the natural soil (profile 1) was medium. The sufficiency of mobile phosphourus in the urban soils (profiles 2 and 3) was estimated as high to very high. This enrichment with P is likely caused by the presence of household waste in the soils.

Heavy metals content. The main inorganic pollutants on the urban and suburban territory were Cu, Pb, Zn, Ni. HM content in the soils presented in Table 2.

The content of HM in the soil of the recreation zone (pro-

Table 1. Chemical properties of soils

Horizon	Depth, cm	pH H_2O	TOC, %	P_2O_5	K_2O
				mg/kg	
P.1 Gleyed soddy podzol (Umbric Albic Gleic Podzol)					
AY	2-17	4.0	1.81	90	37
AYe	17-30	3.5	0.90	100	14
E	25-30	4.1	0.42	85	1
BF	30-44	3.6	1.31	84	2
BCg	44-82	3.9	0.33	186	1
Cg	82-94	4.3	0.32	96	1
P.2 Urbostratozem (Urbic Technosol) on buried gray-humus gley soil (Umbric Gleysol)					
UR1ay	3-15	6.0	3.83	181	32
UR2	20-30	6.2	2.64	98	40
UR2	30-47	7.8	3.11	168	78
UR3rt	47-60	7.9	6.91	156	67
UR4	60-99	8.2	2.63	151	20
UR5rt	99-112	8.3	5.22	145	53
UR6	124-140	8.4	4.48	132	38
UR7	140-155	8.1	2.85	n.d.*	n.d.
[AY]	155-171	7.7	1.93	105	90
CG	171-176	8.2	0.51	n.d.	n.d.
P.3 Technogenic urbostratozem (Urbic Technosol)					
UR1	2-19	6.5	4.71	164	63
UR2	19-22	6.4	1.45	153	23
UR3	22-28	6.6	3.44	113	42
TCH	28-30	6.2	3.93	121	18
UR4	30-60	6.9	4.31	280	35

**n.d. – no data*

Table 2. Heavy metal concentration in soils, mg/kg

Horizon	Depth, cm	Zn	Pb	Cu	Ni
P.1 Gleyed soddy podzol (Umbric Albic Gleic Podzol)					
AY	2-7	13	26	11	13
AY	7-17	14	14	5	11
AYe	17-30	14	15	4	8
P.2 Urbostratozem (Urbic Technosol) on buried gray-humus gley soil (Umbric Gleysol)					
UR1ay	3-15	**178**	**46**	29	18
UR2	20-30	**179**	**55**	**43**	19
UR2	30-47	**120**	**100**	30	12

UR3rt	47-60	**80**	**150**	30	10
UR4	60-99	**80**	**60**	25	8
UR5rt	99-112	**80**	**100**	25	10
UR6	124-140	**60**	**50**	20	8
UR7	140-155	50	20	20	12
[AY]	155-171	40	20	**50**	12
CG	171-176	30	20	2	12
P.3 Technogenic urbostratozem (Urbic Technosol)					
UR1	2-19	**244**	**64**	**84**	**38**
UR2	19-22	40	18	24	11
UR3	22-28	**73**	**35**	**44**	20
TCH	28-30	**211**	**77**	**168**	**45**
UR4	30-60	**202**	**70**	**160**	**33**
Approximate permissible concentrations (APC), mg/kg (HN 2.1.7.2511-09)		55	32	33	20
Background values*		51	16	15	11

**Median value for sandy and loamy-sandy soils of North-West Russia, humus horizon (Matinian et al. 2007)*
***The values exceeding the APC are in bold*

file 1) corresponded to the background level, excluding Zn. The Zn content in this profile substantially lower than the background level. This phenomenon could be explained by genesis and chemical composition of the soil forming material (limnoglacial sands which are depleted in many elements). The content of HM in urbostratozems (profiles 2 and 3) significantly exceeded background values, not only in the upper horizons, but also in the middle and lower parts of the profile. The HM pollution was associated with the presence of anthropogenic material in soil horizons. The HM concentrations exceeding the allowed levels were found in both upper and lower parts of profile of the industrial zone soil. Technogenic urbostratozem was characterized by high content of copper (5 times higher than the value of the APC) and nickel (2 times higher than the value of the APC), which was not observed in the urban soil of the residential zone. Most likely, this type of pollution is associated with trains'repair and maintenance. The profile distribution of total HM concentrations in urbostratozems mainly depends on the composition of bulk layers and the level of anthropogenic impact at different stages of the soil formation.

Soil respiration. The actual respiration of the upper horizons of the studied soils was 3-6 mg CO_2/100 g of soil/24 h. Potential respiration (after addition of the readily available substrate – 10% glucose solution) was two or more times higher than the actual respiration. This indicates the deficiency of labile organic substances in the urbostratozems and the natural soil.

The highest value of CO_2 emission per soil mass was found in the soil of industrial zone (sample 3), while the natural soil (sample 1) were characterized by the lowest intensity of respiration. The calculation of the biological activity per 1 g of carbon (Fig. 1) allows to specify the utilization of soil organic matter by microorganisms. The natural soil exhibited significantly higher activity than urbostratozems, which may indicate the presence of stable organic matter in these soils. Thus, as the anthropogenic impact increases, there is a tendency to a decrease in soil respiration rate. Other researchers also had shown significant reducing of soil respiration rates in metal-contaminated soils in comparison with unpolluted soils (Papa et al., 2010).

Bacteria and fungi diversity. The least number of microorganisms was found in the natural soil of the recreation zone (Oranienbaum Park, sample 1, Table 3). The number of fungi was similar in the soil of the recreational and industrial zones, while in the soil of the residential zone it was significantly lower. The maximum number of prokaryotes was found in the humus horizon of urbanozem of the residential

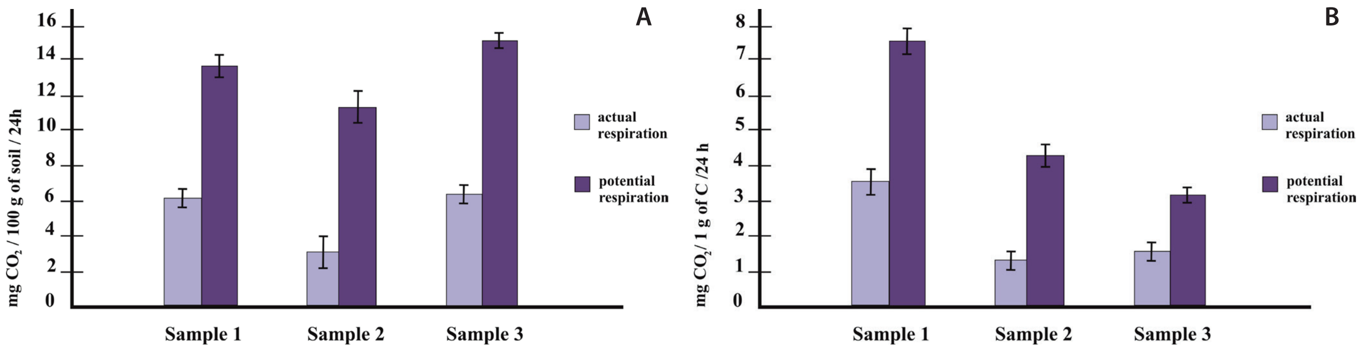


Fig. 1. Soil respiration rates (actual and potential): a – mg CO_2/100g soil/24h; b - mg CO_2/1g C/24h

zone (Polish Garden, sample 2). In the soil of the industrial zone (Elektrodepo, sample 3), the number of bacterial MCT was only slightly lower than in the soil of the residential zone. The neutral reaction of the soils, the high content of organic matter and the availability of mineral nutrition elements in the surface horizon of the urbostratozems contribute to active development of bacteria in these soils. Additionally, the stimulating effect of metals on soil microflora had been reported previously (Dai et al. 2004; Stefanowicz et al. 2012; Naylo et al. 2019).

In the upper horizon of the urbostratozem of the residential zone (sample 2), the diversity expands due to the appearance of MCT 3. At the same time, the structure of the bacterial cenosis also changes: if MCT 1 dominated in natural soil, MCT 2 dominated in the urbostratozem. In the ubostratozem (sample 2) MCT 1 and MCT 3 were subdominants, occupying positions of frequently encountered microorganisms. Frequency, abundance and high numbers indicate a relatively stable, though impoverished in the number of species cenosis.

Wide diversity and improvement of the structure of the bacterial cenosis in the urbostratozem of industrial zone (sample 3) compared with previous soils were observed. MCT 1 and MCT 2 were dominants. Also in this sample a significant number of rare and accidental MCTs were found, which characterizes this soil as the most favorable for the formation of wide diversity of bacteria.

The fungi were well adapted to the conditions of all the studied soils (Table 5). The structure of the fungal communities of the urbostratozems was close to each other and differed significantly from the natural soil due to the emergence of new MCT. Only MCT 2 and MCT 4 were common to all the studied soils. The most favorable structure of the fungal cenosis was observed in the humus horizon of the urbostratozem of the Polish Garden (sample 2). The horizon was characterized by a significant number of dominants, subdominants, and also rare species. The high fungal species diversity in urban soils was explained by the heterogeneity of their properties (Marfenina et al. 2017). The higher diversity of bacterial and fungal communities in uban parks than in the forest was found by other researchers (Hui et al. 2017).

The soil of the industrial zone (sample 3) was characterized by a not very favorable structure of the fungal cenosis (the presence of one dominant and a small number of rare MCTs). There was a low diversity of fungi (the presence of two dominants and four subdominants) in the soddy podzol (sample 1).

Analysis of similarity coefficients (Table 6) showed that the greatest difference in the structure of the bacterial and fungal communities was observed between the natural soil of the recreational zone (sample 1) and the urbostratozem of the industrial zone (sample 3). Presumably, this is due to the fact that under the conditions of intense technogenic impact, new dominant MCT appeared. This phenomenon is consistent with conclusions of other researchers (Bityukova et al., 2016), which showed that in urbanozems species resistant to urban environmental conditions were predominate.

Algae are considered to be the most sensitive to anthropogenic effects. In polluted soils algal communities are degraded. The number and diversity of cenoses decreases, the presence of unicellular algae is characteristic (Zenova and Shtina 1991). In the studied soils, algae were found only in the soil of the Polish Garden where their number was 500 CFU/g of dry soil. The algae found there belong to the Ochrophyta and Chlorophyta, and are unicellular (Fig. 2). Algae were absent in the soddy podzol, likely as a result of a strongly acidic reaction of soil solution. The absence of algae in Electrodepo soil could be explained by the greater dryness of the soil and the more intense insolation of its surface. Indicator role of diatoms, which was shown in some publications (Dorokhova 2007), in our study was not revealed due to differences in the physicochemical characteristics of soils not related to pollution.

Enzymatic activity. The number of producers of the enzymes was determined for indirect assessment of the soil enzymatic activity (Table 7).

Table 3. The number of bacteria and fungi in the soils

Sample	Bacteria, CFU* 10^4/ g of dry soil	Fungi, CFU* 10^3/ g of dry soil
1	15.4 ±5.65	53.1±3.63
2	561.3±42.9	9.0±2.9
3	312.2±24.0	75.3±3.63

**CFU- colony-forming unit*

Table 4. Structure of bacterial community in the soils

MCT	Sample 1		Sample 2		Sample 3	
	Abundance, %	Frequency, %	Abundance, %	Frequency, %	Abundance, %	Frequency, %
1	66.6	100	25	100	46.1	100
2	33.3	50	50	100	42.3	100
3	0	0	25	80	3.8	40
4	0	0	0	0	2.56	80
5	0	0	0	0	1.28	40
6	0	0	0	0	2.56	40
7	0	0	0	0	0.6	20
8	0	0	0	0	0.6	20

Table 5. Structure of fungi community in the soils

MCT	Sample 1		Sample 2		Sample 3	
	Abundance, %	Frequency, %	Abundance, %	Frequency, %	Abundance, %	Frequency, %
1	12.5	33.3	3.7	33.3	0	0
2	6.25	33.3	25.9	66.6	2.21	100
3	6.25	33.3	0	0	0	0
4	31.25	66.6	11.1	33.3	2.21	33.3
5	37.5	66.6	3.7	33.3	0	0
6	6.25	33.3	0	0	0	0
7	0	0	14.81	100	0	0
8	0	0	14.81	66.6	0	0
9	0	0	14.81	66.6	0	0
10	0	0	7.4	66.6	0	0
11	0	0	3.7	33.3	3.53	66.6
12	0	0	0	0	1.76	100
13	0	0	0	0	90.2	100

Table 6. Similarity coefficients of Sørensen for bacterial/fungal communities for the soils of different functional zones

Samples	Sample 1	Sample 2	Sample 3
Sample 1	-	0.80/0.53	0.40/0.36
Sample 2	0.80/0.53	-	0.54/0.43
Sample 3	0.40/0.36	0.54/0.43	-

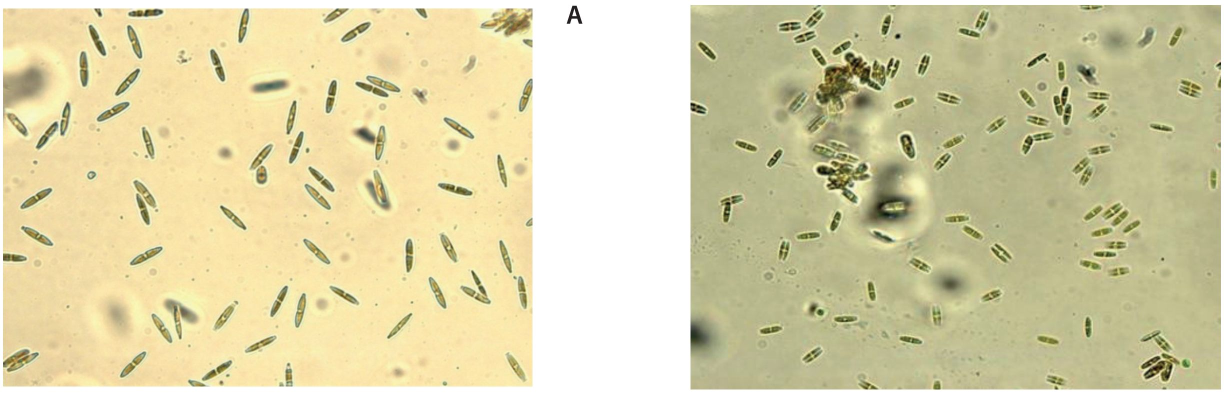


Fig. 2. Representatives of Ochrophyta algae in sample 2: a - *Navicula Bory*; b - *Pinnularia Ehrenberg*

Table 4. Structure of bacterial community in the soils

Number of enzyme producers	Sample 1		Sample 2		Sample 3	
	CFU*10^4/ g of dry soil	%	CFU*10^4 /g of dry soil	%	CFU*10^4/g of dry soil	%
Cellulase producers	1.33+0.15		1.16+0.13		2.36+0.19	
Lipase producers	165.0+0.97		175.0+0.97		180.0+0.89	
Protease producers						
Active	3.3+0.65	11.1	56.6+2.61	45.9	186.0+17.2	74.6
Weak	27.0+1.72	88.9	66.6+2.35	54.1	63.3+2.84	25.4
Amilase producers						
Very active	16.6+1.30	19.3	53.3+2.35	37.2	14.3+1.73	48.8
Active	33.0+0.65	3.8	36.6+0.65	25.6	73.3+0.65	25.0
Weak	66.0+0.65	76.9	53.3+1.13	37.2	76.6+3.26	26.2

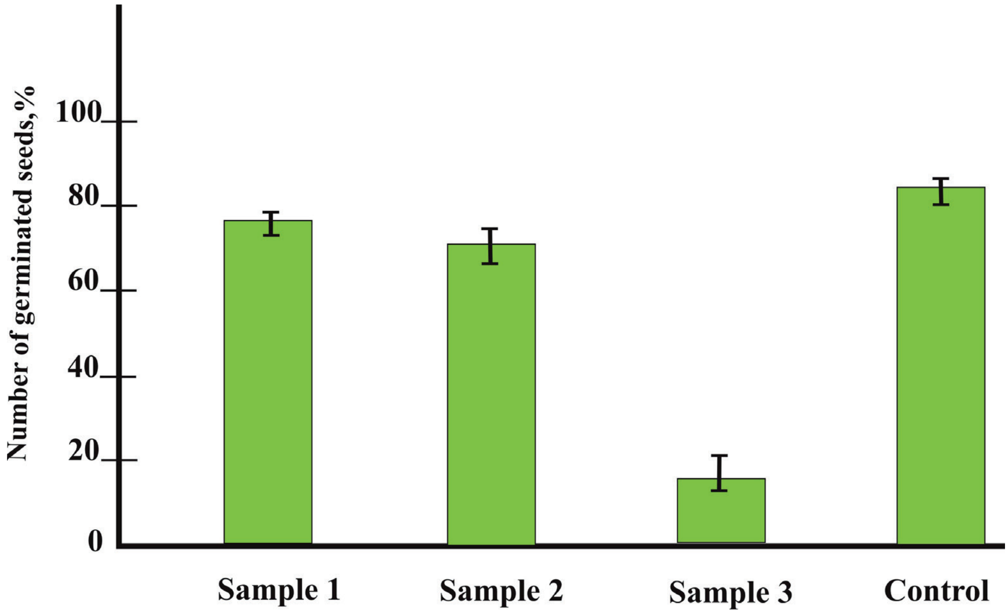


Fig. 3. Assessment of phytotoxity: number of germinated seeds of cress in soil samples against control

Table 8. Medium length of roots and stems of cress on different soils

Sample	Medium length of root, mm	Medium length of stem, mm
1	18.5+1.5	3.5+0.5
2	5.0+0.5	3.0+0.5
3	6.0+0.5	2.0+0.5
control	25.0+3.5	4.0+0.5

The minimum number of producers of all enzymes, except for cellulase, was observed in the soddy podzol in the recreational zone (sample 1) due to low concent of organic matter and acid pH. Both characteristics are typical for natural soils of the region. The maximum number of protease and amylase producers was found in the soil of the industrial zone (sample 3). A significant number of microorganisms with high enzymatic activity in the soil of the industrial zone may be due to the fact that in this soil there is a maximum content of organic carbon, an optimal pH value and a high supply of nutrients (table 1), which is consistent with data on the determining role of these soil properties for microorganisms (Marfenina 2005). The soil of the residential zone (sample 2) was intermediate in enzymatic activity. Lipolytic activity was almost the same in all samples. Differences in the enzymatic activity of two urban soils can also be explained by differences in the structure of bacterial and fungal communities (table 5, 6). The analysis of potential enzymatic activity confirmed our assumptions about a significant difference in the biological parameters of soils in different functional zones of the city.

Phytotoxity. The germination of cress (Lepidium sativum L.) seeds was used for assessment of the soil phytotoxicity. Seed germination, the average length of the root and the stem in the experiment compared with the control were choosen as the criteria for phytotoxicity. Seed germination was found to be the best indicator of phytotoxicity (Fig.3). Inhibition of seed development was observed in all soil samples studied, but the soil of the industrial zone was the most phytotoxic (sample 3) which was also consistent with our hypothesis. In addition to a 4-fold decrease in seed germination, a 5-fold decrease in the length of the root and a 2-fold decrease in the length of the stem were observed in this soil in comparison with the control (Table 8).

CONCLUSIONS

In this study, the urban soils of residential and industrial zones of Saint Petersburg were compared with the natural soil of recreational zone. The urban soils of residential and industrial zones differ from the natural soil by neutral and alkaline reaction and higher content of organic carbon. Both urban soils were polluted by heavy metals (Pb, Cu and Zn). Heavy metal pollution was more pronounced in the soil of the industrial zone. The soil of the industrial zone differs in the structure of fungal and bacterial communities and in potential enzymatic activity from natural soil and from the soil of the residential zone as well. Contrary to the initial hypothesis about the negative effect of heavy metal pollution on soil microbiota, a greater number of morpho-cultural types of fungi and the highest number of microorganisms producing cellulase, amylase and protease were found in the soil of the industrial zone. Thus, assessment of the enzymatic activity of the soils by culturing of microorganisms-producers of protease, amylase, cellulase and lipase on special media used in the study was not sufficient to demonstrate the negative effect of the increased concentration of heavy metals in urban soils. However, in soils with a high content of heavy metals (Polish Garden and Electrodepo) a decrease in cress seed germination and a sharp reduction in the length of roots in cress seedlings were found. Therefore, phytotesting seems a promising and easy to use method for the comparative assessment of soils contaminated with heavy metals. The culture method used in our study did not allow to identify the response of the soil microbial community to heavy metal pollution. Therefore, it is necessary to use alternative methods to search for indicators of the activity and diversity of microorganisms in the environmental assessment of urban soils contaminated with heavy metals.

REFERENCES

Aleksandrovskiy A., Dolgikh A. and Aleksandrovskaya E. (2012). Pedogenetic features of habitation deposits in ancient towns of European Russia and their alteration under different natural conditions. Boletin de la Sociedad Geologica Mexicana, 64(1), 71-77. DOI: 10.18268/BSGM2012v64n1a6.

Bityukova V., Vlasov D., Dorokhova M., Kasimov N., Kislyakova N., Kirillov P., Kosheleva N., Nikiforova E., Petukhova N., Ryzhov A., Savoskul M., Saul'skaya T. and Shartova N. (2016). East – West of Moscow: a spatial analysis of social and environmental issues. Ed. by Kasimov N. Moscow: Faculty of Geography, Lomonosov Moscow State University.

Bouchez T., Blieux A., Dequiedt S., Domaizion I., Dufresne A., Ferreira S., Godon J., Hellal J., Joulian C., Quaiser A., Martin-Laurent F., Mauffret A., Monier J., Peyret P., Schmitt-Koplin P., Sibourg O., D'oiron E., Bispo A., Deportes I., Grand C., Cuny P., Maron P. and Ranjard L (2016). Molecular microbiology methods for environmental diagnosis. Environmental Chemistry Letters, 14, 423-441. DOI: 10.1007/s10311-016-0581-3.

Braun B., Böckelmann U., Grohmann E. and Szewzyk U. (2006). Polyphasic characterization of the bacterial community in an urban soil profile with in situ and culture-dependent methods. Applied Soil Ecology, 31, 267-279. DOI:10.1016/j.apsoil.200505.003.

Burghardt W., Morel J.and Zhang G.-L. (2015). Development of the soil research about urban, industrial, traffic, mining and military areas (SUITMA). Soil Science and Plant Nutrition, 61, supl. 1, 3-21. DOI: 10.1080/00380768.2015.1046136.

Calvarro L.M., de Santiago-Martin A., Gomez G.Q., Gonsallez-Huecas C., Quintana J.R., Vázquez A. et al. (2014). Biological activity in metal contaminated calcareous agricultural soils: the role of the organic matter composition and the particle size distribution. Environmental Science and Pollution Researche, 21(9), 6176-6187. DOI: 10.1007/s11356-014-2561-0.

Castaldi S., Butigliano F. andVirzo de Santo A. (2004) Suitability of soil microbial parameters as indicators of heavy metal pollution. Water, Air, and Soil Pollution, 158, 21-35. DOI: 10.1023/B:WATE.0000044824.88079.d9.

Charzynski P., Markiewicz M. and Switoniak M., eds. (2013). Technogenic soils atlas. Torun: Polish Society of Soil Science.

Classification System of Russian Soils (2004). Moscow: Pochv. Inst. im. Dokuchaeva. (in Russian).

Dai J., Becquer T., Rouiller J., Reversat G., Bernhardt-Reversat F. and Lavelle P. (2004). Influence of heavy metals on C and N mineralization and microbial biomass in Zn-Pb-Cu- and Cd-contaminated soils. Applied Soil Ecology, 2005, 99-109. DOI: 10.1016/j.apsoil.2003.09.003.

Dorokhova M. (2007) Diatoms as indicators of soil conditions in oil production regions. Oceanological and Hydrobiological Studies (Poland), 36, suppl.1, 129-135.

Hagmann D.F., Goodey N. M., Mathieu C., Evans J., Aronson M.F.J., Gallagher F. and Krumins J.A. (2015). Effect of metal contamination on microbial enzymatic activity in soil. Soil Biology and Biochemistry, 91, 291-297. DOI: 10.1016/j.soilbio.2015.09.012.

Huot H., Joyner J., Córdoba A., Shaw R., Wilson M., Walker R., Muth T. and Cheng Z. (2017). Characterizing urban soils in New York City: profile properties and bacterial communities. Journal of Soils and Sediments, 17, 393-407. DOI: 10.1007/s11368-016-1552-9.

Hui N., Jumpponen A., Francini G., Kotze D.J., Liu X., Romantschul M., Srömmer R. and Setal H. (2017). Soil microbial communities are shaped by vegetation type and park age in cities under cold climate. Environmental Microbiology 19(3), 1281-1295. DOI: 10.1111/1462-2920.13660.

Hygienic Standard 2.1.7.2042-06 (2006). Approximate permissible concentrations of chemical substances in soil. (in Russian).

Greinert A. (2015). The heterogeneity of urban soils in the light of their properties. Journal of Soils and Sediments, 15, 1725-1737. DOI: 10.1007/s11368-014-1054-6.

IUSS Working group WRB (2006). World reference base for soil resources. World Soil Recourses Reports № 103. FAO, Rome.

IUSS Working Group WRB (2014). World Reference Base for Soil Resources 2014. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.

Ivashchenko K., Ananyeva N., Vasenev V., Sushko S., Seleznyova A. and Kudeyarov V. (2019). Microbial C-availability and organic matter decomposition in urban soils of megapolis depend on functional zoning. Soil Environ. 38(1), 31-41. DOI: 10.25252/SE/19/61524.

Labutova N.(2008). Methods of studing of soil microorganisms. SPb.: izd. SPbSU. (in Russian).

Labutova N. and Bankina T. (2013). Basics of soil enzymology. SPb.: izd. SPbSU. (in Russian).

Lehmann A. and Stahr K. (2007). Nature and significance of anthropogenic urban soils. Journal of Soils and Sediments, 7(4), 247-260. DOI: 10.1065/jss2007.06.235.

Marfenina O. (1987). Micromycetes as indicators of soil pollution by heavy metals. In: The influence of industry on environment. Moscow: Nauka, 189-196. (in Russian).

Marfenina O. (2005). Anthropogenic ecology of soil fungi. Moscow: Meditsina dlya vsekh. (in Russian).

Marfenina O., Lysak L., Ivanova A., Glushakova A., Kachalkin A., Nikolaeva V., Karlsen A. and Tepeeva A. (2017). Biodiversity in urban soils: threats and opportunities (on the example of cultivated microorganisms). Soils of Urban, Industrial, Traffic, Mining and Military Areas (SUITMA 9) 22-26 May 2017, Moscow. Abstracts, 109-111.

Matinian N.N., Reimann K., Bakhmatova K.A. and Rusakov A.V. (2007). The background concentration of heavy metals and As in arable soils of North-West of Russia (based on the materials of international geochemical atlas). Vestnik of Saint Petersburg state university, Seriya Biologia, 3, 123-134. (in Russian).

Microbiology laboratory manual (2005). Ed.: Netrusov A.I. Moscow: Academia publishing. 608 p. (in Russian).

Muchlbachova G., Sagova-Maeckova M., Omelka M., Szakova J. and Tlustos P. (2015). The influence of soil organic carbon on interactions between microbial parameters and metal concentrations at a long-term contaminated site. Science of the Total Environment, 502, 218-223. DOI: 10.1016/j.scitoenv.2014.08079.

Nannipieri P., Ascher J., Ceccherini M.T., Landi L., Pietramellara G. and Renella G. (2003). Microbial diversity and soil function. European Journal of Soil Science, 54, 655-670. DOI: 10.1046/j.1365-2389.2003.00556.x.

Naylo A., Almeida Pereira S.I., Benidire L., El Khalil H., Castro P.M.L., Ouvrard S., Schwartz C. and Boularbah A. (2019). Trace and major element contents, microbial communities, and enzymatic activities of urban soils of Marrakech city along an anthropization gradient. Journal of Soils and Sediments, 19, 2153-2165. DOI: 10.1007/s11368-018-2221-y.

Papa S., Bartoli G., Pellegrino A. and Fioretto A. (2010). Microbial activities and trace element contents in an urban soil. Environment Monitoring Assessment, 165, 193-203. DOI: 10.1017/s10661-009-0938-1.

Piotrowska-Dlugosz A. and Charzyńsky P. (2015). The impact of the soil sealing degree on microbial biomass, enzymatic activity, and physicochemical properties in the Ecranic Technosols of Toruń (Poland). Journal of Soils and Sediments, 15, 47-59. DOI: 10.1007/s11368-014-0963-8.

Prokof´eva, T., Gerasimova M., Bezuglova O., Bakhmatova K., Gol'eva A., Gorbov S., Zharikova E., Matinyan N., Nakvasina E. and Sivtseva N. (2014). Inclusion of soils and soil-like bodies of urban territories into the Russian soil classification system. Eurasian Soil Science, 47(10), 959-967. DOI: 10.7868/S0032180X14100104.

Prokof"eva T. and Poputnikov V. (2010). Anthropogenic transformation of soils in the Pokrovskoe-Streshnevo Park (Moscow) and adjacent residential areas, Eurasian Soil Science, 43 (6), 701-711. DOI: 10.1134/S1064229310060116.

Rabotnova I. and Pozmogova I. (1979). Chemostat cultivation and inhibition of microbial growth. M.: izd. Nauka. (in Russian).

Rozanova M., Prokof"eva T., Lysak L. and Rakhleeva A. (2016). Soil organic matter in the Moscow State University botanical garden on the Vorob'evy Hills. Eurasian Soil Science, 49 (9), 1013-1025. DOI: 10.1134/S106422931609012X.

Schindelbeck R., van Es H., Abawi G., Wolfe D., Whitlow T., Cugino B., Idowu O. and Moebius-Clune B. (2008). Comprehensive assessment of soil quality for landscape and urban management. Landscape and Urban Planning, 88, 73-80. DOI: 10.1016/j.landurbplan.2008.08.006.

Stefanovicz A., Kapusta P. and Szarek-Lukaszewska G. (2010). Pine forest and grassland differently influence the response of soil microbial communities to metal contamination. Science of the Total Environment, 408, 6134-6141. DOI: 10.1016/j.scitotenv.2010.08.056.

Stefanovicz A.., Kapusta P., Szarek-Lukaszewska G., Groudzińska K., Niklińska M. and Vogt R.(2012). Soil fertility and plant diversity enhance microbial performance in metal-polluted soils. Science of the Total Environment, 439, 211-219. DOI: 10.1016/j.scitotenv.2012.09.030.

Szegi J. (1983). Methods of soil microbiology. Moscow: "Kolos" Publ. (in Russian).

Terekhova V., Pukalchik M. and Yakovlev A.(2014). The triad approach to ecological assessment of urban soils. Eurasian Soil Science, 47(9), 952-958. DOI: 10.1134/S1064229314090129.

Vorobieva L., ed. (2006). Theory and practice chemical analysis of soils. Moscow: izd. MGU. (in Russian).

Yuangen Y., Campbell C., Clark L., Cameron C. and Paterson E. (2006). Microbial indicators of heavy metal contamination in urban and rural soils. Chemosphere, 63, 1942-1952. DOI: 10.1016/j.chemosphere.2005.10.009.

Yuangen Y., Paterson E. and Campbell C. (2001). Urban soil microbial features and their environmental significance as exemplified by Aberdeen City, UK. Chinese Journal of Geochemistry, 20(1), 34-44. DOI: 10.1007/BF03166847.

Zenova G., Shtina E., Dedysh S.et al. (1995). Ecological connections of algae in biogeocoenoses. Microbiology, 64(2), 149-164. (in Russian).

Zhigareva T., Ratnikov A., Sviridenko D., Popova G., Petrov K., Kas'yanenko A., Chernykh N. and Kartuzova M. (2006). Investigation of behavior of Cd and Zn and their influence on soil microbocenoces. Vestnik of RUDN. Ser. Ecology and life safety, 1(13), 34-40. (in Russian).

Zvyagintsev D., Bab'eva I. and Zenova G. (2005). Soil biology. Moscow: izd. MGU. (in Russian).

Received on June 15th, 2020 Accepted on January 21th, 2020

Nataliia V. Terekhina[1*], Margarita D. Ufimtseva[1]
[1] Institute of Earth Sciences, Saint-Petersburg State University, 33/35, 10 line, Vasileostrovsky Island, Saint-Petersburg, 199178, Russia
***Corresponding author:** n.terehina@spbu.ru

LEAVES OF TREES AND SHRUBS AS BIOINDICATORS OF AIR POLLUTION BY PARTICULATE MATTER IN SAINT PETERSBURG

ABSTRACT. Accumulation of chemical elements by leaves of trees and shrubs in urban (Central District of St. Petersburg) and background habitats were studied. To determine proportion of pollutants accumulating on the surface of leaves, chemical content of washed and unwashed leaves were analyzed. The results of the study showed that big part (19-62%) of pollutants is deposited on the surface of leaves of urban lindens, and only 10% on the surface of leaves from background places. Average difference between quantity of particulate matter for them is 4 times. *Tilia cordata* and *Ulmus laevis* has the highest value of ash content between washing and washing leaves. The level of contamination (Kk) showed high values for Fe (8.83), Co (7.47), Cr (5.62), Pb (4.31), Zn (3.04) for unwashed leaves of urban lindens; for the washed leaves this index slightly increased only for Fe (3.12) and Pb (2.13). Accumulative ability depends on the structure of leaf blade of each species, and the ecological situation of the habitat. *Ulmus laevis, Tilia cordata, Populus* sp., and Rosa rugosa accumulate more pollutants, and can be recommended for protective green plantings. *Tilia cordata*, as the most common species in the city green spaces, can be used as an indicator of the level of atmospheric pollution.

KEY WORDS: phytoindication, air pollution, tree leaves, heavy metals

CITATION: Nataliia V. Terekhina, Margarita D. Ufimtseva (2020) Leaves of trees and shrubs as bioindicators of air pollution by particulate matter in Saint Petersburg. Geography, Environment, Sustainability, Vol.13, No 1, p. 224- 232
DOI-10.24057/2071-9388-2019-65

INTRODUCTION

The growth of urbanization is accompanied by an increase in atmospheric air pollution, the source of which is thermal power plants, industrial enterprises and, often to a greater extent, emissions of motor vehicles. This is especially pronounced in countries where the control and management of environmental impact processes are at a low level. Air pollution negatively affects public health, thereby worsening the socio-economic situation in cities (Ghio et al. 2001).

In St. Petersburg over the past 30 years, especially since 2009, the emissions of pollutants into the atmospheric air have significantly increased, which was accompanied by an increase in the incidence of the population. The Central District, characterized by a high density of urban development and large road traffic flows, has the highest non-carcinogenic risk index for the period 2010–2015 among all districts of the city (Movchan et al. 2018). The negative influence of particulate pollution on public health and on the urban trees considered in the article Beckett K.P. et al. (1998). For the assessment and monitoring of air pollution in St. Petersburg, there is an automated system that includes 25 automatic stations (3 of them are in Central District), 2 stationary observation posts, and 3 mobile laboratories. The stations are located in all 18 administrative districts of St. Petersburg, providing daily information on the level of air pollution in the city. Recorded data for a number of chemical compounds: CO, NO, NO_2, SO_2, O_3, PM_{10} – particulate matter with a diameter of less than 10 microns, PM2.5 – particulate matter with a diameter of less than 2.5 microns. The average annual concentration of PM10 in the central part of the city over the past 10 years had declined from 1.1 to 0.2 MPC – maximum permissible concentration of pollutants in the air of populated areas (GN 2.1.6.2604-10); the maximum single concentrations of PM10 vary from 2.1 to 1.0 MPC (Ecological portal of St. Petersburg 2019). More detailed information on the quality and quantity of particulate matter is not available.

Indirect assessment of air pollution is obtained by measuring the amount of pollutants in the snow cover. In the works of Nesterov EM et al. (2009), Vorontsova A.V., Nesterov E.M. (2012), Vorontsova A.V. (2013), Yufereva L.M. (2013) it is shown that the high concentrations of chemical elements (Zn, Cu, Fe, V), which are accumulated in the snow cover of the Central District, do not exceed the MPC for chemicals in the water of water bodies for drinking and cultural and domestic water use. It is also noted that in winter the atmospheric air is cleaner.

The shortcomings of the available methods for assessing the degree of atmospheric pollution make us look for other approaches. Phytoindication is one of such methods that allows obtaining information on the qualitative characteristics and chemical composition of particles deposited on the surface of plant leaves (Tomašević et al. 2005; Sæbø et al. 2012; Nowak et al. 2013; Yang et al. 2015). To do this, samples of leaves are divided into two parts, one of which is washed. The difference in the concentrations of chemical elements in unwashed and washed leaves will show what part of them was deposited on the surface of the leaves. We chose this method, despite the fact that it has a number of difficulties in application. Overview of problems related with interaction between urban vegetation and particle air pollution is presented in article by Janhäll S. (2015). In particular, there are different ways to wash leaves: mechanical cleaning, washing through solvents, weak acid solutions, but also sample washing with distilled water (Ugolini et al. 2013). Previous studies have shown that these approaches may give different results (Alfani 1996; Aksoy et al. 1997; Palmieri et al. 2005; Tomašević et al. 2005; De Nicola et al. 2008; Dzierżanowski et al. 2011). The washing procedure prior to chemical analysis is a critical point in biomonitoring studies, since removal of particles from a leaf surface strongly depends on the washing treatment and its duration (Tomašević et al. 2011). Another problem with the use of this method is that differences in the accumulation of pollutants are determined by the structure of the leaf surface of different plant species, leaf area, and their life cycle (evergreen or deciduous plants) (Popek et al. 2013; Ugolini et al. 2013). Anićić M. et al. (2011) note the elements content in the leaves does not reflect atmospheric deposition directly. However, many scientists recommend using urban woody plants for phytoindication and remediation (Janhäll 2015; Yang et al. 2015; Bargagli et al. 2019).

Our previous data (Ufimtseva, Terekhina 2014, 2017) give an idea of soil and plant contamination in the district, but do not reflect the contribution of atmospheric pollutants, therefore the goal of this study are to find out what proportion of air pollutants is included in the structure of leaves of different trees and shrubs, and which is deposited on their surface; also which species should be used to assess the quality of the urban environment, and which ones should be recommended for the creation of protective green spaces.

MATERIALS AND METHODS

Objects of research in the Central District of St. Petersburg are the tree species widely used in city landscaping: *Tilia cordata* (10 samples), *Populus* sp. (4), as well as less common trees and shrubs: *Quercus robur* (1), *Ulmus laevis* (2), *Syringa vulgaris* (1), *Cotoneaster lucidus* (1), *Berberis vulgaris* (1), *Syringa josikaea* (1), *Rosa rugosa* (1 sample). In total, 22 samples of urban plants were studied. Shrubs grew on the Field of Mars (Marsovo Polye), woody plants grew in different types of urban planting: gardens, parks, street plantings. Sampling was carried out in August 2006, when for a long time there was no rain. Medium samples of leaves without petioles were taken, at a height of 1-2 m from the soil surface along the entire perimeter of crowns from 3-5 neighboring trees or shrubs. Samples were divided into 2 parts, one part was washed in in running tap water for 10 s. We assume that a certain amount of substances can be washed out of leaf tissues when washing, but we hope that these are not significant values. Then the samples were placed in paper bags with labels and dried to air dry condition.
Primary sample preparation of plant samples was carried out according to generally accepted methods (Guidelines 1972). Ashing of plant samples was carried out by the dry method in a muffle furnace at a temperature of 400-450 ° C. The ash content was calculated using the formula: % ash = (M / p) * 100, where M – the weight of the ash, g; p – the weight of the sample of dry matter, g.

The determination of the chemical composition of the plant ash was carried out in the spectral analysis laboratory of the Russian Geological Research Institute (VSEGEI) by the atomic emission method with inductively coupled plasma (ICP AES) for the following elements: Fe, Mn, Cu, Zn, Pb, Ni, Cr, Co, Cd, Ba, Sr.

The results of spectral analyzes were processed statistically. The indicators of washout for each element were calculated: the difference between the content of the element in unwashed and washed leaves, expressed as a percentage of the weight of unwashed leaves. Also, the washout index was calculated for the content of ash substances in the leaves. Comparing heavy metal contents of washed and unwashed plants, a paired t-test was performed in order to determine significance of difference between them. To characterize the intensity of accumulation of chemical elements by plants and the intensity of transport and industrial pollution of the studied areas, a concentration coefficient (Kk) was calculated representing the ratio of the chemical element content in a plant growing in the studied area to its content in plants of background conditions. As a in regional background residential habitats for *Tilia cordata*, 4 samples were taken in small settlements of the Leningrad Region; for other species, the average values of the chemical elements were calculated using our previously obtained data on the chemical composition of the leaves of trees in the background habitats and literature data (Paribok et all. 1982; Ufimtseva, Terekhina 2005; Drozdova et al. 2015).

The total indicator of pollution, commonly used to assess the state of the soil (Saet et al. 1988), was calculated: $Zc = \Sigma Kk / (n-1)$, where n – the number of elements for which Kk> 1, and included in the calculation.

Data were processed by using the software Statistica. Difference on heavy metals deposition between washed and unwashed leaves within each site assessed by t-test for independent samples. Descriptive statistics calculated using Microsoft Excel.

RESULTS AND DISCUSSION

Values of ash content and concentrations of chemical elements in leaves of Tilia cordata from Central District of St. Petersburg and background places are presented in table 1.

Ash content as an indicator of the content of inorganic leaves in leaves varies from 6.88 to 16.03% for unwashed leaves and from 6.9 to 10.18% for washed leaves of city lindens. Moreover, the minimum values of both washed and unwashed leaves belong to sample № 9 taken from a healthy tree in Mitropolichiy garden, protected from highways by buildings. The maximum values of ash content for washed and unwashed leaves belong to sample № 7, taken on Embankment of the Obvodny Canal, geographically not very remote from the Mitropolichiy garden, but subject to heavy traffic load. In the first case, the difference between washed and unwashed leaves is not significant, and in the second it is 5.85% – almost a third of the ash content of unwashed leaves. The difference between the ash content of unwashed and washed leaves of city lindens is on average 2.44%, for leaves of background linens – 0.66%, which indicates the presence of a significant amount of ash substances on the surface of the leaves of plants in the city. There is not much information in the literature about changes in the ash content when working with washed and unwashed leaves. The article of Aksenova Yu.E. (2017) provides data for poplar in Tomsk (11 samples): the value of ash content for unwashed leaves varies from 14.07 to 26.6%, for washed leaves is 12.88-18.87%. The difference between the mean values is 2.58%, which corresponds with our data.

The average values of the content of chemical elements in the leaves of *Tilia cordata* presented in fig. 1, they show the highest concentrations for Fe, Mn, Zn, Cu. The amount of such biophilic element as Mn in leaves of background conditions is higher than in urban areas, while all other elements in urban plants are higher than in background ones, which indicates a violation of the zonal biological circulation of trace elements due to anthropogenic effects. Paired t-test demonstrated significance of difference between washed and unwashed leaves for all investigated elements excluded Sr. For a number of chemical elements (Ni, Cu, Zn, Pb) the results are similar to the data for Nevsky Prospect (Slepyan, 1997). Our data are consistent with data by unwashed and washed leaves of *Tilia tomentosa* in Cracow (Czaja et al. 2014), and *Tilia platyphyllos* from south-west Poland (Piczak et al. 2003), but significantly superior data by *Tilia cordata* for Huszlew and Lyublin (Chwil et al. 2015).

Table 1. Concentrations of chemical elements (μg*g^{-1} of dry matter) and ach content (%) in unwashed and washed leaves of *Tilia cordata* from Central District of St. Petersburg and from background places in Leningrad region.

№	ash, %	Fe	Mn	Cr	Ni	Cu	Zn	Pb	Cd	Sr	Ba	Co
Tilia cordata in city, unwashed leaves												
1	9.37	1098.27	49.01	4.54	1.70	13.59	31.78	5.55	0.06	53.44	40.27	0.49
2	10.32	1577.20	48.87	5.39	1.96	16.28	39.92	6.45	0.07	52.57	41.75	0.76
3	12.32	3040.68	63.91	7.73	2.83	24.63	79.07	9.93	0.14	47.54	51.36	1.77
4	9.26	1431.02	40.87	4.18	1.76	14.63	40.27	5.93	0.09	32.96	28.42	0.69
5	8.93	1130.73	26.98	2.93	1.46	13.22	36.98	5.48	0.10	39.12	30.19	0.54
6	12.47	3653.77	64.70	5.42	2.71	25.68	62.34	13.84	0.08	32.54	57.48	1.58
7	16.03	4203.73	85.65	8.13	5.08	25.48	98.09	15.74	0.23	52.41	72.77	2.24
8	11.59	1466.63	32.30	4.39	2.79	14.25	39.97	7.88	0.12	49.93	34.87	0.77
9	6.88	437.94	31.98	1.73	1.53	8.88	21.06	2.91	0.06	47.07	22.16	0.20
10	14.74	3061.78	52.51	5.87	3.01	21.08	87.41	10.08	0.10	45.69	59.11	1.62
average	11.19	2110.18	49.68	5.03	2.48	17.77	53.69	8.38	0.11	45.33	43.84	1.07
SD	2.80	1268.25	18.07	1.96	1.09	5.98	26.27	4.02	0.05	7.84	16.00	0.68
Tilia cordata in city, washed leaves												
1	8.53	298.42	33.05	1.72	0.83	8.70	17.66	2.59	0.05	44.12	23.47	0.15
2	8.94	350.02	22.15	1.87	0.72	7.54	21.72	2.96	0.04	43.70	15.37	0.18
3	8.40	352.34	18.86	1.91	0.54	7.00	19.98	2.54	0.03	31.07	9.49	0.19
4	7.60	403.91	21.78	1.71	0.78	8.59	24.85	2.96	0.08	27.36	12.54	0.19
5	8.20	321.10	20.32	1.41	0.93	10.74	28.53	2.83	0.06	60.01	18.94	0.13
6	8.74	384.89	17.59	1.83	0.68	9.08	21.14	2.94	0.03	27.87	10.92	0.23
7	10.18	398.60	29.95	1.68	1.07	7.62	25.44	3.08	0.05	38.27	11.09	0.22
8	10.04	245.71	17.88	1.47	0.86	7.54	21.98	4.64	0.11	61.63	18.27	0.15
9	6.90	188.22	33.14	1.20	1.02	7.80	17.25	1.99	0.06	49.89	17.46	0.08
10	9.95	452.42	20.81	1.58	0.77	7.37	26.77	3.01	0.04	43.79	14.23	0.23
average	8.75	339.56	23.55	1.64	0.82	8.20	22.53	2.95	0.06	42.77	15.18	0.18
SD	1.07	79.19	6.11	0.23	0.16	1.11	3.78	0.68	0.02	12.13	4.38	0.05
Tilia cordata in country, unwashed leaves												
11	7.53	110.56	69.96	1.16	0.73	4.37	14.75	1.69	0.04	20.70	13.40	0.10
12	8.32	372.41	238.42	0.78	1.62	6.17	15.23	2.09	0.06	44.26	57.66	0.20
13	6.41	250.88	74.42	0.94	0.95	5.09	21.52	2.34	0.02	35.36	63.29	0.12
14	8.12	221.59	660.64	0.70	3.56	8.77	19.17	1.65	0.05	72.30	53.05	0.14
average	7.59	238.86	260.86	0.89	1.71	6.10	17.67	1.94	0.04	43.16	46.85	0.14
Tilia cordata in country, washed leaves												
11	7.51	89.30	75.62	1.26	0.87	4.87	15.10	1.46	0.06	19.98	14.72	0.11
12	7.53	121.08	221.52	0.53	1.29	5.55	12.65	1.60	0.04	42.60	42.60	0.09
13	6.00	75.52	78.98	0.67	0.50	5.20	20.88	1.12	0.03	43.13	71.57	0.09
14	6.69	149.80	648.00	0.58	3.34	7.63	17.60	1.38	0.07	78.98	54.55	0.10
average	6.93	108.93	256.03	0.76	1.50	5.81	16.56	1.39	0.05	46.17	45.86	0.10

Notes. Addresses of sampling points: 1 – Tavricheskiy Sad, center, 2 – The middle edge of Tavricheskiy Sad along Kirochnaya St., 3 – Preobrazhenskaya Square, 4 – Klenovaya St., the alley on the left when driving from Inzhenernaya St., outer side, 5 – the same place, inside, 6 – Garden of the Winter Palace, roadside, 7 – Embankment of the Obvodny Canal, 7A, 8 – Mitropolichiy Sad, 9 – the same place, 10 – Novgorodskaya St., 1

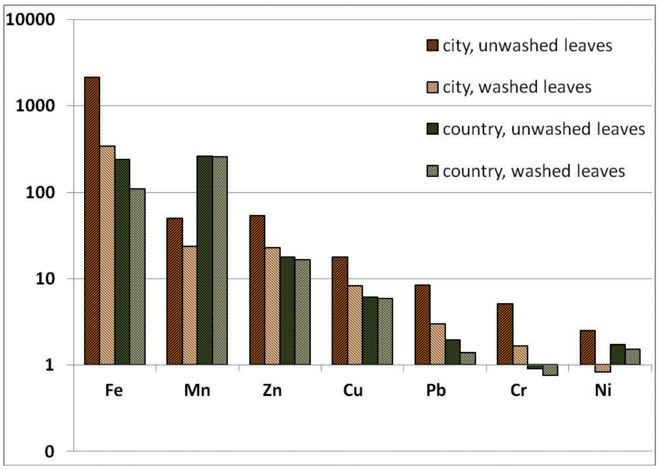


Fig. 1. Average values of the content of chemical elements in the leaves of *Tilia cordata* in urban and background habitats (mg / kg of dry matter; logarithmic scale)

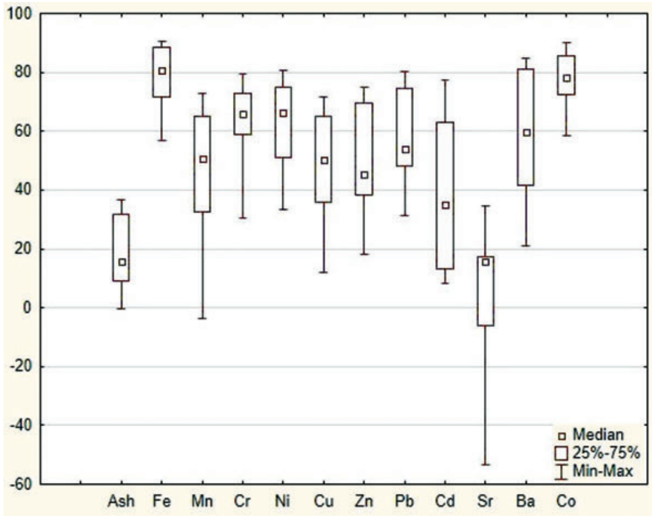


Fig. 2. Average percentages of washout of chemical elements from *Tilia cordata* leaves for urban habitats

Washout – the difference between the concentrations of chemical elements of unwashed and washed leaves of *Tilia cordata*, expressed as a percentage of the weight of these elements in unwashed samples, – is very significant. For all elements except Sr, it is close to 50% or exceeds this value (Fig. 2). This means that most of the chemical elements are on the surface of the leaves and washed away from it. The predominant elements washed out from the leaves of urban lindens are Fe (78.8%), Co (78.5%), Cr (62.9%), Ni (61.8%), Ba (59.0%), Pb (58.2%), which are the main pollutant elements. Biophilic elements have slightly lower, but also high washout values: Zn (49.8%), Mn (46.8%), Cu (48.0%), due to their technogenic penetration. Sr, despite the high content in the leaves, varies greatly in index of washing, so its average value is the lowest (4.9%). For all elements except Sr, the difference in the t-test of comparing means is significant between washed and unwashed leaves.

As for the background plants, the comparison of washed and unwashed leaves shows low values of the percentage of elements washout (an average of 10%). The maximum values are Fe (47.3%), Pb (26.4%), Co (26.4%), Cr (17.5%), Ni (13.5%), which is probably due to sampling within settlements with anthropogenic stress. In some samples, the excess of elements such as Cd, Sr, Mn in the washed leaves over their content in unwashed leaves is within the limits of the analysis error and indicates that these elements are not pollutants in the background conditions.

According to the data (Tomašević et al. 2011), washout indices were calculated, based on which it can be concluded that the average difference in the concentrations of chemical elements between washed and unwashed leaves of *Tilia cordata*, gathered in three parks of Belgrade, is 26.3 %, which is 2 times less than our indices, but the washed elements are represented mainly by the same pollutants: Fe (49.7%), Cr (38.1%), Pb (30.8%), Ni (21.5%).

Coefficient of concentration Kk (Table 2) was used to assess the contamination of leaves of urban plants. High average values are characteristic for unwashed leaves of Tilia cordata in Fe (8.83), Co (7.47), Cr (5, 62), Pb (4.31), Zn (3.04). The total indicator of pollution Zc varies from 1,7 to 7.91 with average value 4.59. For washed leaves such high values of Kk are not observed; slightly higher values are only for Fe (3.12) and Pb (2.13), Zc varies from 1.57 to 2.9, with average value 2.34, which confirms the fact that most of the pollutants are contained in dust-like particles on the leaf surface.

Values of ash content and concentrations of chemical elements in leaves of different trees and shrubs are presented in table 3. Ash content in leaves of various plants is different and depends on species, character of leaves surface and ecological situation of habitat (Weerakkody et al. 2018). The differences of this indicator for the studied species are presented in Fig. 3, but statistically they are not supported due to the insufficient number of samples. *Populus* sp. has the highest value of ash content as for unwashed (14.4%) and washed leaves (14.6%). It is related with xeromorphic nature of leaves, therefore big difference between ash content for unwashed and washed leaves is absent. *Ulmus laevis* and *Rosa* rugosa have rough surface of leaves that leads to accumulation of dust; other reason for *Ulmus* – these trees grow on the streets with heavy traffic. *Tilia cordata* also has the high value of ash content, but usually it is related with wide leaves covered excreta of insects contributed to sticking dust. The next species *Cotoneaster lucidus* has hard xeromorphic leaves and relatively high ash content. Other species has ash content about 6-7%. The biggest difference between these figures for unwashed and washed leaves (19.2%) is for *Ulmus laevis* and *Tilia cordata*. For two species Syringa *josikaea* and *Berberis vulgaris* this difference is negative, because their habitat – on The Field of Mars – was relatively clean place, more than 10 m away from the roadways, with good blowing.

Content of Fe, Cr, Ni, Cu, Pb, Ba is maximum for leaves of *Ulmus laevis* and *Rosa rugosa*. Zn, Cd, Sr have the maximum values for leaves of *Poplar* sp., this is the species features of accumulation. For all elements, except Zn, Cd, and Sr, the difference in the t-test of comparing means is significant between washed and unwashed leaves of investigated species.

The literature contains data on the accumulation of particulate matter by leaves of woody plants in different countries: for Poland and Norway (Sæbø et al. 2012), for USA (Nowak et al. 2013), for Serbia (Urošević et al. 2019), for Iran (Norouzi et al. 2015; Kardel et al. 2018), for China (Chen et al. 2016; Liu et al. 2017); also a review of this subject is presented in an article by Yang, J., Chang, Y., Yan P. (2015). The series of plant species that characterize the decrease in the accumulation of suspended particles vary greatly depending on the local urban flora, so their use in different regions requires additional research.

Comparison of the obtained data with the content of chemical elements in a "Reference plant" (Markert 1992) showed that the Fe content for all samples exceeds the baseline level, which is explained by a significant anthropogenic load. Low level of Mn is observed in leaves of urban trees and shrubs and demonstrates violation of the ratio Fe / Mn – the excess Fe intake leads to a deficiency of Mn (Kabata-Pendias

Table 2. Coefficient of concentration and total index of contamination for unwashed and washed leaves of *Tilia cordata*

№	ash	Fe	Mn	Cr	Ni	Cu	Zn	Pb	Cd	Sr	Ba	Co	Zc
unwashed leaves													
1	1.12	4.60	0.19	5.08	0.99	2.23	1.80	2.86	1.39	1.24	0.86	3.43	3.23
2	1.18	6.60	0.19	6.03	1.14	2.67	2.26	3.32	1.61	1.22	0.89	5.29	3.77
3	1.11	12.73	0.24	8.64	1.65	4.04	4.48	5.11	3.23	1.10	1.10	12.40	6.05
4	1.00	5.99	0.16	4.67	1.03	2.40	2.28	3.05	2.17	0.76	0.61	4.86	3.78
5	1.08	4.73	0.10	3.27	0.85	2.17	2.09	2.82	2.22	0.91	0.64	3.78	3.51
6	1.15	15.30	0.25	6.06	1.58	4.21	3.53	7.13	1.89	0.75	1.23	11.07	6.50
7	1.34	17.60	0.33	9.08	2.96	4.18	5.55	8.10	5.27	1.21	1.55	15.69	7.91
8	1.32	6.14	0.12	4.91	1.63	2.34	2.26	4.06	2.82	1.16	0.74	5.39	3.84
9	0.91	1.83	0.12	1.93	0.90	1.45	1.19	1.50	1.47	1.09	0.47	1.40	1.70
10	1.31	12.82	0.20	6.56	1.75	3.45	4.95	5.19	2.37	1.06	1.26	11.34	5.64
average	1.15	8.83	0.19	5.62	1.45	2.91	3.04	4.31	2.45	1.05	0.94	7.47	4.59
SD	0.14	5.31	0.07	2.19	0.63	0.98	1.49	2.07	1.15	0.18	0.34	4.75	1.86
washed leaves													
1	1.23	2.74	0.13	0.26	0.55	1.50	1.07	1.86	0.94	0.96	0.50	1.49	2.16
2	1.29	3.21	0.09	0.28	0.48	1.30	1.31	2.13	0.93	0.95	0.33	1.82	2.44
3	1.21	3.23	0.07	0.29	0.36	1.20	1.21	1.83	0.66	0.67	0.20	1.99	2.36
4	1.10	3.71	0.09	0.26	0.52	1.48	1.50	2.13	1.70	0.59	0.27	1.94	2.49
5	1.18	2.95	0.08	0.22	0.62	1.85	1.72	2.04	1.33	1.30	0.40	1.30	2.08
6	1.26	3.53	0.07	0.28	0.45	1.56	1.28	2.11	0.63	0.60	0.23	2.33	2.71
7	1.47	3.66	0.12	0.26	0.71	1.31	1.54	2.22	1.10	0.83	0.24	2.28	2.42
8	1.45	2.26	0.07	0.22	0.57	1.30	1.33	3.34	2.23	1.33	0.39	1.58	2.23
9	1.00	1.73	0.13	0.18	0.68	1.34	1.04	1.44	1.22	1.08	0.37	0.85	1.57
10	1.44	4.15	0.08	0.24	0.51	1.27	1.62	2.16	0.91	0.95	0.30	2.39	2.90
average	1.26	3.12	0.09	0.25	0.55	1.41	1.36	2.13	1.16	0.93	0.32	1.80	2.34
SD	0.16	0.73	0.02	0.03	0.11	0.19	0.23	0.49	0.49	0.26	0.09	0.50	0.36

et al. 2001). For some samples, an excess of Cr, Co and Cd is noted. Poplar leaves accumulate Zn, which is its species feature. Paired t-test demonstrated significance of difference between washed and unwashed leaves for all investigated elements excluded Zn, Cd, and Sr.

Comparison of the contents of chemical elements in the leaves of linden and poplar, taken at the same points, showed that in general unwashed leaves of linden contain 2-4 times more elements (especially Fe, Cr, Pb, Ba); washed leaves of both species contain approximately the same amount of elements (except for Zn, Co, Cd, which in both cases prevail in poplar leaves). Thus, significantly more particulate matters are deposited on the surface of linden leaves than on the surface of poplar leaves, due to the fact that, as mentioned above, the sticky secretions of insects living on lindens serve as a suitable surface for the accumulation of pollutants.

Indicators of the percentage of substances washed away from the surface of leaves of investigated species of trees and shrubs are presented in Fig. 4. The average values of these indicators for all studied plants for each chemical element are arranged in the following row: Fe (63.8%), Co (57.4%), Cr (37.3%), Pb (36.7%), Ni (35.7%), Cu (34.7%), Ba (28.9%), Zn (25.2%), Mn (17.9%), Cd (11.6%), Sr (3.7%). The maximum values of the average washout index for all investigated chemical elements were noted in *Tilia cordata* (53.6%) and *Ulmus laevis* (52.3%). *Rosa rugosa* (37.2%) and *Quercus robur* (35.6%) are significantly behind them. Minimum washout value is for *Cotoneaster lucidus* (15.6%), with smooth and hard leaves. For other species, this index varies from 21 to 28%. Comparison of our data with the data of other researchers yielded the following results: for a number of evergreen plant species in Palermo (Italy) (Olivia et al. 2004), the percentage of washout for Cr is similar to ours, but for other elements we have much higher rates. In the city of Amman (Jordan), the washout for Pb, Cd, Zn, Cu is higher for *Pinus eldarica* than ours (Al-Alawi et al. 2007), and for Bojnourd (Iran) washing the leaves for *Fraxinus* excelsior removed of Cu, Zn, Pb about 20–46% (Solgi et al. 2020) which is similar to our data. Probably, these indicators depend both on species features and on the degree of air pollution in these cities.

Coefficients of concentration and total index of contamination for unwashed and washed leaves of different species of investigated trees and shrubs are presented in table 4. Pb,

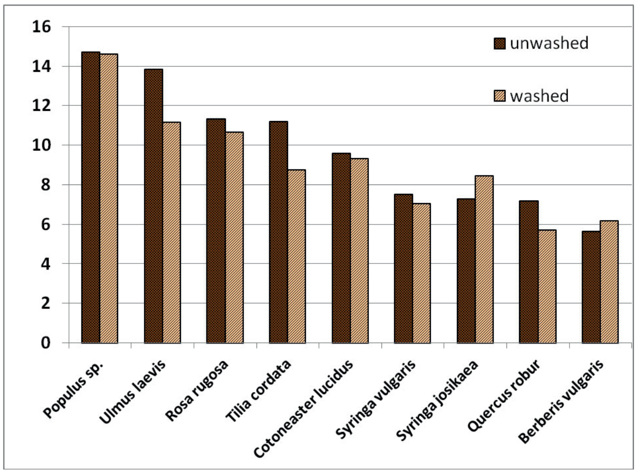


Fig. 3. Ash content (% of dry matter) in unwashed and washed leaves of investigated plants

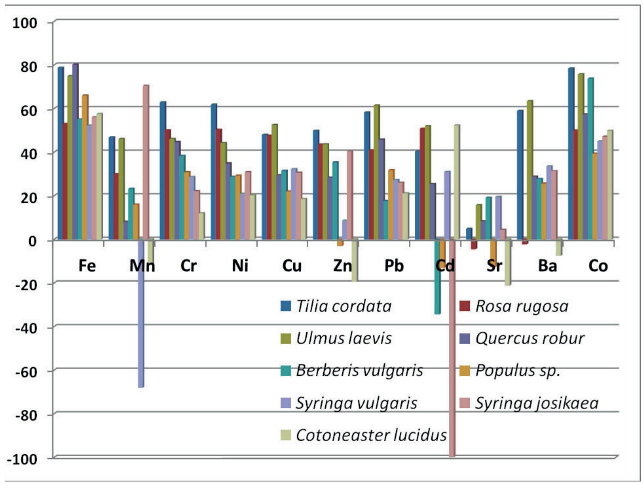


Fig. 4. Percentage of chemical elements, which were washed away from the leaves of different plant specie

Note: Negative values of washout show that the content of chemical elements in washed leaves is higher than in unwashed ones.

Table 3. Concentrations of chemical elements (μg*g^{-1} of dry matter) and ach content (%) in unwashed and washed (w) leaves of shrubs and trees from Central District of St. Petersburg

№	ash, %	Fe	Mn	Cr	Ni	Cu	Zn	Pb	Cd	Sr	Ba	Co
Syringa vulgaris												
15	7.49	330.20	19.15	1.43	0.54	9.07	45.56	2.38	0.07	20.31	12.81	0.15
15w	7.04	157.62	32.18	1.02	0.42	6.14	41.62	1.73	0.05	16.34	8.52	0.08
Cotoneaster lucidus												
16	9.58	535.85	29.67	2.05	0.73	9.51	29.78	3.06	0.06	36.20	52.29	0.26
16w	9.29	227.52	33.11	1.80	0.59	7.74	35.60	2.42	0.03	43.87	56.23	0.13
Berberis vulgaris												
17	5.63	480.58	43.62	1.63	0.52	10.64	23.15	2.17	0.02	23.60	13.52	0.20
17w	6.18	216.09	33.50	1.01	0.37	7.29	14.95	1.79	0.03	19.09	9.76	0.05
Syringa josikaea												
18	7.26	604.63	73.14	2.11	0.84	11.55	73.37	2.96	0.02	33.85	31.96	0.25
18w	8.45	265.88	21.59	1.64	0.58	8.01	43.76	2.19	0.05	32.36	21.96	0.13
Rosa rugosa												
19	11.31	2247.08	44.69	5.67	2.35	19.57	43.67	7.22	0.10	36.54	48.98	1.11
19w	10.64	1056.34	31.30	2.83	1.17	10.26	24.68	4.27	0.05	38.18	50.10	0.55
Ulmus laevis												
20	13.89	3051.15	60.26	6.35	4.46	32.23	86.97	11.31	0.11	34.04	54.46	1.43
20w	10.88	684.79	27.81	2.52	1.83	12.18	47.54	3.98	0.05	25.02	17.95	0.33
21	13.74	2134.14	61.74	4.14	2.89	19.38	65.01	9.00	0.14	43.57	43.85	1.04
21w	11.45	592.66	38.14	2.82	2.04	11.08	37.79	3.77	0.07	41.45	17.63	0.26
Quercus robur												
22	7.16	756.56	22.19	2.47	1.01	7.16	24.36	3.67	0.04	12.54	17.84	0.31
22w	6.77	411.97	15.73	1.75	0.72	5.50	18.28	2.55	0.03	11.78	12.59	0.19
Populus sp.												
23	11.58	753.30	31.39	2.21	2.29	11.93	141.29	3.76	0.39	29.19	17.26	0.89

23w	10.45	329.03	31.58	1.56	0.79	9.21	189.23	2.61	0.51	33.98	12.23	0.79
24	12.38	1740.43	41.23	3.89	3.42	16.96	137.42	6.36	0.44	45.06	38.01	1.14
24w	12.19	392.29	27.39	1.90	2.38	12.44	141.44	3.06	0.52	49.51	21.22	0.56
25	13.89	388.47	37.64	2.58	1.21	10.08	144.41	3.74	0.49	55.26	18.61	0.48
25w	13.90	174.95	33.36	2.56	1.45	9.48	145.91	2.65	0.51	63.79	17.09	0.35
26	20.96	1055.43	56.81	3.42	1.84	16.41	440.14	6.54	1.18	92.43	44.85	0.69
26w	21.78	258.95	45.54	1.99	1.11	11.04	326.69	5.51	1.15	101.06	35.28	0.23
Regional background		272.88	215.90	0.74	1.65	5.72	31.02	0.84	0.07	33.20	25.00	0.25
Reference plant (by Markert, 1992)		150	200	1.5	1.5	10	50	1	0.05	50	40	0.2

Notes. Addresses of sampling points: 15-19 – Field of Mars, 20 – Admiralty Embankment near the Palace Bridge, 21 – Garden of the Winter Palace, roadside, 22 – Field of Mars, 23 – Preobrazhenskaya Square, 24 – Embankment of the Obvodny Canal, 7A, 25 – Mitropolichiy Sad, 26 – Novgorodskaya St., 1

Table 4. Coefficients of concentration and total index of contamination for unwashed and washed (w) leaves of different species of trees and shrubs (bold text indicates the highest values for unwashed leaves)

species	Fe	Mn	Cr	Ni	Cu	Zn	Pb	Cd	Sr	Ba	Co	Zc
Syringa vulgaris	1.21	0.09	1.94	0.33	1.59	1.47	2.83	1.02	0.61	0.51	0.61	2.61
Syringa vulgaris (w)	0.58	0.15	1.39	0.26	1.07	1.34	2.06	0.70	0.49	0.34	0.34	2.39
Cotoneaster lucidus	1.96	0.14	2.78	0.45	1.66	0.96	3.64	0.78	1.09	**2.09**	1.03	2.68
Cotoneaster lucidus (w)	0.83	0.15	2.45	0.36	1.35	1.15	2.87	0.37	1.32	2.25	0.52	2.10
Berberis vulgaris	1.76	0.20	2.22	0.32	1.86	0.75	2.58	0.31	0.71	0.54	0.80	2.81
Berberis vulgaris (w)	0.79	0.16	1.37	0.23	1.28	0.48	2.12	0.42	0.58	0.39	0.21	2.38
Syringa josikaea	2.22	0.34	2.86	0.51	2.02	**2.37**	3.51	0.32	1.02	1.28	1.02	2.85
Syringa josikaea (w)	0.97	0.10	2.23	0.35	1.40	1.41	2.60	0.64	0.97	0.88	0.54	2.55
Rosa rugosa	**8.23**	0.21	**7.70**	1.43	**3.42**	1.41	**8.58**	1.35	1.10	**1.96**	**4.42**	**4.44**
Rosa rugosa (w)	3.87	0.14	3.84	0.71	1.80	0.80	5.07	0.66	1.15	2.00	2.21	3.05
Ulmus laevis	**9.50**	0.28	**7.12**	**2.23**	**4.52**	**2.45**	**12.1**	**1.73**	1.17	**1.97**	**4.95**	**5.21**
Ulmus laevis (w)	2.34	0.15	3.63	1.17	2.04	1.38	4.61	0.83	1.00	0.71	1.19	2.56
Quercus robur	2.77	0.10	3.36	0.61	1.25	0.79	4.36	0.54	0.38	0.71	1.25	3.25
Quercus robur (w)	1.51	0.07	2.37	0.44	0.96	0.59	3.03	0.42	0.35	0.50	0.77	3.46
Populus sp.	**3.61**	0.19	**4.11**	1.33	**2.42**	**6.96**	**6.06**	**8.66**	**1.67**	1.19	**3.20**	**4.39**
Populus sp. (w)	1.06	0.16	2.72	0.87	1.84	6.47	4.11	9.35	1.87	0.86	1.92	4.21

Cr, Cu show Kk>1 for unwashed and washed leaves of all species, whereas other elements can have Kk both more and less than 1. It should be noted that the main pollutants of urban soils in the Central Region are Zn, Pb, Cd, Cu (Ufimtseva et al. 2014). Here is also three species – *Ulmus laevis, Populus* sp., and *Rosa rugosa* – demonstrate high values by concentration of many chemical elements and total index of contamination.

CONCLUSIONS

As a result of the research conducted to assess the influence of atmospheric pollution on the chemical composition of leaves of urban plants, the following conclusions are obtained:

– The average values of the difference between the ash content of unwashed and washed leaves for city and background lindens differ by 4 times, which indicates the increasing of particulate matter quantity on the surface of the leaves of urban plants. *Tilia cordata* and *Ulmus laevis* has the highest value of ash content between washing and washing leaves.

– High concentrations of Fe, Zn, Cu and low Mn values for *Tilia cordata* unwashed leaves in city indicate a violation of the biological circulation of microelements due to anthropogenic effects.

– More than half of the content of Fe, Co, Cr, Ni, Ba, Pb was washed off from the leaves of urban *Tilia cordata*. Washout for background *Tilia cordata* leaves averages 10%. High washout values are for *Ulmus laevis* leaves (52.3%), slightly lower – for *Rosa rugosa* (37.2%) and *Quercus robur* (35.6%).

– Sr content in the leaves of all studied species is high, but urban values differ little from background values, that indicates a natural source of its entry into plants. The washout for this element is minimal.

– The level of plants contamination (Kk) showed high values for Fe (8.83), Co (7.47), Cr (5.62), Pb (4.31), Zn (3.04) for unwashed leaves of Tilia *cordata*. For the washed leaves, slightly increased values of Kk were only for Fe (3.12), Cr (2.16) and Pb (2.13). For all other species of plants Pb, Cr, Cu show Kk>1 for unwashed and washed leaves. *Ulmus laevis, Populus* sp., and *Rosa rugosa* accumulate more chemical elements, then other species.

Results of the study showed the most of pollutants – Fe, Co, Cr, Ni, Ba, Pb – deposited on the surface of the leaves and can be washed off into the soil with precipitation. Different types of trees and shrubs accumulate pollutants in different ways, most of all they are accumulated by leaves of *Ulmus laevis, Tilia cordata, Populus* sp., *Rosa rugosa*. One can recommend them for use in green areas, creating protective vegetation strips along the obviously dangerous point and line technical objects, with the aim of improving the ecological situation in general and protecting the health of the local population in particular. *Tilia cordata* can be used as bioindicator of city air pollution by particulate matter. ■

REFERENCES

Aksenova Yu.E. (2017). Estimation of the impact of vehicles on the state of atmospheric air by the method of biogeochemical indication. Proceedings of the conference Problems of Geology and Mineral Resources. Tomsk, 684-686 (in Russian).

Aksoy A., Ozturk M.A. (1997). Nerium oleander L. as a biomonitor of lead and other heavy metal pollution in Mediterranean environments. Science of the Total Environment, 205, 145-150. DOI: 10.1016/S0048-9697(97)00195-2.

Al-Alawi Mu'taz M., Khakhathi L. Mandiwana (2007). The use of Aleppo pine needles as a bio-monitor of heavy metals in the atmosphere. Journal of Hazardous Materials, 148(1-2), 43-46. DOI: 10.1016/j.jhazmat.2007.02.001.

Alfani A., Maisto G., Iovieno P., Rutigliano F.A., Bartoli G. (1996). Leaf contamination by atmospheric pollutants as assessed by elemental analysis of leaf tissue, leaf surface deposit and soil. Journal of Plant Physiology, 148(1-2), 243-248. DOI: 10.1016/S0176-1617(96)80321-X.

Anićić M., Spasić T., Tomašević M., Rajšić S., Tasić, M. (2011). Trace elements accumulation and temporal trends in leaves of urban deciduous trees (Aesculus hippocastanum and Tilia spp.). Ecological Indicators, 11, 824-830. DOI: 10.1016/j.ecolind.2010.10.009.

Aničić Urošević, M., Jovanović, G., Stević, N., Deljanin, I., Nikolić, M., Tomašević, M., Samson, R. (2019). Leaves of common urban tree species (Aesculus hippocastanum, Acer platanoides, Betula pendula and Tilia cordata) as a measure of particle and particle-bound pollution: a 4-year study. Air Quality, Atmosphere and Health, 12(9), 1081-1090. DOI: 10.1007/s11869-019-00724-6.

Bargagli R., Ancora S., Bianchi N., Rota E. (2019). Deposition, abatement and environmental fate of pollutants in urban green ecosystems: Suggestions from long-term studies in Siena (Central Italy). Urban forestry and urban greening, 46, 126483. DOI: 10.1016/j.ufug.2019.126483.

Beckett, K.P., Freer-Smith P.H., Taylor G. (1998). Urban woodlands: their role in reducing the effects of particulate pollution. Environmental Pollution, 99, 347-360. DOI: 10.1016/S0269-7491(98)00016-5.

Chen L., Liu C., Zou R., Yang M., Zhang Z. (2016). Experimental examination of effectiveness of vegetation as bio-filter of particulate matters in the urban environment. Environmental Pollution, 208, 198-208. DOI: 10.1016/j.envpol.2015.09.006.

Chwil S., Kozłowska-Strawska J., Tkaczyk P., Chwil P., Matraszek R. (2015). Assessment of air pollutants in an urban agglomeration in Poland made by the biomonitoring of trees. Journal of Elementology, 20(4), 813-826. DOI: 10.5601/jelem.2015.20.1.742.

Czaja M., Kołton A., Baran A., Muras P. (2014). Influence of urban transport on heavy metals accumulation in the leaves of lime trees in Cracow. Logistyka, 4, 4193-4197.

De Nicola F., Maisto G., Prati M.V., Alfani A. (2008). Leaf accumulation of trace elements and polycyclic aromatic hydrocarbons (PAHs) in Quercus ilex L. Environmental Pollution, 153, 376-383. DOI: 10.1016/j.envpol.2007.08.008.

Drozdova I.V., Alekseeva-Popova N.V., Beliaeva A.I. (2015). The content of heavy metals in the soil-plant system of some specially protected natural areas of St. Petersburg and the Leningrad region. Environmental problems of industrial cities. Proc. of 7th All-Russia scientific-practical conf. with international participation, Part 1, ed. Prof. E.I. Tikhomirova, Saratov, 251-254 (in Russian).

Dzierżanowski K, Gawroński S.W. (2011). Use of trees for reducing particulate matter pollution in air. Challenges of Modern Technology, 1(2), 69-73. DOI: 10.1007/s12273-014-0180-9.

Ecological portal of St. Petersburg Available at: www.infoeco.ru/index.php?id=53 [Accessed 10 May. 2019].

Ghio A, Devlin R. (2001). Inflammatory lung injury after bronchial instillation of air pollution particles. American Journal of Respiratory and Critical Care Medicine, 164, 704-708. DOI: 10.1164/ajrccm.164.4.2011089.

GN 2.1.6.2604-10, Hygienic standards Supplement N 8 to GN 2.1.6.1338-03 «Maximum permissible concentrations (MPC) of pollutants in the atmospheric air of populated areas» (in Russian).

Guidelines for geological survey scale 1: 50 000 Biogeochemical and geobotanical studies. (1972). Leningrad: Nedra (in Russian).

Janhäll S. (2015). Review on urban vegetation and particle air pollution – Deposition and dispersion. Atmospheric Environment, 105, 130-137. DOI: 10.1016/j.atmosenv.2015.01.052.

Kabata-Pendias Alina, Pendias Henryk (2001).Trace elements in soils and plants. 3rd ed. Boca Raton FL: CRC Press LLC. .

Kardel F., Wuyts K., De Wael K., Samson R. (2018). Biomonitoring of atmospheric particulate pollution via chemical composition and magnetic properties of roadside tree leaves. Environmental Science and Pollution Research, 25-26, 25994-26004. DOI: 10.1007/s11356-018-2592-z.

Liu Y., Yang Z., Zhu M., Yin J. (2017). Role of plant leaves in removing airborne dust and associated metals on Beijing roadsides. Aerosol and Air Quality Research, 17, 2566-2584. DOI: 10.4209/aaqr.2016.11.0474.

Markert B. (1992). Establishing of "Reference Plant" for inorganic characterization of different plant species by chemical fingerprinting. Water Air and Soil Pollution, 64, 533-538. DOI: 10.1007/BF00483363.

Movchan V.N., Zubkova P.S., Kalinina I.K., Kuznetsova M.A., Sheinerman N.A. (2018). Assessment and forecast of the ecological situation in St. Petersburg in terms of air pollution and public health indicators. Vestnik SPSU. Earth Sciences, 63(2), 178-193. (in Russian with English summary). DOI: 10.21638/11701/spbu07.2018.204.

Nesterov E.M., Zarina LM, Piskunova M.A. (2009). Monitoring of behaviour of heavy metals in snow and soil covers of the central part of St.-Petersburg. Bulletin of the Moscow Region State University. Series: Natural Sciences, 1, 27-34. (in Russian with English summary).

Norouzi S., Khademi H. (2015). Source identification of heavy metals in atmospheric dust using Platanus orientalis L. leaves as bioindicator. Eurasian Journal of Soil Science, 4(3), 144-152. DOI: 10.18393/ejss.2015.3.144-152.

Nowak D.J., Hirabayashi S., Bodine A., Hoehn R. (2013). Modeled PM2.5 removal by trees in ten U.S. cities and associated health effects. Environmental Pollution, 178, 395-402. DOI: 10.1016/j.envpol.2013.03.050.

Oliva S.R., Valdés B. (2004). Influence of washing on metal concentrations in leaf tissue. Communications in Soil Science and Plant Analysis, 35(11-12), 1543-1552. DOI: 10.1081/CSS-120038553.

Palmieri R.M., La Pera L., Di Bella G., Dugo G. (2005). Simultaneous determination of Cd(II), Cu(II), Pb(II) and Zn(II) by derivative stripping chronopotentiometry in Pittosporum tobira leaves: a measurement of local atmospheric pollution in Messina (Sicily, Italy). Chemosphere, 59, 1161-1168. DOI: 10.1016/j.chemosphere.2004.11.066.

Paribok T.A., Sazykina N.A., Temp G.A., Troitskaya E.A., Leina GD, Chervyakova E.G. (1982). Metal content in the leaves of trees in the city. Botanicheskii Zhurnal, 67(11), 1533-1539. (in Russian).

Piczak K., Leśniewicz A., Zyrnicki W. (2003). Metal concentrations in deciduous tree leaves from urban areas in Poland. Environmental Monitoring and Assessment, 86, 273-287. DOI: 10.1023/A:1024076504099.

Popek R., Gawrońska H., Wrochna M. et al. (2013). Particulate matter on foliage of 13 woody species: deposition on surfaces and phytostabilisation in waxes – a 3-Year Study. International Journal of Phytoremediation, 15, 245-256. DOI: 10.1080/15226514.2012.694498.

Saet Yu.E., Revich B.A. (1988). Ecological and geochemical approaches to the development of criteria for the regulatory assessment of the state of the urban environment. Bulletin of the Russian Academy of Sciences, Geographical Series, 4, 37-46. (in Russian).

Slepyan E.I., Voloshko L.N., Dzyuba O.F. (1997). Flora of Nevsky Prospect and the natural environment of the historical center of St. Petersburg. Life and safety, 2-3, 406-453. (in Russian).

Solgi E., Keramaty M., Solgi M. (2020). Biomonitoring of airborne Cu, Pb, and Zn in an urban area employing a broad leaved and a conifer tree species. Journal of Geochemical Exploration, 208, 106400. DOI: 10.1016/j.gexplo.2019.106400.

Sæbø A., Popek R., Nawrot B., Hanslin H.M., Gawronska H., Gawronski S.W. (2012). Plant species differences in particulate matter accumulation on leaf surfaces. Science of the Total Environment, 427-428, 347-354. DOI: 10.1016/j.scitotenv.2012.03.084.

Tomašević M., Vukmirović Z., Rajšić S., Tasić M., Stevanović B. (2005). Characterization of trace metal particles deposited on some deciduous tree leaves in an urban area. Chemosphere, 61, 753-760. DOI: 10.1016/j.chemosphere.2005.03.077.

Tomašević M., Aničić M., Jovanović Lj., Perić-Grujć A., Ristić M. (2011). Deciduous tree leaves in trace elements biomonitoring: A contribution to methodology. Ecological indicators, 11, 689-695. DOI: 10.1016/j.ecolind.2011.04.017.

Ufimtseva M.D., Terekhina N.V. (2005). Phytoindication of ecological state of urban ecosystems in St. Petersburg. St. Petersburg: Nauka (in Russian with English summary).

Ufimtseva M.D., Terekhina N.V. (2014). Ecological-geochemical assessment of soil condition in historical center of St. Petersburg. Vestnik of Saint-Petersburg University. Earth Sciences, 2, 122-136. (in Russian with English summary). DOI: 10.21638/11701/spbu07.2017.206.

Ufimtseva M.D., Terekhina N.V. (2017). Assessment of ecological status of Central District (Saint-Petersburg) on the basis of ecophytoindication. Vestnik of Saint-Petersburg University. Earth Sciences, 62(2), 209-217. (in Russian with English summary). DOI: 10.21638/11701/spbu07.2017.206.

Ugolini F., Tognetti R., Raschi A., Bacci L. (2013). Quercus ilex L. as bioaccumulator for heavy metals in urban areas: Effectiveness of leaf washing with distilled water and considerations on the trees distance from traffic. Urban Forestry & Urban Greening, 12, 576-584. DOI: 10.1016/j.ufug.2013.05.007.

Vorontsova A.V., Nesterov E.M. (2012). Geochemistry of snow cover in urban environment. Izvestia: Herzen University Journal of Humanities & Science, 147, 125-132. (in Russian with English summary).

Vorontsova A.V. (2013). Features of the behavior of pollutants in the snow cover of St. Petersburg and their impact on the urban environment. PhD Dissertation. St. Petersburg: Herzen University (in Russian with English summary).

Weerakkody U., Dover J.W., Mitchell P., Reiling K. (2018). Evaluating the impact of individual leaf traits on atmospheric particulate matter accumulation using natural and synthetic leaves. Urban Forestry & Urban Greening, 30, 98-107. DOI: 10.1016/j.ufug.2018.01.001.

Yang, J., Chang, Y., Yan, P. (2015). Ranking the suitability of common urban tree species for controlling PM2.5 pollution. Atmospheric Pollution Research, 6(2), 267-277. DOI: 0.5094/APR.2015.031.

Yufereva LM, Gavrilova A.A., Yuferev M.Yu. (2013). Investigation of pollution by heavy metals of the snow cover of garden-park zones in the center of St. Petersburg. Proceedings of Petersburg Transport University, 4, 157-162. (in Russian).

Received on June 16th, 2019 Accepted on January 21th, 2020

Olga B. Popovicheva[1*], Elena Volpert [2], Nikolay M. Sitnikov[3], Marina A. Chichaeva[4], Sara Padoan[5]
[1] SINP Moscow State University, Moscow, Russia
[2] Geographical Faculty, Moscow State University, Moscow, Russia
[3] Central Aerological Observatory, Dolgoprudny, Moscow region, Russia
[4] Peoples Friendship University of Russia (RUDN University), Moscow, Russia
[5] Universität der Bundeswehr München, Neubiberg, Germany
***Corresponding author:** polga@mics.msu.su

BLACK CARBON IN SPRING AEROSOLS OF MOSCOW URBAN BACKGROUND

ABSTRACT. Air quality in megacities is recognized as the most important environmental problem. Aerosol pollution by combustion emissions is remaining to be uncertain. Measurements of particulate black carbon (BC) were conducted at the urban background site of Meteorological Observatory (MO) MSU during the spring period of 2017 and 2018. BC mass concentrations ranged from 0.1 to 10 μg m^{-3}, on average 1.5±1.3 and 1.1±0.9 μg/m^3, in 2017 and 2018, respectively. Mean BC concentrations displayed significant diurnal variations with poorly prominent morning peak and minimum at day time. BC mass concentrations are higher at night time due the shallow boundary layer and intensive diesel traffic which results in trapping of pollutants. Wind speed and direction are found to be important meteorological factors affected BC concentrations. BC pollution rose identifies the North as the direction of the preferable pollution. A negative correlation between BC concentrations and wind speed confirms the pollution accumulation preferably in stable weather days. Relation of BC pollution to a number of agriculture fires is distinguishable by air mass transportation from South and South-Est of Russia and Western Europe. Mean season BC concentrations at rural and remote sites in different world locations are discussed.

KEY WORDS: air quality, black carbon, pollution, megacity

CITATION: Olga B. Popovicheva, Elena Volpert, Nikolay M. Sitnikov, Marina A. Chichaeva, Sara Padoan (2020) Black Carbon In Spring Aerosols Of Moscow Urban Background. Geography, Environment, Sustainability, Vol. 13, No 1, p. 233-143
DOI-10.24057/2071-9388-2019-90

INTRODUCTION

At present, air pollution by the smallest suspended particles (aerosols) is recognized as one of the most important environmental problems. Particulate matter characterized by a small aerodynamic diameter of less than 10 μm (PM10) contains environmentally hazardous components and it is recommended by the World Health Organization (WHO) as an air quality indicator. In accordance with the WHO air quality standards for PM10, a daily mean maximum permissible mass concentration of 50 μg/m^3 is established (WHO 2005). Black carbon (BC) is the light-absorbing component of aerosols, warming the atmosphere and thus affecting the Earth's radiation balance on a global scale (Bond et al. 2013). Estimates suggest that BC is the second most critical factor, following by carbon dioxide, contributing to global warming by direct forcing (Jacobson 2010). BC has drawn a considerable attention because of its environmental significance (Ramachandran and Rajesh 2007; Ahmed et al. 2014; Diapouli et al. 2017).

BC is mainly present in aerosols in urban environment (Mousavi et al. 2018), and thus it concerns local air quality and population health (Janssen et al. 2011). BC-containing particles are small enough to be readily inhaled into the human body and affects the respiratory system leading to exacerbate respiratory, cardiovascular, and allergic diseases, thus relating to adverse health effects (Pope III and Dockery 2006; Steiner et al. 2013). Epidemiological evidence links the exposure to BC with cardiopulmonary hospital admissions and mortality (WHO 2012).

Assessments of the ability for BC to affect the environment and health require a comprehensive study of the aerosol composition as well as a deep understanding of aerosol-related impacts. The comprehensive characterization of aerosols in the urban atmosphere makes it possible to identify sources of pollution and assess the consequences of effects of their emissions on air quality and human health. BC is increasingly recognized as the most important pollution contributor originating from anthropogenic activities, such as fossil fuel combustion (transport, energy production, residential heating) and biomass burning (domestic and wildfires). Diesel exhaust, which comprises high amounts of BC, is classified as a carcinogen for humans by the International Agency for Research on Cancer (IARC). Impact of traditional agriculture biomass burning (BB) activities on regional air quality is a major environmental concern, indicating that it may profoundly affect public health in urban areas (Popovicheva et al. 2017b).

BC emissions are found particularly large in most urbanized Asian cities where the rapid industrial and economic development has been accompanied by serious fine particle pollution of the atmosphere (Ohara et al. 2007; Chen et al. 2014). For US and European cities the source apportionment approach is intensively developed in order to estimate the most significant and dangerous combustion sources (Herich et al. 2011; Healy et al. 2017). An actual review on the status of ambient pollution in global megacities showed that the five most polluted megacities are Delhi, Cairo, Xi'an, Tianjin and Chengdu, all of which had an annual average concentration of PM2.5 greater than 89 μg/m^3 (Cheng et al. 2016). European megacities (London, Moscow and Paris) had much lower annual average concentrations between 18 and 21 µg/m^3. However, these concentrations still are above the air quality guideline set by WHO to 10 µg/m^3 (WHO 2005).

Moscow megacity is the leader among all large cities of Russia on total number of inhabitants. However, for Moscow the situation is currently complicated by the existing lack of the aerosol composition and BC pollution assessments. High anthropogenic emissions of PM in Moscow can occur due to traffic, industry, heating, waste recycling, and construction. Additionally, secondary aerosol formation and long transportation can impact the particulate loading and composition, as it was occurred during an extreme smoke events related to intensive wildfires near the city (Popovicheva et al. 2014; Popovicheva et al. 2019).

However, Environmental Protection Agency "Mosecomonitoring" conducts the measurements for only PM10 and PM2.5 mass concentrations at some monitoring stations. The analysis of measured daily mean PM10 concentrations and calculated ones by chemical transport model CHIMERE showed the high temporal variability reflecting the influence of local sources and atmospheric processes (Kuznetsova et al. 2011; Gubanova et al. 2018). Presently, BC is not a subject of the consistent monitoring, it is remaining as a topic of a few scientific research performed in the Moscow center (Golitsyn et al. 2015; Kopeikin et al. 2018).

The purpose of this study is the analysis of the aerosol pollution related to BC in total PM10 mass in the urban background of the Moscow megacity. Spring season is chosen because this time is characterized by an elevated number of BC sources including the seasonal agriculture fires in surrounding areas and residential BB around the city. We measure and examine the diurnal variability of BC concentrations under real conditions of the diversity of urban sources of emissions and meteorological factors. Wind speed and direction affect BC concentrations while air mass transportation from fire-affected regions support the increased level of BC pollution in spring in the urban background.

MATERIALS AND METHODS

Measurement site

Moscow megacity covers an area of 2 561 km^2 and has a registered population exceeding 12.5 million in 2018. It is the largest city of Russia, and generally represents a typical urban area. Presently, in Moscow megacity 26 gaseous and particulate pollutants (PM10 and PM2.5) are under continuous measurements (Kul'bachevskii 2017). Around 630 industrial enterprises of various branches of mechanical engineering and metal working, power engineering, chemistry and petrochemistry, light and food industry, production of construction materials (including 30 000 stationary emissions sources) are registered. Around 50% of all pollutant emissions from industrial sources are emitted by enterprises producing and redistributing energy, gas, and water. Gaseous Automobile transport exhaust composes 95% of total city emissions. According the reports from the Moscow Committee of environmental protection and natural resources and the Department of the Federal State Statistics Service, industrial production zones occupy around 17% of the city area, from stationary sources fourteen biggest enterprises provide up to 85% of gross pollution (Bityukova and Saulskaya 2017) while transport emissions compose up to 93 % of gross pollution from mobile sources (Kul'bachevskii 2017). In total industrial emissions the processing productions, the production of oil products, and the production and power distribution of gas and water are dominated while gas composes 96.7% of fuel consumption. The biggest fractions of industrial emissions are NO_x (50%), CO (15%,), SO_2 (16%) and 11% of volatile organic compounds. The Moscow city often faces serious traffic congestion problems because of the increased vehicle numbers; the total count of vehicles was registered as much as 4.6 million by the end of 2017 (Kul'bachevskii 2017). High speed of construction in Moscow additionally increases the aerosol loading in the atmosphere.

Black carbon mass concentration measurements were conducted at the rooftop of two–story building of the Meteorological Observatory of Moscow State University (MO MSU). MO is located at the territory of the MSU campus, southwest of the Moscow city (55°07'N, 37°52'E) (Fig. 1).

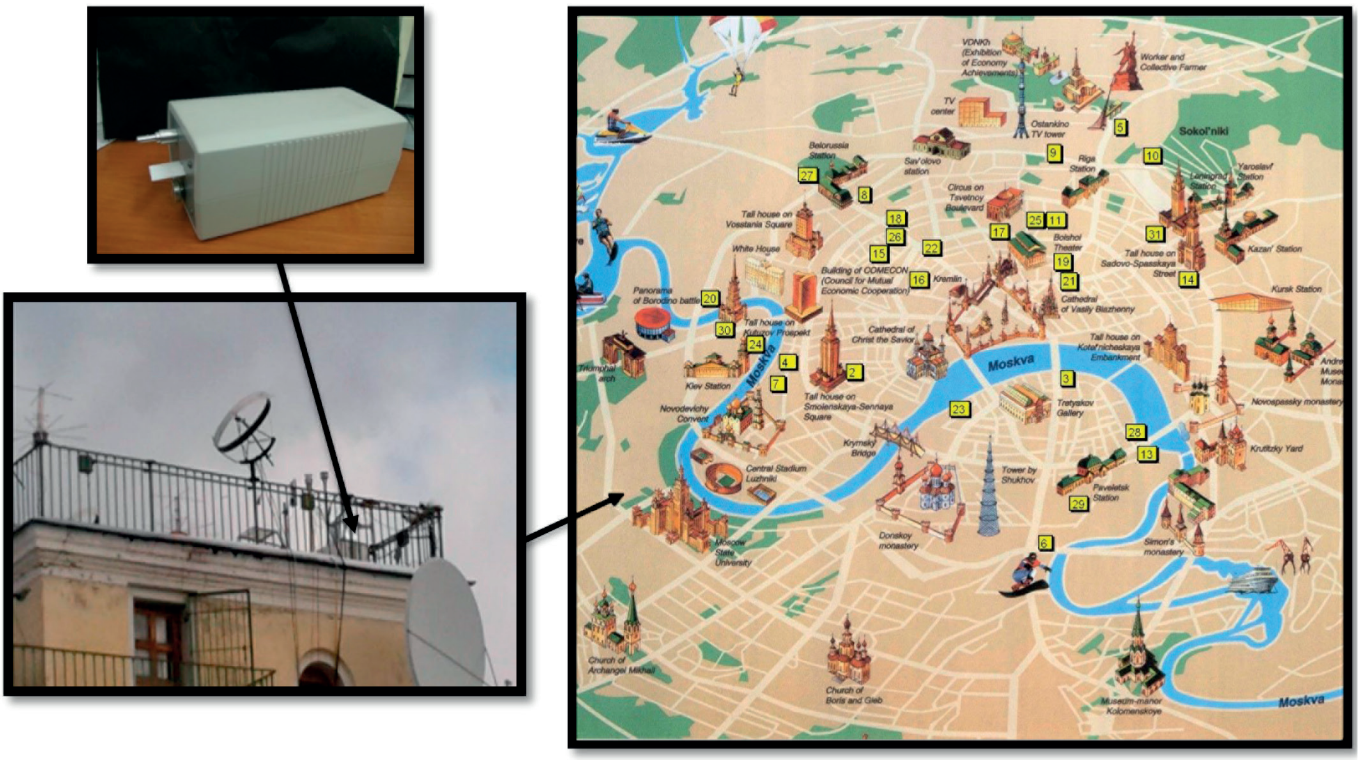


Fig. 1. Location of the portable aethalometer on the roof of Meteorological Observatory (MO) of MSU during spring campaigns of 2017 and 2018. MO MSU is on a Moscow city map

It takes place in an area of hills Vorobievy Gory which is well ventilated. There is no any industrial plants or commercial areas nearby (Fig. 2). At about 800 m to the north from the MO MSU there is the residential area and to the north-west the Lomonosovsky prospect highway is located. The closed industrial enterprises are central and quarter heating stations, and industrial areas which take place at the distance of 3 km and longer from the MSU. Therefore, MO MSU is considered operated as an urban background site.

Data collection

Aerosol equivalent BC concentrations (eBC) were measured using the portable aethalometer custom made by MSU/CAO. In this instrument the light attenuation caused by the particles depositing on a quartz fiber was analyzed at three wavelengths (450, 550, and 650 nm). eBC concentrations were determined by converting the time-resolved light attenuation to EBC mass at 650 nm and characterized by a specific mean mass attenuation coefficient as described elsewhere (Popovicheva et al. 2017a). Calibration parameter for quantification eBC mass was derived during parallel long-term measurements against an AE33 aethalometer (Magee Scientific) that operates at the same three wavelengths.

Attenuation coefficient b_{atn} is defined as

$$b_{atn} = A\delta ATN / V \quad (1)$$

where *A* is the filter exposed area, *V* is the volume of air sampled, and *δATN* is the light attenuation defined as follows:

$$\delta ATN = 1n(I_0 / I) \quad (2)$$

where I_0 and I is the light intensity transmitted through unexposed and exposed parts of the filter, respectively. Good linear correlation between the aethalometer's attenuation coefficient b_{atn} and the eBC concentrations calculated with the AE33 aethalometer (at 660 nm) was achieved ($R^2 = 0.92$). This allowed the estimation of eBC mass concentrations using the regression slope and intercept between b_{atn} at 650 nm and eBC of the AE33 aethalometer at 660 nm:

$$EBC = 3.3 \cdot 10^5 \cdot A \cdot \delta ATN / V \quad (3)$$

where 3.3×10^5 is the correction factor that includes the specific mass absorption coefficient for the MSU aethalometer calibrated against the AE33 aethalometer. Equation (3) assumes the mass absorption cross-section (MAC) adopted by AE33 equal to 9.89 m^2 g, values for A and V used to be in m^2 and m^3. The constant MAC value adopted here is an approximation, assuming a uniform state of mixing for BC in atmospheric aerosol. This can be considered as a valid assumption in case of background aerosol measurements performed in this study. The level of uncertainty of eBC measurements was 30 ng/m^3 for a 6 min integration time. Aethalometer filters were changed manually, when ATN approach the threshold value around 70 that corresponds the twice decreasing of light intensity transmitted.

Measurements were performed from 17 April to 25 May of 2017 and from 19 April to 23 May of 2018. Continuous measurements of meteorological parameters (temperature, relative humidity, pressure, precipitation, wind speed and wind direction) were performed each 3 hours routinely by MO MSU meteorological service. PM10 mass concentrations were collected by Mosecomonitoring using the tampered element oscillating microbalance TEOM 1400a (Thermo Environmental Instruments Inc., USA). These data were used for the comprehensive database of whole sampling runtime.

In spring time fires are usually observed on the South of the Moscow city, the agriculture practice with a purpose to remove the last year grass on the fields is widespread in this season. Biomass burning in residential areas around city is

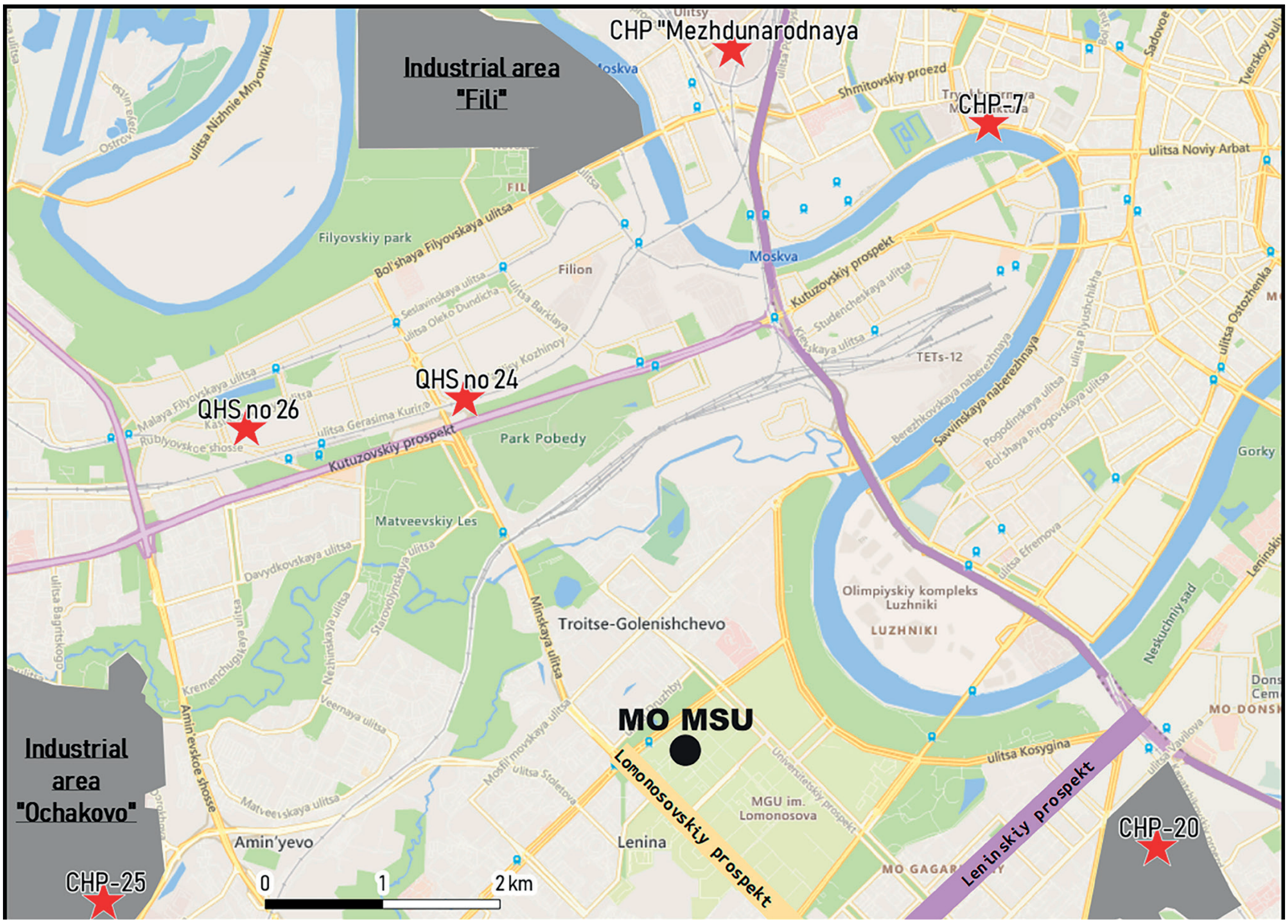


Fig. 2. The map of industrial areas and heating plans in Moscow around MSU. The black dot marked location of MO MSU. CHP – Central Heating Plan, DHS – District Heating Station

pronounced, especially during May holidays from 1 to 10 May. Therefore, air mass arriving to urban area due to long-range transportation may impact the air quality of the city, especially if direction of their transportation well correlates with fire-affected regions. To evaluate the air mass transportation impact the backward trajectories (BWT) were generated using NOAA HYbrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model of the Air Resources Laboratory (ARL) (Stein et al. 2015) with coordinate resolution equal to 1° × 1° of latitude and longitude. The potential source areas were investigated using 2-day backward trajectories for air masses arriving each 12 hours to the MO MSU at 500 m heights above sea level (A.S.L.). Fire information was obtained from Resource Management System (FIRMS), operated by the NASA/GSFC Earth Science Data Information System (ESDIS). It is based on satellite observations which register the open flaming with temperature above 2000 K. Daily maps were related to the computed trajectories, providing a clear picture of the geographical location of fires, with the several kilometer resolutions. A number of fires which could affect air masses transported to the MO MSU was calculated as a sum of fires occurred at distance 0.5° on both latitude and longitude from BTW. The number of fires passed by BWT was estimated as the sum of the amounts of all points of fire caught on the back trajectories points or in their neighborhood of no further than 0.5° along the latitude and longitude. Number of fires indicates the BB-influenced days.

RESULTS AND DISCUSSION

Diurnal BC patterns

The diurnal eBC mass concentrations measured in spring seasons of 2017 and 2018 are shown in Figure 3. On a daily basis of 2017, there is an increase in concentrations starting from 04:00, and the relatively prominent morning peak occurs at 7:00-09:00. eBC concentrations are found to be the smallest during midday until the evening, around 19:00. After this time the gradual increase of eBC is observed, reaching the maximum at 21:00-22:00. During 2018, the mean diurnal BC mass concentration is found to be constant from midnight to 04:00, the poorly noticeable peak occurs at around 05:00, and after the gradual decrease is marked until 20:00. Morning peak similar observed in spring of 2017 is not observed anymore in spring of 2018. The diurnal variation in mean BC mass concentrations is appeared around 1 μg/m^3, approaching maximum 2 and 1.8 and minimum 0.9 and 0.5 μg/m^3 during spring of 2017 and 2018, respectively.

The atmospheric abundance of BC is affected by both the stability of the boundary layer and anthropogenic activities in an urban environment (Ramachandran and Rajesh 2007). The nocturnal boundary layer depth is shallower than its daytime counterpart by a factor of 3–5. The development of a well–mixed layer height begins from 08:00, reaching maximum value around 14:00, decreases after 17:00, and later drops to the lowest level at night. The surface inversion after sunset results in the accumulation of BC, causing even higher concentrations in the late evening. It should be note that in many urban areas the evening peak is happened between 21:00 and 22:00 (Ramachandran and Rajesh 2007; Kozlov et al. 2011; Chen et al. 2014).

In large cites the BC morning peak is well observed for diurnal mass concentrations; it is attributed to the combined influence of the lower mixing layer height and vehicle traffic enhancement in the morning (Ramachandran and Rajesh 2007; Kozlov et al. 2011; Chen et al. 2014). The minimum BC concentrations are found during midday when there are fewer anthropogenic BC emissions, while the deeper boundary layer leads to a faster dispersion resulting in a dilution of BC concentrations at midday.

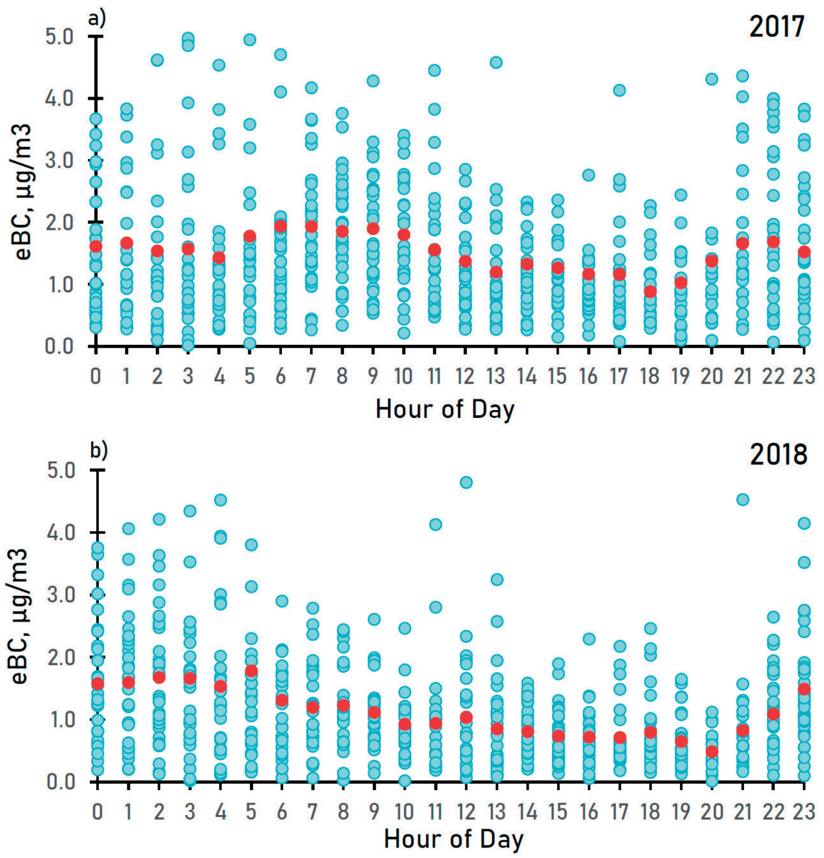


Fig. 3. Diurnal variations of hourly eBC concentrations during the sampling periods of a) 2017 and b) 2018 spring seasons. Plotted red points are the average values for aggregated data from the subsequent one hour on available days

Diesel and petrol are the major fossil fuels used for road transportation, the main part of the diesel vehicles are heavy-duty vehicles. Diesel vehicles are known to produce much more BC than the vehicles which run on gasoline (Weingartner et al. 1997). Heavy duty (e.g., trucks, buses etc.) diesel vehicle can contribute an average 42% BC to particulate matter (Reddy and Venkataraman 2002). Vehicles which run on leaded petrol and unleaded petrol (which do not have catalytic convertors) can give rise to 6 and 23% of BC in PM mass. They include two-wheelers such as scooters, motor bikes in addition to cars. In Moscow, a mixture of all vehicles runs on the roads and contributes to BC. In 2017 the Uniform interdepartmental information and statistical system (www.fedstat.ru) reported around 4640 000 vehicles, including 90.4 and 8.5 % of light and heavy duty cars and trucks, respectively, as well as 1,1 % of buses.

Stable high mean BC concentration from midnight to pre–dawn is observed especially prominent in 2018 (Fig. 3). Such diurnal profile is different from observed in large Russian city Tomsk (Kozlov et al. 2011) and Indian city Ahmedabad (Ramachandran and Rajesh 2007), and in megacity Athens (Diapouli et al. 2017). One reason may be related to increased emissions from diesel trucks (especially after 22:00) which could contribute to the higher BC concentration in the late evening (Garland et al. 2008). Here we should confirm that there is a restriction for heavy duty cargo vehicles to entry the Moscow city center during daytime, they are allowed to bring cargo only after midnight. Therefore, probably stable high BC observed later evening and during night time is due to the significant impact of diesel running vehicle emissions (almost from heavy duty cargo transport) in Moscow megacity.

Relationships between BC and meteorological parameters

The physical and chemical processes of BC emission, transformation, transportation, and accumulation in the atmosphere relate to meteorological conditions, i.e. wind speed and direction, transport of air masses, and the intensity of precipitation. The wind roses at MO MSU in both 2017 and 2018 spring seasons show the relatively homogeneous distribution of the wind direction during the sampling period, with the prevailing south-west direction and with a repeatability frequency up to 10% in this direction (Fig. 4). This finding confirms the observations of the South-West prevailing wind direction in the air layer from 40 to 500 m for the period of 2004–2014 at the MO MSU (Lokoshchenko 2015). The western component of the wind strengthens in spring and summer, that means that the South- Southwest direction is registered more rarely, and the West-Southwest one, more frequently.

The highest frequency of eBC pollution, around 13%, is found during northern wind, according to the eBC pollution roses (Fig. 5), when eBC mass concentrations approached 7 and 5 μg/m^3 in 2017 and 2018, respectively. Other directions of SW (225°), NW (315°), and NE (45°) relating to the high frequency of 6, 5, and 8%, respectively, indicate the most relevant emission source locations. Analysis of the geographical locations on a map shows that in the direction of SW (270–225°) the industrial area "Ochakovo" is located, and two quarter thermal stations take place in the direction of NNW and NW (around 335°), central heating plants can pollute MO MSU from north and NE (45°) (Fig. 2).

Wind speed is an important factor affecting BC concentrations: higher wind speeds contribute to stronger BC dispersion. It is probable that due to higher wind speeds BC produced from local sources are transported to other locations. Such phenomenon is typically observed in polluted urban area (Chen et al. 2014). It was found from BC measurements made in urban sites in Toronto that low wind speeds lead to much less dispersion of BC and the traffic emissions remain concentrated around the emission site (Sharma et al. 2002). It was also noted that as the traffic density was relatively constant through the measurement period, higher wind speeds have a dilution effect on BC. Correlation between wind speeds and BC mass concentrations may give an indication of the proximity of BC sources at the measurement site, while the absence of significant correlation shows that the BC originates from distant sources.

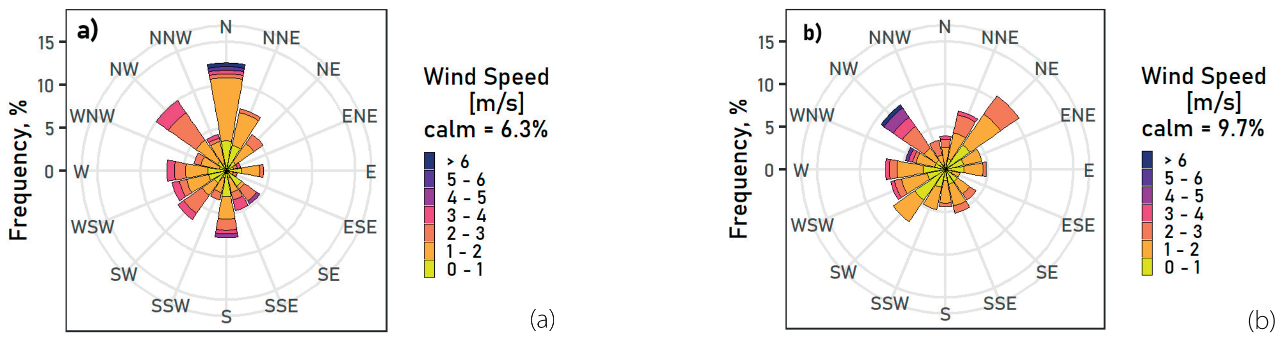


Fig. 4. Wind rose during spring season of a) 2017 and b) 2018

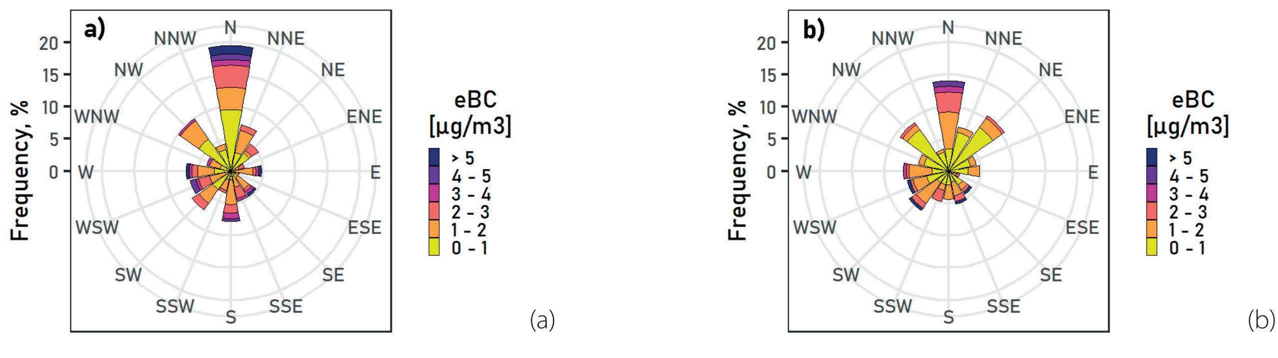


Fig. 5. eBC pollution rose during spring season of a) 2017 and b) 2018

A regression of BC mass concentrations with wind speed at MO MSU during spring seasons yield a negative regression coefficient (slope) -0.51 and -0.41, respectively (Fig. 6). This indicates that with increase in wind speeds the BC mass concentrations would decrease, as seen. Such correlation between wind speeds and BC mass concentrations gives an indication of the intensive BC sources in a megacity.

Seasonal PM10 variations

The time series of PM10 mass concentrations over the whole period of observation is depicted in Fig. 7 and 8 for two-year spring seasons. In 2017, 1h mean PM10 concentrations showed a strong variation from lowest of 8 μg/m^3 up to the highest value of 82 μg/m^3, on average 22±16 μg/m^3. In 2018, the PM10 variations from 2 to 90 μg/m^3 were even more frequent but on short-term, on average 26±16 μg/m^3. These values are comparable to PM10 concentrations averaged over warm season (May-September) obtained from the data for 236 urban background monitors in Europe in the periods of 2003–2006, on averaged amounts to 19.9 g/m^3 (Konovalov et al. 2009). For spring periods of 2014 and 2016 PM10 concentrations was 30.0 and 25.4 μg/m^3, and 22.5 and 17.8 μg/m^3 in the Moscow center and at the suburban station, respectively; the difference of 26% between their average values is caused by an addition from urban sources (Kopeikin et al. 2018).

During spring seasons of our study PM10 concentrations exceeded the daily mean maximum permissible mass concentrations from one to seven times in 2017 and 2018, respectively. The longest episode of the highest PM10 was observed at MO MSU from 30 April until 3 May of 2017, on average 55±19 μg/m^3. Almost at the same days, from 29 April until 2 May the ambient temperature has approached an abnormally high level for this season, + 21°C (Fig. 9). In observation period of 2018, the permissive maximum PM10 values were exceeded seven times. the ambient temperature has approached + 21°C twice, on 1 and 18 May.

Such exceeding of the air quality standards can occur during meteorological conditions unfavorable for pollutant dispersion, passage of cold atmospheric fronts, and arrival of air masses polluted by fires around the city (Kuznetsova et al. 2011). Analysis of location impact showed that the maximum PM10 concentration was observed at the stations in the eastern sector of the Moscow city and less often in the west of Moscow, at the MO MSU station (Kuznetsova et

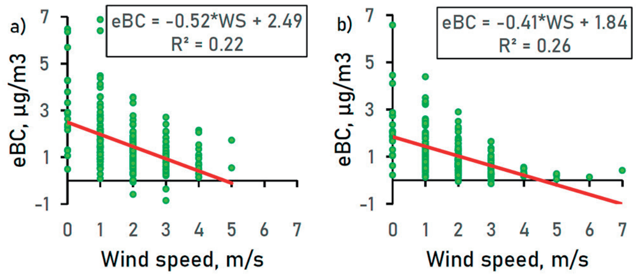


Fig. 6. Correlation of eBC mass concentration and wind speed (WS) during the sampling period of a) 2017, and b) 2018 spring seasons. Negative scope is obtained for linear fit of measurement points

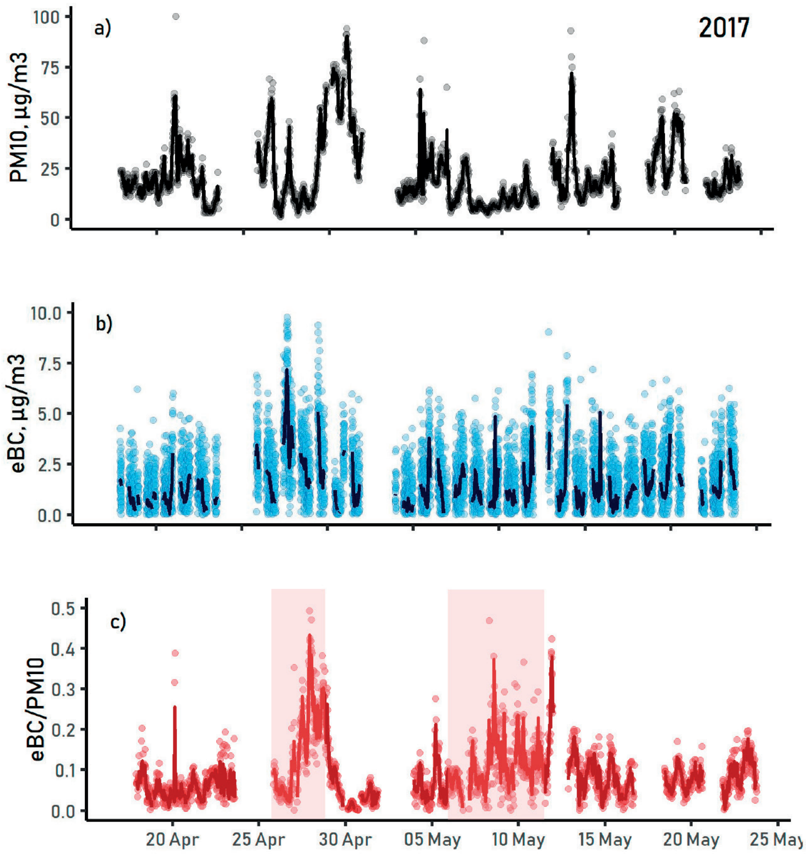


Fig. 7. 1-h mean (curve) and 1 min mean (dots) a) PM10 mass concentrations, b) eBC mass concentrations, and c) eBC/PM10 ratio in sampling period 2017 on the MO MSU. Pollution episodes are indicated.

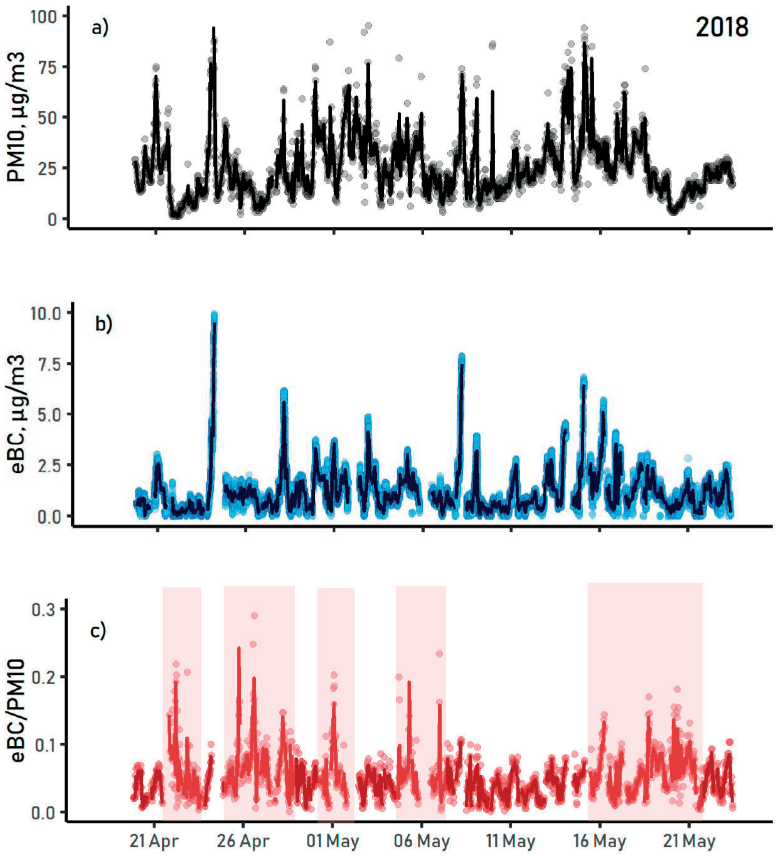


Fig. 8. 1-h mean (curve) and 1 min mean (dots) a) PM10 mass concentrations, b) eBC mass concentrations, and c) eBC/PM10 ratio in sampling period 2018 on the MO MSU. Pollution episodes are indicated.

al. 2011). This apparently reflects both local sources and a windward position in respect to the prevailing urban transfer. In the Moscow suburban station in Zelenograd city, both measured and calculated daily mean PM10 concentration was almost always lower than at all urban stations, which confirms that a megalopolis is a significant source of aerosol emissions. Observation of PM10 in summer showed a distinct correlation between short-term episodes of high PM10 with the local temperature maximums (Kuznetsova et al. 2011).

A substantial positive correlation between temperature variations and PM10 measurements at the European stations in the summer months was also established from the numerical analysis (Konovalov et al. 2009). The existence of this correlation can be attributed to the fact that, as a rule, high-temperature periods are followed by an anticyclonic (blocking) situations promoting the build-up of the atmospheric pollutants (Kuznetsova et al. 2011). The PM10 concentration maximum is reached directly before or at the moment of the passage of cold front after the period of stagnation. A high PM10 concentration can also be a result of the wind uplift of aerosols due to wind strengthening before the front passage in a warm sector of a cycle with sufficient precipitation and the air mass change, whereas the situation was preceded by dry weather favorable for the soil erosion.

In our study we observe relatively good correlation between high PM10 and the local temperature maximums, with correlation coefficient R2 equal 0.68 and 0.57 for 2017 and 2018, respectively (Fig. 10). It is expected that intensive precipitation may lead to strengthened wet aerosol deposition. Analysis of 24-h PM10 variation with daily precipitation, shown on Fig. 10, indicates the decrease of PM10 in relation with days of rain. The most prominently this phenomenon is observed for 18-20 May when during the period of the most intensive rains (more than 7.4 mm) we observed the lowest PM10 mass (around 13 µg/m^3).

Seasonal BC and level of pollution variations

The time series of 1-h mean eBC concentrations over the spring periods of sampling show the high eBC variability, ranged from 0.1 up to 5 and 10 µg/m^3, on average 1.5±1.3 and 1.1±0.9 µg/m^3, in 2017 and 2018, respectively (Fig. 7 and 8). Observations performed in the Moscow center showed the averaged BC concentrations of 4.4 and 1.7 µg/m^3 while on suburban background station they were consistently less than 3.0 and 1.05 µg/m^3, in spring 2014 and 2016 (Kopeikin et al. 2018).

Table 1 presents the mean BC concentrations obtained during the spring seasons at MO MSU, and rural and remote sites in different world locations. Comparisons show BC at MO MSU is lower than in European large cities like London (2.7 µg/m^3) (Kendall et al. 2001) and Budapest (2.95 µg/m3) (Salma et al. 2004). However, the sites in these cities are classified as urban ones taking place in the city center. The closest to the measured BC value for MO MSU we find with the spring mean BC in the less polluted city in the Europe Helsinki, 1.03 µg/m^3 (Järvi et al. 2008). The biggest BC was observed in the Indian and Chinese cities: around 5.5 and 8.8 µg/m^3 in Dhanbad, the coal capital of India, and Guangzhou, the biggest megacity of South China (Singh et al. 2015) and (Wu et al. 2013), respectively.

Episodes of BC pollution can be identified by the increased ratio of eBC to PM10 (eBC/PM10). In 2017, the highest ratio more than 40% was observed from 27 April to 2 May, and around 9 and 13 May (Fig. 7). High temperature in the period of observations approaching a maximum at those days (Fig. 9) could relate to intensive biomass burning in the region around a city. Analyses of BWT arrived to the MO MSU during sampling period shows the highest number (48.9%) originating from North-West of the Moscow city. The remaining BWT arrived from North-East (28.1%), North - West (16.9%), and South - East (6.2%). Analyses of days of air mass transportation through fire areas demonstrates that the

Table 1. Mean season BC concentrations (μg/m³) at rural and remote sites in different world locations

Site	Sampling period	BC	Reference
URBAN			
MO MSU, Moscow, Russia	April-May	1.1	this work
Helsinki, Finland	2017 – 2018	1.0	Jarvi et al. 2008
Budapest, Hungary	Spring 2004 – 2005	2.9	Salma et al. 2004
London 1, United Kingdom	April-May 2002	2.4	Kendall et al. 2001
Zurich, Switzerland	Spring 1996	1.2	Herich et al. 2011
La Reunion Island, France	Summer 2009 – 2010	0.5	Bhugwant and Brémaud 2001
Dhanbad, India	March-April 1998	5.5	Singh et al. 2015
Nanjing, China	March-April 2012	4.0	Wang et al. 2017
Mexico City, Mexico	April-May 2014	3.4	Salcedo et al. 2006
Guangzhou, China	April 2003	8.9	Wu et al. 2013
RURAL			
Yongxing Island, China	May-June 2008	0.5	Wu et al. 2013
Near Guangzhou, China	May-June 2008	2.6	Wu et al. 2013
Xilinhot, China	April 2005	2.0	Niu and Zhang 2010
Payerne, Switzerland	Summer 2008 – 2010	0.4	Herich et al. 2011

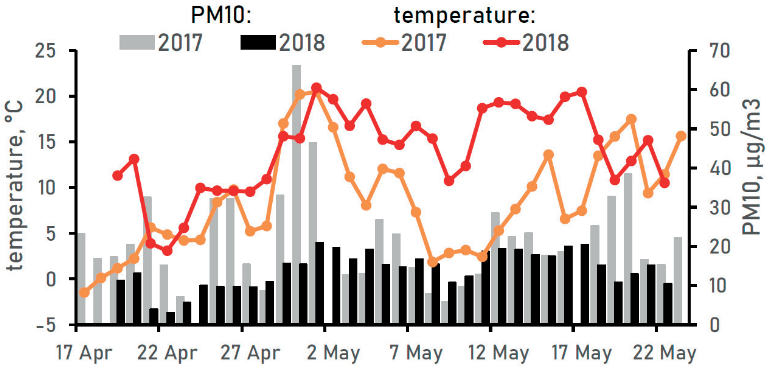


Fig. 9. 24-h mean temperature during the sampling period in 2017 (orange line) and 2018 (red line) and 24- h average PM10 mass concentrations in 2017 (gray) and 2018 (black)

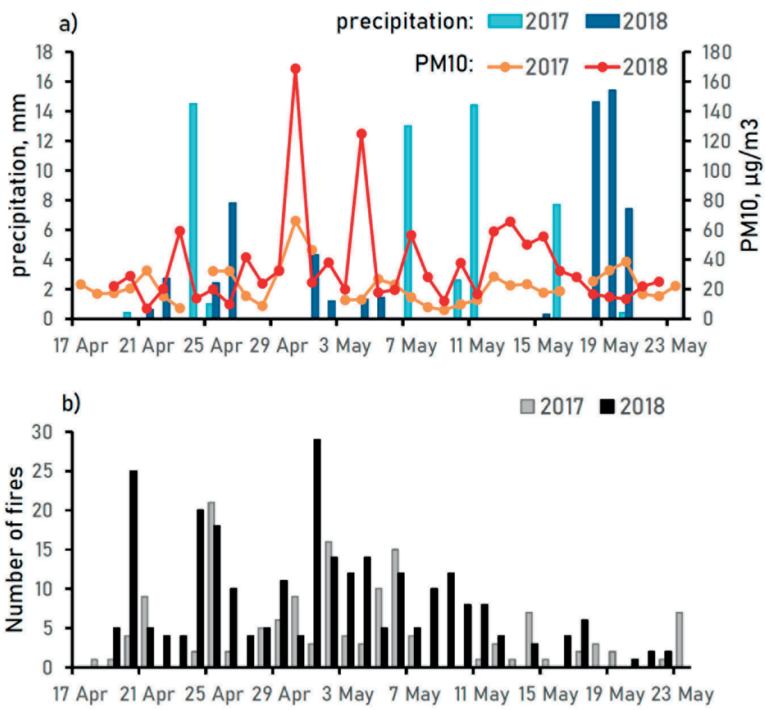


Fig. 10. a) Precipitation and 24-h average PM10 mass concentrations and b) number of fires during the spring seasons of 2017 and 2018

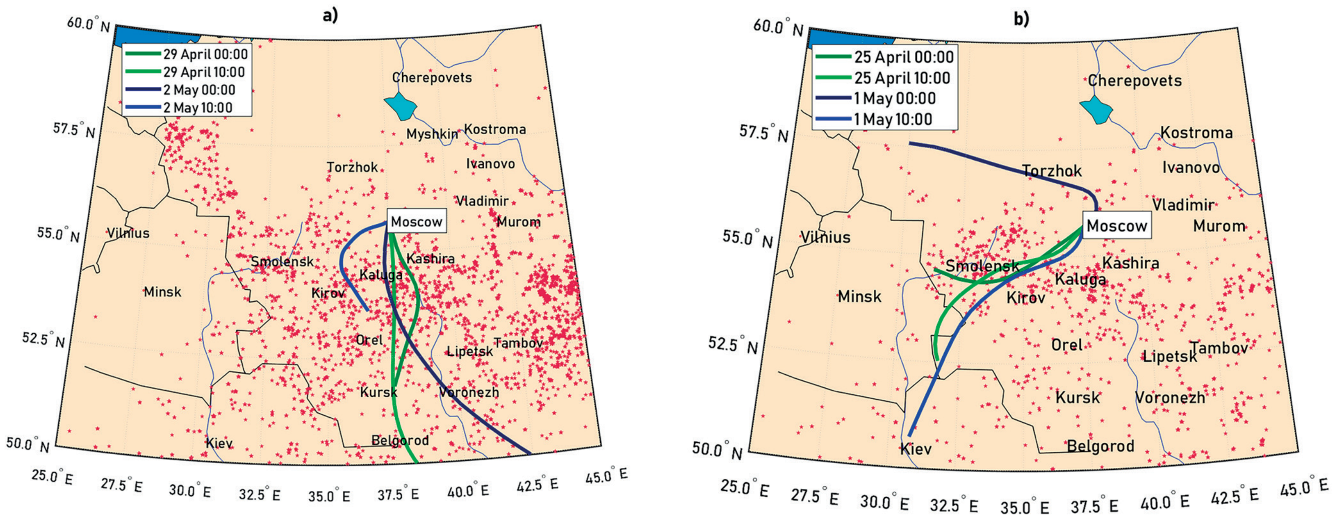


Fig. 11. 2-day backward air mass transportation from HYSPLIT model in the days of highest fires observed by FIRMS, 500 m a.s.l., a) 29 April and 2 May 2017, and b) 25 April and 1 May 2018. There are two trajectories for each day: one finished at 00:00 and second at 10:00 of the day.

biggest number of open flaming fires occurred on 26 April, 1-5 May (Fig. 10). Figure 11 shows the BWT arrived on 29 April and 2 May from areas with most intensive agriculture fires in south of Russia and Western Europe. Relation of abnormal temperature and the biggest number of open fires to the highest eBC/PM10 at the days around 2 May confirms the reason of the BC pollution episode on urban background level. In the days of the biggest precipitations (07, 11, and 16 May) the low eBC/PM10 rations around 3-5 % were observed, relating to the washout effect.

In 2018, the level of pollution related to eBC/PM10 showed the periodically repeated increased values of eBC/PM10, up to 10 % (Fig. 8). Analyses of BWT arrived to the MO MSU during sampling period shows the highest number (34%) originating from north-east of the Moscow city. The remaining BWT arrived from South - West and North - West (25%), and South - East (16%). The biggest number of open flaming fires occurred in the South-East direction from the Moscow area at the days 20, 24-25 April, and 1 May (Fig. 10). Figure 11 shows the BWT arrived on 25 April and 1 May from areas of most intensive agriculture fires in South of Russia and Western Europe. Such observations confirm that high atmospheric pollutions in Moscow are mainly caused by the transport of pollutions from the regions with intense emissions over a distance of hundred kilometers (Golitsyn et al. 2015).

CONCLUSIONS

Measurements performed at the Meteorological Observatory MSU improve the luck of the analysis for the aerosol pollution related with BC as the most important pollution contributor in total PM10 mass in the urban background of Moscow megacity. The highest level of BC pollution, up to 40%, was observed in the days of the May holiday in Russia, related with abnormal high temperature and intensive agriculture fires in the south of Russia and biomass burning around the city. Diurnal profile in the spring season is observed to be different from other large cities: we note stable high BC at night, the less pronounced morning peak in comparison with nocturne BC concentrations, and significant increase of mean diurnal BC late in the evening. As Moscow is an urban, industrialized and densely populated city, the diurnal variations in BC mass concentrations can mainly be attributed to the vehicular emissions and transport from the surrounding regions. Mean spring BC concentrations at urban background of Moscow megacity are observed similar to one in European cities and significantly lower than in industrialized Asian cities, in well relations with intensive transport and using of gas, diesel, and gasoline instead of coal in Asia.

ACKNOWLEDGEMENTS

The authors thank the financial support from RSF project №18-17-00149 for BC measurement performance at MO MSU and air mass transportation modeling. Financial support from RSF project № 19-77-30004 for aethalometric methodic development and data analyses is acknowledged.

REFERENCES

Ahmed T., Dutkiewicz V.A., Khan A.J. and Husain L. (2014). Long term trends in black carbon concentrations in the Northeastern United States. Atmospheric Research, 137, 49-57. doi: 10.1016/j.atmosres.2013.10.003

Bhugwant C. and Brémaud P. (2001). Simultaneous measurements of black carbon, PM10, ozone and NOx variability at a locally polluted island in the Southern tropics. Journal of Atmospheric Chemistry, 39, 261-280. doi: 10.1023/A:1010692201459

Bityukova V. and Saulskaya T. (2017). Changes of the anthropogenic impact of Moscow industrial zones during the recent decades. Vestnik Moskovskogo Unviersiteta, Seriya Geografiya, 3, 24-33 (in Russian with English summary).

Bond T.C., Doherty S.J., Fahey D.W., Forster P.M., et al.(2013). Bounding the role of black carbon in the climate system: A scientific assessment. Journal of Geophysical Research: Atmospheres, 118, 5380-5552. doi: 10.1002/jgrd.50171

Chen X., Zhang Z., Engling G., Zhang R., et al. (2014). Characterization of fine particulate black carbon in Guangzhou, a megacity of South China. Atmospheric Pollution Research, 5, 361-370. doi: 10.5094/APR.2014.042

Cheng Z., Luo L., Wang S., Wang Y., et al. (2016). Status and characteristics of ambient PM2.5 pollution in global megacities. Environment International, 89-90, 212-221. doi: 10.1016/j.envint.2016.02.003

Diapouli E., Kalogridis A.-C., Markantonaki C., Vratolis S., et al.(2017). Annual variability of black carbon concentrations originating from biomass and fossil fuel combustion for the suburban aerosol in Athens, Greece. Atmosphere, 8, 34. doi: 10.3390/atmos8120234

Dismuke C. and Egede L. (2015). The impact of cognitive, social and physical limitations on income in community dwelling adults with chronic medical and mental disorders. Global Journal of Health Science, 7(5), 183-195. doi: 10.5539/gjhs.v7n5p183

Garland R., Yang H., Schmid O., Rose D., et al. (2008). Aerosol optical properties in a rural environment near the mega-city Guangzhou, China: implications for regional air pollution, radiative forcing and remote sensing. Atmospheric Chemistry and Physics, 8, 5161-5186. doi: 10.5194/acp-8-5161-2008

Golitsyn G.S., Grechko E.I., Wang G., Wang P., et al. (2015). Studying the pollution of Moscow and Beijing atmospheres with carbon monoxide and aerosol. Izvestiya Atmospheric and Oceanic Physics, 51, 1-11. doi: 10.1134/S0001433815010041

Gubanova D., Belikov I., Elansky N., Skorokhod A. and Chubarova N. (2018). Variations in PM2.5 surface concentration in Moscow according to observations at MSU Meteorological Observatory. Atmospheric and Oceanic Optics, 31, 290-299. doi: 10.1134/S1024856018030065

Healy R., Sofowote U., Su Y., Debosz J., et al. (2017). Ambient measurements and source apportionment of fossil fuel and biomass burning black carbon in Ontario. Atmospheric Environment, 161, 34-47. doi: 10.1016/j.atmosenv.2017.04.034

Herich H., Hueglin C. and Buchmann B. (2011). A 2.5 year's source apportionment study of black carbon from wood burning and fossil fuel combustion at urban and rural sites in Switzerland. Atmospheric Measurement Techniques, 4, pp. 1409-1420. doi: 10.5194/amt-4-1409-2011

Jacobson M.Z. (2010). Short-term effects of controlling fossil-fuel soot, biofuel soot and gases, and methane on climate, arctic ice, and air pollution health. Journal of Geophysical Research: Atmospheres, 115. doi: 10.1029/2009JD013795

Janssen N.A.H., Hoek G., Simic-Lawson M., Fischer P., et al. (2011). Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environmental Health Perspectives, 119, 1691-1699. doi: 10.1289/ehp.1003369

Järvi L., Junninen H., Karppinen A., Hillamo R., et al. (2008). Temporal variations in black carbon concentrations with different time scales in Helsinki during 1996–2005. Atmospheric Chemistry and Physics, 8, 1017-1027. doi: 10.5194/acp-8-1017-2008

Kendall M., Hamilton R., Watt J. and Williams I. (2001). Characterisation of selected speciated organic compounds associated with particulate matter in London. Atmospheric Environment, 35, 2483-2495. doi: 10.1016/S1352-2310(00)00431-3

Konovalov I., Beekmann M., Meleux F., Dutot A. and Foret G. (2009). Combining deterministic and statistical approaches for PM10 forecasting in Europe. Atmospheric Environment, 43, 6425-6434. doi: 10.1016/j.atmosenv.2009.06.039

Kopeikin V., Emilenko A., Isakov A., Loskutova O. and Ponomareva T.Y. (2018). Variability of soot and fine aerosol in the Moscow region in 2014–2016. Atmospheric and Oceanic Optics, 31, 243-249. doi: 10.1134/S1024856018030089

Kozlov V., Panchenko M. and Yausheva E. (2011). Diurnal variations of the submicron aerosol and black carbon in the near-ground layer. Atmospheric and Oceanic Optics, 24, 30-38. doi: 10.1134/S102485601101009X

Kul'bachevskii A.O. (Ed.) (2017). Report on the state of the environment in Moscow in 2016. NIiPI IGSP, Moscow. (in Russian).

Kuznetsova I., Konovalov I., Glazkova A., Nakhaev M., et al. (2011). Observed and calculated variability of the particulate matter concentration in Moscow and in Zelenograd. Russian Meteorology and Hydrology, 36, 175-184. doi: 10.3103/S1068373911030046

Lokoshchenko M. (2015). Wind direction in Moscow. Russian Meteorology and Hydrology, 40, 639-646. doi: 10.3103/S1068373915100015

Mousavi A., Sowlat M.H., Hasheminassab S., Polidori A. and Sioutas C. (2018). Spatio-temporal trends and source apportionment of fossil fuel and biomass burning black carbon (BC) in the Los Angeles basin. Science of the Total Environment, 640, 1231-1240. doi: 10.1016/j.scitotenv.2018.06.022

Niu S. and Zhang Q. (2010). Scattering and absorption coefficients of aerosols in a semi-arid area in China: Diurnal cycle, seasonal variability and dust events. Asia-Pacific Journal of Atmospheric Sciences, 46, 65-71. doi: 10.1007/s13143-010-0007-2

Ohara T., Akimoto H., Kurokawa J.-I., Horii N., Yamaji K., Yan X. and Hayasaka T. (2007). An Asian emission inventory of anthropogenic emission sources for the period 1980–2020. Atmospheric Chemistry and Physics, 7, 4419-4444. doi: 10.5194/acp-7-4419-2007

Pope III C.A. and Dockery D.W. (2006). Health effects of fine particulate air pollution: lines that connect. Journal of the Air & Waste Management Association, 56, 709-742. doi: 10.1080/10473289.2006.10464485

Popovicheva O., Kistler M., Kireeva E., Persiantseva N., et al. (2014). Physicochemical characterization of smoke aerosol during large-scale wildfires: extreme event of August 2010 in Moscow. Atmospheric Environment, 96, 405-414. doi: 10.1016/j.atmosenv.2014.03.026

Popovicheva O.B., Engling G., Ku I.-T., Timofeev M.A. and Shonija N.K. (2019). Aerosol emissions from long-lasting smoldering of boreal peatlands: chemical composition, markers, and microstructure. Aerosol and Air Quality Research, 19, 484-503. doi: 10.4209/aaqr.2018.08.0302

Popovicheva O.B., Evangeliou N., Eleftheriadis K., Kalogridis A.C., et al. (2017a). Black carbon sources constrained by observations in the Russian high Arctic. Environmental Science & Technology, 51, 3871-3879. doi: 10.1021/acs.est.6b05832

Popovicheva O.B., Shonija N.K., Persiantseva N., Timofeev M., et al. (2017b). Aerosol pollutants during agricultural biomass burning: A case study in Ba Vi region in Hanoi, Vietnam. Aerosol and Air Quality Research, 17, 2762-2779. doi: 10.4209/aaqr.2017.03.0111

Ramachandran S. and Rajesh T. (2007). Black carbon aerosol mass concentrations over Ahmedabad, an urban location in Western India: comparison with urban sites in Asia, Europe, Canada, and the United States. Journal of Geophysical Research: Atmospheres, 112. doi: 10.1029/2006JD007488

Reddy M.S. and Venkataraman C. (2002). Inventory of aerosol and sulphur dioxide emissions from India: I—fossil fuel combustion. Atmospheric Environment, 36, 677-697. doi: 10.1016/S1352-2310(01)00463-0

Salcedo D., Onasch T.B., Dzepina K., Canagaratna M.R., et al. (2006). Characterization of ambient aerosols in Mexico City during the MCMA-2003 campaign with Aerosol Mass Spectrometry: results from the CENICA Supersite. Atmospheric Chemistry and Physics, 6(4), 925-946. doi: 10.5194/acp-6-925-2006

Salma I., Chi X. and Maenhaut W. (2004). Elemental and organic carbon in urban canyon and background environments in Budapest, Hungary. Atmospheric Environment, 38, 27-36. doi: 10.1016/j.atmosenv.2003.09.047

Sharma S., Brook J., Cachier H., Chow J., et al. (2002). Light absorption and thermal measurements of black carbon in different regions of Canada. Journal of Geophysical Research: Atmospheres, 107, AAC 11-1-AAC 11-11. doi: 10.1029/2002JD002496

Singh S., Tiwari S., Gond D., Dumka U., et al. (2015). Intra-seasonal variability of black carbon aerosols over a coal field area at Dhanbad, India. Atmospheric Research, 161, 25-35. doi: 10.1016/j.atmosres.2015.03.015

Stein A., Draxler R., Rolph G., Stunder B., et al. (2015). Noaa's Hysplit Atmospheric Transport and Dispersion Modeling System. Bulletin of the American Meteorological Society, 96, 2059-2077. doi: 10.1175/BAMS-D-14-00110.1

Steiner S., Czerwinski J., Comte P., Popovicheva O., et al. (2013). Comparison of the toxicity of diesel exhaust produced by bio-and fossil diesel combustion in human lung cells in vitro. Atmospheric Environment, 81, 380-388. doi: 10.1016/j.atmosenv.2013.08.059

Weingartner E., Keller C., Stahel W., Burtscher H. and Baltensperger U. (1997). Aerosol emission in a road tunnel. Atmospheric Environment, 31, 451-462. doi: 10.1016/S1352-2310(96)00193-8

WHO (2005). WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide - Global Update 2005.

WHO (2012). Health effects of black carbon. WHO.

Wu D., Wu C., Liao B., Chen H., et al. (2013). Black carbon over the South China Sea and in various continental locations in South China. Atmospheric Chemistry and Physics, 13, 12257-12270. doi: 10.5194/acp-13-12257-2013

Received on July 13th, 2019 Accepted on November 7th, 2019

Tatiana Gladenkova[1*]
[1] Faculty of Geography, M.V. Lomonosov Moscow State University, Moscow, Russia
***Corresponding author:** tatianagladenkova@gmail.com

BEAUTY AND PERSONAL CARE TRANSNATIONALIZATION: MAIN CHANGES IN ITS SPATIAL STRUCTURE

ABSTRACT. At the turn of 21st century global beauty and personal care industry underwent dramatic changes in its territorial structure. The main factors of that changes were world economics institutional changes, primarily – the international trade liberalization, as well as the R&D progress. During abolition of tariffs or tariff cut the competition in beauty and personal care ratcheted up sharply. That was accompanied by sea changes in its macrogeography, particularly, by the manufacturing transnationalization (mainly, in the form of its «drift» to developing countries) and the general expansion of the range of countries specializing in the beauty and personal care production. The main drive-forces of beauty and personal care transnationalization and the resulting territorial changes in the industry are discussed.

KEY WORDS: beauty and personal care industry, transnationalization, transnational corporations (TNCs), setting up manufacturing strategies

CITATION: Tatiana Gladenkova (2019) Beauty And Personal Care Transnationalization: Main Changes In Its Spatial Structure. Geography, Environment, Sustainability, Vol.13, No 1, p. 244-250 DOI-10.24057/2071-9388-2019-94

INTRODUCTION

For the end of the 20th – the beginning of the 21st centuries the growth of open economic borders of countries as well as the growth of international capital movements, the liberalization of national trade regimes, the formation of a global financial market and a global information network were main characteristic occurrences (Gorbanev 2014; Rodionova 2010, 2016). As a result, it is becoming increasingly difficult to define a clear boundary between the activities of national economies and transnational corporations.

The process of transnationalization of industrial production, which in a general sense implies the movement of capital across the borders of the country of origin and the organization of production abroad, has an objective basis. The key factors for the development of this process are the overcoming of customs barriers, reduction of production costs, among which the cost of transport, labor and raw materials compose a significant share, and others (Sokolsky 2010).

The beauty and personal care industry stands out significantly against the background of the main consumer sectors. The specifics of this industry are consumer character of its products, its transportability, as well as high involvement in world trade, helped by the reduction of WTO customs tariffs. Thus, since 2014, the average tariff for chemicals for the European Union member countries was 4.54%, for the USA – 2.76%, for Russia – 5.74%, for China – 6.5%, for Latin American countries – 5-10% (Wto.org 2014).

Until the end of the 20th century the main producers and consumers of beauty and personal care products were developed countries. However, by the end of the 20th century in connection with the active investment of TNCs in the beauty and personal care of developing countries, the latter loudly declared themselves as suppliers of beauty and personal care products: in 2000, almost half of beauty and personal care products net exporters belonged to developing countries (Comtrade.un.org 2015); and also demonstrated an active growth in demand for goods of the industry (in 2015, the share of the Asia-Pacific region countries accounted for about 35% of world consumption of perfumes and cosmetics (against 20% in 2000)).

The development of the beauty and personal care industry in developing countries took place mainly «from the outside», with the participation of the top TNCs. The exceptions are India, Brazil and the Republic of Korea, where the industry has mainly developed on the basis of local manufacturing competencies and resources.

In comparison with the other industries, the beauty and personal care does not have a high potential for its transnationalization (the products of the industry are characterized by high transportability, while the reduction and abolition of customs tariffs contributed to an increase in the volume of trade in its goods). Nevertheless, the tendency to its involvement in this process is becoming more and more obvious. From this perspective, it seems relevant to consider two key issues: 1) what is the manifestation of the beauty and personal care transnationalization and 2) what are its main causes, forms and consequences.

MATERIALS AND METHODS

The paper includes an analysis of changes in the volume of global beauty and personal care products output from 1990 to 2015, as well as an analysis of the top beauty and personal care companies turnovers from 1977 to 2015 and their setting up manufacturing strategies for the period from 1920 to 2016. Selected data were collected for each country and TNCs, subsequently consolidated into a common base.

The main sources of statistical and analytical information of the presented paper were databases, reports and analytical notes of international organizations (UN Comtrade, UNCTAD, WTO), annual reports of the studied companies and data of annual ratings of the top corporations. During our research the general scientific system approach and various methods of scientific knowledge were widely used: analysis, synthesis, historical modeling.

RESULTS AND DISCUSSION

TNCs as the core of modern beauty and personal care industry

The intensification of the process of the global beauty and personal care industry transnationalization can be indirectly judged by the growth in the number of TNCs that are the main initiators of international capital transfers. Since the end of the 20th century, the number of foreign affiliates of multinational corporations has increased significantly, and their turnovers, incomes and assets are also growing (McCann and Iammarino 2013).

The level of beauty and personal care transnationalization is largely manifested in the growth of the number of countries where TNCs operate. The main centers of the corporate power of the beauty and personal care industry remain North America, Western Europe and Japan, where the vast majority of head units of the top TNCs are concentrated. The main production facilities are also located mostly in developed countries. However, since the end of the 20th century, there has been a tendency to place beauty and personal care manufacturing sites in developing countries, the main reason for which was the motivation of beauty and personal care companies to enter new unsaturated markets (Weber et al. 2002). The consumer nature of the industry's products predetermined the dominance of factors such as the conquest of new markets (to expand the geography of sales and compensate for the slowdown in the growth rate of demand for beauty and personal care products in developed countries), as well as to overcome institutional barriers. Against the background of active investment by TNCs in the beauty and personal care industry of developing countries, local national manufacturers significantly strengthened their positions. As a result of these two parallel processes, the output of the beauty and personal care products from 1990 to 2015 in Asia increased 20 times (in China – 114 times, in India – 40 times), Latin America – 11 times (in Brazil – 9 times, in Mexico – 6 times) (Fig. 1).

Currently, developing countries are beginning to make an increasing contribution to the development of the beauty and personal care industry with their dynamically growing markets, among which stand out (in order of market size) the BRICS countries, Mexico, Argentina, Saudi Arabia, Thailand, Indonesia, Poland, Turkey, Iran, Colombia, Nigeria, UAE, Malaysia. As the volumes of these markets increase, their value will also increase. Developing beauty and personal care markets are significantly faster than developed ones in terms of growth rates – more than twice; which provides a significant share of the growth of the global industry market. Against the backdrop of the stagnation of developed markets, key developing markets already occupy leading positions in top ratings in terms of growth rates and absolute sizes: since 2010, China has held a strong second place after the United States, overtaking Japan, and Brazil since 2008 has held forth place by taken over UK, Germany and France. It is expected that by 2025, Brazil will take third place in this ranking. As a result, the regional structure of the global beauty and personal care market is changing significantly. In 2015, emerging markets together accounted for about 47% of the global consumption of beauty and personal care products, versus 24% in 2000.

It is worth paying attention to the fact that with the increase in the share of the middle class in developing countries, there is an active growth in demand not only for products for mass consumption, but also for premium products produced by major TNCs. Thus, the markets of developing countries are acquiring a qualitatively new role in the modern world beauty and personal care market and in the activities of TNCs in particular.

In recent years, companies originating from developing countries are becoming more confident in the global beauty and personal care market and are successfully competing with the beauty and personal care industry's veterans (in the 1970s, among the leading international firms there was not one originating from developing countries) (see Table 1) (Matusow et al. 2016). The products from East and South-East Asia and Latin America have become particularly popular in both European and North American markets (Kcia.or.kr 2018; Premiumbeautynews.com 2018).

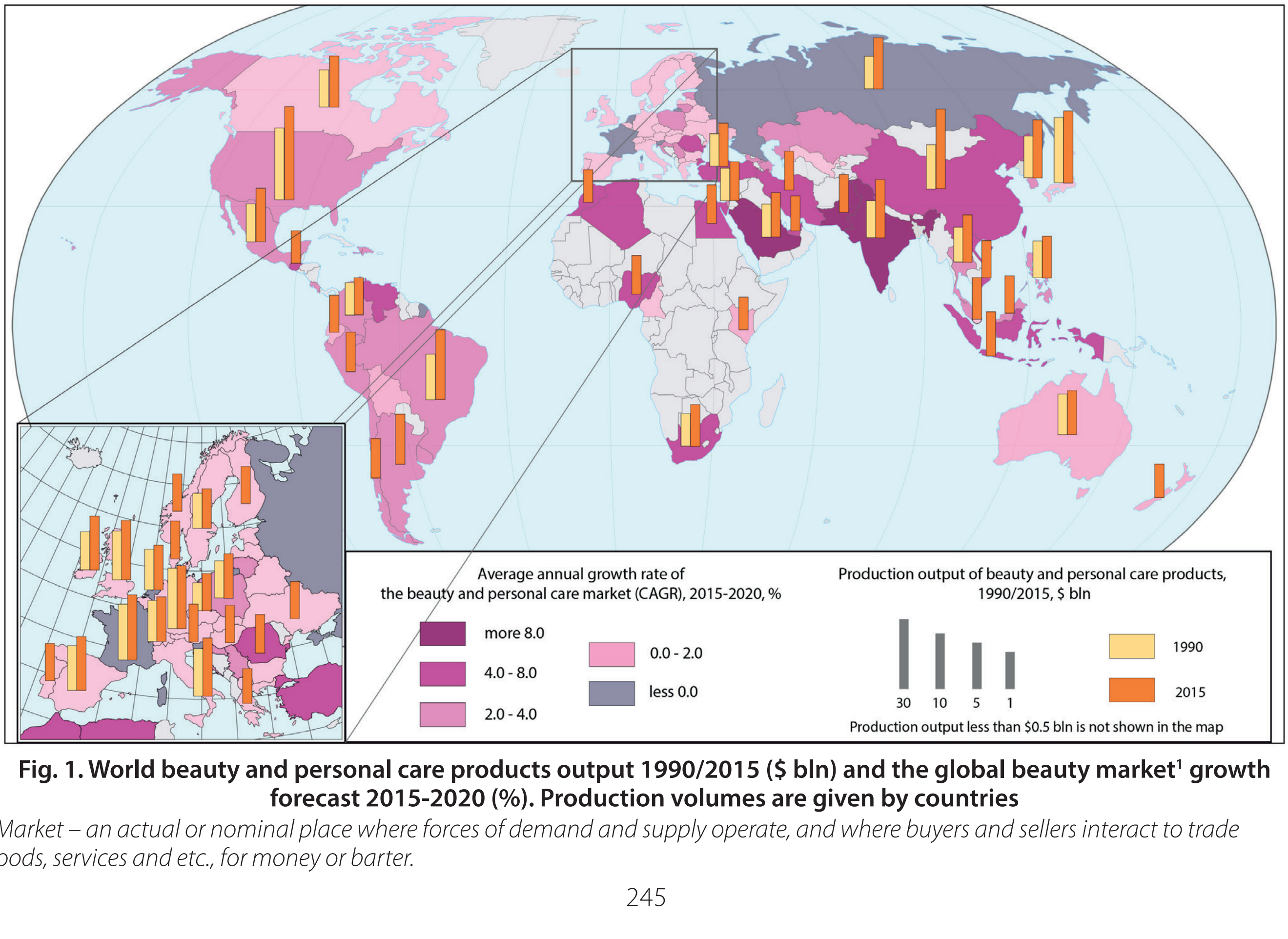


Fig. 1. World beauty and personal care products output 1990/2015 ($ bln) and the global beauty market[1] growth forecast 2015-2020 (%). Production volumes are given by countries

[1] *Market – an actual or nominal place where forces of demand and supply operate, and where buyers and sellers interact to trade goods, services and etc., for money or barter.*

Table 1. Top beauty and personal care transnational corporations, 1977, 2018

№	Company	Turnover, 1977 ($ bln)	Home country	№	Company	Turnover, 2018 ($ bln)	Home country
1	Colgate-Palmolive	2.5	USA	1	Procter & Gamble	44.6	USA
2	Avon	1.4	USA	2	L'Oréal	32.4	France
3	Shiseido	0.9	Japan	3	Unilever	24.3	UK/ Netherlands
4	Revlon	0.9	USA	4	Kimberly-Clark	18.5	USA
5	L'Oréal	0.8	France	5	SCA	15.8	Sweden
6	Bristol-Myers	0.7	USA	6	Estée Lauder	13.7	USA
7	Unilever	0.7	UK/ Netherlands	7	Colgate-Palmolive	11.2	USA
8	Procter & Gamble	0.6	USA	8	Johnson & Johnson	8.6	USA
9	Chesebrough-Pond's	0.5	USA	9	Kao	10.8	Japan
10	Wella	0.4	Germany	10	Shiseido	10.3	Japan
11	Johnson & Johnson	0.4	USA	11	Coty Inc.	9.4	USA
12	Gillette	0.4	USA	12	Beiersdorf	7.4	Germany
13	Schwarzkopf	0.4	Germany	13	AmorePacific	6.2	South Korea
14	Norton Simon (Max Factor)	0.4	USA	14	LVMH	6.1	France
15	American Cynamid	0.3	USA	15	Avon	5.7	USA
16	Beiersdorf	0.3	Germany	16	Unicharm	5.5	Japan
17	Kanebo	0.3	Japan	17	LG H&H	5.3	South Korea
18	Beecham	0.2	UK	18	L Brands	4.8	USA
19	Faberge	0.2	USA	19	Henkel	4.5	Germany
20	Estee Lauder	0.2	USA	20	Mary Kay	3.5	USA
World turnover		15		World turnover		545.5	

Thus, one of the important geographical consequences of the transnationalization of the beauty and personal care industry is a significant increase in the number of countries where TNCs have their own manufacturing facilities.

The share of TNCs in the beauty and personal care industry of foreign countries is also steadily growing. This can be illustrated by China, where in 2016, 80% of the beauty and personal care goods were produced by foreign companies or joint ventures (Trade.gov 2016). In many ways, an increase in FDI flows to China has been facilitated by «open door» policies and reforms that began in the 1980s. At the same time, SEZs and various industrial clusters organized under the new policy became important drivers of China's economic growth. In organized SEZs and clusters, TNCs not only organized their production, thereby providing new jobs to the local population, but were also able to create ties with local firms. Thanks to a favorable business environment, knowledge flows: TNCs exchange their knowledge, experience and technologies with national producers. As a result, the level of development of local production has grown significantly. The largest transnational corporations, such as P&G, Unilever, L' Oréal, took advantage of the «first move»: they were the first to enter the unsaturated Chinese market, subsequently becoming its permanent leaders. At the same time, for TNCs, China serves as a production site serving the markets of the Asia-Pacific region. For example, more than half of the products manufactured at the L' Oréal plant in Yichang are exported abroad to the countries of Southeast Asia, Japan, South Korea and New Zealand. In 2016, in Indonesia, Unilever Indonesia, P&G Home Products Indonesia and L' Oréal Indonesia produced 50% of the total beauty and personal care products of the country (Trade.gov 2016). A number of international brands are also produced in Indonesia under license from Unilever, L' Oréal and Revlon (Trade.gov 2015).

Despite the fact that over the past forty years, the share of the top TNCs in world turnover of beauty and personal care products has declined significantly (from 1977 to 2018 the share of the top-20 TNCs decreased from 83% to 45.6%), they continue playing an important role in the modern global beauty and personal care (see Table 1)[1]. Thus, the transnationalization index for such major manufacturers as P&G (in 2016 the share of beauty and personal care products accounted for 23.4% of total sales (Pg.com 2017)) and Unilever (in 2016 the share of beauty and personal care products accounted for 38.3% of total sales (Unilever.com 2017)), from 1998 to 2012 for the first company increased from 40.3% to 61%, and for the second decreased from 89.3% to 87.1% (which may be due to faster growth in demand for the company's products in the parent country) (Unctad.org 1998, 2012, 2015).

[1] *The reduction in the share of the top TNCs in total turnover of beauty and personal care products is a consequence of changes in the corporate structure of the industry, which manifested themselves in an increase in the number of narrowly specialized firms. In that respect, recently more and more new manufacturers appear in the world market, including from developing countries (Sokolsky 2015).*

The main types of setting up manufacturing strategies of TNCs in developing countries

By the nature of the industrial management the strategies used by beauty and personal care TNCs in developing countries can be divided into two main types. *The first type* of strategy is based on the orientation of companies towards cheap raw materials and labor. Corporations produce products for mass consumption, which are subsequently exported (from 70% to 100%) to industrialized countries and partially sold within the region of the producing country. Prerequisites for this type of industry development strategy have been the accumulation of knowledge and the intensive growth of the economy. The transportation costs of companies are largely compensated by reducing the costs of labor and raw materials.

In addition to these factors, the active development of this type of TNC strategy has largely contributed to the provision by developing countries of the benefits of multinational companies (various grants, tax incentives and subsidies, as well as privileges from countries participating in various international regional trade blocs).

Type of strategy with a focus on raw materials and labor, mainly implemented in the countries of South-East and South Asia. As part of transnational vertical integration, regional states specialize in different stages of the value-added chain. Each plant located in a particular country performs one of the successive stages of raw material processing. Or each production site specializes in a specific commodity unit within the framework of the node-based division of labor.

An example is the P&G's production development model, whose diversified activities cover all stages of the production chain (Fig. 2). The P&G production chain is structured as follows. At the first stage, raw materials from Indonesia and Malaysia are processed at the oleochemical plant in Malaysia (Kuantan). The resulting chemicals in the form of intermediates are transported to other company facilities in ASEAN countries and around the world. At the next stage, from Indonesia, where the largest TNC's Southeast Asian plants are located (Jakarta and Karavang), raw materials and intermediate products are delivered throughout the region. From Singapore, where the company's only Asian perfume production center is located, fragrances and flavors are delivered to other TNC's factories throughout the Asian region. The final stage of the production chain is carried out in Thailand, which has the world's largest export-oriented P&G factory for the production of hair products, and in the Philippines (Kabuyao), where the largest entire Asian region multifunctional factory of the company is located.

Another example – L' Oréal's strategy, which began developing the region twenty years ago by launching plants in India (Pune) in 1994, and in China (Pudong), in 1997 (Loreal.com 2017). Production is carried out mainly using local raw materials, that significantly reduces the cost of final products.

The second type of the setting up manufacturing strategy of the TNCs in developing countries is characterized by a focus on the local market. Companies place their production facilities in foreign countries for use them as a platforms, providing regional market with the produced products. The main prerequisites for the development of this type of strategy were high rates of economic development and, as a result, an increase in the purchasing power of the local population, as well as a facilitated process of penetration into the local market for TNCs.

This type of strategy was developed at the beginning of the 20th century in Latin America and only at the beginning of the 21st century began to be applied in South, East and South-East Asia (in combination with the strategy of the first type). This can be explained, firstly, by the fact that Latin America was originally a more developed region. Secondly, at the beginning of the 20th century, when the beauty and personal care industry was just beginning to acquire its present character, the location of enterprises in South and Central America allowed European and American companies to save significantly on value added taxes (Jones 2006).

A pioneer in the application of this model was Colgate-Palmolive, which in the 1920s organized the production of toilet soap «Palmolive», firstly in Mexico, and later in Brazil and Argentina (Jones 2006). In 1948, Revlon opened a factory in Mexico. In 1954, Avon located its factories in Puerto Rico and Venezuela, and in 1958 – in Cuba, Mexico and Brazil. L' Oréal opened the first two plants in South America in 1959 in Rio de Janeiro and São Paulo. In the 1960s Helena Rubenstein opened seven plants in Latin America (Jones et al. 2005).

At the beginning of the 21st century many TNCs have begun to apply a similar strategy for the countries of South, East and South-East Asia. In 2012, L' Oréal opened its largest plant in Indonesia (in Jababeka), serving the markets of South-East Asia: 75% of its products are exported to countries in the region, 25% are sent to the domestic market. In 2013, Unilever announced the opening of a new plant in India (Hamgaon) to provide markets in South and Southeast Asia.

Such a form of expansion as mergers and acquisitions for a long time was mainly used by TNCs only in developed countries. The purpose of M&A transactions, in addition to territorial expansion and assortment expansion of the product portfolio,

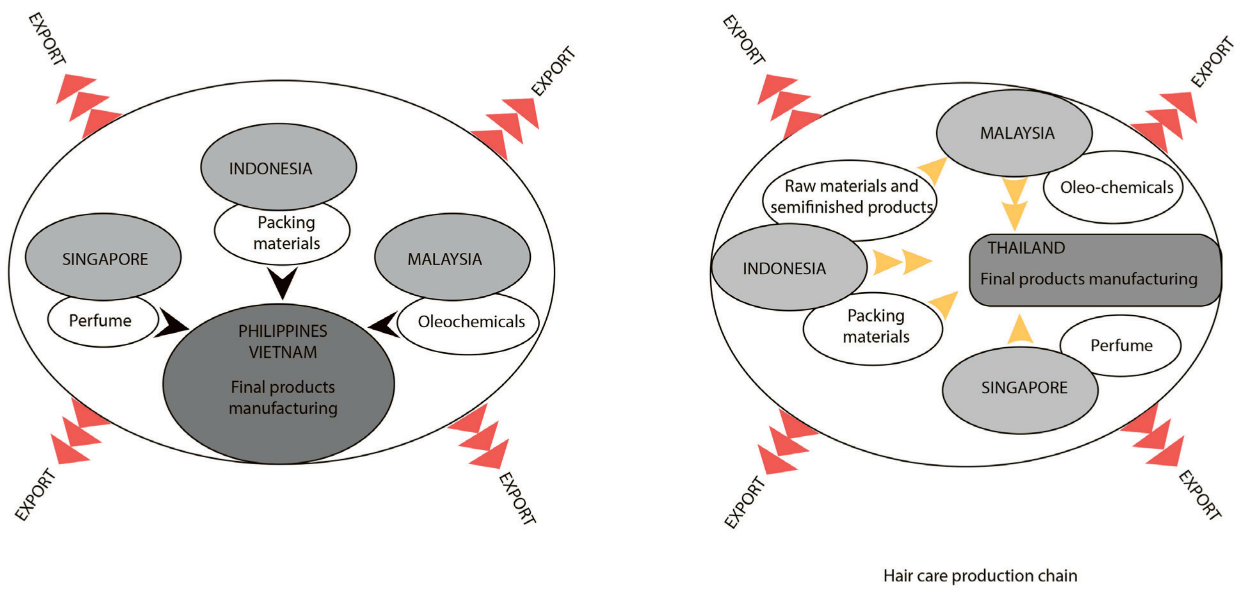


Fig. 2. Procter & Gamble production chain in ASEAN countries

is to acquire the accumulated knowledge of other companies, as well as faster access to new markets due to the high recognition of local brands. Recently, in view of strengthening the position of national producers of developing countries, the largest transnational corporations not only actively interact with local firms in the format of contract manufacturing, but also acquire local companies.

Expansion of the beauty and personal care production geography

Depending on the time of siting of beauty and personal care manufacturing facilities in developing countries, these countries can be divided into several main groups. *The «First wave» of beauty and personal care manufacturers* – the countries of Latin America (Argentina, Brazil and Mexico), in which TNCs began to locate their production facilities in the 1920s-1950s.

Later, in the 1990s – 2000s TNCs began to locate their production sites in the countries of the Middle East, as well as East, Southeast and South Asia (India, China, Malaysia, Thailand, Indonesia, Vietnam, the Philippines) – *the «Second wave» of beauty and personal care manufacturers*.

However, starting from the 2010s, due to the loss of competitive advantages in these countries due to higher costs, TNCs have to look for new «launching grounds» to accommodate their production facilities. Thus, *the beauty and personal care manufacturers of the «Third wave»* become the poorest countries of South-East and South Asia: Bangladesh, Nepal, Myanmar, Laos, Cambodia, Sri Lanka, Pakistan; as well as Africa (excl. South Africa), Colombia, Bolivia and Peru.

Localization of the beauty and personal care production facilities depends largely on innovation and the R&D in general. In view of this, nowadays, an important trend in the development of the beauty and personal care industry is the location of R&D centers of the top TNCs in the first and second wave countries. One of the main reasons of this trend is that in order to compete in the new market it's necessary to adapt global brands to local conditions (the needs and expectations of customers, differences in phenotypes and etc.). Employees of organized R&D centers actively interact with local scientists and research institutes of the host countries, that also ensures the flow of knowledge.

A telling illustration of the «drift» of the beauty and personal care industry is China. Foreign beauty and personal care companies, which began to locate their production sites here in the 1990s, were attracted by the presence of rich natural resources, cheap and fairly skilled labor, as well as high intensity of labor exploitation.

However, over the past decade, China has acquired the status of a state with a high growth rate of purchases of products from developed countries. In view of the rising living standards in China, the country has become very attractive for many foreign firms. Starting from the end of the 20th century, top TNCs actively conquered the local market (Cosmeticsbusiness.com 2017). Their success is largely due to the «fashion» factor: due to the low level of Chinese consumer confidence in local products, the population is increasingly choosing import brands (Euromonitor.com 2015; Ey.com 2016). A specific feature of the transnational business is the ability of TNCs to take into account the needs of the world market and to form a significant demand for their products before the start of their production.

In general, over the past ten years, the growth rate of Chinese exports decreased from 31% to 8% per year, which was partly due to the rise in labor costs by 14% over the past fifteen years (Hktdc.com 2017). While the average Chinese worker earns $ 650-700/month, the average wage of a Vietnamese worker is $ 250/month, and the Cambodian worker earns $ 130/month (Tradingeconomics.com 2016) (see Table 2).

Among the South-East Asian countries, Vietnam is of particular interest to companies originating from developed countries. The country's high level of literacy, which is about 94% of the total adult population, and a large proportion of young people are very attractive for the top TNCs[1].

Table 2. Average annual wages in industry 1980, 1990, 2016 ($ thousand)

Country	1980	1990	2016
South Korea	9.0	19.6	36.0
Turkey	0.5	5.1	18.3
Malaysia	6.8	7.5	9.0
China	0.8	2.3	8.4
Brazil	3.3	4.2	8.1
Mexico	3.3	3.9	5.8
Argentina	4.4	5.6	9.6
Thailand	0.3	2.1	4.7
India	0.7	0.9	2.7
Vietnam	0.1	0.5	2.5
Indonesia	0.4	0.6	2.2
Philippines	1.0	1.1	1.8
Cambodia	0.08	0.3	1.6
Laos	0.07	0.1	1.3
Myanmar	0.07	0.09	1.2

[1] *In 2015, Nike produced 42% of its products in Vietnam, while China produced only 30% of its products. Microsoft also moved its Nokia manufacturing division from China to Vietnam. Many global corporations, such as Adidas, Samsung, and Intel, are actively investing in the Vietnamese industry.*

Thus, one of the possible scenarios for the further development of beauty and personal care TNCs, by analogy with low-tech industries (textile and footwear), can be the almost complete transfer of industry to countries with even lower production costs (the principle of «cascades»).

At the moment, Africa is characterized by the advantages that were previously characteristic of Latin America and South and South-East Asia. Among the main factors that make the region attractive for the top TNCs, we can highlight the rapid population growth of the continent, 60% of which is younger than 25 years, the rapid growth of the middle class[1], increasing urbanization (Un.org 2017), as well as the improvement of the business regulation system (Euromonitor.com 2015).

In addition to the leading economies – Egypt, Nigeria and South Africa; Kenya, Ethiopia, Tanzania and Uganda have the greatest potential for the development of the beauty and personal care industry in the African region. For the three top TNCs, L' Oréal, Unilever and P&G, these countries are strategically important because the industrial siting within their territories provides access to the Middle East markets, where currently one of the highest growth rates of demand for perfumery and cosmetic products are observed (Wordpress.com 2016; McDougall 2015).

In 2016, the group of world beauty and personal care products net exporters included 8 African countries – Ghana, Togo, Senegal, Côte d'Ivoire, Egypt, Madagascar, Comoros and Swaziland. The participation of these countries in the group of net exporters may indirectly indicate their potential for placing TNCs production facilities within their territories in order to use them as «launching grounds» with an export orientation.

Assessing the future prospects for the transnationalization of the beauty and personal care, it can be stated that the development strategies of TNCs will increasingly be a reference point for developing countries that are becoming more attractive to Western companies not only because of low production costs, but also the growth of consumer potential in them, and the active participation of the public sector in these countries in attracting foreign investment.

As for the general forecast regarding the transnationalization of the global beauty and personal care industry, it can be assumed that this process will continue to increase, but will manifest itself not in the growth of the number of TNCs subsidiaries, but in the increase in the number of mergers and acquisitions of other companies. The organizational form of TNCs will also change significantly: recently along with the holding system, joint ventures with companies in host countries, contract agreements, including licensing, strategic alliances and outsourcing, have been developing rapidly. The increase in the value of industrial cooperation is typical not only for the beauty and personal care industry, but also for the global industry as a whole, which is due to the optimizing the production process.

CONCLUSIONS

1. At the beginning of the 20th century in the conditions of the unliberated world market, the process of manufacturing transnationalization in the beauty and personal care industry was driven by the motivation of TNCs to overcome customs barriers for the disposal of their goods in deep markets of the industrialized countries. As a result, during that period there was an increase in the concentration of the beauty and personal care in developed countries, which predetermined their absolute dominance in the industry.

2. In the last quarter of the 20th century due to the liberalization of foreign trade and increased competition between firms in the industry, the main incentive for beauty and personal care companies became the maximum production cost-effectiveness. In this regard, there is a powerful «drift» of the industry to developing countries with their cheap raw materials and labor. Today the balance of power between developed and developing countries in the global beauty and personal care industry is almost equivalent. This suggests a steady destruction of the core-periphery paradigm.

3. In the course of the beauty and personal care transnationalization, TNCs evolved: if in the early stages of development multinational companies were focused on the raw materials of less developed countries, then recently they have been actively interacting with the local authorities of the countries where the «drift» of production occurs to adapt their production systems to local conditions, involving home firms. In parallel, there is a process of moving the R&D activity to developing regions in centers with high scientific potential.

4. As the technological and human resources of the most industrialized developing countries increase, as well as in the process of increasing production costs in them, one of the possible strategies for developing the activities of TNCs is to move beauty and personal care manufacturing to countries with a lower level of technical and economic development, including the poorest countries of South-East Asia, South America, and also the countries of Africa.

[1] As defined by the African Development Bank, the middle class is "the one who spends from $2 to $20 per day on purchases, based on purchasing power parity". According to bank estimates, more than 34% of Africans (more than 300 million people) fall under this description, and their number is expected to increase to 42% (more than 1 billion people) by 2060. With more than 300 million middle class, the African region is practically caught up with the Indian and Chinese markets for this indicator.

REFERENCES

Comtrade.un.org (2015). United Nations Statistics Division Official Website. [online] Available at: www.comtrade.un.org/ [Accessed 15 Oct. 2016].

Cosmeticsbusiness.com (2017). Cosmetics Business Official Website. [online] Available at: www.cosmeticsbusiness.com [Accessed 5 Jan. 2018].

Euromonitor.com, (2015). Beauty and Personal Care. [online] Available at: www.euromonitor.com/beauty-and-personal-care [Accessed 15 Oct. 2016].

Ey.com (2016). Ernst & Young The luxury and cosmetics financial factbook 2016 edition: Keep calm and care about your consumer. [online] Available at: www.ey.com/Publication/vwLUAssets/EY-the-luxury-and-cosmetics-financial-factbook-2016/$FILE/EY-the-luxury-and-cosmetics-financial-factbook-2016.pdf [Accessed 5 Jan. 2018].

Gorbanev V.A. (2014). Human Geography of Russia and rest of the world. Moscow: UNITY-DANA (in Russian).

Hktdc.com (2017). Hong Kong Trade Development Council (HKTDC) Official Website. [online] Available at: www.china-trade-research.hktdc.com/business-news/article/China-Consumer-Market/China-s-Cosmetics-Market/ccm/en/1/1X000000/1X002L09.htm [Accessed 5 Jan. 2018].

International Labour Office (2002). Key Indicators of the Labour Market, 2001-2002. Psychology Press.

International Labour Organization Official Website. [online] Available at: www.ilostat.ilo.org [Accessed 5 Nov. 2019].

Jones G.G. (2006). Globalizing the Beauty Business before 1980. [online] Working knowledge. Harvard Business School. HBS: 06-056. Available at: www.hbswk.hbs.edu/item/globalizing-the-beauty-business-before-1980 [Accessed 15 May 2016].

Jones G.G., Kiron D., Dessain V., Sjoman A. (2005). L'Oréal and the Globalization of American Beauty. [online] Faculty and Research. Harvard Business School. HBS: 805-086. Available at: www.hbs.edu/faculty/Pages/item.aspx?num=32300 [Accessed 1 Jan. 2016].

Kcia.or.kr (2018). Korean Cosmetic Association Official Website. [online] Available at: www.kcia.or.kr [Accessed 5 Jan. 2018].

Loreal.com (2017). L'Oreal Official Website. [online] Available at: https://www.loreal.com [Accessed 14 Dec. 2017].

Matusow J., Cosgrove J. (2016). Top 20 Global Beauty Companies 2016. [online] Top 20 Companies. Beauty Packaging. Available at: www.beautypackaging.com [Accessed 15 Oct. 2016].

McCann P. and Iammarino S. (2013). Multinationals and Economic Geography: Location, Technology and Innovation. U.K.: Edward Elgar.

McDougall A. (2015). L'Oreal aims to be №1 for beauty in Africa with new partnership. [online] News & Analysis on Cosmetics Formulations & Packaging in Europe. Cosmetics Design Europe. Available at: www.cosmeticsdesign-europe.com/Article/2015/03/10/L-Oreal-aims-to-be-no.-1-for-beauty-in-Africa-with-new-partnership [Accessed 15 Oct. 2016].

Pg.com (2017). Procter & Gamble Official Website. [online] Available at: https://www.pg.com [Accessed 14 Dec. 2017].

Premiumbeautynews.com (2018). Premium Beauty News Official Website. [online] Available at: www.premiumbeautynews.com [Accessed 5 Jan. 2018].

Rodionova I.A. (2010). World economy: industrial sector. Moscow: RUDN (in Russian).

Rodionova I.A. (2016). Metamorphosis in the spatial organization of the world economy at the beginning of the XXI century. Moscow: University Book (in Russian).

Sokolsky V.M. (2010). International division of labor in the chemical industry: theory and facts. Moscow University Vestnik, Series 5, Geography 5, 25-34 (in Russian).

Sokolsky V.M. (2015). Structural and territorial transformation of the global chemical industry as the indicator of «a third industrial divide». Regional Research 2 (48), 153-163 (in Russian).

Trade.gov (2015). The U.S. Commercial Service and Industry & Analysis (I&A). International Trade Administration of the U.S. Department of Commerce. [online] Available at: www.trade.gov/industry/materials/asiacosmeticsmarketguide.pdf [Accessed 15 Oct. 2016].

Trade.gov (2016). Market Research Library. International Trade Administration of the U.S. Department of Commerce. [online] Available at: www.trade.gov/industry/materials/ITA.FSC.Cosmoprof.2016_final2.pdf. Accessed 21 Dec 2016.

Tradingeconomics.com (2016). Wages in Manufacturing. Trading Economics Official Website. [online] Available at: www.tradingeconomics.com/country-list/wages-in-manufacturing [Accessed 15 Oct. 2016].

Un.org (2017). Population Division. United Nations, Department of Economic and Social Affairs. [online] Available at: www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html [Accessed 3 Nov. 2017].

Unctad.org (1998). UNCTAD World Investment Report, 1998: Trends and determinants. [online] Available at: www.unctad.org/en/Docs/wir1998_en.pdf [Accessed 15 May 2016].

Unctad.org (2012). UNCTAD World Investment Report, 2012. Towards a New Generation of Investment Policies. [online] Available at: http://unctad.org/en/Publications Library/wir2012_embargoed_en.pdf [Accessed 15 May 2016].

Unctad.org (2015). UNCTAD World Investment Report, 2015: Reforming International Investment Governance. [online] Available at: www.iberglobal.com/files/2015/wir2015_overview.pdf [Accessed 15 May 2016].

Unilever.com (2017). Unilever Official Website. [online] Available at: www.unilever.com [Accessed 14 Dec. 2017].

Weber J.M. and Villebonne J.C. (2002). Differences in purchase behavior between France and the USA: the cosmetic industry. Journal of Fashion Marketing and Management, 6(4), 396-407.

Wordpress.com, (2016). Fashion Avenue News Magazine. [online] Available at: www.btetv.wordpress.com/2016/01/11/africa-the-new-eldorado-of-beauty/ [Accessed 17 Jan. 2016].

Wto.org, (2014). International Trade Statistics. World Trade Organization (WTO) Official Website. [online] Available at: www.wto.org/english/res_e/statis_e/its2014_e/its2014_e.pdf [Accessed 15 May 2016]

Received on July 18th, 2019 Accepted on January 21th, 2020

Oxana A. Klimanova[1*], Olga I. Illarionova[1]
[1] Lomonosov Moscow State University, Leninskiye Gory 1, Moscow, 119991, Russia
***Corresponding author:** oxkl@yandex.ru

GREEN INFRASTRUCTURE INDICATORS FOR URBAN PLANNING: APPLYING THE INTEGRATED APPROACH FOR RUSSIAN LARGEST CITIES

ABSTRACT. Modern approaches to urban planning assume the dualistic nature of urban green infrastructure (GI). On the one hand, green infrastructure is as an integrated network of natural and semi-natural areas, featuring a delivery of various benefits to humans. On the other hand, GI is multifunctional and provides the residents by complex of ecosystem services to be user-oriented. Most official reports and programs use common indicators that do not characterize distribution, dynamics or state of GI. In our research, we assessed the quality of GI in 15 largest Russian cities by using an integrated assessment of 13 indicators that make up three groups: the ones 1) characterizing general GI availability; 2) supporting a comfortable urban environment («recreational indicators»); and 3) forming a stable ecosystem («integrity indicators»). The cities were ranked by values of every indicator from 1 to 15 and then the results were summed and normalized to get a total mark (max. 100). To assess the development of GI elements of each group, we also ranked cities separately by values of different groups indicators. Thus, our study revealed that satisfactory marks for both recreational and integrity indicators have Ufa, Nizhny Novgorod, Kazan, Ekaterinburg, Perm and Voronezh. In contrast, Saint Petersburg, being a densely built-up city in an auspicious natural zone, got the worst result. According to the final assessment, the quality of green infrastructure in Krasnoyarsk, a large industrial city, and four cities from the steppe zone (Rostov-on-Don, Samara, Omsk, Novosibirsk) is also unsatisfactory. Our method does not cover all GI aspects (like vegetation health) and since it is based solely on remote sensing data and statistics data, there is definitely a room for improvement. However, this method, while being relatively quick and simple to accomplish, allows to assess not only general availability of GI, but its quality and distribution as well, which are essential for urban spatial planning.

KEY WORDS: green infrastructure, assessment, indicator, sustainability, urban planning

CITATION: Oxana A. Klimanova, Olga I. Illarionova (2020) Green infrastructure indicators for urban planning: applying the integrated approach for russian largest cities. Geography, Environment, Sustainability, Vol.13, No 1, p. 251-259
DOI-10.24057/2071-9388-2019-123

INTRODUCTION

Urbanization of megacities meets a lot of different environmental, economic and social problems and risks (Kötter et al. 2009). As a remedy to some of these negative consequences of urbanization, the installation of green infrastructure as opposed to grey infrastructure is identified as an alternative nature-based and cost-effective solution for improving the sustainability of the urban development (Ahern 2013; Alberti 2008).

Green infrastructure is described as an integrated network of natural and semi-natural areas and features which deliver a variety of benefits to humans (Naumann 2011). Modern approaches to urban planning assume the dualistic nature of urban green infrastructure (Benedict and McMahon 2012; Austin 2014; Lafortezza et al. 2013). In the context of landscape ecology, green infrastructure preserves its integrity due to the reasonable structure and internal links and interactions (Alberti 2008; Forman 2015). At the same time GI is multifunctional and provides the residents by complex of ecosystem services to be user-oriented (Gill et al. 2007; Hansen et al. 2014).

In the past years, an increasing number of studies have been conducted on green infrastructure indicators, which can be generally categorized into two groups. The first one includes «general» indicators that generally characterize the quantity and types of urban green space (Green infrastructure and territorial cohesion 2011; Gomez et al. 2011; Kabisch et al. 2013; Huang 2017). Data for them, mainly the share of vegetated area in the city, can be collected from remote sensing (Huang et al. 2017; Yang et al. 2014; Ahmad 2014). Remote sensing enables us to evaluate not only a current share of different green infrastructure types within the city, but to reveal its changes as well (Cheng et al. 2011; Li et al. 2015).

Based on remote sensing data, land cover models can be interpreted as a variation of this indicator (Cheng et al. 2011; Zeng 2016; Huang et al. 2017). When combined with land use statistics, they are used to assess ecosystem services (Green Infrastructure and Territorial cohesion 2011). Projects like Urban Atlas and CORINE Land Cover offer an open access for this data on European cities. Such algorithm was applied to evaluate similar indicators for Moscow (Klimanova et al. 2018). Some other countries develop national standards to regulate urban land use in functional zones. Thus, in China, permitted ranges of the percentages of urban land are allocated for different uses are established for residential, industrial, administrative use and green spaces (MOHURD 2010).

Usually the indicators of this group are not regulated by general international standards, but rather by local state or city government. For example, London Department for Communities and Local Government requires 40% of green area within the city, including green roofs and household gardens. This indicator – the percentage of urban green infrastructure – is often referred to by different European manuals for urban planning (like Handley 2003, Good Practice Guidance for Green Infrastructure and Biodiversity, TCPA 2012), Urban Green Infrastructure Planning: A Guide for Practioners (Hansen 2017). In China, key goals for the national garden city scheme was to ensure that the percentage of green spaces in urban areas should be higher than 36% and have potential opportunities to achieve more than 40% in the future (MOHURD 2010). Russia also has its national standards for urban green infrastructure. According to them, green space must occupy at least 40% of the city's area (Sanitary Regulations and Standards 2003).

Another «general» indicator from the first group is urban green infrastructure per capita. World Health Organization recommends the minimum of 9 m^2 green space per capita, while an optimum is about 50 m^2 per capita (WHO 2016). These regulations may vary in different countries (Russia, China), depending on nature zones, but they must not be below international recommendations.

The second group of «special» indicators describes the quality of green infrastructure. Generally speaking, it determines how well urban green areas fulfill their functions. It is common knowledge that vegetation removes air pollutants, reduces urban heat island effect and street noise. However, the efficiency of these functions depends on the vegetation health, particularly on the density and condition of the leaf cover. These parameters are assessed by matching NDVI with field data (Ahmad 2014; Chen 2002). Another common indicator to assess green infrastructure quality is Leaf Area Index (LAI), varying from 1 to 10. It is defined as the green leaf area per unit ground surface area, and measured by using either field date at the local scale or remote sensing for large territories (Fang 2014).

The accessibility of green areas promotes active lifestyle and outdoor activities, positively effecting physical and mental health (Annerstedt et al. 2016; Ekkel et al. 2017). To assess how much green infrastructure influences the conditions for urban recreation, *the availability of green space within easy walking distance from residential area* is measured. A number of works, including (WHO 2016) reports, turn to Accessible Natural Greenspace Standard (ANGSt) (Ma 2016), according to which the accessibility indicator can be assessed at three levels: 5 minutes walk from home – 400 m, 10 minutes – 800 m, 15 minutes – 1200 m (Pauleit 2003). Moreover, ANGSt recommends that citizens should have at least 1 accessible ha of green area within 300 m of residential area, 20 ha within 2000 m, 100 ha within 5000 m, and 500 ha within 10000 m (Handley 2003). General green infrastructure accessibility is defined as a ratio of the population with an access to green area in the 300 m buffer to the total city's population (WHO 2016). This standard was also partially used to evaluate accessibility in Google Earth Engine system (2005–2015) for 28 world megacities (Huang et al. 2017).

Special fragmentation/cohesion indicators are used to assess the integrity of urban green infrastructure (Andersson et al. 2008; Werner 2011). Basically, fragmentation is a process in which large elements are divided into smaller isolated segments with their geometric shape becoming more complex due to the economic activity (McGarigal 1995). As fragmentation disrupts the interchange of energy and matter between green areas, it primarily affects supporting and regulating ecosystem services (Zhang 2017). Of course, urban forests lack the sustainability of natural ecosystems and are more vulnerable. Still, GI connectivity and integrity allows species (mostly birds and insects) to move through all the green network and thus to create stable populations (Shavnin 2011). At the same time, however, high fragmentation means an even distribution of green elements in the city, and that is important for GI recreational role. Thus, when dealing with GI fragmentation, both sides of its interpretation should be considered.

European Environmental Agency (EEA 2017) also uses a similar fragmentation indicator in their reports – *a distribution of green urban areas (GUA)*, that is defined as the relationship between green area boundaries (edges) and all the other elements present in the city. The total length of the edges is compared with the city's urban area. The edge density provides an indication of the distribution of GUAs. A high edge density in a city indicates a relatively high number of green areas that border residential, commercial, industrial or other public buildings. Consequently, a higher value for the indicator may be due to a long boundary length, i.e. more small patches. The agency, however, does not establish certain values of indicators.

Finally, the green infrastructure integrated assessment is often considered a part of urban sustainability (or urban ecosystems health) research (Bell et al. 2008; Mell 2009; Pakzad et al. 2016; Zeng 2016). For instance, urban ecosystems health of Chinese megacities is described by parameters like per capita green area, built-up area green coverage rate, percentage of different type of land system (cropland, forest cover, grassland, water area and unutilized land) (Zeng et al. 2016). All of them, though, belong to the assessment of the facility and land systems.

In most cases, the common method for assessing urban green infrastructure incorporates an index system and weight determination for a final value that reflects the green infrastructure status of each city. Other works have attempted (Cheng 2011; Yang 2014; Klimanova 2018) to combine the advancements made in urban natural ecosystem assessment with the distinct purposes for recreation of urban green areas. Measures of availability and accessibility have been integrated, whereas the link between these components and GI fragmentation is still weak. The assessment of urban green infrastructure requires both groups of indicators and also the accounting of spatial distribution of green spaces.

In this study, we aimed to assess the current state of urban green infrastructure by establishing an indicator system associating GI general indicators their quality. These approaches were applied to the 15 largest Russian cities to evaluate their status quo and facilitate future urban management.

MATERIALS AND METHODS

Study area

The objects of our research are the largest Russian cities with population over one million people. Eight out of fifteen cities are situated in the European part of the country, four in the Urals and three in Siberia. Their social-economic data is shown in Table 1.

The largest and most populated city is the capital – Moscow. It is the only Russian city that, according to the international indicators, can be considered a megacity and that is one of top-25 world cities with the highest GDP and GDP growth (Dobbs 2011). Saint Petersburg is the second most populated city in Russia and like Moscow is also a city of federal importance. The highest rate of population increase is found in Voronezh and Krasnoyarsk, while in Nizhny Novgorod the population is rapidly dwelling. Population growth is an important factor to consider when investigating urban GI and particularly its dynamics, since green area may not change a lot, but considering a significant population increase, it is not a positive result. A growing population means a growing demand for green infrastructure, so the «protection and conservation policy» is not sufficient. More measures for green area enhancement should be applied.

Omsk, Chelyabinsk, Ufa and Krasnoyarsk are big industrial cities with significant emissions from point source. The share of transport emissions varies from 34% in Krasnoyarsk to 94 in Moscow (Rosstat 2016). Unfortunately, while industrial centers need regulating and purifying green infrastructure first and foremost, they lack it most than others.

All of the studied cities are administrative centers of federal subjects of Russia, with most of them being the main drivers of regional economic development. They undoubtedly demonstrate the best practice of urban planning in their regions, but they still face a lot of environmental problems and need a number of measures to improve urban green infrastructure.

METHODOLOGY

The system of indicators

To assess the present state of urban green infrastructure, it is basically crucial to develop a system of indicators. Notably, indicators for urban planning should have some standard values to serve as reference when deciding on measures for improvement. It was the basic principle during the selection of indicators in our research. Moreover, indicators should be quite easy to evaluate, interpret and regulate. The aim is to choose comprehensible instruments that can characterize both strong and weak points of urban green infrastructure, indicate the most and the least effective strategy for its enhancement.

Our systemization of indicators bases on the idea that green infrastructure does not only fulfill the citizens' needs, but also supports itself as a system of integrated elements. Thus, we divided all indicators into three groups: those that 1) characterize general green infrastructure availability; 2) support a comfortable urban environment; 3) form a stable integrated system of green infrastructure. The groups and their indicators are present in Table 2.

Table 1. Socioeconomic development of largest cities in 2016

City	Natural zone	Area (km^2)	Population (million)	Annual population growth, 2007-2017, per cent	Population density (prs per km^2)	GDP. (billion USD) (2015)
Moscow	Mixed forest	2432	12.381	+1.05	5091	213.15
St Petersburg	Coniferous forest	1439	5.282	+1.4	3670	61.5
Novosibirsk	Forest-steppe	481	1.603	+1.38	3333	11.5
Ekaterinburg	Coniferous forest	401	1.456	+0.97	3631	13.8
Perm	Coniferous forest	806	1.456	+0.53	1806	8.1
Nizhny Novgorod	Mixed forest	317	1.262	-0.11	3981	11.6
Kazan	Mixed forest	635	1.232	+0.94	1940	8.5
Chelyabinsk	Forest-steppe	504	1.178	+0.89	2337	8.5
Samara	Forest-steppe	543	1.170	+0.24	2155	9.2
Rostov-on-Don	Steppe	355	1.125	+0.64	3169	9.3
Ufa	Mixed forest	667	1.116	+0.83	1673	8.6
Krasnoyarsk	Coniferous forest	378	1.083	+1.52	2865	8.9
Voronezh	Forest-steppe	601	1.040	+2.15	1730	6.5
Omsk	Steppe	580	1.016	+0.35	1752	7.3
Volgograd	Steppe	861	1.015	+0.27	1179	6.0

Source: Russian Federal State Statistics Service (2016) and Institute of Urban Economics (2015)

Table 2. Indicator system for green infrastructure integrated assessment

Group	#	Indicators and units of measurement
1. General green infrastructure availability	1.1	Total urban green infrastructure, per cent
	1.2	Green infrastructure availability, sq. m per capita
	1.3	Green infrastructure availability of the least vegetated city district, sq. m per capita
	1.4	Difference between urban green infrastructure per capita in the most and the least vegetated city districts, sq. m per capita
	1.5	Change of the percentage of total green infrastructure, per cent
2. Comfortable urban environment formation	2.1	Urban green infrastructure for recreation availability, sq. m per capita
	2.2	Share of residential area without green infrastructure in the 300-meter buffer, per cent
	2.3	Share of residential area without green infrastructure in the 800-meter buffer, per cent
3. Green infrastructure sustainability	3.1	Tree vegetation in green infrastructure, per cent
	3.2	Agricultural lands in green infrastructure, per cent
	3.3	Green infrastructure fragmentation index
	3.4	Urban protected areas, per cent
	3.5	Change of tree cover share in green infrastructure2, per cent

Indicators assessment

Integrated assessment consists of three stages (fig. 1). Stage one includes the preparation of spatial data for the evaluation. We used materials on tree cover within cities for 2000-2016 (Hansen 2013) and land use categories (from OpenStreetMap (OSM) to define agricultural lands and green areas. The administrative borders for the cities are also taken from corresponding OSM layers. To assess the per capita green infrastructure for recreation, we used the following OSM layers and categories: vegetation (wood), landuse / vegetation (forest, orchard, grass, meadow, village_green, recreation_ground), landuse (garden, park, allotments, cemetery).

To assess how much green infrastructure influences the conditions for urban recreation, the availability of green space within easy walking distance from residential area is measured. A number of works, including (WHO 2016) reports, turn to Accessible Natural Greenspace Standard (ANGSt) (Ma 2016), according to which the accessibility indicator can be assessed at three levels: 5 minutes walk from home – 400 m, 10 minutes – 800 m, 15 minutes – 1200 m (Pauleit 2003). Moreover, ANGSt recommends that citizens should have at least 1 accessible ha of green area within 300 m of residential area, 20 ha within 2000 m, 100 ha within 5000 m, and 500 ha within 10000 m (Handley 2003). In our study we use 300 and 800 m buffers.

The most common parameters to fragmentation measurement are the number of patches, patch density (patches per 100 ha), mean patch size, total edge length, edge density (the sum of all green elements' edges, divided by total urban area – m/ha), total core area and percentage of core area (the one that is not influenced by edge effect) and Euclidean nearest neighbor distance (the shortest distance between close patches) (McGarigal 2014). Fragmentation parameters were evaluated in Fragstats by inputting Hansen tree cover rasters with 30 meter spatial resolution and using six operations from «class metrics» block: Edge Density (ED), Largest Patch Index (LPI), Patch Density (PD), Number of Patches (NP), Total Core Area (TCA), Euclidean Nearest Neighbor Distance (ENN). For the «background value» we assigned the class of non-tree vegetation of a classified raster image. To do so, we input a special matrix into Fragstats, so that a value of pixels with tree vegetation would be considered true and non-tree vegetation – false. We also verified the results by doing the same operation with NDVI classified images that we created from Landsat 5, 7 and 8 images for 2000 and 2016. Four classes were defined: water surface; unvegetated surface, sparse vegetation cover and dense vegetation cover. The first class includes NDVI values below 0; the second class – up to 0,2; the third – up to 0.35; and the last one consists of all higher values. The edge depth defines the distance of edge effect and consequently the size of core area, so we took 120 meters for this parameter and created a matrix for the operation as well.

On stage two we assessed the indicators. To evaluate per capita ones, we used the official statistical data on urban population. The summed values of indicators for the cities were later ranked from 1 to 15. The same operation was done separately for the first, second and third groups of indicators. Fragmentation index is also a result of ranking cities by six Fragstats parameters we have described earlier. Their values were summed and then ranked again. As a result, the maximum value of fragmentation index is 15, and the minimum is 1.

Stage three is a final integrated assessment of urban green infrastructure, based on all chosen indicators, and the following normalization of its values.

RESULTS

Statistics of the indicators.

Table 3 shows variabilities for 15 cities.

First group indicators. The highest percentage of total green infrastructure (#1.1) is found in Perm and Moscow, while Saint Petersburg, Volgograd and Samara the lowest total share of green areas. As to the per capita green infrastructure (#1.2), the best results are discovered in Volgograd and Ufa, and the worst in Saint Petersburg. Per capita green infrastructure of the least vegetated city district (#1.3) is another per capita indicator that, unlike the previous one, characterizes the spatial homogeneity of green

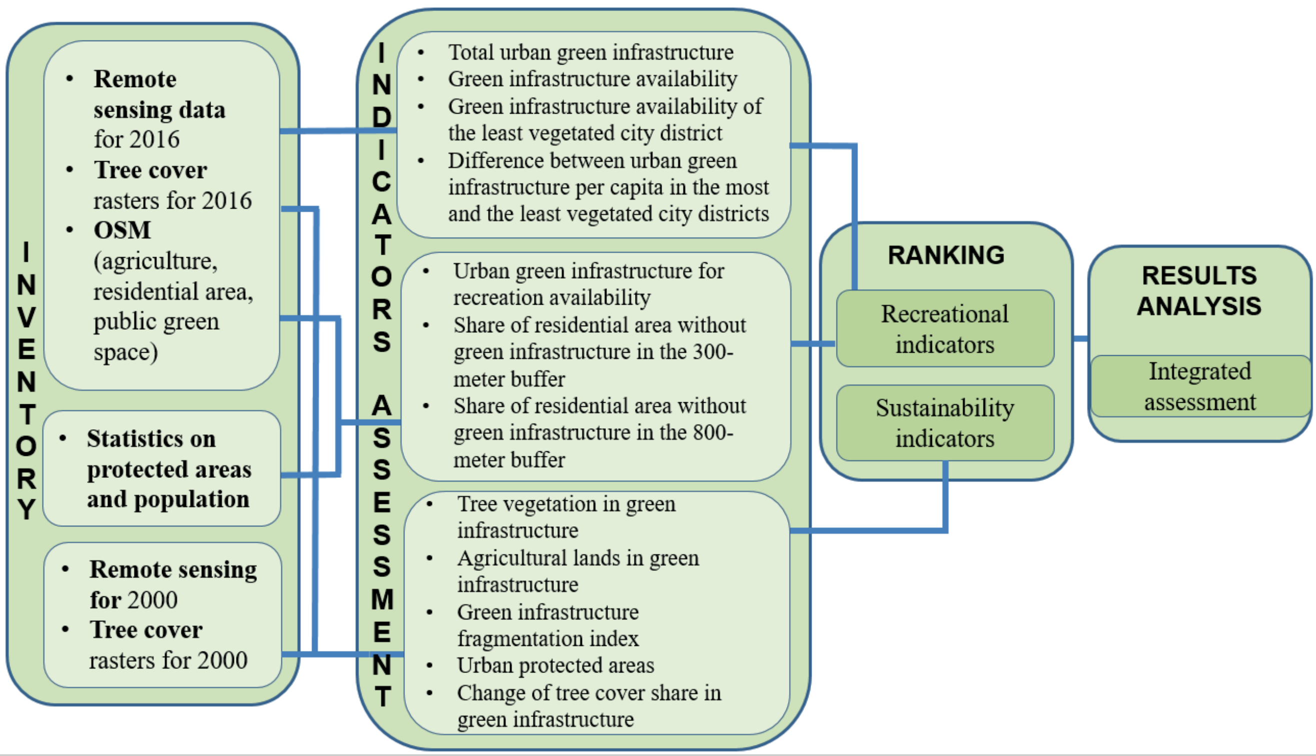


Fig. 1. Three stages of study

infrastructure. Thus, the least vegetated districts of Saint Petersburg, Moscow and Volgograd are the worst among other cities, while the best are in Novosibirsk, Kazan and Perm. The difference between urban green infrastructure per capita in the most and the least vegetated city districts (#1.4) is especially high in Moscow, Saint Petersburg and Samara, and less significant in Rostov-on-Don and Omsk. The change of the percentage of total green infrastructure (#1.5) is the last «general» indicator, that has shown the most distinctive decrease of green area in Saint Petersburg, and a particularly big increase in Nizhny Novgorod and Novosibirsk.

Second group indicators. Per capita urban green infrastructure for recreation (#2.1) is the highest in Perm and Voronezh, and the lowest in Rostov-on-Don and Omsk. The largest share of residential area without green infrastructure in the 300-meter buffer (#2.2) is found in Volgograd and Ekaterinburg, while Chelyabinsk and Omsk have the best accessibility to urban green areas. As to the availability of green infrastructure in the 800-meter buffer (#2.3), the most provided cities are Saint Petersburg, Ekaterinburg, Moscow and Kazan, and the least ones are Rostov-na-Dony and Ufa.

Third group indicators. Tree vegetation takes the largest part of urban green infrastructure (#3.1) in Perm and Voronezh, while in Chelyabinsk, Omsk and Volgograd green area mostly consists of non-tree vegetation. Agricultural lands occupy the most significant share of urban area (#3.2) in Volgograd, Saint Petersburg and Kazan, and the smallest percentage in Krasnoyarsk and Novosibirsk. Green infrastructure fragmentation index (#3.3) is maximum in Samara and Kazan, and minimum in Saint Petersburg. As to the urban protected areas (#3.4), their share is distinctively high in Chelyabinsk and Voronezh, but quite low in Krasnoyarsk and Omsk. In most cases, the change of tree cover is insignificant. However, the positive dynamics in Voronezh and Samara, and the negative one in Ekaterinburg are worth mentioning.

The variance ratio reflects the gaps among these cities, which is calculated by dividing the variance by the mean value of each variable. It demonstrates how even or uneven green infrastructure aspects are developed in different cities. As we can see, variation ratio is high for most indicators, particularly for the difference between urban green infrastructure per capita in the most and the least vegetated city districts, meaning that results for this one vary greatly in different cities – while districts in Rostov-on-Don and Omsk are quite evenly vegetated, the situation in Saint Petersburg and Moscow is 100 times worse. On the whole, most cities are vegetated quite differently. In only four cases – total green infrastructure percentage, tree vegetation in green infrastructure percentage and «change» indicators – most cities got results that are relatively close to mean values. Moreover, the mean value of indicator 1.1 shows that most cities have the share of green area above recommended 40%. The same applies to the per capita indicators since there is more than 9 m2 GI per capita in every city. Though, the same cannot be said about the least vegetated districts in Ufa, Saint Petersburg, Kazan, Krasnoyarsk, Moscow, Novosibirsk and Rostov-on-Don, where central parts lack green elements.

Relation between indicators of different groups

As we have mentioned in part two, for every indicator the cities were ranked from 1 to 15. So, a total sum of all indicator-ranks for each city is actually a value of its green infrastructure integrated assessment. Later, we also evaluated the sum of ranking points for each group of indicators separately. Since the number of indicators differs in three groups and thus their summed values cannot be compared, we normalized the results by using the following formula: $I_n = I_x \times 100 / I_{max}$, where I_n is a normalized value, I_x – a sum of ranking points for a city, I_{max} – a maximum possible sum of ranking points in a group. In case of dynamics indicators (1.5 and 3.5), we give a city 0 points if the change is negative. So, for the first and third group I_{max} is 63, for the second one – 45. Normalized values vary between 1 and 100.

Figure 2 demonstrates the relation between general, comfort and integrity indicators. When comparing results of the first two groups («recreational» and «integrity» indicators), we can divide cities into three sub-groups, depending on their interrelation:

- a total mark for «recreational» indicators is higher than a total mark for «integrity» indicators;
- a total mark for «integrity» indicators is higher than a total mark for «recreational» indicators;
- marks for «recreational» and «integrity» indicators are similar.

Table 3. Statistics of the variables for 15 cities

#	Maximum value	Minimum value	Mean value	Variance ratio
1.1	73	36	53.6	1.98
1.2	466	137	247.87	45.72
1.3	59	1	17.13	17.00
1.4	5123	47	800.20	2066.39
1.5	11.6	-11.7	2.34	14.69
2.1	406	33	131.07	73.49
2.2	74	25	46.20	3.91
2.3	33	2	9.60	7.65
3.1	62	12	34.9	5.48
3.2	24.5	0.2	6.40	9.48
3.3	61	20	48.00	1.90
3.4	30	0.3	8.68	7.93
3.5	8.7	-1.8	0.70	12.44

The first sub-group includes Ufa, Rostov-on-Don, Omsk, Novosibirsk, Nizhny Novgorod and Krasnoyarsk. Chelyabinsk, Perm, Saint Petersburg, Moscow, Voronezh, Kazan and Volgograd belong to the second sub-group, while Ekaterinburg and Samara have approximately equal normalized values of both indicators groups.

Integrated green infrastructure assessment

Besides getting normalized values for three groups, the same operation was done for all 13 indicators. Thus, the research has revealed that Voronezh, Ekaterinburg and Perm have the most effective and sustainable green infrastructure (Fig. 3). The first two also have the largest share of protected green areas, while the percentage of total green area is one of the highest in the latter. These cities got 65-75 normalized points out of 100, so the quality of their green infrastructure can be described as «good'. In contrast, Saint Petersburg, being a densely built-up city with limited open space, got the worst result. According to the final assessment, the quality of green infrastructure in Krasnoyarsk, a large industrial city, and four cities from the steppe zone (Rostov-on-Don, Samara, Omsk, Novosibirsk) is also unsatisfactory. Their total mark is about 40. Other cities got 50-65 points, and since their marks are above average, the quality of their green infrastructure can be defined as satisfactory.

Since our integrated assessment is based on ranking, the final marks do not reveal the results' difference between the

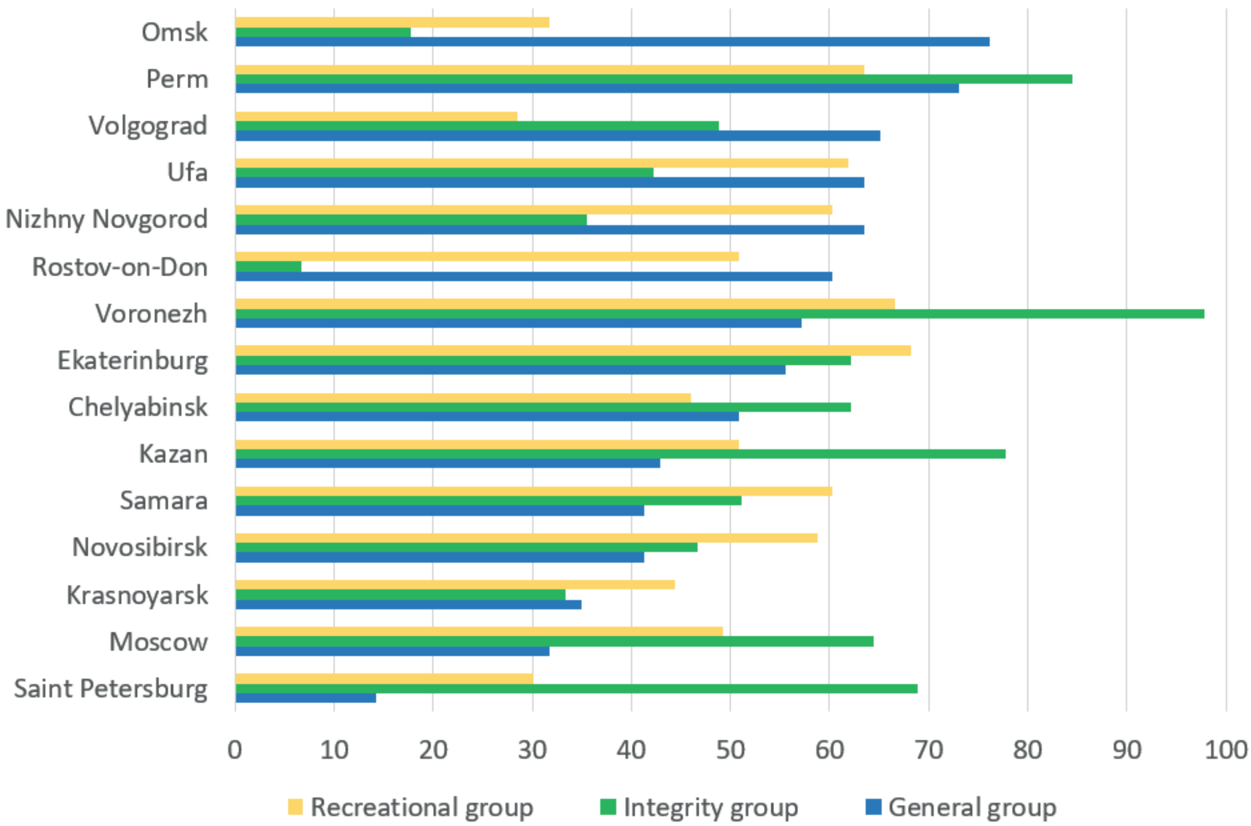


Fig. 2. Normalized values of indicators from 3 groups

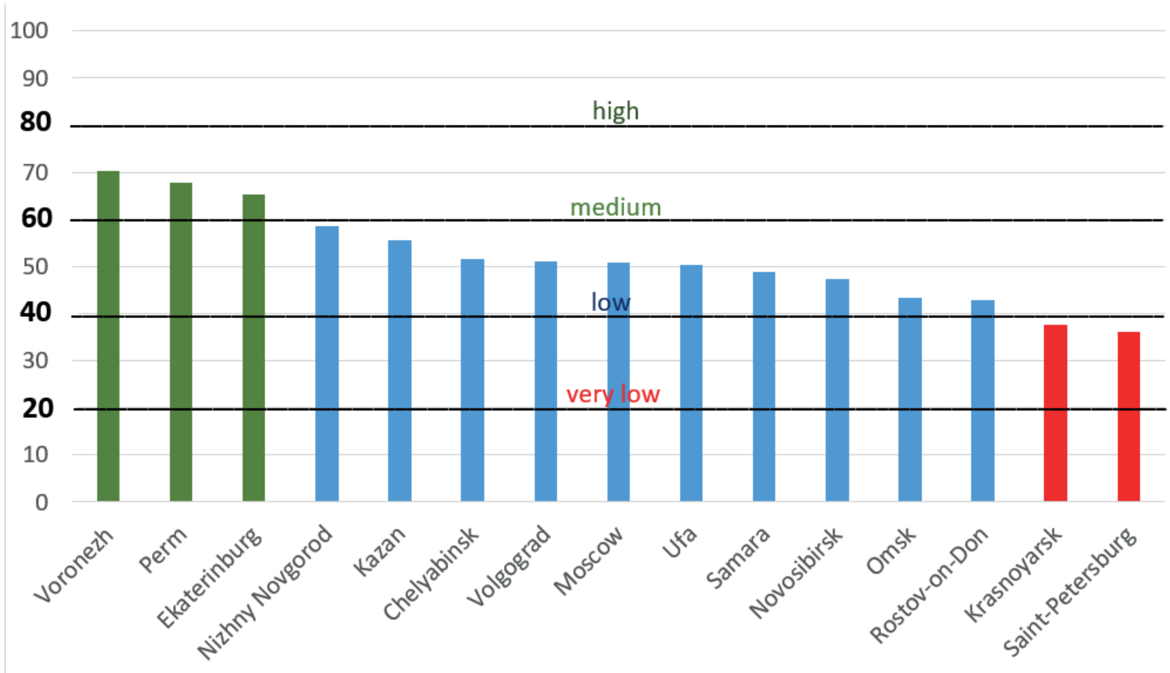


Fig. 3. Normalized values and marks for green infrastructure state

cities and we cannot fully comprehend the weight behind the normalized values. When there is a small variance of minimum and maximum values, it does not mean that the city that got 1 point for this indicator has «unsatisfactory» results. For instance, for per capita green infrastructure Volgograd got 15 points and Saint Petersburg – 1. However, the leader's absolute values for this indicator are only 4 times better than the absolute values of the last one, while in some cases (like urban protected areas percentage) the absolute values of the first city can be 30 times better.

DISCUSSION

Green infrastructure plays a key role in a formation of a comfortable urban environment, but at the same time it is affected by intense processes of urbanization, like surface sealing and active building, transport and industrial activity. However, cities are not only the cause of negative influence, but also the main decision-making centers for green infrastructure improvement.

The development of green infrastructure is a necessary measure for keeping urban sustainability, particularly in the context of the global environment change. Sustainable cities and communities is among the priority UN sustainable development goals (Sustainable Development Goals 2015). Green infrastructure planning is an essential part of urban development policy, as it combines both nature conservation measures and plans for the formation of a comfortable urban environment. The conflict between contradicting environmental and users' approaches is the major problem of urban nature protection, that is especially keen in large cities.

The Russian national program «Housing and urban environment» (2019) includes 36 indicators, but only six of them describe green infrastructure: 1) the percentage of public green area, 2) total share of urban green area, 3) the state of green area, 4) attractiveness of green area, 5) the variety of services of green area, 6) the percentage of population with access to public green area. However, most of these indicators do not describe spatial heterogeneity and GI sustainability.

Our research on 15 cities gives an integrated method of green infrastructure assessment, which offers two main contributions.

The first one is a tested integrated system of 13 indicators for the current state and quality of urban green infrastructure. Among other common «traditional» indicators like the percentage of total green area and total green area per capita, we also take into account the spatial differences of green infrastructure and consider tree vegetation more valuable in the context of ecosystem services. In our study we used data from both open source geoportals and official statistics base, that guarantees its consistency. Obviously, the suggested method is not perfect. The full assessment of green infrastructure quality requires a big variety of data sources, many of which are not available for all 15 cities at the moment. At the same time, the method we suggest is relatively simple in evaluation and use, and can be applied for cities in other countries as well.

The second contribution is a method, that allows to combine two blocks of green infrastructure indicators – the ones characterizing the influence on the urban environment, and those defining the integrity of green infrastructure – in one assessment. Today, the integrated multifunctional approach to green space organization should be considered the best urban planning practice. However, in most cases there are still quite a few natural areas within the city borders in Russia that actually continue to exist as self-sustaining ecosystems. Sometimes, this phenomenon is found in some other cities, particularly in Africa and South America: Rio-de-Janeiro (Illarionova 2018), Dar-es-Salaam (Lindley 2015), Caracas (Werner 2011). Our method offers a convenient way to assess whether measures directed at the formation of a comfortable urban environment are developed proportionally to the measures for green infrastructure sustainability improvement.

By assessing the change of green area in 15 largest cities in 2000 and 2016, we reveal that was insignificant (under 3%), with the most noticeable decrease concerning non-tree vegetation that covers former agricultural lands on the outskirts and wastelands in central districts. A positive aspect of this result is a conclusion that forest area remains relatively stable and does not shrink under the cities' growth. There are, though, a few cases of local forest cutting in some areas, for instance in Saint Petersburg.

We did not discover, though, a direct correlation between the sustainability of urban green infrastructure and nature zones or industrial activity. One should expect in «arid» cities a higher mark for socially significant «recreational» indicators since green infrastructure is meant to mitigate the effects of the harsh environment, while natural fragmentation should lower the marks of «integrity» indicators there. This contradiction, found in all cities of steppe zones (excluding Voronezh), only proves that green infrastructure development is not sufficient there.

Moscow, Saint Petersburg and Krasnoyarsk, even being in an auspicious forest zone with favorable conditions for vegetation, got low marks for integrity indicators, meaning that their green infrastructure is especially vulnerable. Satisfactory marks for both recreational and integrity indicators have Ufa, Perm, Nizhny Novgorod, Kazan, Ekaterinburg and Voronezh.

The most cities of steppe zone with unfavorable environmental conditions indeed belong to the group with a low-quality green infrastructure. Moreover, industrial cities like Ekaterinburg, Perm and Nizhny Novgorod do not always get the worst results, meaning an effective local urban planning policy and intense measures for improvement.

Not all indicators involved can help realize a better urban green infrastructure, but they still provide a reference point for making specific regulations to control human activity and improve eco-environment management.

CONCLUSIONS

In this research, we used an integrated method to assess the current state of green infrastructure in 15 Russian most populated cities. The major contributions and conclusions can be summarized as follows.

We offered a system of 13 indicators to assess green infrastructure quality that includes two sub-systems to characterize 1) the influence of green areas on the comfort of urban environment, and 2) the integrity of green space as an interconnected ecosystem.

Besides two sub-systems that characterize the indicators' assessment focus (social or environmental), we marked out «general» and «special» indicators: the common ones that are often used in reports on urban planning, but actually show an average picture of the situation; and those that are less «popular», but allow to assess spatial distribution and homogeneity.
We divided cities into three sub-groups, depending on the interrelation between 'sustainability' and 'recreational' indicators. That way we found out that six cities have green infrastructure that lacks recreational functions, while green areas in six other cities is most vulnerable.

Thus, cities that got the lowest marks (below 50) for integrity indicators should dedicate more attention to the restoration and improvement measures to reduce fragmentation. In contrast, cities, ranking last in the assessment of green infrastructure mitigating effects, are better to review their urban planning strategies to achieve an optimal spatial distribution of green areas and a comfortable urban environment.

REFERENCES

Ahern J. (2013). Urban landscape sustainability and resilience: The promise and challenges of integrating ecology with urban planning and design. Landscape Ecology, 28(6), 1203-1212. DOI: 10.1007%2Fs10980-012-9799-z.

Ahmad A. (2014). Temporal changes in urban green space based on normalized difference vegetation index. Applied Mathematical Sciences, 8(55), 2743-2751. DOI: 10.12988/ams.2014.432230.

Alberti M. (2008). Advances in urban ecology. Integrating humans and ecological processes in urban ecosystems. New York: Springer.

Andersson E. and Bodin Ö. (2008). Practical tool for landscape planning? An empirical investigation of network based models of habitat fragmentation. Ecography, 32(1), 123-132. DOI: 10.1111/j.1600-0587.2008.05435.x.

Annerstedt van den Bosch M., Mudu P, Uscila V., Barrdahl M., Kulinkina A., Staatsen B., Swart W., Kruize H., Zurlyte I. and Egorov A.I. (2016). Development of an urban green space indicator and the public health rationale. Scandinavian Journal of Public Health, 44, 159-167. DOI: 10.1177/1403494815615444.

Austin G. (2014). Green Infrastructure for landscape planning: Integrating human and natural systems. New York: Routledge Press.

Bell S. and Morse S. (2008). Sustainability indicators: Measuring the immeasurable? Earthscan, London: Sterling.

Benedict M.A. and McMahon E.T. (2012). Green infrastructure: Linking landscapes and communities. Chicago: Island Press.

Chen X. (2002). Relationships among phenological growing season, time–integrated normalized difference vegetation index and climate forcing in the temperate region of Eastern China. International Journal of Climatology: A Journal of the Royal Meteorological Society, 22 (14), 1781-1792. DOI: 10.1002/joc.823.

Chen B. and Wang R. (2014). Integrated ecological indicators for sustainable urban ecosystem evaluation and management. Ecological Indicators, 47, 1-4. DOI: 10.1016/j.ecolind.2014.09.020.

Cheng L., Li F. and Deng H.F. (2011). Dynamic of land use and its ecosystem services in China's megacities. Acta Ecologica Sinica, 31(20), 6194-6203.

Dobbs R. (2011). Urban world: mapping the economic power of cities. New York: McKinsley Global Institute.

Ekkel E.D. and de Vries S. (2017). Nearby green space and human health: evaluating accessibility metrics. Landscape and Urban Planning, 157, 214-220. DOI: 10.1016/j.landurbplan.2016.06.008.

Environmental European Agency (EEA) (2017). Indicators for urban green infrastructure. [online] Environmental European Agency. Available at: www.eea.europa.eu/themes/sustainability-transitions/urban-environment/urban-green-infrastructure/indicators_for_urban-green-infrastructure [Accessed 1 Sep. 2019].

Fang H. (2014). Leaf area index models. Reference Module in Earth Systems and Environmental Sciences, 1-11.

Forman R. (2015). Urban ecology: science of cities. Cambridge: Cambridge University Press.

Gomez F., Jabaloyes J., Montero L., De Vicente V. and Valcuende M. (2011). Green areas, the most significant indicator of the sustainability of cities: research on their utility for urban planning. Urban Planning Development, 311, 311-328. DOI: 10.1061/(ASCE)UP.1943-5444.0000060.

Gill S.E., Handley J.F., Ennos A.R. and Pauleit S. (2007). Adapting cities to climate change: the role of the green infrastructure. Built Environment, 33(1), 115-133. DOI: 10.2148/benv.33.1.115.

Green Infrastructure and Territorial Cohesion. (2011). Technical Report EEA. [online]. Available at: www.eea.europa.eu/publications/green-infrastructure-and-territorial-cohesion. [Accessed 1 Sep. 2019].

Handley J. (2003). Providing accessible natural green space in towns and cities: a practical guide to assessing the resource and implementing local standards for provision. London: Countryside Council for Wales.

Hansen R. (2017). Urban green infrastructure planning: a guide for practitioners. Copenhagen: University of Copenhagen.

Hansen R. and Pauleit S. (2014). From multifunctionality to multiple ecosystem services? A conceptual framework for multifunctionality in green infrastructure planning for urban areas. Ambio, 43, 516-529. DOI: 10.1007/s13280-014-0510-2.

Hansen M.C. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342, 850-853. DOI: 10.1126/science.1244693.

Huang C. (2017). Green spaces as an indicator of urban health: evaluating its changes in 28 mega-cities. Remote Sensing, 9, 12. DOI: 2072-4292/9/12/1266.

Huang H., Chen Y., Clinton N., Wang J., Wang X., Liu C., Gong P., Yang J., Bai Y., Zheng Y. and Zhu Z. (2017). Mapping major land cover dynamics in Beijing using all Landsat images in Google Earth Engine. Remote Sensing of Environment, 202, 166-176. DOI: 10.1016/j.rse.2017.02.021.

Illarionova O. (2018). Transformation of green infrastructure in large cities of South America. Vestnik Moskovskogo Universiteta, 5(3), 23-29 (in Russian). Available at: www.vestnik5.geogr.msu.ru/jour/article/view/414 [Accessed 1 Sep. 2019].

Institute of Urban Economics (2015). Ranking of Russian Central Cities. [online] Available at: http://www.urbaneconomics.ru/centr-obshchestvennyh-svyazey/news/reyting-stolichnyh-gorodov-rossii-ot-fonda-institut-ekonomiki (in Russian) [Accessed 1 Sep. 2019].

Kabisch N. and Haase D. (2013). Green spaces of European cities revisited for 1990–2006. Landscape and Urban Planning, 110, 113-122. DOI: 10.1016/j.landurbplan.2012.10.017.

Klimanova O. (2018). Impacts of urbanization on green infrastructure ecosystem services: the case study of post-soviet Moscow. Belgeo. Revue belge de géographie, 4. Available at: www.journals.openedition.org/belgeo/30889 [Accessed 1 Sep. 2019]. DOI: 10.4000/belgeo.30889.

Kötter T. and Friesecke F. (2009). Developing urban indicators for managing mega cities. Land governance in support of the MDGs: Responding to New Challenges, 4-21.

Lafortezza R., Davies C., Sanesi G. and Konijnendijk C.C. (2013). Green infrastructure as a tool to support spatial planning in European urban regions. iForest-Biogeosciences and Forestry, 6, 102-108. DOI: 10.3832/ifor0723-006.

Li X., Gong, P. and Liang L. (2015). A 30-Year (1984–2013) record of annual urban dynamics of Beijing city derived from Landsat data. Remote Sensing of Environment, 166, 78-90. DOI: 10.1016/j.rse.2015.06.007.

Lindley S. (2015). Green infrastructure for climate adaptation in African cities. Urban vulnerability and climate change in Africa. Springer International Publishing.

Ma J. (2016). Green infrastructure and urban liveability: Measuring accessibility and equity. Doctoral dissertation. [online]. Available at: www.researchspace.auckland.ac.nz/handle/2292/31220 [Accessed 1 Sep. 2019].

McGarigal K. (2014). FRAGSTATS Help. Documentation for FRAGSTATS, 4.

Mell I. (2009). Can green infrastructure promote urban sustainability? Proceedings of the Institution of Civil Engineers – Engineering Sustainability, 162(1), 23-34. DOI: 10.1680/ensu.2009.162.1.23.

Ministry of Housing and Urban-Rural Development of the People's Republic of China. MOHURD. (2010). Eco-City Implementation Guideline for Green Building. Beijing.

Naumann S. (2011). Design, Implementation and Cost Elements of Green Infrastructure Projects. Final report of European Commission. [online]. Available at: www.ec.europa.eu/environment/enveco/biodiversity/pdf/GI_DICE_FinalReport.pdf. [Accessed 1 Sep. 2019].

Niemeijer D. and de Groot R.S. (2008). A conceptual framework for selecting environmental indicator sets. Ecological Indicators, 8, 14-25. DOI: 10.1016/j.ecolind.2006.11.012.

Pakzad P. and Osmond P. (2016). Developing a sustainability indicator set for measuring green infrastructure performance. Procedia, Social and Behavioral Sciences, 216, 68-79. DOI: 10.1016/j.sbspro.2015.12.009

Pauleit S. (2003). Promoting the natural green structure of towns and cities: English nature's accessible natural greenspace standards model. Built Environment, 29(2), 157-170. DOI: 10.13140/RG.2.1.2343.1121.

Russian Federal State Statistics Service. Rosstat. (2016). Environment. [online]. Available at: www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/environment/ [Accessed 1 Sep. 2019].

Sanitary Regulations and Standards (2003). SANPIN 2.2.1/2.1.1.1200-03 Sanitary Protection Zones and Sanitary Classification of Industrial, Residential and Other Objects. Moscow: Standartinform.

Shavnin S.A. (2011). Spatial structure of urban forest ecosystems as an indicator of their sustainability. Izvestiya OGAU, 29(1), 12-16.

Sustainable Development Goals (2015). Transforming Our World. General Assembly. [online] Available at: www.unfpa.org/resources/transforming-our-world-2030-agenda-sustainable-development. [Accessed 1 Sep. 2019].

Werner P. (2011). The ecology of urban areas and their functions for species diversity. Landscape and Ecological Engineering, 7(2), 231-240. DOI: 10.1007/s11355-011-0153-4.

World Health Organization (WHO) (2016). Urban Green Spaces and Health: A Review of Evidence; World Health Organization: Geneva, Switzerland. [online] Available at: www.euro.who.int/en/health-topics/environment-and-health/urban-health/publications/2016/urban-green-spaces-and-health-a-review-of-evidence-2016 [Accessed 1 Sep. 2019].

Yang J., Huang C., Zhang Z. and Wang L. (2014). The temporal trend of urban green coverage in major Chinese cities between 1990 and 2010. Urban Forestry & Urban Greening, 13(1), 19-27. DOI: 10.1016/j.ufug.2013.10.002.

Zeng C., Deng X., Xu Sh., Wang Y. and Cui J. (2016). An integrated approach for assessing the urban ecosystem health of megacities in China. Cities, 53, 110-119. DOI: 0.1016/j.cities.2016.01.010.

Received on September 19th, 2019 Accepted on January 21th, 2020

Marina A. Chichaeva[1*], Mikhail Yu. Lychagin[2], Anton V. Syroeshkin[1], Olga V. Chernitsova[2]
[1] Peoples Friendship University of Russia (RUDN University), Moscow, Russia
[2] Lomonosov Moscow State University, Moscow, Russian Federation
***Corresponding author:** machichaeva@gmail.com

HEAVY METALS IN MARINE AEROSOLS OF THE AZOV SEA

ABSTRACT. The content of heavy metals and Al in the aerosol matter over the Sea of Azov has been studied. According to the special test the vast majority of samples were attributed to the type of marine aerosol. The ranges of contents were determined as following: Fe (200 – 2000 ng / m^3), Al (20 – 200 ng / m^3), Zn (10 – 280 ng / m^3), Cu (2 – 23 ng / m^3), Ni (1 – 16 ng / m^3), Pb (3 -30 ng / m^3), Cd (0.4 –2.8 ng / m^3); Mn (3 – 23 ng / m^3), Cr (1 – 15 ng / m^3). The spatial distribution of HMs in the marine aerosol of the Sea of Azov depends on the influence of the river-sea geochemical barrier zone in the Taganrog Bay and the anthropogenic impact of the coastal industrial cities. HM concentrations decrease from the northern coast of the bay and the mouth of the Don River towards the open sea. The maximum HM content in marine aerosol observed in the mouth area of the Don River. It may be associated with the HM accumulation at the river-sea geochemical barrier, and also with the anthropogenic impact of the cities of Rostov-on-Don, Azov and Taganrog. Anthropogenic impact of the city of Mariupol cause the maximum values of Fe, Cr, and Cd in marine aerosol matter of the western part of the Taganrog Bay.

KEY WORDS: heavy metals, spatial distribution, aerosol, micro particles, sea surface microlayer, pollution, Don River, Taganrog Bay

CITATION: Marina A. Chichaeva, Mikhail Yu. Lychagin, Anton V. Syroeshkin, Olga V. Chernitsova (2020) Heavy Metals In Marine Aerosols Of The Azov Sea. Geography, Environment, Sustainability, Vol.13, No 2, p. 1-8
DOI-10.24057/2071-9388-2020-11

INTRODUCTION

The aerosol of the surface layer of the atmosphere, including over water areas, is a mixture of particles of various genesis – from the local sources to the long-range atmospheric transport. Calculations show that over the territory of Europe, marine aerosol accounts for about 25-36% of the total aerosol matter (Georgoulias et al. 2016). Aerosol transport from the continent to water bodies is currently well reported (Marín-Beltrán et al. 2019; Akinori et al. 2019; Mahowald, 2018; Von Glasow, 2013; Jordi et al. 2012; Paytan et al. 2009). The phenomena of transcontinental transport of aerosols and abundant atmospheric deposition to the surface of the ocean in the Arctic zone, over the waters of the North Atlantic are known (Vinogradova et al. 2019; Maslennikova et al. 2018; Sullivan et al. 2017; Lukashin et al. 2018; Vinogradova et al. 2017). It is shown that among the sources of marine aerosol generation (marine, dust, urban, black carbon, volcanic), the ocean surface is the most powerful, delivering up to 10^{16} kg per year (Kondratiev et al. 1999). The reverse process of transporting aerosol matter of marine origin towards coastal territories is also being investigated.

The concentrations of chemical elements in the Sea Surface Microlayer (SSM), their interaction and removal with marine aerosol have been reported in a number of papers (Li et al. 2019; Li et al. 2018, Rastelli et al. 2017). It has been shown that marine aerosol is formed mainly from SSM during the destruction of air bubbles that occur in the thickness of the sea water during gas evolution in the dispersed phase. In addition, it is formed by a direct wind breakdown of water droplets from the sea surface when waves collapse. Aerosol particles formed under these conditions are referred by dimension to the PM 10 group. The mechanisms of marine aerosol formation suggest a significant similarity between the chemical composition of aerosol particles and the surface microlayer (Lapshin et al. 2002; O´Dowd et al. 2007; O´Dowd et al. 2002).

The climatic role of marine aerosols is significant (Van Dolah et al. 2000; Li et al. 2019). Cloudiness is strongly modulated by the emission of aerosols from the sea surface, affecting its formation and microphysical properties. The mechanism of the effect of aerosol on cloudiness in general can be described as follows. An increase in the concentration of aerosol in the atmosphere changes the water content of clouds and the size of cloud droplets, prolonging the lifetime of the cloud and thus reducing the amount of precipitation (Brooks et al. 2017; Chandrakar et al. 2016; Buseck et al. 1999; Takata et al. 2009).

Many studies are currently devoted to studying the concentration of heavy metals (HM) in aerosol particles. Scientific research is carried out within the framework of the UN programs: GESAMP, National Institute of Health, USA – Marine biotoxins. The mechanisms of toxicant concentrating in SSM and their further aerosol transportation in coastal areas are studied (Li et al. 2018). The accumulation and transfer of toxic substances at the ocean-atmosphere boundary can lead to significant pollution of the near-water layer of the atmosphere (Song et al. 2019; Qureshi et. al., 2009; Kolesnikov et al. 2005). The studies are carried out by groups of scientists in various areas of the world (Furness et al. 2017), the North Sea (Salomons et al. 1988), the Arctic Seas (Shevchenko et al. 2003; Golubeva et al. 2011), the Atlantic and Indian Oceans (Witt et al. 2006; Rädlein et al. 1992), Antarctic seas (Tuohy et al. 2015). HMs were found in aerosol matter above the Ross Sea off the Antarctic coast.

HMs as part of aerosol particles are involved in global atmospheric circulation processes, and therefore HM monitoring in marine aerosol considered urgent. The European Monitoring and Evaluation Program (EMEP; http://www.emep.int) and the Acid Deposition Monitoring Network in East Asia (EANET; http://www.eanet.cc) internationally cover the origin, spatial distribution aerosols, monitor their chemical composition, use methods of computer simulation of migration flows (Sajeev et al. 2014). The EMEP and EANET monitoring programs rely on satellite-based atmospheric sounding data, as well as data on the chemical composition of aerosols (including HM concentrations) from a network of reference stations in Eurasia, where aerosol sampling is carried out. Computer simulation is corrected by observations on the ground. Thus, remote monitoring methods rely on routine monitoring of aerosol composition.

Toxic substances entering the atmosphere with marine aerosol microparticles, along with urban aerosol sources, pose a serious potential threat to public health in the coastal zone of the seas and oceans (Mahowald et al. 2018; Aryasree et al. 2015; Van Dolah et al. 2000; Syroeshkin 2002; Lapshin 2002). HM entering the body through the respiratory system, penetrating the skin, inhalation and swallowing pathways (Wang et al. 2018) can increase morbidity and mortality from cardiovascular, respiratory diseases, cirrhosis, anemia, neuropathy (Liu et al. 2018), and also affect the kidneys, liver and gastrointestinal tract, causing cancer (Csavina et al. 2013; Izhar et al. 2016; Sánchez-Rodas 2017). There is an increased risk of pulmonary (asthma) and allergic diseases (Walsh et al. 2017; Kirkpatrick et al. 2011; Fleming et al. 2009). The role of marine aerosol as a carrier of palitoxins and endotoxins to the coastal zone has been shown (Patocka et al. 2018; Lang-Yana et al. 2014).

The ecological state of the Sea of Azov, despite the long decline in production activity that began in the 1990s, remains tense. The sea has a relatively small area and depth, there are large industrial centers along its shores, and thereby it is one of the most polluted seas in Russia (Klenkin et al. 2009). The mouth area of the Don River presents hydrodynamic and geochemical barrier. That's why the Sea of Azov considered an interesting and extraordinary subject for studying aerosol matter in the lowest atmosphere. The purpose of this work was to establish the levels of content and spatial distribution of HMs in the marine aerosol microparticles of the Sea of Azov.

MATERIALS AND METHODS

The paper presents data on the concentration of heavy metals in aerosol microparticles of the Sea of Azov obtained by the Azov-2006 complex ship expedition from the Taganrog Bay to the port of Temryuk. The studies were carried out in 2 stages: July 16 – 25 and September 25 – October 1, 2006. Each aerosol sampling period lasted 3-5 hours. Points of change of aerosol filters are marked at the map (Fig. 1). Simultaneously, the content of HM in the sea surface microlayers of 0.2 and 1 mm thick was studied. It has been previously shown that the main source of pollution of the atmospheric near-water layer is aerosol emission from the sea surface, which enables enrichment of the SSM due to the heterogeneous convection (Lapshin et al. 2005; Kolesnikov et al. 2005). A total of 54 aerosol filters and 168 samples of the bulk water and sea surface micro layer were collected.

Aerosol sampling was carried out on AFA-RMP-3 analytical filters using a Karcher NT 351 ECO vacuum cleaner with a maximum air flow rate of 78 dm^3/s. The basis of the AFA filter is Petryanov's filter cloth made of perchlorovinyl fiber, in accordance with state specifications TU 951892-89. This is a fibrous material using mechanical methods of particle capture, as well as the electrostatic attraction of aerosol particles to charged filter fibers. Due to this, the material is characterized by high capture efficiency. The canvas is characterized by an irregular arrangement of polymer fibers of the different thickness, so the filter has a different pore diameter. Using the sampling technique methodology, these filters provide collecting 99.9% of aerosol particles with a linear size more than 0.3 microns. 3 filters were exposed simultaneously, the exposure time was 3-5 hours. The volume of pumped air was 16.0±0.1 m^3/h, the height of the filter holder was not more than 10 m above the water edge in the headwind. Aerosols were not collected in the rain, and the sampler was not visibly sprayed. During the selection, the following meteorological parameters were controlled: wind direction and speed, air temperature, humidity, pressure, cloud cover (Syroeshkin et al. 2005).

For the elemental analysis, filters were packaged in a special snap-on polypropylene bags. After being delivered to the laboratory, the filters were incubated in Teflon bombs in 10 ml of aqua regia for 1 day. The samples were mineralized under pressure in an MDS2000 microwave oven in the following mode: 2 min. 20 sec – at 80% power, 5 min. – at 100% power. In all experiments, processing and subsequent analysis of the control filter were carried out.

The concentrations of HM in aerosol samples were determined using a SpectrAA-800 atomic absorption spectrometer with electrothermal atomization and the Zeeman effect according to the Varian protocol with modifications according to the results of international intercalibration with the IAEA MEL laboratory (Coquery et al. 2001; Kolesnikov 2005).

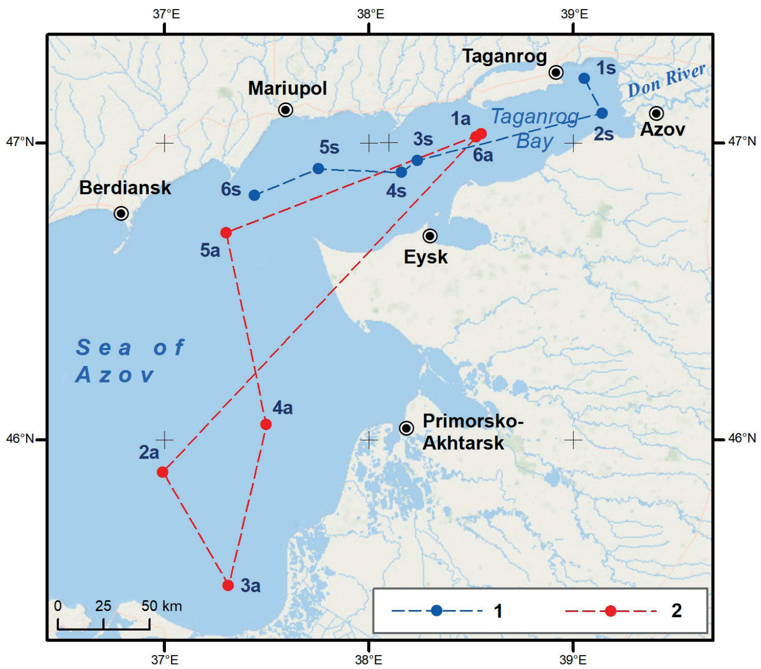


Fig. 1. Aerosol filter change points in the expedition Azov-2006: 1 – July, 2 – September – October

In parallel, an analysis of 3-5 samples was carried out. The content of elements in the reference sample, see Appendix (Table 2). In all series of measurements, the background content of elements in the AFA control filters was taken into account. The volume of air pumped through the filters allowed the elements to accumulate in an amount significantly exceeding the background content. The average relative error in determining parallel samples did not exceed 20% with a confidence level of 0.95 (Syroeshkin et al. 2005).

The size spectrum of aerosol particles was not studied. However, according to the published data, it is possible to assume the size of aerosol matter generated by the sea surface. The amount of aerosol particles in the drive layer of air sharply increases at winds of about 7-10 m/s. This is due to the emergence of a direct wind failure of water droplets from the ridges of sea waves. This mechanism was noted as one of the first. O'Dowd in the article "Marine aerosol production: A review of the current knowledge" gives a graph of the dependence of the generation of the aerosol substance on the wind speeds for the particles of Aitken and for particles from 10 to 100 nm (O'Dowd et al. 2007). Process of destruction of the air bubbles on the surface of the sea is the most important for the generation of marine aerosol on a global scale producing two types of droplets of film (0.9 microns) and reactive (2 – 2.5 microns) (Syroeshkin et al, 2005; Syroeshkin et al. 2006; Syroeshkin et al. 2014).

RESULTS AND DISCUSSION

To determine the genesis of aerosol particles, a method was used based on their dispersed composition and element content normalized to Al content. The test showed mainly marine origin of the aerosol collected during the studies (Kolesnikov et al. 2005; Syroeshkin et al. 2006; Syroeshkin 2005).

It is known that heavy metals can be found in natural waters in dissolved and suspended forms and have both natural and anthropogenic origin. A significant part of the suspended forms of HM entering the seas with river runoff is deposited on the river-sea geochemical barrier. The suspended matter remaining in the water column tends to be distributed at the water-bottom (sedimentation process) and water-air (particle flotation) interfaces. Marine aerosol formed from the surface microlayer of the sea inherits its chemical composition. It can include both sea salts and solid microparticles of various origin.

In most cases, the concentration of elements in sea waters is significantly lower than in river waters, where up to 75% of Fe, Mn, Ni, Cu are transported in conjunction with organic substances. Fe, Al, Mn, Ni, and Cu migrate in river waters in forms of colloids, simple and complex ions with a positive and negative charge. The river-sea barrier zone acts as a trap for the most substances (Gordeev 1983).

The main source of pollutants for the Sea of Azov is the runoff of large and small rivers: Don, Kuban, Mius, Eya, Beisug, Kagalnik, etc. (Mikhailenko et al. 2018). The catchment area of the Sea of Azov is about 570 000 km^2, with the Don and Kuban River catchments account for about 85% of the total (Matishov et al. 2002). The ratio of dissolved and suspended forms of HM at the marine edge of the Don River delta is different. For Fe, Mn, Cr, Pb, the predominance of suspended forms is noted. For Zn, Ni, Cu, and Cd, dissolved forms are prevailing. The flows of metals in the lower reaches of the Don River are largely determined by the influence of the city of Rostov-on-Don. Flows of dissolved and suspended forms of Ni, Cu, Zn, Cd and other HMs significantly increase downstream the city (Tkachenko et al. 2017). As shown below, this also affects the chemical composition of aerosols, especially in the mouth area of the Don River.

Spatial distribution of Heavy Metals and Al in marine aerosols

Iron. The Fe distribution is characterized by the presence of two regions with elevated elemental abundances (Fig. 2). Concentration increases from the open central part of the sea towards the northern and northeastern parts, especially sharply in the Taganrog Bay. On average, the concentration of Fe in the open sea is 200-1000 ng/m^3. The gulf zone is marked by Fe values of about 1800-2000 ng/m^3, which is an order of magnitude higher than the average concentration over the entire water area.

Perhaps this fact is explained by the presence of a dynamic geochemical barrier in the mixing zone of fresh and salt waters, the stirring up of bottom sediments and the lifting of suspended particles to the sea surface, where they can be captured in marine aerosol during wind-wave disruption and other processes of formation of marine aerosol. The zone of

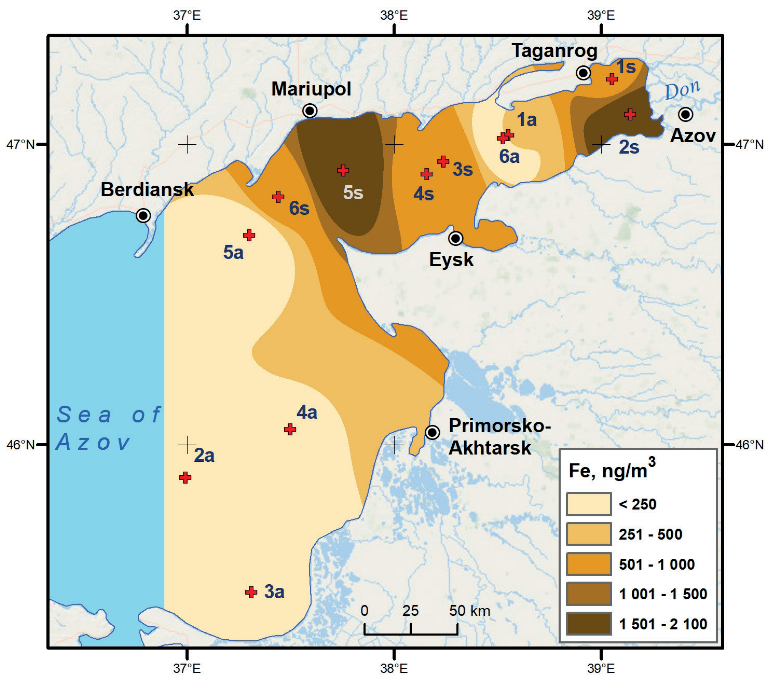


Fig. 2. Distribution of Fe in marine aerosols

the northern coast also stands out, in the area of Mariupol, Fe is about 1800 ng / m³, which can be associated with the technogenic influence of the city. Pollutants can also come from the precipitation, as well as from the coastal abrasion, which provides the terrigenous material (Mikhailenko 2018).

Aluminum. The average concentration of Al in marine aerosols in most of the water area is 20–50 ng / m³ (Fig. 3). It rises up to 90-100 ng / m³ at the northern coast in the region of Mariupol, and even higher up to 200 ng / m³ in the Don River mouth area and the Taganrog Bay. This is an order of magnitude higher than the average concentration over the sea. Probably, the maximum in the Taganrog Bay can be explained by the secondary mobilization of aluminosilicate particles and their lift to the surface due to the turbulent mixing and flotation processes. Aerosol particles are enriched with Al, since their main source is the sea surface. The maximum of Al in the area of Mariupol is less pronounced than for the most part of chemical elements, which indicates a low Al content in wastewaters.

Lead. The distribution of Pb in the marine aerosol is characterized by a maximum content at the mouth of the Don River, where it reaches 30 ng / m³. Moving towards the sea, the Pb content at first sharply (5 times), and then gradually decreases (Fig. 4). The high content of Pb in the aerosol of the Taganrog Bay may be due to the anthropogenic impact of Taganrog, which is an industrial city, port and center of ferrous metallurgy.

Cadmium. The average Cd value in marine aerosols over the Sea of Azov is about 0.4 – 0.6 ng / m³. Its content increases up to 2.2 ng / m³ in the mouth of the Don River and Taganrog Bay. The concentration of Cd reaches maximum values up to 2.8 ng / m³ in the north-western part of the sea nearby Ukrainian town of Mariupol (Fig. 5). Since the greatest values are comparable and located close to the industrial centers of the northern coast of the sea, it can be assumed that the supplier of Cd is mainly atmospheric emissions and wastewater from industrial cities.

Zinc. In the open part of the Sea of Azov, the Zn content in aerosol particles is from 10 to 60 ng / m³ (Fig. 6). For the Zn distribution, as for Cd, 2 peaks were noted – in the area of Mariupol (100-110 ng / m³) and in the mouth of the Don River. The second maximum is more contrasting, the values here are 200-280 ng / m³. It is caused by water pollution in the lower reaches of the river by wastewaters of Rostov-on-Don and Azov town, which is noted by many authors (Tkachenko et al. 2017; Bufetova et al. 2019; and others).

Copper. The average Cu content in the open part of the Sea of Azov is 2–8 ng / m³. The Cu concentration increases as it approaches the coast of Ukraine, reaching 12-16 ng / m³ in the area of Mariupol (Fig. 7). Taganrog Bay is a zone of

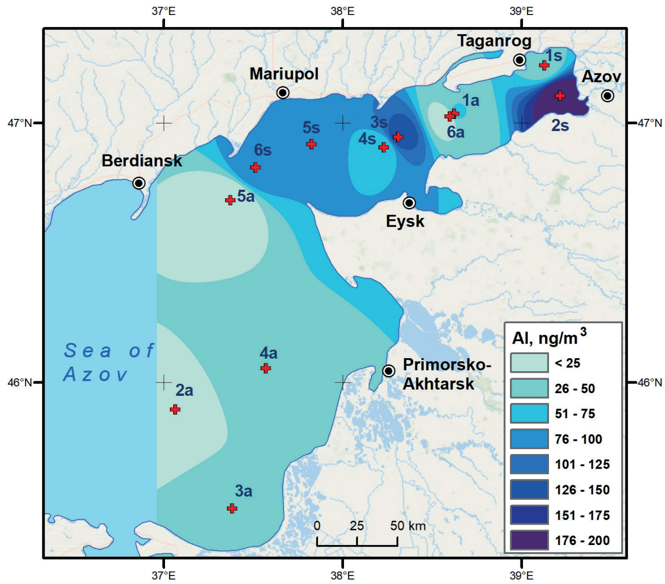


Fig. 3. Distribution of Al in marine aerosols

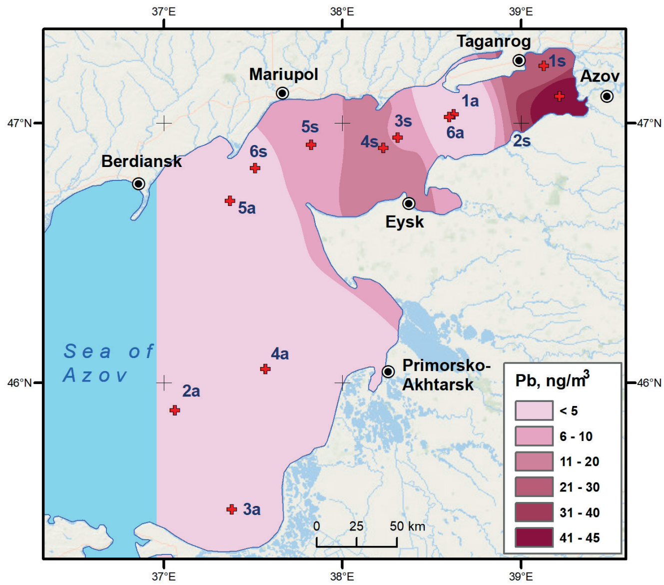


Fig. 4. Distribution of Pb in marine aerosols

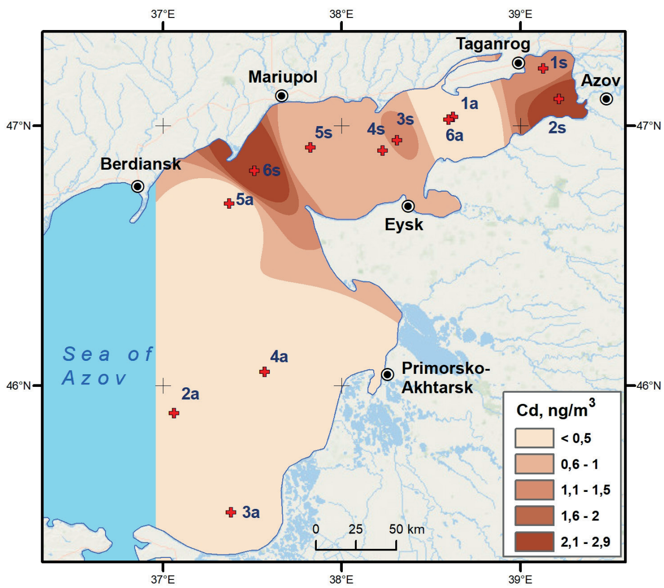


Fig. 5. Distribution of Cd in marine aerosols

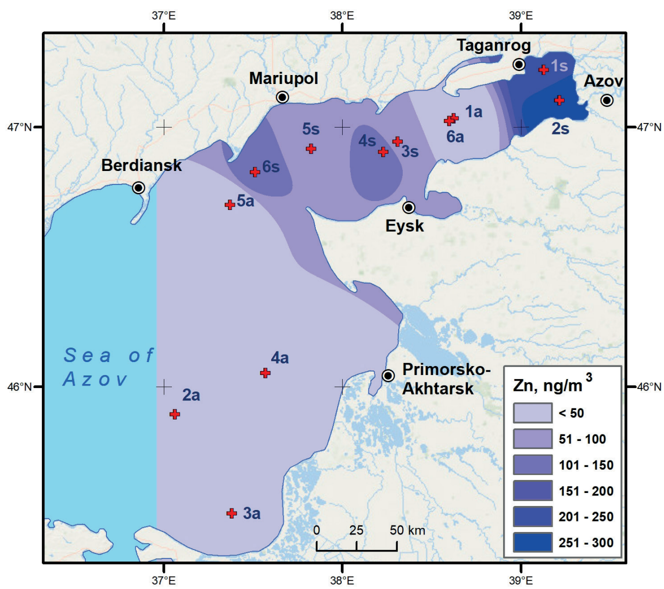


Fig. 6. Distribution of Zn in marine aerosols

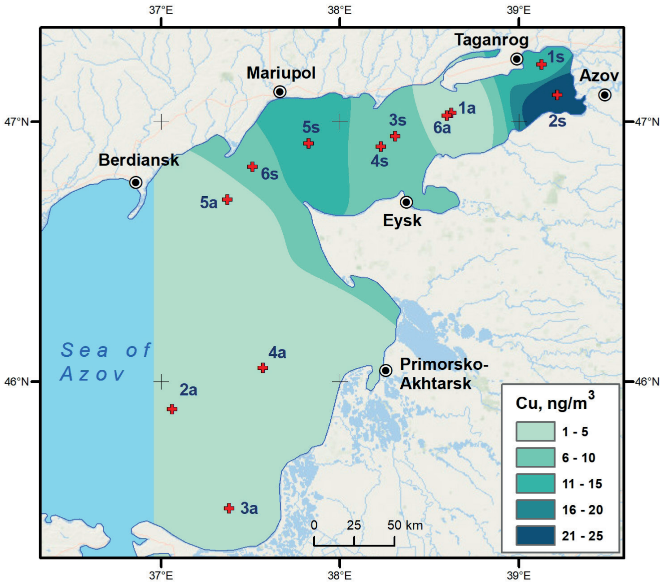


Fig. 7. Distribution of Cd in marine aerosols

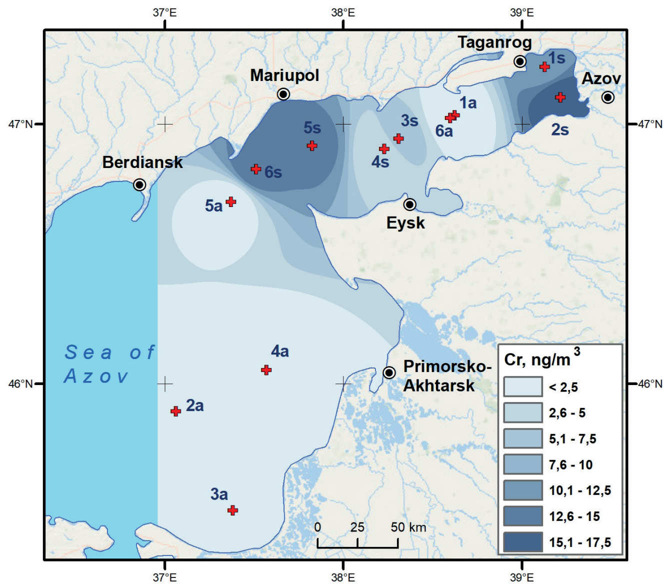


Fig. 8. Distribution of Cr in marine aerosols

high Cu content in marine aerosol. Here, the concentration increases to 20-23 ng / m^3, which can be attributed to the influence of the city of Taganrog, as well as the barrier zone of mixing of salt waters of the Sea of Azov and fresh water of the Don River, which acts as the main source of Cu for the adjusting water area (Bufetova et al. 2019).

Cromium. The Cr distribution in the marine aerosol (Fig. 8) is characterized by a slow increase from the southern part of the open sea, where it is 1-5 ng / m^3 to the northern (15 ng / m^3 in the region of Mariupol) and the northeastern part up to the entrance to the Taganrog Bay, where it reaches 18 ng / m^3. This is consistent with trends in concentrations of the most HMs. High contents are confined to industrial centers on the north and north-east coast of the Sea of Azov (Mariupol, Taganrog), as well as to the water area of the Taganrog Bay.

Thus, for most of the heavy metals, the concentrations increase as they move from the open part of the sea towards the northern coast and the mouth of the Don River, where under the influence of industrial centers and the geochemical river-sea barrier, local maximums of values are associated with increase in anthropogenic load. Maximum contents of For Fe, Cr, and Cd are confined to the mouth part of the Don River, and the city of Mariupol. For Pb, Zn and Cu, the main maximum content in the aerosol is observed in the eastern part of Taganrog Bay and the mouth of the Don River.

A large role in the pollution of the Sea of Azov belongs to the cities located on the coast and in the Don River Delta: Azov, Taganrog, Yeysk, Primorsko-Akhtarsk, Temryuk – due to the discharge of insufficiently purified "conditionally clean" water from enterprises directly into water bodies and streams (Khovansky et al. 1990). It is also worth noting the contribution of ports, shipping, landfills and dumping.

Taganrog is one of the leading industrial centers on the coast of the Sea of Azov, in which industrial enterprises are located: car assembly enterprises (TagAZ), steel manufacturing, pipe production (TagMet), energy and heating boilers, repair and reconstruction of ships, aircraft, etc. There are 51 organizations that have emissions of harmful substances into the atmosphere. The general indicators of pollutant emissions into the atmosphere of Taganrog range from 3.6% to 5.1% of the pollutant emissions of the entire region (Nechipurenko et al. 2019).

Mariupol is one of the most disadvantaged cities in Ukraine in terms of air pollution in Pb, Zn, Cu, Ni, Mn, Cr and Cd compounds. Two large metallurgical enterprises are located on the territory of the city, which are powerful sources of atmospheric emissions (Grishchenko et al. 2018; Voityuk et al. 2018). Industrial dust includes toxic oxides of Fe, Si, Al, Mg, Mn, Cr, etc. (Monin et al. 2012).

The behavior and distribution of trace elements is influenced by the mixing of fresh and salt waters of the rivers and the Sea of Azov, where the great majority (70–95%) of river suspended matter precipitates at the the river-sea geochemical barrier, capturing pollutants brought by the river, including heavy metals (Mikhailenko et al. 2018). Shallow-water Taganrog Bay also belongs to the such zones. Here HM content in the marine aerosol rises sharply due to the enrichment SSM with pollutants. Concentrations of HM in the water of the Sea of Azov in places exceed MPC, especially in the Taganrog Bay. The sum of multiples of the MPC TM for the Taganrog Bay in 2005 amounted to 4.7 MPC, while for the open sea it is less (3.7 MPC), which is consistent with the data obtained during the study of the composition of the sea aerosol of the Sea of Azov (Klenkin et al. 2009).

Based on the comparison with the data from other regions, the Sea of Azov can be attributed to the group of seas with the medium or high HM content in marine aerosol. The research of aerosol particles was carried out: 2002-2007 in Black Sea (Yablokov et al. 2002; Lapshin et al. 2003; Syroeshkin et al. 2004; Syroeshkin et al. 2014; Syroeshkin et al. 2006), 2005– 2008 in the Kara, Barents and White Seas (Syroeshkin et al. 2010), 2001– 2004 in the Baltic Sea (Syroeshkin et al. 2004), 2006 – 2008 in the Atlantic Ocean (Syroeshkin et al. 2012), 2007 – 2009 in the Arctic Ocean (Lapshin et al. 2010). The concentration levels of Ni and Cr are shown in Fig. 9 A. It is clearly seen that aerosols of the Sea of Azov contain these metals in low concentrations 10-100 ng / m^3 comparable with the Arctic seas. The concentration range of Cd and Mn (Fig. 9 B) is 2 orders of magnitude with the lowest levels in the Arctic Ocean, and the highest values in the Baltic, Azov and Mediterranean Seas. The content of Pb in the marine aerosols (Fig. 9 C) varies from 0.1 – 10 ng / m^3 in the group of subjects with a relatively low anthropogenic impact (White, Barents, and Kara seas, Arctic and Atlantic oceans), to 10-100 ng / m^3 in the group of southern seas (Black, Azov, Mediterranean, and Baltic seas), subjected to the strong anthropogenic pressure.

In general, the HM content in aerosol microparticles is determined by a number of factors. The first group includes the natural factors, primarily climate related and associated with the geographical position of seas, such as precipitation, evaporation, water salinity, etc. The second group is associated with the human economic activity. The Mediterranean

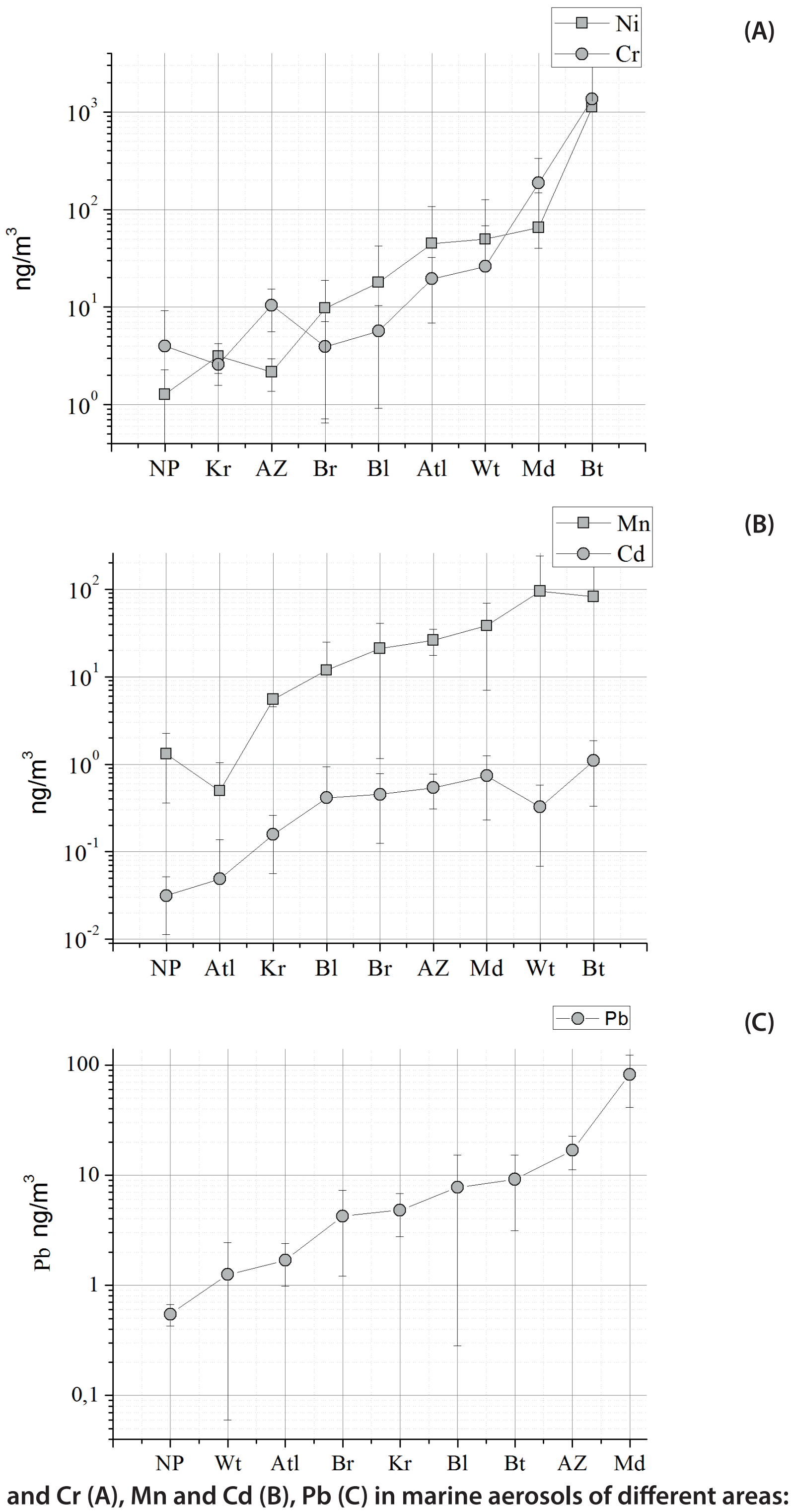


Fig. 9. Concentration of Ni and Cr (A), Mn and Cd (B), Pb (C) in marine aerosols of different areas: NP- North Pole; Wt- White Sea; Atl- South Atlantic; Br – the Barents Sea; Kr – Kara Sea; Bl – Black Sea; Bt – the Baltic Sea; Az.- Sea of Azov; Md – the Mediterranean Sea;

Sea, the Black Sea, and in particular the shallow enclosed Sea of Azov are subjected to the high anthropogenic pressure, which affect the content of various elements in the sea water, and therefore in the sea surface microlayers and marine aerosols. The natural distribution of elements plays a subordinate role here. It is worth considering that it is almost impossible to determine the specific pollution sources while studying aerosol particles in the vast water areas.

CONCLUSIONS

1. The spatial distribution of heavy metals in the marine aerosol of the Sea of Azov is determined by the influence of the river-sea geochemical barrier zone in the Taganrog Bay and the anthropogenic impact of the coastal industrial cities. HM concentrations increase from the open part of the sea towards the northern coast and the mouth of the Don River.
2. The maximum content of HM in marine aerosol was observed in the mouth area of the Don River in the eastern part of the Taganrog Bay, which can be associated with the anthropogenic impact of the cities of Rostov-on-Don, Azov and Taganrog. High contents of Fe, Cr, and Cd were also found in the western part of the Taganrog Bay, due to the technogenic impact of the city of Mariupol.
3. The Sea of Azov can be attributed to the group of seas with the moderate or high HM content in marine aerosol in comparison with the data from other regions.

REFERENCES

Akinori I., Myriokefalitakis S., Kanakidou M., Mahowald N. M., et al. (2019). Pyrogenic iron: The missing link to high iron solubility in aerosols. Science Advances, 5(5), 7671, DOI: 10.1126/sciadv.aau7671 .

Aryasree S, Nair PR, Girach IA, Jacob S. (2015). Winter time chemical characteristics of aerosols over the Bay of Bengal: continental influence. Environmental Science and Pollution Research, 22(19), 14901-14918, DOI: 10.1007/s11356-015-4700-7.

Brooks S.D., Thornton D.C.. (2018). Marine Aerosols and Clouds. Annual Review of Marine Science, 3(10), 289-313. DOI: 10.1146/annurev-marine-121916-063148.

Buseck P. R., Posfai M. (1999) Airborne minerals and related aerosol particles: Effects on climate and the environment. Proceedings of the National Academy of Sciences, 96(7), 3372–3379. DOI:10.1073/pnas.96.7.3372.

Chandrakar K.K., Cantrell W., Chang K., Ciochetto D., et al. (2016). Aerosol indirect effect from turbulence-induced broadening of cloud-droplet size distributions. Proceedings of the National Academy of Sciences USA, 113(50), 14243-14248. DOI: 10.1175/JAS-D-18-0006.1.

Coquery M., Villeneuve J. P. (2001). Final report on the split sampling exercises and quality assurance activities. EU Project № ENV RM S9602, Amsterdam, ICWC, 51.

Csavina J., Field J., Taylor M.P., Gao S., et al. (2012). A Review on the Importance of Metals and Metalloids in Atmospheric Dust and Aerosol from Mining Operations. Science of the Total Environment, 433, 58-73. DOI: 10.1016/j.scitotenv.2012.06.013.

Fleming L.E., Bean J.A., Kirkpatrick B., Cheng Y.S., et al. (2009). Exposure and effect assessment of aerosolized red tide toxins (brevetoxins) and asthma. Environmental Health Perspectives, 117(7), 1095-100. DOI: 10.1289/ehp.0900673.

Furness R.W. (2017). Heavy Metals in the Marine Environment. CRC press.

Garrett W.D. (1965) Collection of slick-forming materials from the sea surface. // Limnol. Oceanog, 10, 602-605. .

Georgoulias A.K., Alexandri G., Kourtidis K.A., Lelieveld J., et al. (2016). Spatiotemporal variability and contribution of different aerosol types to the Aerosol Optical Depth over the Eastern Mediterranean. Atmospheric Chemistry And Physics, 16(21), 13853-13884. DOI: 10.5194/acp-16-13853-2016.

Golubeva N.I., Burtseva L.V., Gromov S.A. (2011). Heavy Metals in Atmospheric Air in the Kara Sea Water Area in September–October. Oceanology, 58, 870-878. DOI: doi.org/10.1134/S000143701806005X.

Goncharuk V.V., Lapshin V.B., Chichaeva M.A., Samsoni-Todorov A.O., et al. (2012). Heavy metals, aluminum and arsenic in aerosols of the oceans. Chemistry and Water Technology, 34(1), 1-15. DOI: 10.3103/S1063455X12010018.

Gordeev V.V. (1983). River runoff into the ocean and features of its geochemistry. Moscow: Nauka.

Grebennikova T.V., Syroeshkin A.V., Chichaeva M.A., Esper S.A., Lvov D.K. (2017). Natural foci of influenza A in the western Arctic. Voprosy Virusologii, 62(1), 11-17. DOI: 10.18821/0507-4088-2017-62-1-11-17.

Grishchenko S.V., Grishchenko I.S., Kostenko V.S., Basenko I.N., et al. (2018). Hygienic assessment of atmospheric air pollution of populated areas of Donbass by heavy metals. Bulletin of Hygiene and Epidemiology, 22 (1), 11-15.

Izhar S., Goel A., Chakraborty A., Gupta T. Annual trends in occurrence of submicron particles in ambient air and health risk posed by particle bound metals. (2016) Chemosphere, 146, 582-590. DOI: 10.1016/j.chemosphere.2015.12.039.

Jordi A., Basterretxea G., Tovar-Sánchez A., Alastuey A., Querol X. (2012) Copper aerosols inhibit phytoplankton growth in the Mediterranean Sea. Proceedings of the National Academy of Sciences USA, 109(52), 21246-21249. DOI: 10.1073/pnas.1207567110.

Kirkpatrick B., Fleming L.E., Bean J.A., Nierenberg K., et al. (2011). Aerosolized Red Tide Toxins (Brevetoxins) and Asthma: Continued health effects after 1 hour beach exposure. Harmful Algae, 10(2), 138-143. DOI:10.1016/j.hal.2010.08.005.

Kolesnikov M.V., Matveeva I.S., Lapshin V.B., Pletenev S.S., et al. (2005). Heavy metals in marine aerosols of the Russian part of the Black sea. Oceanology, 45(1), 102-111.

Lang-Yona N., Lehahn Y., Herut B., Burshtein N., Rudich Y.. (2014) Marine aerosol as a possible source for endotoxins in coastal areas. Science of the Total Environment, 499, 311-8.DOI: 10.1016/j.scitotenv.2014.08.054.

Li J., Han Z., Yao X., Xie Z., Tan S. (2019). The distributions and direct radiative effects of marine aerosols over East Asia in springtime. Sci Total Environ. 651(2), 1913-1925. DOI: 10.1016/j.scitotenv.2018.09.368.

Li S., Du L., Tsona N.T., Wang W. (2018). The interaction of trace heavy metal with lipid monolayer in the sea surface microlayer. Chemosphere. 196, 323-330. DOI: 10.1016/j.chemosphere.2017.12.157.

Liu Y., Li S., Sun C., Qi M., et al. (2018) Pollution Level and Health Risk Assessment of PM2.5-Bound Metals in Baoding City Before and After the Heating Period. International Journal of Environmental Research and Public Health, 15(10), 2286. DOI: 10.3390/ijerph15102286.

Mahowald N.M., Hamilton D.S., Mackey K.R., Moore J.K., et al. (2018). Aerosol trace metal leaching and impacts on marine microorganisms. Nature Communications, 9(1), 2614. DOI: 10.1038/s41467-018-04970-7.

Marín-Beltrán I., Logue J.B., Andersson A.F., Peters F.. (2019). Atmospheric Deposition Impact on Bacterial Community Composition in the NW Mediterranean. Frontiers in Microbiology, 10(858), 1-14. DOI: 10.3389/fmicb.2019.00858.

O'Dowd C.D., Jimenez J.L., Bahreini R., Flagan R.C., et al. (2002). Marine aerosol formation from biogenic iodine emissions. Nature, 417 (6889), 597-598. DOI: 10.1038/nature00775.

O'Dowd C.D., de Leeuw G. (2007). Marine aerosol production: a review of the current knowledge. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 365(1856), 1753-1774. DOI: 10.1098/rsta.2007.2043.

Patocka J., Nepovimova E., Wu Q., Kuca K. (2018). Palytoxin congeners. Archives of Toxicology, 92(1), 143-156. DOI: 10.1007/s00204-017-2105-8.

Paytan A., Mackey K.R.M., Chen Y., Lima I.D., et al.(2009). Toxicity of atmospheric aerosols on marine phytoplankton. Proceedings of the National Academy of Sciences USA, 106(12), 4601-4605. DOI: 10.1073/pnas.0811486106.

Qureshi A, MacLeod M, Hungerbühler K. (2009). Modeling aerosol suspension from soils and oceans as sources of micropollutants to air. Chemosphere, 77(4), 495-500. DOI: 10.1016/j.chemosphere.2009.07.051.

Rädlein N., Heumann K.G. (2006) Trace Analysis of Heavy Metals in Aerosols Over the Atlantic Ocean from Antarctica to Europe. International Journal of Environmental Analytical Chemistry, 48, 127-150.

Rastelli E., Corinaldesi C., Dell'Anno A., Lo Martire M., et al. (2017). Transfer of labile organic matter and microbes from the ocean surface to the marine aerosol: an experimental approach. Scientific Reports, 7, 11475. DOI: 10.1038/s41598-017-10563-z.

Sajeev P., Randall V.M., Van Donkelaar A., et al. (2014). Global Chemical Composition of Ambient Fine Particulate Matter for Exposure Assessment Environmental Science & Technology, 48(22), 13060-13068. DOI: 10.1021/es502965b.

Salomons W., Bayne B. L., Duursma E.K., Förstner U. (1988). Pollution of the North Sea. Springer, Berlin, Heidelberg, 300-347.

Sánchez-Rodas D., Alsioufi, L., Sánchez de la Campa A. M., González-Castanedo Y. (2017). Antimony speciation as geochemical tracer for anthropogenic emissions of atmospheric particulate matter, Journal of Hazardous Materials, 324(B), 213-220. DOI: 10.1016/j.jhazmat.2016.10.051.

Song S.K., Shon Z.H., Choi Y.N., Son Y.B., et al. (2019) Global trend analysis in primary and secondary production of marine aerosol and aerosol optical depth during 2000-2015. Chemosphere, 224, 417-427. DOI: 10.1016/j.chemosphere.2019.02.152.

Sullivan R.C., Levy R.C., da Silva A.M., Pryor S.C.. (2017) Developing and diagnosing climate change indicators of regional aerosol optical properties. Scientific Reports, 7(1), 1-13. DOI: 10.1038/s41598-017-18402-x. .

Syroeshkin A.V., Chichaeva M.A. (2010). Concentaration level of heavy metals within marine aerosols of Western Arctic seas, Southern Atlantic and Arctic ocean. Trace elements in medicine, 11(2), 15.

Takata K., Saito K., Yasunari T. (2009). Changes in the Asian monsoon climate during 1700-1850 induced by preindustrial cultivation. Proceedings of the National Academy of Sciences USA, 106(24), 9586-9589. DOI: 10.1073/pnas.0807346106.

Tuohy A., Bertler N., Neff P., Edwards R., Emanuelsson D., et al. (2015). Transport and deposition of heavy metals in the Ross Sea Region, Antarctica. Journal of Geophysical Research: Atmospheres, 120, 10, 996-11,011. DOI: 10.1002/2015JD023293.

Van Dolah F.M.. (2000). Marine algal toxins: origins, health effects, and their increased occurrence. Environmental Health Perspectives, 108(1), 133-41. DOI: 10.1289/ehp.00108s1133.

Von Glasow R., Jickells T.D., Baklanov A., Carmichael G.R., et al. (2013). Megacities and Large Urban Agglomerations in the Coastal Zone: Interactions Between Atmosphere, Land, and Marine Ecosystems. AMBIO A Journal of the Human Environment, 42(1), 13-28. DOI: 10.1007/s13280-012-0343-9.

Vinogradova A.A., Ivanova Yu.A. (2017). The transfer of air masses and pollution to the Arctic islands of Russia (1986-2016): long-term, interannual and seasonal variations. Geophysical processes and biosphere. 16(4), 5-20. DOI: 10.21455/GPB2017.4-1.

Vinogradova A.A., Kotova E.I. (2019). Heavy metal pollution of the northern seas of Russia: flow from the atmosphere and river runoff. Geophysical processes and biosphere. 18 (1), 22-32 DOI: 10.21455/GPB2019.1-3.

Voityuk Yu. Yu., Kuraeva I.V., Loktionova E.P. (2018). Sources of heavy metal pollution in the territories of industrial agglomerations of ferrous metallurgy. Materials of the VI International Scientific Conference. Modern problems of landscape science and geoecology (on the 100th anniversary of the birth of Professor V.A. Dementiev). Ed. A.N. Witchenko.

Walsh J.J., Lenes J.M., Weisberg R.H., Zheng L., et al. (2017) More surprises in the global greenhouse: Human health impacts from recent toxicmarine aerosol formations, due to centennial alterations of world-wide coastal food webs. Marine Pollution Bulletin, 116(1-2), 9-40. DOI: 10.1016/j.marpolbul.2016.12.053.

Wang X., He S., Chen, Zhang Y, et al. (2018). Spatiotemporal Characteristics and Health Risk Assessment of Heavy Metals in PM2.5 in Zhejiang Province. International Journal of environmental research and public health, 15(4), 583. DOI: 10.3390/ijerph15040583.

Witt M., Baker A. R., Jickells T. D. Atmospheric trace metals over the Atlantic and South Indian Oceans: Investigation of metal concentrations and lead isotope ratios in coastal and remote marine aerosols (2006). Atmospheric Environment, 40, 5435-5451. DOI: 10.1016/j.atmosenv.2006.04.041.

Zatsepa S.N., Lapshin V.B., Oradovsky S.G., Simonov A.I. A device for sampling water from a surface microlayer. Copyright certificate No. 1375974 of December 3, 1985.

Klyonkin A.A., Korablina I.V., Korpakova I.G. (2007). Description of the current level of pollution of water and bottom sediments of the Sea of Azov by heavy metals. Ecology and industry of Russia, 5, 30-33.

Kondratiev K. Ya., Moskalenko N.I., Pozdnyakov D.V. (1983). Atmospheric aerosol. Leningrad: Hydrometeoizdat.

Lapshin V.B., Chichaeva M.A., Matveeva I.S., Chichaev A.N., et al. (2010). Heavy metals, aluminum and arsenic in aerosols of the Atlantic, Arctic Oceans and European seas of Russia. Investigated in Russia, 34, 393-403.

Lapshin V.B., Yablokov M.Yu., Matveeva I.S. et al. (2002). Are marine aerosols toxic? Investigated in Russia, 118, 1302-1316.

Lukashin V.N., Klyuvitkin A.A., Bobrov V.A., Dara O.M., Shevchenko V.P. (2018). The chemical composition of the aerosols of the North Atlantic. Oceanology 58(5), 781-791. DOI: 10.1134/S0030157418050052.

Maslennikova A.V., Shevchenko V.P., Belogub E.V., Maslov A.V., Blinov I.A. (2018). Sedimentary material from the drifting ices of the Ermak Plateau and the Fram Strait: new data on spore-pollen spectra, mineralogy, and geochemistry. Mineralogy. 4(4), 102-118.

Matishov G.G. (2002). Ecosystem studies of the Sea of Azov and the coast. Volume 4, Apatity: Publ. KSC RAS.

Mikhailenko A.V., Fedorov Yu.A., Dotsenko I.V. (2018). Heavy metals in the components of the landscape of the Sea of Azov. Rastov-on-Don: Southern Federal University.

Monin V.L. (2012). Monitoring of technogenic air pollution in the city of Mariupol. Bulletin of the Azov State Technical University, 24, 327-334.

Nechipurenko V.V., Merinova Yu.Yu. Industrial complex as a factor in the pollution of urban systems of the Rostov region. (2019) Collection of scientific papers based on the materials of the 9th International Scientific and Practical Conference. Edited by E.I. Tikhomirova. 81-85.

Savenko V.S. The geochemistry of ocean aerosol. (1998). Bulletin of Moscow University. ser. 5. Geography, 1, 28-32.

Smirnov A.N., Lapshin V.B., Balyshev A.V., Lebedev I.M., Syroeshkin A.V. (2004). Supranadmolecular complexes of water. Investigated in Russia, 38, 413-421.

Smirnov A.N., Lapshin V.B., Balyshev A.V., Popov P.I., Lebedev I.M., Syroeshkin A.V. (2003). Cooperative anisotropic motion of the dispersed phase in aqueous solutions. Investigated in Russia, 39, 422-425.

Syroeshkin A.V., Popov P.I. (2005) «Marine Aerosols. Toxicity, research methods», Moscow: RUDN.

Syroeshkin A.V., Smirnov A.N., Goncharuk V.V., Uspenskaya E.V., et al. (2006). Water as a heterogeneous structure. Investigated in Russia, 88, 843-854.

Syroeshkin A.V., Chichaeva M.A., Matveeva I.S. (2014). Repeatability of one-to-one relationships between concentrations of heavy metals and dispersion of marine aerosol (as exemplified by two expeditionary studies on the Black Sea. Heliogeophysical studies, 10(10), 113-127.

Tkachenko A.N., Tkachenko O.V., Lychagin M.Yu., Kasimov N.S. (2017). Heavy metal flows in aquatic systems of the Don and Kuban deltas. Doklady Academy of Sciences, 474(2), 234-237.

Khovansky A.D. (1990). Geochemical assessment of the state of the river system of the Lower Don, Rostov-on-Don: Publishing House of the University of Rostov.

Received on January 1th, 2020 Accepted on January 21th, 2020

Olga A. Samonova[1], Elena N. Aseyeva[1*]
[1] Faculty of Geography, Lomonosov Moscow State University, Leninskiye Gori, 1, Moscow, 119991, Russia
***Corresponding author:** aseyeva@mail.ru

PARTICLE SIZE PARTITIONING OF METALS IN HUMUS HORIZONS OF TWO SMALL EROSIONAL LANDFORMS IN THE MIDDLE PROTVA BASIN – A COMPARATIVE STUDY

ABSTRACT. Partitioning of metals in soil particles of various size classes has been receiving greater significance due to the necessity to predict the behaviour and pathways of the potentially toxic elements in the environment. In this study the analysis of metals' distribution in various particle size fractions was performed to characterize and compare geochemical features of the topsoil horizons of two small erosional landforms located in uncontaminated area of the central part of European Russia (the Middle Protva basin, mixed forest zone). The landforms represent two typical lithological types of gullies in the study area. Soil samples were fractionated and the concentrations of Fe, Mn, Ti, Zr, Ni, Co, Cr, Zn, Cu, Pb were determined in five particle size fractions: 1–0.25, 0.25-0.05, 0.05–0.01, 0.01–0.001 and <0.001 mm. The metals' concentrations and their distribution in various particle sizes were found to be related to gully litho-type. The contribution of the clay to the total amount of metals was the greatest for Mn, Zn, Ni and Co in both systems. The highest mass loading for Ti, Zr and Cr came from the coarse silt, while for Cu and Pb it was made by different particle size fractions: the medium and fine silt or the coarse silt. The highest contribution of Fe also came from different fractions, either from the coarse sand or the clay depending on the system.

KEY WORDS: soil, humus horizons, distribution, heavy metals, particle size fractions, gully

CITATION: Olga A. Samonova, Elena N. Aseyeva (2020) Particle Size Partitioning Of Metals In Humus Horizons Of Two Small Erosional Landforms In The Middle Protva Basin – A Comparative Study.
Geography, Environment, Sustainability, Vol.13, No 1, p. 260-271
DOI-10.24057/2071-9388-2019-116

INTRODUCTION

Geochemical analysis of small erosional landforms has been in the focus of the recent studies because gullies play a vital role in the sediment transfer within river basins and sediment-associated redistribution of potentially hazardous elements such as heavy metals (Golosov 2006). During the transport of solid material, various physical, physico-chemical, chemical and biochemical processes occur and cause changes in distribution of elements within the mineral assemblages and among particle size fractions. The topsoil represents the most active layer, where such changes are more evident in comparison with deeper soil horizons (Samonova and Aseeva 2006). Most studies dealing with geochemical analysis of soil-geochemical systems have concentrated on total levels or mobile phases of trace elements (Semenkov and Koroleva 2019) with very little emphasis on their partitioning among soil particle size fractions. However, this type of analysis contributes to better understanding of size distribution functions of potentially hazardous substances and serves as the basis for accurate risk assessment, soil remediation and estimation of metal fluxes in basin-type landscape-geochemical systems.

The studies that have been conducted to determine heavy metals in different particle size fractions mainly focus on urban or roadside soil, road/urban dust and sediment (Sutherland 2003; Ljung et al. 2006; Ajmone-Marsan et al. 2008; Acosta et al. 2009; Vlasov et al. 2015; Khademi et al. 2019; Zhang et al. 2019), while very little data are available on the distribution of heavy metals in various particle size fractions from agricultural (Qian et al. 1996; Gong et al. 2014) and natural (Acosta et al. 2011; Hardy and Cornu 2006; Samonova and Aseeva 2008; Samonova and Aseyeva 2013) soil.

The majority of the studies have shown that fine soil particles tend to concentrate and retain higher amounts of metals and have a stronger ability to carry potentially hazardous substances due to their high specific area, more organic matter and Fe/Mn/Al oxides content (Förstner 1982; Hardy and Cornu 2006; Gong et al. 2014; Liu et al. 2018). In coarser particle size fractions, with a lower sorption capacity, the concentrations of metals fully depend on the mineralogical composition, in particular the proportion of heavy minerals and neoformations (Protasova 2003). Consequently, the association of metals with various particle sizes is determined by parent material, weathering and soil formation pathway (Pobedintseva 1975; Dobrovolskii 1983; Protasova 2003; Hardy and Cornu 2006; Acosta et al. 2011; Kabata-Pendias 2011). The effect of different parent materials and their mineralogical assemblages was studied in the detail by Acosta et al. (2011) for the soils located in the Betic Cordilleras (SE Spain). The results showed that the distribution of metals in particle size fractions separated from humus horizons and derived from igneous, metamorphic and sedimentary rocks significantly depended on parent rock. No distinct trends of accumulation in specific fractions were observed for Fe and Ti, while for trace elements two trends were revealed. Some elements, such as Pb, Zn and Cu, accumulated preferentially in the clay fraction regardless of parent rock, however some metals, such as Ni, Co, and Cr, were enriched either in the clay or sand fractions depending on rock types. The study performed in France by Hardy and Cornu (2006) for natural and agricultural soils on silty parent material (classified as Podzoluvisols) revealed the preferential association of trace metals with the clay fraction. Qian et al. (1996) found that in moderately and slightly polluted agricultural tropical soils in China the majority of the metals accumulated in the clay fraction and also in the coarse and medium sand. For the urban soil the major carriers of heavy metals are fine fractions, including the silt and clay, with particle sizes < 0.05mm (Ljung et al. 2006) and < 0.01mm (Ajmone-Marsan et al. 2008). These fractions, having high metal levels, might affect human health and environmental quality since they are easily transported by suspension, can remain in air for a long period and enter organisms, including humans, by ingestion and inhalation (Ljung et al. 2006; De Miguel et al. 1997). The coarser fractions in urban soils are relatively depleted in metals, however in some European cities they reveal high contents of certain metals, in particular geogenic Cr and Ni, due to their additions from specific parent lithologies (Ajmone-Marsan et al. 2008). Urban dust and aerosol are characterized by higher total concentrations of trace metals (Pb, Zn, Cu, Cd, and Cr) than urban soil (Banerjee 2003; Khademi et al. 2019) because they mostly consist of fine particles. The study of roadside dust in Moscow (Vlasov et al. 2015) showed that heavy metals accumulate in the medium and fine silt and also in the clay fraction. The study of metals' partitioning in aerosol fractions performed in an industrial city in southern Poland (Zwozdziak et al. 2017) also revealed that concentrations of metals vary with aerosol particle size: the maximum supply of Fe was related to coarse aerosol fractions (3μm-5μm), while the addition of Zn and Pb was associated with finer material (0.30μm-0.67μm). The supply of Mn, Cr, and Cu originating from both natural and man-made sources did not show any particular association with particle sizes.

Taking into account the importance of background data on particle size effect on metals' concentrations, in this paper we present the results of a soil-geochemical study of two small erosional landforms located 100 km to southwest from Moscow. The objectives of the study were (i) to investigate and compare the concentrations and distribution of Fe, Mn, Ti, Zr, Ni, Co, Cr, Zn, Pb in various particle size fractions (1–0.25mm, 0.25–0.05mm, 0.05–0.01mm, 0.01-0.001mm and <0.001mm) separated from the landforms' humus horizons, and (ii) to assess the contribution of particle size classes to total metal concentrations, considering the relative abundances of individual particle size fractions. It was hypothesized that comparing the soils of the two erosional landforms will reveal the effect of lithologies on the studied parameters because the landforms belong to different lithological types. In addition to this, the study will provide background geochemical data on natural sources of dust and aerosol to make a sound basis for pollution assessment of these components in the urban landscapes of Moscow and New Moscow area.

MATERIALS AND METHODS

Study area and study objects

The study area (55°12-13'N, 36°21-22'E) is located in the south-eastern part of the Smolensk-Moscow Upland (314 m a.s.l.), in the Middle Protva basin, 100 km to the southwest from Moscow (Fig. 1). Geomorphologically, it belongs to a marginal area of the Middle Pleistocene (MIS 6) glaciation with moraine topography modified by post-glacial erosional and fluvial processes. Climate is humid temperate continental characterized by moderately moist and warm summers (T_{July}=17.5°C), cold winters (mean $T_{January}$=–9.9°C) and mean annual precipitation about 600 mm. Natural vegetation is variable because the territory is located in the transition zone from mixed to deciduous forests. The forest communities cover about 60% of the studied area, while cultivated lands and meadows occupy about 40% of its territory. The major soil type is Retisol (World reference base 2015). In the Russian classification system, it corresponds to sod-podzolic soils. In the study area such soils occupy sub-horizontal surfaces and gentle slopes of the interfluve areas and develop under mixed forest (oak-spruce, lime-spruce, birch-spruce) vegetation mostly on mantle (loessial) loam, whose origin has still not been finally settled. Composed predominantly of silt-size particles, the mantle loam could have different origin (Kolevatykh 2011): glaciofluvial, eluvial, subaerial, polygenetic. It is assumed that it contains an aeolian component (Buggle et al. 2008; Muratov 1953) modified to some degree by local reworking and biological processes. About 50% of the soils in the interfluve area developed on the mantle loam in the nearest past were subjected to arable farming. The steepest relief elements in the study area such as sides of the river valleys or gullies represent erosional sites and are occupied by soddy soils (Regosols, according to World reference base 2015).

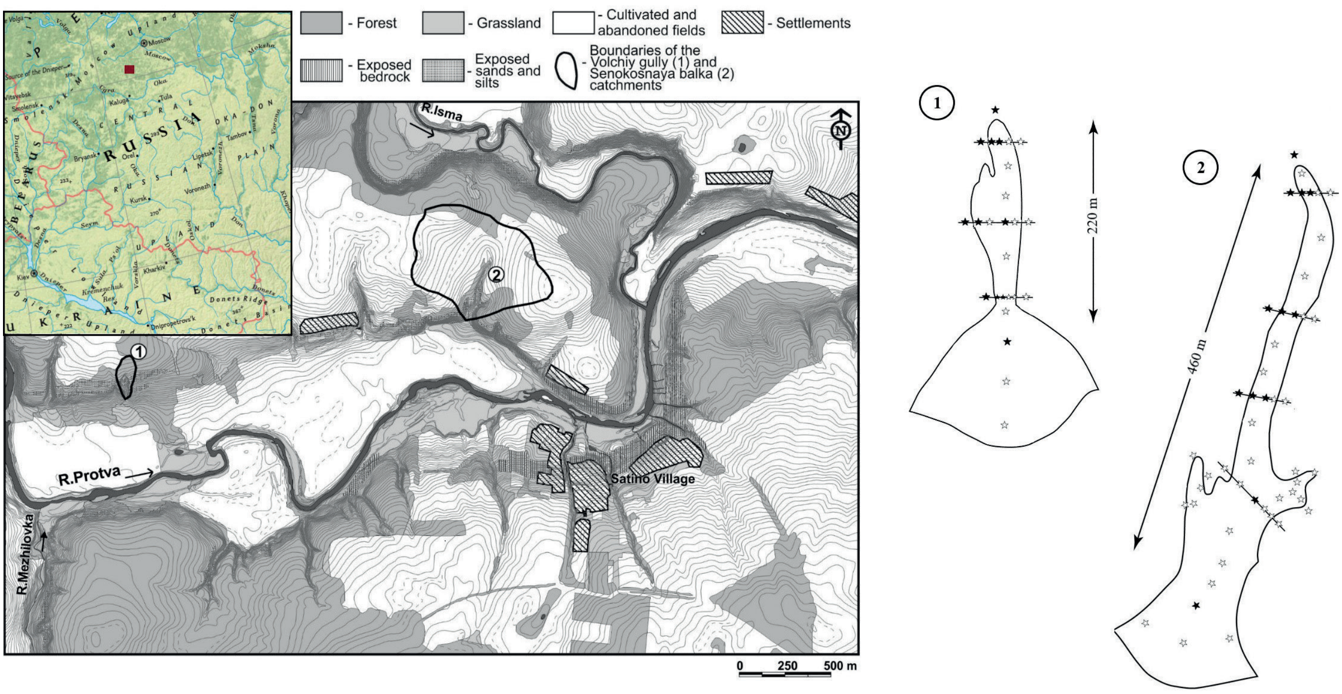


Fig. 1. The map of the study area (modified after Panin et al. 2009) with locations and schemes of the study objects – the gully (1) and the balka (2). Sampling sites are indicated with five-pointed stars

Black stars are the positions where the concentrations of metals were measured not only in a bulk soil sample but also in the separated particle size fractions.

The present study deals with the topsoil horizons of a gully and a small "U"-shaped dry valley (called balka in Russian) incised into the left side of the Protva river valley (Fig. 1). The gully and the balka represent two common types of small erosional systems typical of the study area. The gully is a smaller and younger Holocene landform whose formation was initiated in between 4.3- 4.6 ka BP (cal.) probably due to a local forest fire, coincided with sharp changes of hydroclimatic conditions (Panin et al. 2011). The gully has well developed fan, a concave form of longitudinal profile and mostly "V"-shaped cross-sections characterized by sharp edges and straight sides, 20-50° steep. Similar to other Holocene gullies of the study area, very little sediment storages are found in its bottom (Fig. 2).

The soil cover of the gully's catchment is dominated by sod-podzolic soils under forest communities while in the bottom and on the sides of the landform soddy soils with weakly differentiated A-AC-C profile are developed. The gully's detrital fan is covered by soddy soils with a relatively thick humus horizon developing under herbaceous meadow communities. The soils are formed on different parent materials since the system cuts through various lithologies including glaciofluvial sands, highly erodible glacio-lacustrine silts and fine-grained sands with a limited involvement of loessial (mantle) loams and boulder clays (Fig. 2).

The balka is morphologically different and an older system whose formation was initiated in the Pleistocene and continued in the Holocene during phases and episodes of high erosion (Panin et al. 2009). The balka is incised entirely into loamy deposits, such as Late Pleistocene loessial loams, which cover significant share of interfluves area in the region and also Middle Pleistocene glacial sediments.

It has a smoothed longitudinal profile (Fig. 2), well developed fan and "U" -shaped cross-sections. The bottom of the balka is covered by loamy sediments (Fig. 2) accumulated during the periods of low erosion therefore the parent material for soil formation is more homogeneous than in the gully. The soddy soils of the balka's sides and bottom are developing on loamy deposits under grass and large shrub vegetation. The large proportion of the balka's catchment area with sod-podzolic soils was used as tillage and recently is occupied by grass vegetation. Thus, the gully and the balka, having different morphology, belong to different lithological types, which allows us to evaluate the effect of parent lithology on particle size partitioning of metals in their humus horizons: the balka can be viewed as more homogeneous, monolithic, system with loamy deposits serving as parent material and gully belongs to a sandier and less homogeneous, heterolithic type.

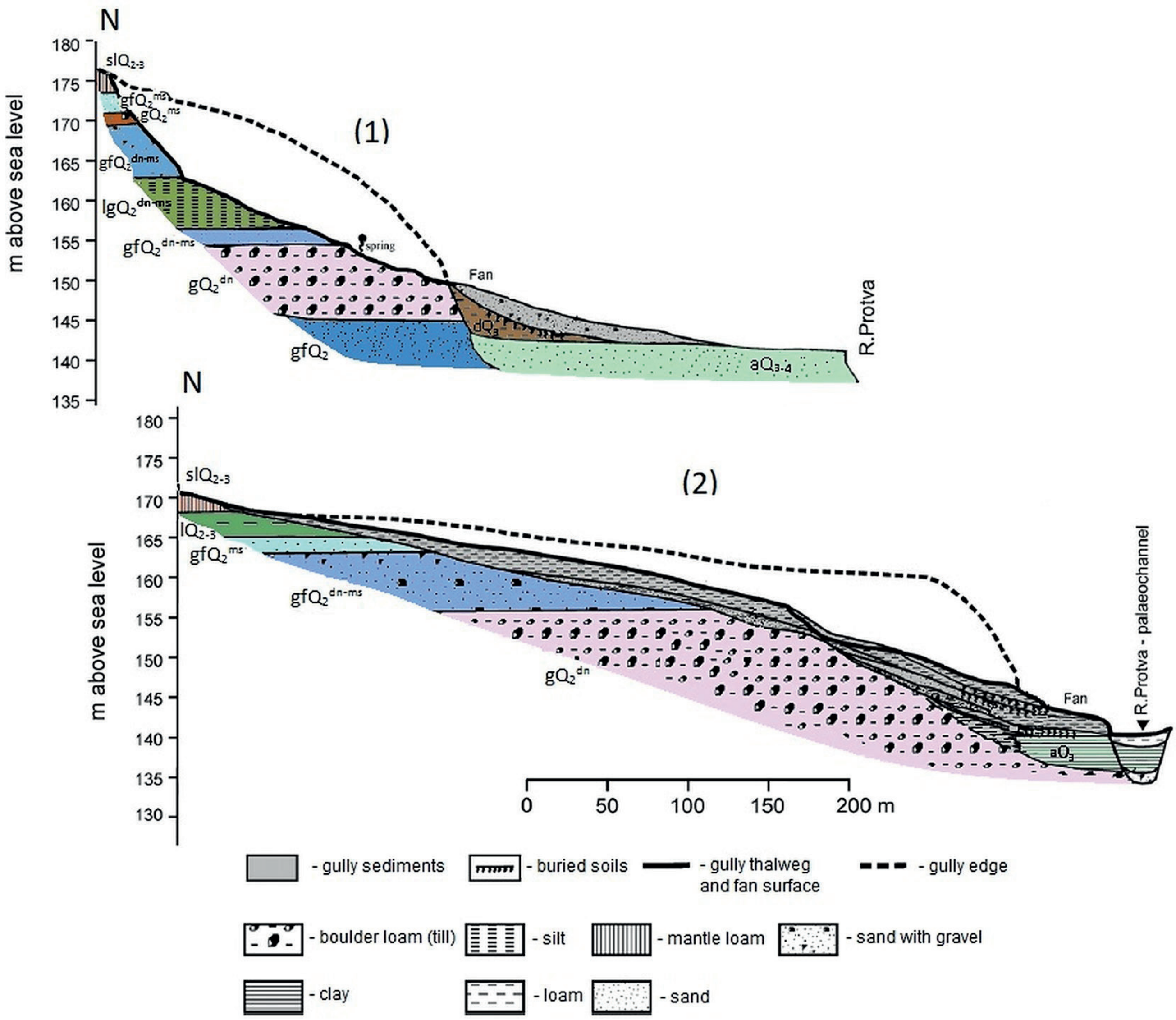


Fig. 2. Geological sections along the gully (1) and balka (2) incised in different types of Quaternary deposits and showing different sediment storage in their bottoms (from Panin et al. 2009)

Geological indexes: Origin: a – alluvium, sl – mantle loam, g – glacial till, gf– glacio-fluvial, lg – limno –glacial; Age: Q_2 – Middle Pleistocene, dn – Dnieper stage (MIS-8), ms – Moscow stage (MIS-6), Q_3– Late Pleistocene, Q_4 – Holocene.

Sampling and analytical procedures

The samples were taken from the surface A horizon (0-10 cm) according to the scheme displayed in Fig. 1. The sampling locations in each landform represent three major geomorphic units, such as sides, bottom and detrital fan, connected by sediment fluxes as well as adjacent catchment area considered as a source of solid matter (Samonova et al. 2014).

The collected bulk samples were analyzed for exchangeable soil acidity in 1M KCl solution. The concentrations of total organic carbon (TOC) were determined using $K_2Cr_2O_7$ wet-combustion method and then recalculated to humus content (Van Reeuwijk 1992). The total levels of metals and their lateral distribution in the humus horizons of the gully and balka soils in all the sampled locations (n=75) were analyzed in our previous study (Samonova et al. 2014). Particle size analysis and physical fractionation of the bulk soil into five particle size classes was performed for the fine earth fraction <1 mm. Overall, 22 soil samples were subjected to physical fractionation (11 samples for each landform). The boundaries between particle size classes were defined in accordance with the Russian conventional fraction groups (Arinushkina 1992): coarse and medium sand (1–0.25 mm), fine sand (0.25–0.05mm), coarse silt (0.05–0.01 mm), medium and fine silt (0.01–0.001 mm), clay (<0.001 mm). The listed fractions for the convenience were designated by capital latter G and (according to particle sizes) by Arabic numerals: G1, G2, G3, G4-5, G6. The sand fractions (G1 and G2) were separated by wet sieving while the silt fractions (G3 and G4-5) as well as the clay fraction (G6) were obtained by sedimentation and siphoning, during times determined by Stokes' law (Arinushkina 1992). In the individual particle size fractions (3-5 g) and also in the bulk soil material, the concentrations of Mn, Cu, Ni, Co, Cr, Zn, Pb, Ti and Zr were determined by atomic emission spectroscopy with 3-phase dc arc using DFS-458 equipment and the concentrations of Fe were determined by atomic absorption spectrometry. In total 110 samples of the individual fractions were analyzed for the metals' contents. Many of the analysed elements, such as Ni, Co, Cr, Zn, Pb, Cu are often-studied and linked with health risk concerns (Rinklebe et al. 2019; Zwozdziak et al. 2017), while some, namely Ti and Zr, have received little attention in environmental studies as compared to other potentially toxic elements (PTEs). Also, the key (typomorphic) elements, such as Mn and Fe, are necessary in explaining the distribution of PTEs since their (hydr)oxides may incorporate trace elements and thus influence mobilization and accumulation of trace metals in soils.

Data treatment

For two independent subpopulations of the samples ("the balka" and "the gully" datasets, n=11, each), we calculated descriptive statistics and used two-sample t-test and significance level α=0.05 to compare the means between the two groups in order to understand whether there is a difference between the metal concentrations in bulk soil or in a particular particle size fraction on the basis of the landform type. For the statistical treatment the software program SPSSv11.0 was used. The magnitude of the differences between two subpopulations was also evaluated by calculating enrichment ratio. It shows how many times the average metal concentrations in one group of samples (bulk soil or individual particle size fractions) taken from one landform exceeds the average concentrations in the other group. This ratio in the text is given as a numeric index next to the element symbol.

The relative importance of a particular fraction with respect to its contribution to total metal concentrations was assessed on the basis of the soil textural data and the content of a particular metal in each size fractions by calculating average mass loading using the following formula (Sutherland 2003; Gong et al. 2014):

$$MGi = (Mi * Gi) / (\sum_{i=1}^{5}(Mi * Gi) * 100) \quad (1)$$

where *Mi* – is the average content of the metal in a particle size fraction i (mg·kg^{-1}); *Gi* – the average mass percentage of individual fraction.

RESULTS AND DISCUSSION

The particle size distribution and physicochemical characteristics of the bulk soil in the gully and the balka systems

The data on average proportions of individual particle size fractions in humus horizons of various locations of the two systems are shown in Fig. 3. The greatest differences in the granulometry of the topsoil horizons are observed in the gully system, which can be explained mostly by its heterolithic nature. In the catchment area the particle-size distribution is very characteristic for loessic material and resembles one found in the topsoil of the balka's catchment. Towards the gully's fan due to the landform incision into glaciofluvial strata the granulometry of the topsoil changes: total sand share (G1 and G2) increases to 55%, while the proportions of the silt (G3 and G4-5) drop accordingly from 70% to 32% (Fig. 3a). The amount of the clay fraction stays relatively stable, however, on average, the clay is slightly enriched in the soils located on the sides of the landform and in its adjacent catchment area (13-12%). For the longitudinal soil sequence in the gully's bottom, the linear trend estimates performed in our previous study (Samonova et al. 2014) proved that the amount of the sand fractions exhibit an increasing trend (r=0.79, p=0.002), while the silt fractions reveal decreasing trends (r=-0.75, p=0.005), which might indicate redistribution of soil material in longitudinal direction taking place during summer rainstorms or spring snowmelt.

Compared to the gully, the distribution of particle size fractions in soils along the different units of the balka system is more uniform. In all geomorphic positions the balka's topsoil horizons are dominated by the coarse silt, whose proportions

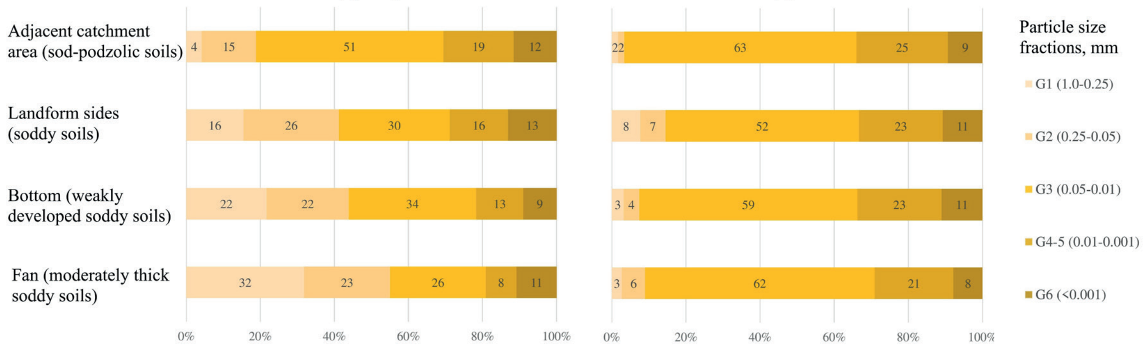


Fig. 3. Average granulometry of the topsoil horizons in the gully (a) and the balka (b) in various soil locations

vary between 52-63% (Fig. 3b). The clay comprises 8-11% of the fine earth fraction. The contents of the sand fractions (G1 and G2) vary within a wider range: in the topsoil of the balka's sides the sand accounts for 14%, which is higher than in the adjacent area (4%) or in the balka's bottom (7%). The estimations of linear trends for the longitudinal soil sequence in the balka's bottom (n=16) (Samonova et al. 2014) showed the patterns similar to those described for the gully.

The humus content in soils of the gully and balka systems varies considerably from very low (1-2%) to very high values (Fig. 4a) due to control of erosional-depositional processes as well as different development of vegetation communities and ground cover producing organic residues and protecting soil from erosion. The highest average levels of soil organic matter in both systems are observed in the soils occupying the sides of the landforms (Fig. 4a) probably due to accumulation of soil material that moved with creep processes (Azhigirov et al. 1987; Gennadiev et al. 2008; Golosov 2006). The minimum humus content in the two systems is found in the bottom locations which reveal decreasing linear trends in longitudinal soil sequences (from the headcuts towards fans) (Samonova et al. 2014).

The data on the soil exchangeable acidity (Fig. 4b) display that the soils of adjacent catchment areas formed on a single parent material mostly have weakly acidic reaction. The topsoil of the internal units in the balka are characterized by weakly acidic or neutral reaction, while in the gully this parameter changes more sharply, from weakly acidic to weakly alkaline reaction. The sharp increase in pH in the landform long profile is accompanied by the rise of the fine sand content in topsoil horizons, which is caused by the incision of the gully landform into calcareous sandy strata.

Comparison of the gully and balka humus horizons on the basis of heavy metal concentrations in the bulk and individual soil particle size fractions

The metal contents calculated for the bulk soil and individual particle size fraction for two subpopulations of samples (grouped on the basis of gully or balka membership) are presented in Appendix (Table A), and Fig. 5. The magnitude of differences between the mean values in two groups of samples was evaluated using the enrichment ratio.

The results of the performed comparison demonstrated that the bulk soil (humus horizons) of the gully and the balka systems have nearly equal levels of Ni and very small differences in the average concentrations of a wide spectrum of the elements including Fe, Zr, Cu, Pb, Co (± 10%). Despite very narrow divergence between metals' contents in the gully and the balka topsoils, the differences are statistically significant (Table 1) for all listed elements except for Cu. The concentrations of Co, Zr are slightly higher in the soils of the balka, while the gully topsoil has higher levels of Pb and Fe. The larger magnitude of

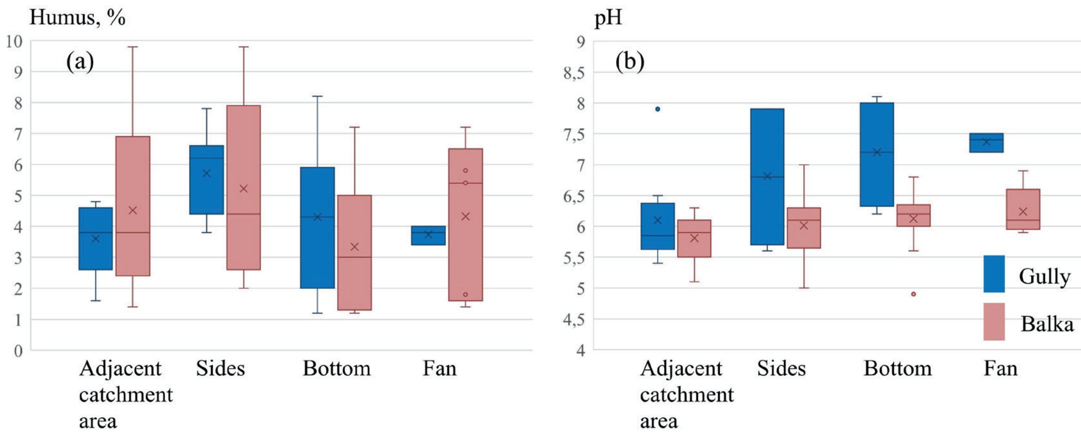


Fig. 4. Statistics on humus content (a) and pH (b) in the topsoil of the gully and balka systems

Table 1. Results of two-sample t-test for metal concentrations in bulk and individual particle size fractions separated from humus horizons of the balka and the gully: two-tailed test p-values (n=11)

Elements	G1 1-0.25	G2 0.25-0.05	G3 0.05-0.01	G4-5 0.01-0.001	G6 <0.001 mm	Bulk
Cu	0.556	0.495	<0.001*	<0.001*	<0.001*	0.474
Zn	0.105	0.482	0.001*	<0.001*	<0.001*	0.861
Pb	0.002*	0.174	<0.001*	<0.001*	<0.001*	0.001*
Co	0.011*	0.897	<0.001*	<0.001*	0.450	<0.001*
Ni	0.841	0.117	<0.001*	<0.001*	0.126	<0.001*
Cr	0.551	0.390	<0.001*	<0.001*	0.001*	<0.001*
Mn	0.006*	0.273	0.012*	0.003*	0.752	<0.001*
Fe	0.156	0.212	<0.001*	<0.001*	0.629	<0.001*
Ti	0.219	1,000	0.312	0.000*	0.017*	<0.001*
Zr	0.081	0.873	0.284	0.001*	<0.001*	<0.001*

** the star sign indicates p-values below the chosen threshold (a=0.05), which means that two sets of data are significantly different from each other. For such cases the font colour (red, blue or black) is used to show the magnitude of the difference. The red or blue fonts are used to demonstrate high magnitude of the divergence between mean concentrations with the enrichment ratio ≥2. The red font shows that the mean concentration of a particular element is higher in the balka, while the blue one – in the gully samples.*

differences between the bulk soil of the two systems were observed for Mn, Cr, Ti, and Zn, showing higher levels in the the balka (with the enrichment ratio equal 2 for Mn, and 1.3-1.4 for other elements). The difference, however, is significant only for Mn, Cr and Ti (Table 1), while for Zn it is not. Thus, the variations in lithology across the balka and the gully systems have clear response in metals concentrations in the bulk soil: the topsoil of the balka system is generally higher in metals than that of the gully except for two metals – Fe and Pb, enriching soils of the gully.

The sand fractions might have different genesis in the two systems. In the gully, the sand could be inherited mainly from glaciofluvial strata, but in the balka sand material is mostly derived from mantle (loessial) loams. According to the obtained data, the coarse and medium sand (G1) separated from the balka's topsoil compared to that in the gully contains equal average amount of Ni, but higher concentrations of $Mn_{2.8}$, $Co_{2.1}$, $Pb_{1.6}$, Ti,$Zr_{1.5}$, Cu,$Zn_{1.2}$, and lower concentrations of $Fe_{1.4}$. and $Cr_{1.2}$. However, statistically these differences are significant only for Mn, Co and Pb (Table 1). The higher content of these three elements in coarse particle size fraction might reflect a relative abundance of manganese coatings or nodules which are usually enriched in several other trace metals, in particular Co and Pb (Kabata-Pendias 2011; Sposito 2008; Barrón and Torrent 2013). The observed metal depletion in the sand sized material in the gully system seems to be related to the fact that glaciofluvial sediments in comparison with mantle loam or till are relatively enriched in quartz (Protasova 2003; Rychagov and Antonov 1992) that does not contain or very few other elements than Si (Hardy and Cornu 2006). The fine sand fraction (G2) of the two systems has less distinct differences than the coarser one. It is equal in terms of Ti, Zr, Zn, and Co levels. However, it showed statistically insignificant differences for other elements whose average concentrations were mostly 1.5-1.2 higher in the balka, except for Cu, enriching this particle size fraction in the gully soil.

The coarse silt (G3), separated from the gully and the balka soils and composed mostly of primary minerals (Schaetzl et al. 2005), does not differ in terms of Ti and Zr levels. At the same time, in the gully it contains significantly higher amounts of Cu,$Fe_{3.4-3.2}$, $Ni_{2.3}$,$Co_{1.7}$,$Zn_{1.6}$,$Pb_{1.5}$,$Mn_{1.4}$,$Cr_{1.2}$. *The medium and fine silt* (G4-5) from the gully soils also demonstrates higher contents of a wide range of metals ($Cu_{5.9}$,$Pb_{2.5}$, $Fe_{2.4}$, $Zn_{2.0}$, $Ni_{1.9}$, $Co_{1.8}$, $Mn_{1.6}$, $Cr_{1.2}$), while the same fraction from the balka's humus horizon is significantly higher in stable elements – $Zr_{1.5}$ and $Ti_{1.3}$. The predominance of micas and chlorite which accumulate a wide spectrum of trace metals (Hardy and Cornu 2006) over other major silt constituents, such as quartz, alkaline feldspars and plagioclases, seems to be a reasonable explanation for higher amounts of trace elements in the silt-sized material of the gully soils. The fine and medium silt fractions in the balka soils probably contain less micas and more Zr and Ti-rich primary accessory minerals (rutile and zircon).

The clay fraction (G6) showed very little (less than 10%) and not significant differences in Fe, Mn, Co, Ni levels, but higher divergence and significant difference in the average concentrations of other elements (Table 1). In particular, $Cu_{3.3}$, $Pb_{1.5}$, $Zn_{1.4}$, $Cr_{1.2}$ enrich the clay of the gully's soil, while $Zr_{2.0}$ and $Ti_{1.1}$ accumulate in the clay of the balka's humus horizons. Some of these patterns obviously are inherited from the silt fractions because clay is formed from silt due to geochemical and mineralogical transformation of unstable primary minerals in secondary minerals and also physical microdivision of primary silt particles (Hardy and Cornu 2006).

Comparison of the gully and balka on the basis of metals' partitioning in various particle size fractions

Metal distribution among soil particle size fractions is shown in Fig. 5. It is obvious that the metals show different patterns due to the different ways in which these elements can occur in the soils. They can be present in soil (a) as structural components in primary stable (Zr, Ti) or ferrous minerals (Fe, Ti, Mn), (b) in secondary minerals such phyllosilicate clay or as (hydr)oxides (Fe, Mn), especially in redox conditions, and (c) as adsorbed ions on clay minerals and/or humus (Pb, Zn, Cu, Ni). Part of the latter elements and elements of ferro-family (Co, Ni) may also be concentrated in the Fe Mn (hydr)oxides, but they are unlikely to form minerals of themselves.

The analysis of the partitioning of metals among different soil particle size fractions in the gully system showed that the studied elements can be combined into the following groups. The first group includes Fe and metals of ferro-family – Mn, Ni and Co, with a minimum average content in the coarse silt and maximums in two hosting fractions – the clay and the coarse and medium sand. Bimodal distribution of elements with concentration maxima in the clay and sand fraction was reported earlier by Qian et al. (1996) and was explained by the presence of several soil minerals and phases having high retention for metals. It is reasonable to suggest that in the clay the listed elements are associated with secondary clay minerals, while in the sand they are bound to newly formed Fe-Mn-hydroxides or ferrous heavy minerals, such as hornblende, garnet, ilmenite and glauconite, which are common in the soil parent materials of the study area (Rychagov et al. 1992).

The second group includes Zn, Pb, Cu, Cr. These trace metals are characterized by a minimum average content in the sand but by the maximum concentrations in finer fractions: the medium and fine silt (Pb, Cu) and the clay (Zn, Cr). The preferential accumulation of metals in the fine fractions was reported earlier both for the natural (Acosta et al. 2009; Hardy and Cornu 2006) and polluted soils (Ljung et al. 2006) suggesting that they are rather present in adsorbed form in the clay minerals and/or humus.

The third group incorporates stable elements Ti and Zr. In the gully system they reveal the maximum concentrations in the silt fractions: in the coarse silt (Zr) and in the medium and fine silt (Ti), which can be explained by their preferential occurrence in detrital grains of primary accessory minerals such as rutile, titanite, zircon (Sposito 2008). These minerals often survive several cycles of weathering and sedimentation (Deer et al. 1997) and usually concentrate in the fine sand to coarse silt fractions.

In the balka system the patterns of distribution shown by some metals are very similar to those found in the gully's soils. This is the case for Mn, Ni, Co, Fe (1), Cr (2) and also for Ti, and Zr (3). The elements from the third group compared to the gully soil, however, have slightly higher levels in the silt and the clay fractions probably due to higher mineralogical maturity of the correspondent fractions of mantle loam deposits associated with the higher relative abundance of stable accessory minerals.

The particle size partitioning of Zn, Pb and Cu seem to be more lithology dependent. The minimum concentrations of Pb, Cu, Zn in the gully on average are observed in the sand fractions (fine sand), while in the balka the lowest levels are observed in the coarse silt; the maximum concentrations of Pb and Cu in the gully are typical of the medium and fine silt, while in the balka these elements tend to accumulate in the clay.

Fig. 5. Comparison between the abundances of metals in particle size fractions in humus horizons in the gully with those in the balka

Arithmetic mean is indicated as a cross, median – as a line across the box. X-axe – particle size fractions

Comparison of the gully and balka on the basis of variability of metals' concentrations in bulk soils and their various particle size fractions

The extent of variability of the metals' concentrations in the bulk samples of the topsoil horizons and in particle size fractions is given in Table A (Appendix). According to the variation coefficient (CV), which is often viewed as a measure of group heterogeneity (Dmitriev 2009) the highest degree of dispersion of concentrations around the mean in bulk samples of the gully system is shown by Mn (CV=50%). Other metals can be arranged into the following groups: Zn, Pb, Zr, Co (CV=20-30%); Fe, Cr, Cu, Ni, Ti (CV=10-20%). The balka has less variable topsoil geochemistry regarding metal concentrations, which can be explained by relatively homogeneous textural and mineralogical composition of the bulk samples in comparison to those in more heterogeneous gully system. In the fraction of the coarse and medium sand separated from the gully soils, the variation coefficient for the majority of the elements exceed 50%, except for Pb (30%). The dispersion around the mean reaches the highest values for Ti and Mn (101% and 128%, respectively) due to the presence of outliers. The same fraction in the balka's topsoil seems to be less variable with respect to Ti concentrations, however, Ni and Cu display higher and Zn, Pb, Cr, Fe nearly equal degree of variation.

Across the sand fractions, the fine sand in both systems reveals lower dispersion of Mn and Ti. The same trend is observed for Cr and Pb in the gully and for Cu, Ni and Fe in the balka. The higher CV values, which might indicate an increase in spatial differentiation of this fraction in comparison to the coarser sand, are displayed by Zn (in both systems), Cu, Ni, Fe (in the gully) and Pb, Cr (in the balka).

The coarse silt, when comparing with sand fraction G2, demonstrates a significant drop in the variation coefficients for the majority of the elements both in the gully (except for Pb) and in the balka (except for Mn), which could be due to the fact that the silt is more homogeneous in its mineralogical composition.

The medium and fine silt and the clay have relatively low CV for most metals. In the lithologically heterogeneous gully system finer particle size fractions G4-5 and G6 reveal the minimal CV values for all elements, except for Pb and Mn, whose lowest variation were observed in the fine sand (Pb) or coarse silt (Mn). In the balka's soils only Mn displayed the minimal dispersion in the sand fraction (G2), while other elements were found to be the least variable in G3 and finer fractions. When comparing fine soil fractions between the two systems, the general trend was revealed: the variations of the metal concentrations are higher in the gully's soil. In the medium and fine silt the difference in CV values for Zn, Pb, Fe is twice higher or nearly three times for Zr. In the clay the same magnitude of differences in CV is shown by Pb and Zr. The lower variations in metal contents in the balka's fine fractions could be due to more homogeneous mineralogical composition of fine soil material and/or its more intensive transformation by soil-geochemical processes.

Contributions of particle size fractions to total metal concentrations in topsoil horizons of the gully and balka system.

The relative importance of particle size fractions in controlling total metal concentrations in topsoil horizons of the gully and balka system was evaluated by considering the joint effect of average abundances of particle size fractions and their metal contents. The results of these calculations are shown in Fig. 6.

In the gully soils (Fig. 6a), the maximum mass loadings of Mn (39%), Zn (34%), Co and Ni (25%) were associated with the clay fraction. This reflects high concentrations of the listed metals in clay-sized particles (Fig. 5) characterized, however, by relatively low average share (11%) in the topsoil horizons of the studied system. For the same reason the maximum contribution of Cu (37%) and Pb (33%) in the gully's soils came from the fraction of the medium and fine silt (Fig. 6a). The coarse silt, which dominates over other particle size fractions, mostly controls the concentrations of stable elements – Zr (62%), Ti (50%) and also Cr (39%), while the contribution of other elements from this fraction does not exceed 22% (29% for Pb). The only element whose maximum mass loading came from the sand fraction (coarse and medium sand) was Fe (26%).

In the balka the clay fraction with average proportion of 10% makes the maximum contribution to the total amount of a wide range of metals: Mn (43%), Fe (35%), Ni (33%), Zn (31%) and Co (30%). The second important fraction is the coarse silt (Fig 6b). Because of high amount in the bulk soil (58%) it makes the maximum mass loading for Zr (73%), Ti (58%), Cr (53%) Pb (42%) and Cu (32%). The coarser sand fractions G1 and G2 due to their very low amount in the surface horizons of the balka's soils (4%) have very little loadings of Ti, Zr, Cr, Pb but make slightly higher contribution to total abundances of ferro-family elements (Mn, Fe, Co, Ni) and also Cu.

CONCLUSIONS

The bulk soil of the gully and balka systems revealed very little differences in average concentrations for the majority of metals except for Mn, Cr and Ti enriching the balka's soils.

Some individual particle size fractions in the two systems have equal average levels of the specific metals: the coarse and medium sand – Ni, the fine sand – Ti, Zr, Zn, Co; the coarse silt – Ti and Zr, the clay fraction – Fe, Mn, Co and Ni.

The silt fractions separated from the gully's and the balka's humus horizons have the most prominent differences regarding the mean metal contents. The silt size particles in the gully's soils on average are more metal-rich than those in the balka's soils. The significant differences in metal concentrations between the two studied systems were also found for the coarse and medium sand: this fraction has higher levels of the majority of metals in the balka's soil.

The variability of the metal content is related to the sizes of particles and decreases from the coarser to finer particle size fractions, which is explained by relatively homogeneous

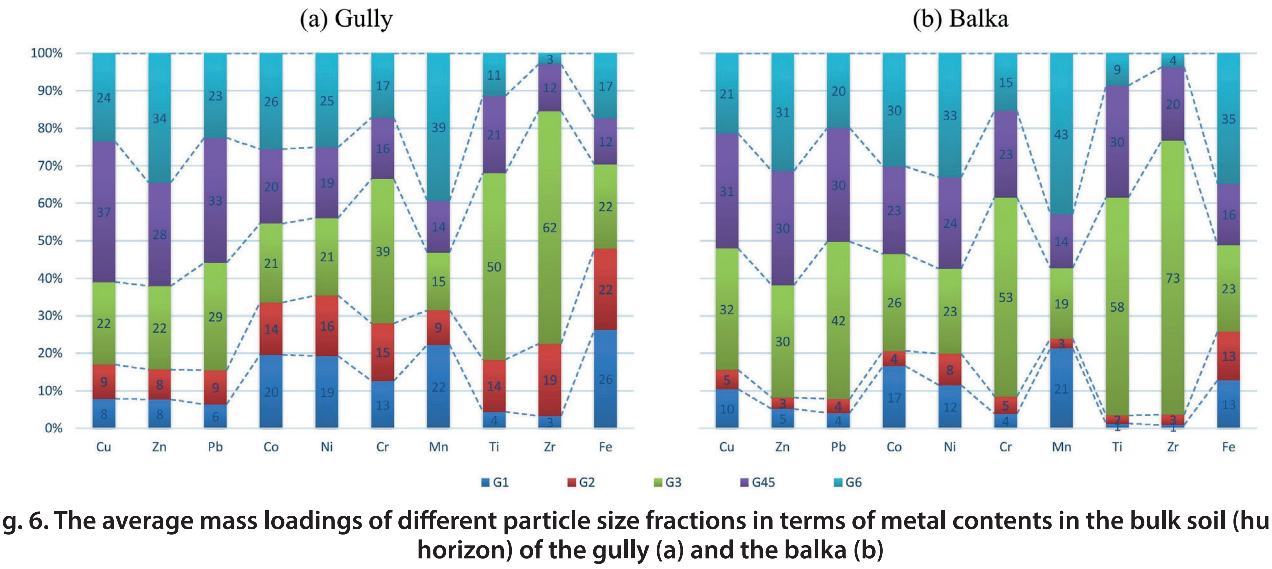


Fig. 6. The average mass loadings of different particle size fractions in terms of metal contents in the bulk soil (humus horizon) of the gully (a) and the balka (b)

mineralogical composition and the predominance of secondary clay minerals over the primary ones in the fine silt and clay fraction.

The topsoil horizons of the two systems have common types of distribution over the particle size fractions for the majority of the studied metals but slightly different ones for Zn, Cu and Pb. Our data provides evidences that preferential association of metals with particle size fractions is not limited to finer fractions. The elements of ferro-family (Fe, Mn, Co, Ni) appear to have bimodal distribution with concentration maxima in the clay and sand fraction. The partitioning of Ti and Zr accumulating in the silt fractions is governed by their incorporation in mineral structure of accessory minerals that are stable during the processes of physicochemical weathering and soil formation.

The contribution of the clay to the total mass loading of metals was the greatest for Mn, Zn, Ni and Co in both systems. The highest mass loading for Ti, Zr and Cr came from the coarse silt, while for Cu and Pb it was made by different particle size fractions: the medium and fine silt in the gully or the coarse silt in the balka. The highest contribution of Fe also came from different fractions, either from the coarse sand in the gully or the clay in the balka.

The genesis and lithogeochemical parameters of the soil parent material might affect the particle size partitioning of metals in the topsoil in the small erosional landforms and control the absolute concentrations of metals in individual particle size fractions.

ACKNOWLEDGEMENTS

This research was supported by the Russian Science Foundation (Project No. 19-77-30004).

REFERENCES

Acosta J., Faz Cano A., Arocena J., Debela F. and Martínez-Martínez S. (2009). Distribution of metals in soil particle size fractions and its implication to risk assessment of playgrounds in Murcia City (Spain). Geoderma, 149, 101-109. DOI: 10.1016/j.geoderma.2008.11.034.

Acosta J., Martínez-Martínez S., Faz A. and Arocena J. (2011). Accumulations of major and trace elements in particle size fractions of soils on eight different parent materials. Geoderma, 161(1), 30-42. DOI:10.1016/j.geoderma.2010.12.001.

Ajmone-Marsan F., Biasioli M., Kralj T., Grčman H., Davidson C.M., Hursthouse A.S., Madrid L. and Rodrigues S. (2008) Metals in particle-size fractions of the soils of five European cities. Environmental Pollution, 152, 73-81. DOI: 10.1016/j.envpol.2007.05.020.

Arinushkina E. (1992). Handbook for chemical analysis of soils. Moscow: Chimiya Publishing house. (in Russian).

Azhigirov A., Golosov V., Dobrovolskaya N., Litvin L. and Samodurova L. (1987). Investigation of the flow of water and sediment along interfluve slopes in the Protva river basin. VINITI, 6389-B87, 51-77. (in Russian).

Barrón V. and Torrent J. (2013). Iron, manganese and aluminium oxides and oxyhydroxides. European Mineralogical Union Notes in Mineralogy, 14, 297-336. DOI: 10.1180/EMU-notes.14.9.

Banerjee A. (2003) Heavy metal levels and solid phase speciation in street dusts of Delhi, India. Environmental Pollution, 123, 95-105. DOI:10.1016/S0269-7491(02)00337-8.

Buggle B., Glaser B., Zoeller L., Hambach U., Markovic S., Glaser I. and Gerasimenko N. (2008). Geochemical characterisation and origin of Southeastern and Eastern European loesses (Serbia, Romania, Ukraine). Quaternary Science Reviews, 27, 1058-1075. DOI: 10.1016/j.quascirev.2008.01.018.

Deer W., Howie R. and Zussman J. (1997) Orthosilicates. In: Rock-forming minerals: Volume 1A, 2nd ed. Bath, UK: Geological Society, 418-466.

De Miguel E., Llamas J., Chacón E., Berg T., Larssen S., Royset O. and Vadset M. (1997). Origin and patterns of distribution of trace elements in street dust: unleaded petrol and urban lead. Atmos. Environ., 31, 2733-2740.

Dmitriev Ye. (2009). Mathematical Statistics in Soil Science, 3rd ed. Moscow: Book House Librokom. (in Russian).

Dobrovolskii V. (1983). Geography of Trace Elements. Global Dispersion. Moscow: Mysl'. (in Russian).

Förstner U. (1982). Chemical forms of metal accumulation in recent sediments. In: G. C. Amstutz, G. Frenzel, C. Kluth, G. Moh, A. Wauschkuhn, R. Zimmermann, A. El Goresy, eds., The Ore genesis. Berlin; Heidelberg: Springer-Verlag, 191-199.

Gennadiev A., Koshovskii T., Zhidkin A. and Kovach R. (2013) Lateral migration of soil solid-phase material within a landscape-geochemical arena detected using the magnetic tracer method. Eurasian Soil Science, 46(10), 983-993.

Golosov V. (2006). The erosional and accumulative processes in river basins in cultivated plains. Moscow: Geos. (in Russian).

Gong C., Ma L., Cheng H., Liu Y., Xu D., Li B., Liu F., Ren Y., Liu Z., Zhao C., Yang K., Nie H. and Lang C. (2014) Characterization of the particle size fraction associated heavy metals in tropical arable soils from Hainan Island, China. Journal of Geochemical Exploration, 139, 109-114. DOI: 10.1016/j.gexplo.2013.01.002.

Hardy M. and Cornu S. (2006). Location of natural trace elements in silty soils using particle-size fractionation. Geoderma, 133, 295-308. DOI: 10.1016/j.geoderma.2005.07.015.

Kabata-Pendias A. (2011) Trace elements in soils and plants, 4th ed. London, New York: CRC Press, Boca Raton.

Khademi H., Gabarrón M., Abbaspour A., Martínez-Martínez S., Faz A. and Acosta J. (2019). Environmental impact assessment of industrial activities on heavy metals distribution in street dust and soil. Chemosphere, 217, 695-705. DOI: 10.1016/j.chemosphere.2018.11.045.

Kolevatykh E. (2011). Geochemistry of landscapes on mantle loams of the Vyatsko-Kama Cis-Urals: synopsis of the PhD theses. Perm. (in Russian).

Liu G., Wang J., Liu X., Liu X., Li X., Ren Y., Wang J. and Dong L. (2018). Partitioning and geochemical fractions of heavy metals from geogenic and anthropogenic sources in various soil particle size fractions. Geoderma, 312, 104-113. DOI: 10.1016/j.geoderma.2017.10.013.

Ljung K., Selinus O., Otabbong E. and Berglund M. (2006). Metal and arsenic distribution in soil particle sizes relevant to soil ingestion by children. Applied Geochemistry, 21, 1613-1624. DOI: 10.1016/j.apgeochem.2006.05.005.

Muratov M. (1953). On the conditions of loam formation in the Quaternary period. Bulletin of the Quaternary Commission, 19, 57-64. (in Russian).

Panin A., Fuzeina Ju. and Belyaev V. (2009). Long-term development of Holocene and Pleistocene gullies in the Protva River basin, Central Russia. Geomorphology, 108, 71-91. DOI: 10.1016/j.geomorph.2008.06.017.

Panin, A., Fuzeina, Y., Karevskaya, I. and Sheremetskaya, E. (2011). Mid-Holocene gullying indicating extreme hydroclimatic events in the centre of the Russian Plain. Geographia Polonica, 84, 95-115. DOI: 10.7163/GPol.2011.1.6.

Pobedinceva I. (1975). The soils on ancient weathering crusts. Moscow: Moscow Univ. Press. (in Russian).

Protasova N. (2003). Rare and trace elements (Mn, Cr, V, Ni, Cu, Zn, Co, Mo, Be, Ti, Zr, Ga, Sr, Ba, I, B) in the parent materials of the Central Chernozemic Belt region. Voronezh University Bulletin. Series Himija, biologija, farmacija, 2, 164-171. (in Russian).

Qian J., Shari X., Wang Z. and Tu Q. (1996). Distribution and plant availability of heavy metals in different particle-size fractions of soil. The Science of the Total Environment, 187, 131-141.

Rinklebe J., Antoniadis V., Shaheen S. M., Rosche O. and Altermann M. (2019) Health risk assessment of potentially toxic elements in soils along the Central Elbe River, Germany. Environment International, 126, 76-88. DOI: 10.1016/j.envint.2019.02.011.

Rychagov G. and Antonov S. (1992). The integrated analysis of the Quaternary deposits of the Satino training station. Moscow: Moscow Univ. Press. (in Russian).

Samonova O. and Aseeva E. (2006). Geochemical transformation of mantle and moraine loams as a result of soil formation in the Central Protva river basin. Vestnik Moskovskogo universiteta. Seriya 5, Geografiya, 6, 67-74. (in Russian with English summary).

Samonova O. and Aseeva E. (2008). The distribution of metals in grain-size fractions of soils in the southeastern part of the Smolensk-Moscow Upland. Vestnik Moskovskogo universiteta. Seriya 5, Geografiya, 3, 32-39. (in Russian with English summary).

Samonova O. and Aseyeva E. (2013). Distribution of metals in particle size fractions in soils of two forested catenas (Smolensk-Moscow upland). Geography, Environment, Sustainability, 6(2), 28-33, DOI: 10.24057/2071-9388-2013-6-2-28-33.

Samonova O., Aseyeva E. and Kasimov N. (2014). Metals in soils of erosional systems in forest zone in the central part of European Russia. J. Geochem. Explor., 144, 247–259. DOI: 10.1016/j.gexplo.2014.03.012.

Semenkov I. and Koroleva T. (2019). The spatial distribution of fractions and the total content of 24 chemical elements in soil catenas within a small gully's catchment area in the Trans Urals, Russia. Applied Geochemistry, 106, 1-6. DOI: 10.1016/j.apgeochem.2019.04.010.

Schaetzl R.J. and Anderson S. (2005). Soils. Genesis and geomorphology. Cambrige university press.

Sposito G. (2008). The chemistry of soils. 2nd ed. New York, USA: Oxford University Press.

Sutherland R. (2003). Lead in grain size of road deposited sediment. Environmental pollution, 121, 229-237, DOI: 10.1016/S0269-7491(02)00219-1.

Van Reeuwijk L. (1992). Procedures for soil analysis. 3rd ed. Wageningen: ISRIC.

Vlasov D., Kasimov N. and Kosheleva N. (2015). Geochemistry of road dust in the Eastern District of Moscow. Vestnik Moskovskogo universiteta. Seriya 5, Geografiya, 1, 23-33. (in Russian with English summary).

World Reference Base for Soil Resources 2014, update 2015. (2015). International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports 106. Rome: Food and Agriculture organization of the United Nation. Available at: www.fao.org/soils-portal/soil-survey/soil-classification/world-reference-base/en/ [Accessed 20 Jan. 2019].

Zhang J., Wu L., Zhang Y., Li F., Fang X. and Mao H. (2019). Elemental composition and risk assessment of heavy metals in the PM10 fractions of road dust and roadside soil. Particuology, 44, 146-152. DOI: 10.1016/j.partic.2018.09.003.

Zwozdziak A., Gini M.I., Samek L., Rogula-Kozlowska W., Sowka I. and Eleftheriadis K. (2017). Implications of the aerosol size distribution modal structure of trace and major elements on human exposure, inhaled dose and relevance to the PM2.5 and PM10 metrics in a European pollution hotspot urban area. Journal of Aerosol Science, 103, 38-52. DOI: 10.1016/j.jaerosci.2016.10.004.

Received on January 1th, 2020 Accepted on January 21th, 2020

APPENDIX

Table A. Metal concentrations in bulk and individual particle size fractions of topsoil horizons in the gully and balka

(n=11; Min: minimum, Max: maximum, VC: variation coefficient, %)

Sample	Statistics	Cu[a]	Zn[a]	Pb[a]	Co[a]	Ni[a]	Cr[a]	Mn[a]	Ti[a]	Zr[a]	Fe[b]
Bulk soil	Gully										
	Min	24	40	21	7	22	40	320	3500	320	1.5
	Max	36	90	44	18	40	63	1300	6200	840	2.0
	Mean	28.2	63.6	31.5	10.7	27.6	52.8	563.6	4281.8	536.4	1.9
	VC	14.2	20.2	22.5	31.2	18.6	14.0	51.7	19.6	29.1	10.6
	Balka										
	Min	20	60	26	11	21	59	600	5100	350	1.5
	Max	34	100	35	14	27	84	1600	6400	720	2.0
	Mean	29.6	82.7	30.8	12.2	24.7	71.7	1107.3	5381.8	570.9	1.8
	VC	12.4	14.4	8.9	10.9	9.1	10.0	29.9	7.7	16.6	14.7
1-0.25mm	Gully										
	Min	31	30	12	6	21	25	420	500	60	2
	Max	150	100	31	40	98	150	6400	5000	290	15
	Mean	72.4	67.5	18.6	16.9	40.5	54.5	1410.9	1246.4	114.5	7.0
	VC	56.4	48.7	31.3	75.6	56.6	61.9	128.2	100.9	55.9	49.5
	Balka										
	Min	17	30	20	11	14	21	520	640	100	2.0
	Max	240	130	50	74	130	120	8600	3600	280	10.0
	Mean	86.6	80.9	30.6	35.3	42.9	46.5	3983.6	1812.7	166	5
	VC	78.0	40.4	31.9	49.6	76.7	59.7	53.9	43.1	38.3	52.5
0.25-0.05mm	Gully										
	Min	19	30	13	3	7	24	250	1600	250	1.5
	Max	200	120	23	24	68	84	860	4200	800	10
	Mean	59.1	56.0	18.7	8.5	24.0	47.1	412.7	2800.0	474.5	4.1
	VC	88.5	66.3	15.7	72.8	73.3	46.6	45.8	27.6	39.5	65.8
	Balka										
	Min	21	25	12	5	15	31	360	1700	200	2.0
	Max	85	150	56	13	62	150	580	3800	610	10.0
	Mean	47.5	51.4	23.9	8.3	35.1	58.6	482.7	2800.0	459.1	5.5
	VC	39.2	68.0	49.4	35.9	39.8	64.2	17.1	23.3	32.5	41.3
0.05-0.01mm	Gully										
	Min	40	30	22	5	9	51	250	4100	470	1.5
	Max	140	100	44	10	25	84	600	6200	1200	4
	Mean	76.9	62.7	31.6	6.9	16.5	63.3	372.7	5363.6	819.1	2.3
	VC	36.0	31.2	22.7	25.4	31.2	14.9	29.2	15.6	33.0	37.3
	Balka										
	Min	17	30	18	3	6	46	180	4800	720	0.5
	Max	31	50	25	5	8	56	410	6200	1200	1.0
	Mean	22.8	38.2	20.4	4.0	7.2	50.5	260.9	5654.5	919.1	0.7
	VC	19.3	22.9	10.1	19.4	10.5	7.1	30.1	7.2	14.5	19.5

0.01-0.001mm	Gully										
	Min	160	120	35	12	24	56	100	4300	250	2
	Max	480	300	130	22	48	74	1200	6300	760	6
	Mean	323.6	192.7	91.5	16.2	37.3	66.6	823.6	5500.0	408.2	3.1
	VC	25.3	30.4	27.4	20.5	22.1	9.5	33.8	13.6	40.1	44.5
	Balka										
	Min	41	80	30	7	15	48	310	5700	540	1.0
	Max	79	130	45	12	26	62	840	8100	800	1.5
	Mean	54.4	98.2	37.2	9.1	19.5	55.7	503.6	7336.4	621.8	1.3
	VC	17.9	16.3	12.9	15.1	20.1	6.3	28.2	10.5	14.7	19.1
<0.001mm	Gully										
	Min	210	200	58	19	49	80	1700	3300	90	4
	Max	380	430	110	42	84	110	6000	5400	200	10
	Mean	280.0	330.9	85.0	28.6	68.2	96.2	3227.3	4190.9	130.0	6.0
	VC	21.2	23.0	22.0	20.4	16.6	8.1	49.6	14.8	23.8	28.9
	Balka										
	Min	58	180	48	17	48	76	1500	3900	210	3.0
	Max	120	300	71	36	92	100	5100	5400	300	8.0
	Mean	86.3	229.1	55.5	26.7	60.0	83.6	3418.2	4790.9	263.6	6.4
	VC	24.2	17.9	10.5	21.6	21.1	9.8	33.8	9.4	12.5	27.5

[a] $mg \cdot kg^{-1}$
[b] %

AUTHOR GUIDELINES

1. Authors are encouraged to submit high-quality, original work: scientific papers according to the scope of the Journal. reviews (only solicited) and brief articles.

2. Papers are accepted in English. Either British or American English spelling and punctuation may be used.

3. All authors of an article are asked to indicate their names (with one forename in full for each author, other forenames being given as initials followed by the surname) and the name and full postal address (including postal code) of the affiliation where the work was done. If there is more than one institution involved in the work, authors' names should be linked to the appropriate institutions by the use of 1, 2, 3 etc superscript. Telephone and fax numbers and e-mail addresses of the authors could be published as well. One author should be identified as a Corresponding Author. The e-mail address of the corresponding author will be published, unless requested otherwise.

EXAMPLE:

Ivan I. Ivanov[1], Anton A. Petrov[1*], David H. Smith[2]
[1]University/Organization name, Street, City, ZIP-code, Country
[2]University/Organization name, Street, City, ZIP-code, Country
*** Corresponding author:** example@example.com

4. The GES Journal style is to include information about the author(s) of an article. Therefore we encourage the authors to submit their portrait photos and short CVs (up to 100 words).

5. Article structure should contain further sections:
- Abstract.
- Keywords (Immediately after the abstract, provide a maximum of 6 keywords, avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.).
- Introduction (State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.).
- Material and methods (Provide sufficient details to allow the work to be reproduced by an independent researcher. Methods that are already published should be summarized, and indicated by a reference. If quoting directly from a previously published method, use quotation marks and also cite the source. Any modifications to existing methods should also be described.
- Results (Results should be clear and concise.).
- Discussion (This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.).
- Conclusions (The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.).
- Acknowledgements (List here those individuals who provided help during the research and funds which provided financial support of your research.).
- References (Make a list of references according to guidance provided below).
- Appendices (If there is more than one appendix, they should be identified as A, B, etc.

Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.).

6. Sections and sub-sections shouldn't be numbered, but divided into sub-levels. If you would like to refer to a particular section, name it without any numbers (for example: see "Materials and methods")

7. Math equations should be submitted as editable text and not as images (use appropriate tool in text editor). Present simple formulae in line with normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

8. Figures and figure captions. Number the illustrations according to their sequence in the text. Regardless of the application used, when your figure is finalized, please 'save as' or convert the images to one of the following formats: TIFF, JPG. Supply files that are optimized for screen use. **The resolution should be at least 300 dpi**. Besides inserting figures into your manuscript, you must also upload appropriate files during paper submission. Ensure that each illustration has a caption. A caption should comprise a brief title (not on the figure itself) and a description of the illustration beginning from "Fig. 1…" (note: no any dots in the end of the capture!). Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.
Example of referring to the figure in text: (Fig. 1)
Example of figure caption: Fig. 1. The distribution of the family Monimiaceae

9. Tables. Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text, place table caption above the table body and place any table notes below the body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules and shading in table cells.

Example of table caption: Table 1. Case studies and used methods

10. The optimum size of a manuscript is about 3,000–5,000 words. Under the decision (or request) of the Editorial Board methodological and problem articles or reviews up to 8,000–10,000 words long can be accepted.

11. To facilitate the editorial assessment and reviewing process authors should submit "full" electronic version of their manuscript with embedded figures of "screen" quality as a PDF or DOC file.

12. We encourage authors to list three potential expert reviewers in their field. The Editorial Board will view these names as suggestions only.

13. Original reviews of submitted manuscripts remain deposited for 3 years.

SCIENTIFIC AND PRACTICAL PEER-REVIEWED JOURNAL “GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY”

FOUNDERS OF THE MAGAZINE: Russian Geographical Society, Faculty of Geography, Lomonosov Moscow State University and Institute of Geography of the Russian Academy of Sciences

The magazine is published with financial support of the Russian Geographical Society.

The magazine is registered in Federal service on supervision of observance of the legislation in sphere of mass communications and protection of a cultural heritage. The certificate of registration: ПИ № ФС77-67752, 2016, December 21.

PUBLISHER
Russian Geographical Society
Moscow, 109012 Russia
Novaya ploshchad, 10, korp. 2
Phone 8-800-700-18-45
E-mail: press@rgo.ru
www.rgo.ru/en

EDITORIAL OFFICE
Lomonosov Moscow State University
Moscow 119991 Russia
Leninskie Gory,
Faculty of Geography, 1806a
Phone 7-495-9391552
Fax 7-495-9391552
E-mail: ges-journal@geogr.msu.ru
www.ges.rgo.ru

DESIGN
Agency «IP TERESHKIN ANTON IGOREVICH»
Moscow, 115088,
26 Simonovsky Val str., bldg. One
Phone: +7 (903) 108-04-44
E-mail: smile.tai@gmail.com

Sent into print 31.03.2020
Order N gi120

Format A4 (210x297mm)

ges.rgo.ru/jour/

ISSN 2071-9388 (Print)

ISSN 2542-1565 (Online)