ISSN 2542-1565

Vol. 19 2026 No. 01

GEOGRAPHY
ENVIRONMENT
SUSTAINABILITY

«The journal GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY was founded in 2008 by Russian Geographical Society, the Lomonosov Moscow State University Geography Department, and the Russian Academy of Sciences Institute of Geography. Since that time the journal publishes **4 issues per year**, containing original research papers and reviews. The journal issues are open source and distributed through subscriptions, library exchanges of leading universities, and via the website through the world»

FOUNDERS OF THE JOURNAL: Russian Geographical Society, Faculty of Geography, Lomonosov Moscow State University and Institute of Geography of the Russian Academy of Sciences

The journal is published with financial support of the Russian Geographical Society.

The journal 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, 1,
Faculty of Geography, 1806a
Phone 7-495-9391552
Fax 7-495-9391552
E-mail: ges-journal@geogr.msu.ru
www.ges.rgo.ru

DESIGN
Layout designer: Tereshkin Anton
Moscow, 115088,
26 Simonovsky Val str., bldg. One
Phone: +7 (903) 108-04-44
E-mail: smile.tai@gmail.com

DOI prefix: 10.24057

Format A4 (210x297mm)

"GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY" is the only original English-language journal in the field of geography and environmental sciences published in Russia. It is supposed to be an outlet from the Russian-speaking countries to Europe and an inlet from Europe to the Russian-speaking countries regarding environmental and Earth sciences, geography and sustainability. The main sections of the journal are the theory of geography and ecology, the theory of sustainable development, use of natural resources, natural resources assessment, global and regional changes of environment and climate, social-economical geography, ecological regional planning, sustainable regional development, applied aspects of geography and ecology, geoinformatics and ecological cartography, ecological problems of oil and gas sector, nature conservations, health and environment, and education for sustainable development.

OPEN ACCESS POLICY. "GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY" is an open access journal. All articles are made freely available to readers immediately upon publication. Our open access policy is in accordance with the Budapest Open Access Initiative (BOAI) definition - it means that articles have free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.

Date of publication: March 31st 2026.

© Russian Geographical Society, 2026

EDITORIAL BOARD

EDITORS-IN-CHIEF:
Kasimov Nikolay S.
Lomonosov Moscow State University,
Faculty of Geography, Russia

Kotlyakov Vladimir M.
Russian Academy of Sciences
Institute of Geography, Russia

DEPUTY EDITORS-IN-CHIEF:
Solomina Olga N. - Russian Academy of Sciences, Institute of Geography, Russia
Tikunov Vladimir S. - Lomonosov Moscow State University, Faculty of Geography, Russia
Vandermotten Christian - Université Libre de Bruxelles Belgium
Chalov Sergei R. - (Secretary-General) Lomonosov Moscow State University, Faculty of Geography, Russia

Alexeeva Nina N. - Lomonosov Moscow State University, Faculty of Geography, Russia
Baklanov Alexander - World Meteorological Organization, Switzerland
Chubarova Natalya E. - Lomonosov Moscow State University, Faculty of Geography, Russia
De Maeyer Philippe - Ghent University, Department of Geography, Belgium
Dobrolubov Sergey A. - Lomonosov Moscow State University, Faculty of Geography, Russia
Ferjan J. Ormeling - University of Amsterdam, Amsterdam, Netherlands
Sven Fuchs - University of Natural Resources and Life Sciences
Haigh Martin - Oxford Brookes University, Department of Social Sciences, UK
Golosov Valentin N. - Lomonosov Moscow State University, Faculty of Geography, Russia
Golubeva Elena I. - Lomonosov Moscow State University, Faculty of Geography, Russia.
Gulev Sergey K. - Russian Academy of Sciences, Institute of Oceanology, Russia
Guo Huadong - Chinese Academy of Sciences, Institute of Remote Sensing and Digital Earth, China
Jarsjö Jerker - Stockholm University, Department of Physical Geography and Quaternary Geography, Sweden
Jeffrey A. Nittrouer - Rice University, Houston, USA
Ivanov Vladimir V. - Arctic and Antarctic Research Institute, Russia
Karthe Daniel - German-Mongolian Institute for Resources and Technology, Germany
Kolosov Vladimir A. - Russian Academy of Sciences, Institute of Geography, Russia
Kosheleva Natalia E. - Lomonosov Moscow State University, Faculty of Geography, Russia
Konečný Milan - Masaryk University, Faculty of Science, Czech Republic
Kroonenberg Salomon - Delft University of Technology, Department of Applied Earth Sciences, The Netherlands
Kulmala Markku - University of Helsinki, Division of Atmospheric Sciences, Finland
Olchev Alexander V. - Lomonosov Moscow State University, Faculty of Geography, Russia
Malkhazova Svetlana M. - Lomonosov Moscow State University, Faculty of Geography, Russia
Maslakov Alexey A. - Lomonosov Moscow State University, Faculty of Geography, Russia
Minkina Tatiana M. - Southern Federal University, Russian Federation
Moreido Vsevolod M. - Russian Academy of Sciences, Water Problems Institute, Russia
Meadows Michael E. - University of Cape Town, Department of Environmental and Geographical Sciences South Africa
O'Loughlin John - University of Colorado at Boulder, Institute of Behavioral Sciences, USA
Paula Santana - University of Coimbra, Portugal
Pedroli Bas - Wageningen University, The Netherlands
Pilyasov Alexander N. - Institute of Regional Consulting, Moscow, Russia
Radovanovic Milan - Serbian Academy of Sciences and Arts, Geographical Institute "Jovan Cvijić", Serbia
Samsonov Timofey E. - Lomonosov Moscow State University, Faculty of Geography, Russia
Sebentsov Alexander B. - Russian Academy of Sciences, Institute of Geography, Russia
Sokratov Sergei A. - Lomonosov Moscow State University, Faculty of Geography, Russia
Tishkov Arkady A. - Russian Academy of Sciences, Institute of Geography, Russia
Wuyi Wang - Chinese Academy of Sciences, Institute of Geographical Sciences and Natural Resources Research, China

EDITORIAL OFFICE

ASSOCIATE EDITOR
Maslakov Alexey A.
Lomonosov Moscow State University,
Faculty of Geography, Russia

ASSISTANT EDITOR
Mozolevskaya Irina V.
Lomonosov Moscow State University,
Faculty of Geography, Russia

PROOF-READER
Syplenkov Anatoly S.
Lomonosov Moscow State University,
Faculty of Geography, Russia

CONTENTS

Disclaimer:

The information and opinions presented in the Journal reflect the views of the authors and not of the Journal or its Editorial Board or the Publisher. The GES Journal has used its best endeavors to ensure that the information is correct and current at the time of publication.

RESEARCH PAPER

MAPPING OF CORAL REEF HABITATS USING SPOT-7 AND SENTINEL-2 SATELLITE IMAGERY IN LIANG AND RAKIT ISLANDS, SALEH BAY, INDONESIA

Syamsul B. Agus[1*], Ananda S. Maulana[1], Taslim Arifin[2], Yulius[2], Aprizon Putra[2], Muhammad Ramdhan[3], Aida Heriati[2], Rinny Rahmania[2], Joko Prihantono[2], Agus Sufyan[4], Eva Mustikasari[4], Ruzkiah Asaf[2], Admi Athirah[2], Hafidzuddin Fihrin[1]

[1]Department of Marine Science and Technology, Faculty of Fisheries and Marine Sciences, IPB University, Jl. Agatis No 1, Bogor 16680, Indonesia
[2]Research Center for Ecology, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta–Bogor Km 46, Cibinong, Bogor 16911, Indonesia
[3]Research Center for Geoinformatics, National Research and Innovation Agency (BRIN), Jl. Sangkuriang, Bandung 40135, Indonesia
[4]Research Center for Biota Systems, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta–Bogor Km 46, Cibinong, Bogor 16911, Indonesia
*Corresponding author: sba_cacul@apps.ipb.ac.id

Received: December 31st 2024 / Accepted: July 28th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3811

ABSTRACT. Coral reef ecosystems are vital for marine biodiversity, coastal protection, and local livelihoods, yet environmental and anthropogenic stressors increasingly threaten them. Accurate and timely mapping of these habitats is essential for effective management and conservation planning. This study aims to evaluate and compare the spatial distribution and extent of coral reef habitats around Liang and Rakit Islands, as well as to assess the classification accuracy of Sentinel-2 Level-2A and the SPOT-7 satellite imagery for the years 2016 and 2021. Both datasets were pre-processed through atmospheric and sun-glint corrections, with SPOT-7 imagery further enhanced by water column correction using the Lyzenga algorithm. Supervised classification in ENVI was used to map four shallow-water habitat classes (coral reef, seagrass, sand, and rubble), and classification accuracy was assessed using confusion matrices to compute Producer Accuracy (PA), User Accuracy (UA), and Overall Accuracy (OA). Across both islands, coral reef and seagrass each occupied roughly one quarter of the mapped shallow-water area, while sand and rubble together accounted for about half of the area, indicating a heterogeneous mosaic of fringing reefs and seagrass meadows interspersed with sandy and rubbly substrates. Maps derived from SPOT-7 imagery produced higher overall accuracies (84% for Liang and 82% for Rakit) than those derived from Sentinel-2A imagery (76% and 70%, respectively). SPOT-7-based classifications consistently delineated larger and more detailed extents of coral reef and seagrass habitats, demonstrating the value of higher-resolution imagery for representing fine-scale features in optically complex coastal environments, whereas Sentinel-2A imagery was more suitable for broad-scale habitat discrimination but tended to underestimate narrow or fragmented patches.

KEYWORDS: coral reef, SPOT-7, Sentinel-2A, coastal habitat, Sumbawa

CITATION: Syamsul B. Agus, Ananda S. Maulana, Taslim Arifin, Yulius, Aprizon Putra, Muhammad Ramdhan, Aida Heriati, Rinny Rahmania, Joko Prihantono, Agus Sufyan, Eva Mustikasari, Ruzkiah Asaf, Admi Athirah, Hafidzuddin Fihrin (2026). Mapping Of Coral Reef Habitats Using Spot-7 And Sentinel-2 Satellite Imagery In Liang And Rakit Islands, Saleh Bay, Indonesia. Geography, Environment, Sustainability, 1 (19), 6-16
https://doi.org/10.24057/2071-9388-2026-3811

ACKNOWLEDGEMENTS: The authors gratefully acknowledge the National Research and Innovation Agency (BRIN) for technical assistance, support, and access to relevant facilities during the study process. We also extend our sincere thanks to the Sumbawa Regency Government for their cooperation and facilitation during field activities. Special appreciation is given to the RIIM G-2/2022 Program, the LPDP Grant from BRIN for funding support under grant number 82/II.7/HK/2022. Their contributions were instrumental in the successful completion of this study.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Saleh Bay, located in West Nusa Tenggara Province, is a marine body situated between the Sumbawa and Dompu Regencies (Kusumawati et al. 2019). Liang and Rakit Islands are among several small islands within Saleh Bay. Mapping coral reef habitats provides essential data to guide regional development and ensure sustainable exploitation of these resources. However, studies on these specific islands remain limited. Saleh Bay's waters host a diverse array of marine and coastal organisms and feature favorable environmental parameters such as cool water temperatures, high salinity, strong currents, high light penetration, neutral pH levels, and elevated concentrations of nutrients notably phosphorus (PO_4) and nitrogen (NO_3) (Yulius et al. 2019; Rahman et al. 2025). Liang and Rakit Islands were selected as study sites due to the lack of comprehensive studies on their coral reef ecosystems, which are vital for maintaining marine biodiversity and ecological sustainability.

Conducting field assessments and analyzing coral reef conditions in these areas can generate scientifically grounded recommendations not only for sustainable marine tourism development but also to support local community empowerment and inform national coastal management policies. For instance, coral reef habitat assessments in Raja Ampat, West Papua, have contributed to the establishment of community-based marine protected areas that balance ecological conservation and tourism (Atmodjo et al. 2020). Similarly, integrated reef monitoring in Bunaken National Park, located north of Manado City, North Sulawesi has informed adaptive management strategies that align with both biodiversity objectives and socioeconomic benefits for local communities (Farhum et al. 2021). These examples underscore how site-specific ecological research, such as that conducted in Saleh Bay, can provide evidence-based inputs for multi-level governance and long-term sustainability planning.

Marine study and exploitation activities are ongoing in Saleh Bay (Azzahra et al. 2025). However, due to coral reef degradation in some areas, including the waters surrounding Moyo, Medang, Rakit, and Ganteng Islands, Saleh Bay is increasingly recognized as a priority area for habitat restoration and ecological enrichment. In light of the long-term consequences of ecologically harmful fishing techniques, reassessing the condition of coral reef environments has become a pressing concern. One of the key goals of this study is to provide a comprehensive inventory of coral reef resources in the area (Edrus et al. 2017), which includes geo-referenced habitat maps, species composition data, live coral cover percentage, substrate types, and reef health indicators such as bleaching level and presence of macroalgae.

The presence of 28 ornamental coral genera in Saleh Bay, including *Euphyllia glabrescens, Euphyllia cristata, Echinopora sp., Goniopora sp.,* and *Lobophyllia sp.,* all of which are commercially valuable species, demonstrates a more even distribution in the mid-shore zones compared to other locations (Herdianti et al. 2025). Overall, the condition of coral reefs in Saleh Bay is considered good, with average live coral cover measured at 53.4%, based on in-situ benthic surveys using Line Intercept Transect (LIT) methods across 12 sampling stations. Better reef health tends to be observed further offshore, likely due to reduced sedimentation. Nevertheless, destructive fishing methods, such as blast fishing and chemical poisons, are still reported in certain locations and continue to compromise coral resilience and recovery (Johan et al. 2018). Accordingly, satellite remote sensing has become an indispensable tool for large-scale and repeated assessment of coral reef conditions, enabling objective detection of spatial and temporal ecological changes. Satellite imagery offers the advantage of wide spatial coverage and frequent revisit times, enabling efficient monitoring and timely detection of changes in reef conditions. Moreover, this method is more cost-effective than extensive field surveys, particularly in large or remote marine areas, making it an effective tool for sustainable coral reef management (Ji et al. 2024). It also facilitates habitat classification and change detection with improved precision, supporting early warnings, restoration planning, and marine spatial management in ecologically sensitive areas.

Computer-based image processing plays a critical role in satellite-based coral reef habitat mapping. Both pixel-based and object-based approaches are commonly applied in remote sensing data analysis. Pixel-based analysis involves classifying each pixel based on spectral values (e.g., color or intensity) without considering spatial relationships among neighboring pixels. This method is widely used for various remote sensing applications, including land use classification and coral reef identification. Its main advantage lies in its simplicity and suitability for high-resolution imagery. However, a key limitation is the lack of contextual spatial information, which can reduce classification accuracy compared to object-based methods that consider pixel groupings and spatial patterns (Agus et al. 2024). One algorithm commonly employed in pixel-based classification is the Lyzenga algorithm (Pratomo et al. 2024).

The Lyzenga algorithm is a well-established technique in satellite image analysis for estimating water depth in shallow marine environments, particularly for seafloor mapping, including coral reef habitats (Borfecchia et al. 2018). Originally developed by Donald Lyzenga in the 1980s, this method is frequently utilized in coral reef mapping based on optical satellite data (Irawan et al. 2017; Tanis & Lyzenga 1981). However, the primary motivation of the present study is not the algorithm itself but the need to better characterize coral reef habitats that are undergoing rapid ecological change in Saleh Bay, an under-studied region with high conservation value. In this context, the objective of this study is to evaluate and compare the spatial distribution and extent of coral reef habitats around Liang and Rakit Islands, as well as to assess the classification accuracy of Sentinel-2 Level-2A and SPOT-7 satellite imagery for mapping shallow-water benthic habitats in this area.

This study offers a novelty by providing the first quantitative comparison of coral reef habitat classification accuracy between SPOT-7 and Sentinel-2A imagery in Saleh Bay, an underexplored region with high ecological potential. This finding advances coral reef mapping practices in turbid and spatially heterogeneous coastal environments and supports the integration of multi-resolution remote sensing for targeted reef conservation planning. It is important to note that the available imagery consists of a 2016 scene for Liang Island and a 2021 scene for Rakit Island. Therefore, this study does not perform a formal change-detection analysis at the same site but rather compares coral reef habitat mapping and classification accuracy between two representative islands and dates.

METHODS

Study Area

The study was conducted in the Saleh Bay region of West Nusa Tenggara, specifically on Rakit Island, located at 118.0°E, 8.4°S, and Liang Island, located at 117.4°E, 8.3°S. These islands were selected as the study sites due to limited existing research on coral reef habitats in this region, which are critical to the sustainability of marine ecosystems.

Additionally, the region has been identified as a potential site for coral reef ecosystem restoration. According to Arifin et al. (2025), temperature, salinity, and pH in Saleh Bay are interconnected parameters that reflect ecosystem stability, while nitrate (NO_3) and ammonia (NH_3) concentrations serve as indicators of elevated anthropogenic influence, all of which directly affect coral reef health by controlling growth, calcification, and susceptibility to bleaching in the bay. These environmental factors can significantly affect the distribution and extent of coral reef habitats, as well as marine biological productivity, both of which are essential for supporting socio-economic activities such as fisheries and coral restoration. For more details, see the map in Fig. 1 below.

Data

SPOT-7 Satellite Imagery

This study utilized SPOT-7 satellite imagery due to its high spatial resolution of 1.5 m in the panchromatic band and 6 m in the multispectral bands. However, the SPOT-7 imagery obtained for this study was not fully pre-processed, particularly with regard to sun glint and atmospheric correction. SPOT-7 imagery is typically available in several processing levels, including Level 1A, Level 1B, and Level 2A (Xie et al. 2008; Turker & Ozdarici 2011). At Level 1A, the imagery undergoes only basic radiometric correction, with no geometric or atmospheric adjustments. In contrast, Level 1B and Level 1C imagery includes geometric corrections but still lacks atmospheric correction, leaving elements such as haze or clouds unaddressed (Amran 2024). For advanced analytical purposes, particularly those requiring accurate surface reflectance data, further corrections must be performed manually. A commonly used method is Dark Object Subtraction (DOS), a straightforward atmospheric correction technique designed to mitigate the effects of atmospheric scattering in imagery. This correction process was performed using the Environment for Visualizing Images (ENVI) 5.3 software.

Study Workflow

The workflow of this study began with downloading the satellite imagery, followed by image mosaicking using ArcMap. Subsequently, atmospheric and sun glint corrections were applied using ENVI. The DOS was employed for atmospheric correction, and the Band Math tool was used to implement the sun glint correction equation. Next, the Lyzenga algorithm was applied in ENVI to correct for the water column, enhancing the spectral differentiation of benthic features. The corrected images were then classified using the classification workflow tool.

Following the classification process, the extent of coral reef habitat distribution was quantified using polygon area calculations within ArcMap, with the measurements extracted from the attribute table (Poursanidis et al. 2018). This study also utilized Sentinel-2A Level-2A satellite imagery, which has a coarser spatial resolution than SPOT-7. Sentinel-2 provides 13 spectral bands with spatial resolutions of 10 m for the visible and near-infrared bands, 20 m for the red-edge and short-wave infrared bands, and 60 m for the coastal aerosol, water-vapor, and cirrus bands. The Level-2A product has undergone systematic atmospheric correction, producing surface reflectance data that are suitable for direct analytical applications. Sentinel-2 imagery is available in several processing levels, such as Level-1C, which is geometrically and radiometrically corrected but still expressed as top-of-atmosphere reflectance, and Level-2A, which includes full atmospheric correction using surface-reflectance algorithms. The Level-2A product corrects for haze, aerosols, and other atmospheric interferences, significantly enhancing data quality and reliability (Flament et al. 2021). As a result, Sentinel-2A Level-2A imagery is widely used for detailed environmental assessments, including mapping coral reef habitat distribution. For more details, see Fig. 2 below.

Fig. 1. Map of the study area in Liang and Rakit Islands, Saleh Bay, Indonesia

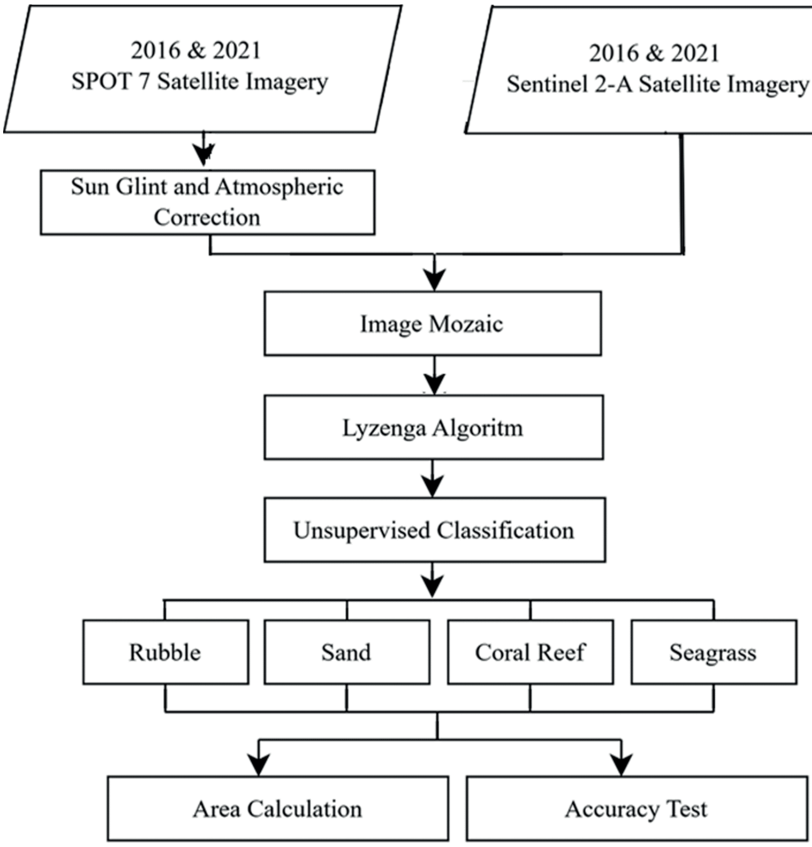


Fig. 2. Workflow of coral reef mapping using SPOT-7 and Sentinel-2A imagery

Atmospheric and Sun Glint Correction

Atmospheric correction was applied using the Dark Object Subtraction (DOS) method implemented in ENVI (Darweesh et al., 2021). Sun-glint was corrected using an empirical line-of-sight correction (ELSC) based on a linear regression between the near-infrared (NIR) band and the visible bands (Hedley et al., 2005; Bernardo et al., 2017; Simarmata et al., 2024), as , as expressed in Eq. (1) (Hedley et al. 2005; Bernardo et al. 2017):

$$R'_i = R_i - b_i\left(R_{NIR} - MinNIR\right) \qquad (1)$$

where R'_i is the deglinted (corrected) pixel value in band *i*; R_i is the original reflectance value in visible band *i*; b_i is the slope of the linear regression between band *i* and the *NIR* band; R_{NIR} is the *NIR* band reflectance for the same pixel; and *MinNIR* is the minimum *NIR* value within the selected deep-water sample used to estimate sun glint.

Image Mosaic

Image mosaicking is a key image processing technique used to merge multiple satellite images into a composite image, offering a more comprehensive spatial view of the study area (Febriandi et al. 2025). Prior to mosaicking, images must be geometrically corrected to ensure spatial accuracy within the same coordinate system. Radiometric adjustments are performed to harmonize color and brightness across tiles, avoiding visible seams (Putra et al. 2026). This process was conducted using Esri ArcGIS Desktop software (ArcMap), which provides a dedicated toolbox for mosaicking and raster processing.

Lyzenga Algorithm

The purpose of water column correction is to enhance satellite imagery by reducing optical interference caused by varying water depths, thereby making submerged objects more distinguishable (Panjaitan et al. 2025). This correction uses the Depth Invariant Index (DII), formulated through the Lyzenga algorithm (Tanis & Lyzenga 1981), expressed as Eq. (2):

$$Y_{ij} = ln\left(L_i\right) - \left[\left(k_i/k_j\right) ln\left(L_j\right)\right] \qquad (2)$$

where L_i is the reflectance value of spectral band *i*; L_j is the reflectance value of spectral band *j*; k_i and k_j are the attenuation coefficients for bands *i* and *j*, respectively.

The attenuation ratio (k_i/k_j) is derived from Eqs. (3) and (4):

$$\frac{k_i}{k_j} = a + \sqrt{a^2} + 1 \qquad (3)$$

$$a = \frac{\left(\sigma_{ii} - \sigma_{jj}\right)}{\left(2\sigma_{ij}\right)} \qquad (4)$$

where σ_{ii} is variance or variant of channel *i*; σ_{jj} is variance or variance of channel *j*; σ_{ij} is the covariance between bands *i* and *j*.

Supervised Classification

Supervised classification is an image-processing technique in which pixels are assigned to classes based on spectral signatures derived from training samples with known labels (Indey et al. 2024). In this study, a supervised classification approach was implemented using the Classification Workflow tool in ENVI to map four benthic habitat classes, namely coral reef, seagrass, sand, and rubble. This approach has been widely applied for coral reef habitat mapping in recent remote-sensing studies (Suasti et al. 2020; Rahmaddeni et al. 2024). Regions of interest representing each class were delineated using field information and visual interpretation of the corrected imagery, and these regions were used to derive class signatures (Putra et al. 2023). The resulting supervised classifier was then applied to the atmospherically and water-column-corrected SPOT-7 and Sentinel-2A images to produce the final habitat maps.

Accuracy Test

An accuracy assessment was conducted to evaluate the effectiveness of the classification results in mapping shallow marine habitats. The confusion matrix (or error matrix) is a standard tool for assessing classification performance by comparing the predicted classes with reference data (Ilyas et al. 2020). The following accuracy metrics were computed, namely 1) Producer Accuracy (PA): the proportion of correctly classified reference pixels; 2) User Accuracy (UA): the reliability of the classified map from the user's perspective; and 3) Overall Accuracy (OA): the total proportion of correctly classified pixels as expressed in Eqs. (5) to (7).

$$PA_i = \frac{n_{ii}}{n_{i+}} \times 100\,\% \qquad (5)$$

$$UA_i = \frac{n_{ii}}{n_{i+}} \times 100\,\% \qquad (6)$$

$$OA = \frac{\sum_{i=1}^{k} n_{ii}}{n} \times 100\,\% \qquad (7)$$

where k is the number of classes in the confusion matrix, n is the total number of reference samples, n_{ii} is the number of correctly classified samples for class i (the i–th diagonal element); n_{i+} is the marginal row total for class i (reference samples of class i); and n_{i+} is the marginal column total for class i (samples classified as class i).

RESULTS

Classification accuracy assessment

To evaluate the classification performance of coral reef habitat mapping, an accuracy assessment was conducted using satellite imagery from Sentinel-2A Level-2A and SPOT-7 for Liang Island in 2016. The assessment was based on a confusion matrix that compares the classified satellite image with reference data from Google Earth. Three key metrics were used to measure classification performance: UA, PA, and OA. These metrics indicate how well each habitat class was identified by the classification algorithm. For more details, see Tables 1 and 2 below.

Based on Table 1 above, the Sentinel-2A Level-2A classification for Liang Island achieved an OA of 76%. The Coral Reef class recorded the highest UA of 85.71%, indicating strong agreement between the classified and reference data. The Sand class had the highest PA of 81.81%, suggesting the classifier successfully captured the actual sand features on the ground. However, the Seagrass class showed a relatively low UA of 66.66%, indicating higher misclassification in that class. In contrast, Table 2 shows the classification accuracy results from SPOT-7 imagery. This classification achieved a higher OA of 84%. The Sand class recorded the highest UA of 91.66%, while the Seagrass class achieved the highest PA of 90%. These results highlight SPOT-7's superior performance in detecting sand and seagrass features. The comparative analysis of both tables above indicates that SPOT-7 imagery outperformed Sentinel-2A Level-2A in terms of classification accuracy for coral reef-related habitats on Liang Island. The higher spatial resolution of SPOT-7 (up to 1.5 meters) enhances its capability to capture detailed habitat boundaries,

Table 1. Sentinel Level 2-A classification accuracy test at Liang Island in 2016

Sentinel-2A Level-2A 2016 Liang Island							
		Rubble	Sand	Coral Reef	Seagrass	Total	UA (%)
Google earth reference	Rubble	9	2	2	0	13	69.23
	Sand	1	9	0	1	11	81.81
	Coral Reef	1	0	12	1	14	85.71
	Seagrass	0	2	2	8	12	66.66
	Total	11	13	16	10	50	
PA (%)		81.81	69.23	75	80		
OA (%)		76					

Table 2. SPOT-7 classification accuracy test at Liang Island in 2016

SPOT-7 2016 Liang Island							
		Rubble	Sand	Coral Reef	Seagrass	Total	UA (%)
Google earth reference	Rubble	10	1	1	0	12	83.33
	Sand	1	11	0	0	12	91.66
	Coral Reef	1	0	12	1	14	85.71
	Seagrass	0	2	1	9	12	75
	Total	12	14	14	10	50	
PA (%)		83.33	78.57	85.71	90		
OA (%)		84					

leading to improved accuracy especially in detecting sand and seagrass areas, which often require finer spatial differentiation. While both datasets performed well in identifying coral reef habitats, Sentinel-2A Level-2A showed limitations in discriminating seagrass, likely due to lower resolution (10–20 m) and mixed pixel effects. The lower user accuracy for seagrass in Sentinel imagery suggests that pixels classified as seagrass often included other substrates like sand or rubble. Therefore, SPOT-7 imagery is more suitable for detailed habitat classification, especially when precision is critical for marine resource management and conservation planning. Similar accuracy assessment was performed for Rakit Island in 2021, using both Sentinel-2A Level-2A and SPOT-7 imagery. For more details, see Tables 3 and 4 below.

Based on Table 3 above, the Sentinel-2A Level-2A classification for Rakit Island achieved an OA of 70%. The highest UA was found in the rubble class of 75%, while the highest PA was recorded in the coral reef class at 83.33%. However, the seagrass class showed a notably low UA of 64.28%, indicating classification difficulty in this category. In contrast, Table 4 shows that SPOT-7 imagery again demonstrated superior classification performance, achieving an overall accuracy of 82%. The rubble class maintained the highest UA of 84.61%, and the seagrass class achieved the highest PA of 88.88%, reflecting strong detection consistency for both benthic habitats. The comparative results from Tables 1 to 4 clearly show the consistent superiority of SPOT-7 imagery over Sentinel-2A Level-2A in classifying benthic habitats, particularly sand, seagrass, and coral reef areas. In both spatial contexts, Liang Island (2016) and Rakit Island (2021), SPOT-7 yielded higher OA (84% and 82%, respectively), compared to Sentinel-2A Level-2A (76% and 70%). SPOT-7's enhanced accuracy is largely attributable to its higher spatial resolution (1.5–6 m), which enables better separation of narrow or spectrally similar benthic features. This was evident in its ability to classify seagrass beds and coral reefs more effectively, as indicated by its consistently higher PA and UA values across both sites. Meanwhile, Sentinel-2A Level-2A, with its coarser resolution (10–20 m), struggled especially with the seagrass category, where spectral confusion with adjacent sandy or rubble substrates led to lower classification reliability. Despite offering broad coverage and corrected reflectance data, its utility appears to be more appropriate for larger-scale or generalized habitat mapping rather than fine-scale coastal ecosystem assessments. In conclusion, the combined accuracy assessment from all four tables confirms that SPOT-7 is a more reliable source for detailed mapping of coastal and shallow marine habitats, making it the preferred option for conservation planning, habitat monitoring, and environmental management in spatially heterogeneous reef environments (Agus et al. 2018).

Spatial Distribution of Coral Reef Habitats

To provide a clear spatial understanding of benthic ecosystem conditions around Liang Island, a comparative analysis of coral reef habitat distribution was conducted using two satellite imagery sources from 2016, namely Sentinel-2A and SPOT-7. These products were selected to evaluate the effectiveness of each in detecting the extent and diversity of shallow marine habitats, particularly coral reefs, seagrass beds, sand, and rubble. This comparison also serves to assess the influence of image resolution and correction techniques on classification accuracy in complex coastal environments. For more details, see the map in Fig. 3 below.

Table 3. Sentinel-2A Level-2A classification accuracy test on Rakit Island in 2021

Sentinel-2A Level-2A 2021 Rakit Island							
		Rubble	Sand	Coral Reef	Seagrass	Total	UA (%)
Google earth reference	Rubble	9	2	1	0	12	75
	Sand	1	7	0	2	10	70
	Coral Reef	1	2	10	1	14	71.42
	Seagrass	1	3	1	9	14	64.28
	Total	12	14	12	12	50	
PA (%)		81.81	75	50	83.33	75	
OA (%)		76	70				

Table 4. SPOT-7 classification accuracy test on Rakit Island in 2021

Sentinel-2A Level-2A 2021 Rakit Island							
		Rubble	Sand	Coral Reef	Seagrass	Total	UA (%)
Google earth reference	Rubble	11	1	1	0	13	84.61
	Sand	1	10	1	0	12	83.33
	Coral Reef	1	0	12	1	14	85.71
	Seagrass	0	2	1	8	11	72.72
	Total	13	13	15	9	50	
PA (%)		81.81	84.61	76.92	80	88.88	
OA (%)		76	82				

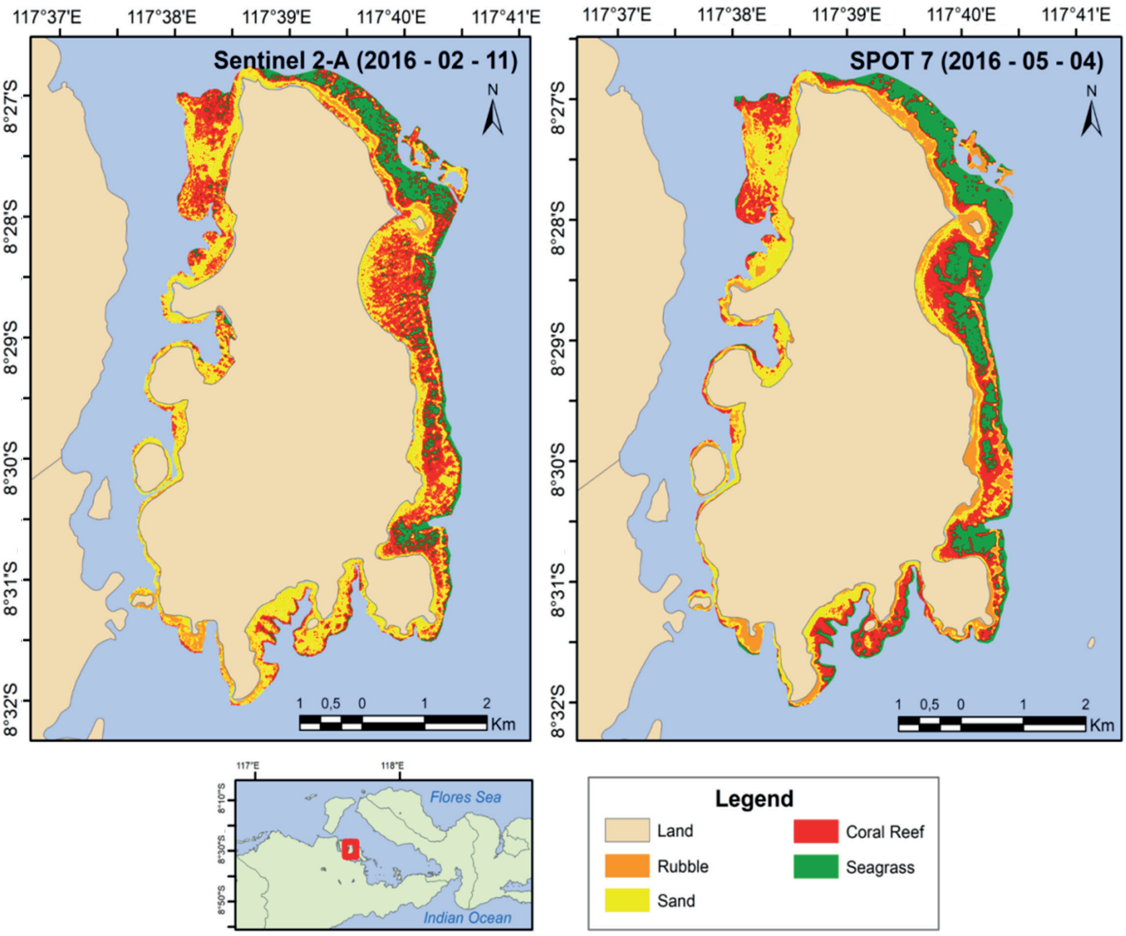


Fig. 3. Spatial distribution of coral reef habitats in Liang Island in 2016 based on Sentinel-2A (left panel) and SPOT-7 (right panel) satellite imagery

Based on the map in Fig. 3 above, a comparison of benthic habitat distribution around Liang Island in 2016, derived from Sentinel-2A imagery (left panel) and SPOT-7 imagery (right panel), shows classification accuracy and habitat detail across coral reef, seagrass, sand, and rubble zones. Both images reveal similar coastal ecosystem patterns, indicating general agreement in habitat identification. However, key differences in spatial precision are evident. The Sentinel-2A image shows broader coral reef coverage, especially along the northern and eastern shores, possibly due to its spectral sensitivity to large-area reflectance. In contrast, SPOT-7's sharper resolution allows for more detailed mapping in complex habitats, as seen in the clearer depiction of seagrass beds along the southern and eastern coasts, which are less visible in Sentinel-2A. SPOT-7's ability to capture smaller and patchier features makes it more effective for mapping fine-scale ecosystems, such as narrow seagrass corridors or fragmented beds. Rubble zones, typically linked to reef degradation, are more concentrated and clearly defined in SPOT-7, supporting accurate identification of disturbed areas. Such detail is vital for restoration planning and reef health monitoring. These findings are consistent with the work of Fakan et al. (2025), who demonstrated that high-resolution satellite imagery, such as SPOT-7, significantly improves the detection of degraded reef structures and enhances the delineation of coastal habitat transitions. While Sentinel-2A is suited for broader regional assessments, SPOT-7 performs better for local-scale analyses where habitat complexity and degradation are management priorities. The datasets are thus complementary: Sentinel-2A for regional trends and SPOT-7 for detailed habitat mapping and conservation planning.

This contrast is further evident in the 2021 classification results over Rakit Island (Fig. 4). Sentinel-2A offers wide regional coverage but lacks spatial detail, while SPOT-7 delivers clearer delineation of reef, seagrass, and rubble features, highlighting its utility for fine-scale habitat monitoring and degradation detection. For more details, see the map in Fig. 4 below.

Based on the map in Fig. 4 above, a comparison of benthic habitat distribution around Rakit Island in 2016, derived from Sentinel-2A imagery (left panel) and SPOT-7 imagery (right panel), is presented. Similar to Fig. 3, both maps delineate four habitat types, namely coral reef, seagrass, sand, and rubble, enabling a visual evaluation of classification accuracy and habitat detail between the two satellite datasets. Both images reveal similar coastal ecosystem patterns, indicating general agreement in identifying habitat zones. In the Sentinel-2A image, the distribution of benthic habitats appears more fragmented and diffusely scattered along the coastline. Although the image provides broad regional coverage and captures seagrass and sand areas relatively well, the delineation of coral reefs in several locations is less distinct, particularly in areas with complex substrate transitions or narrow reef formations. This limitation reflects the moderate spatial resolution of Sentinel-2A, which may not fully capture fine-scale ecosystem variability. By contrast, SPOT-7 imagery offers a more precise and detailed depiction of coastal habitats. Coral reefs and seagrass beds appear more clearly defined and spatially concentrated in certain sections of the coastline, particularly along the eastern and southern fringes. The higher spatial resolution of SPOT-7 allows for enhanced detection of smaller and more heterogeneous features, such as fragmented seagrass patches or narrow reef corridors, which are often overlooked in coarser imagery. Additionally, rubble zones, often associated with

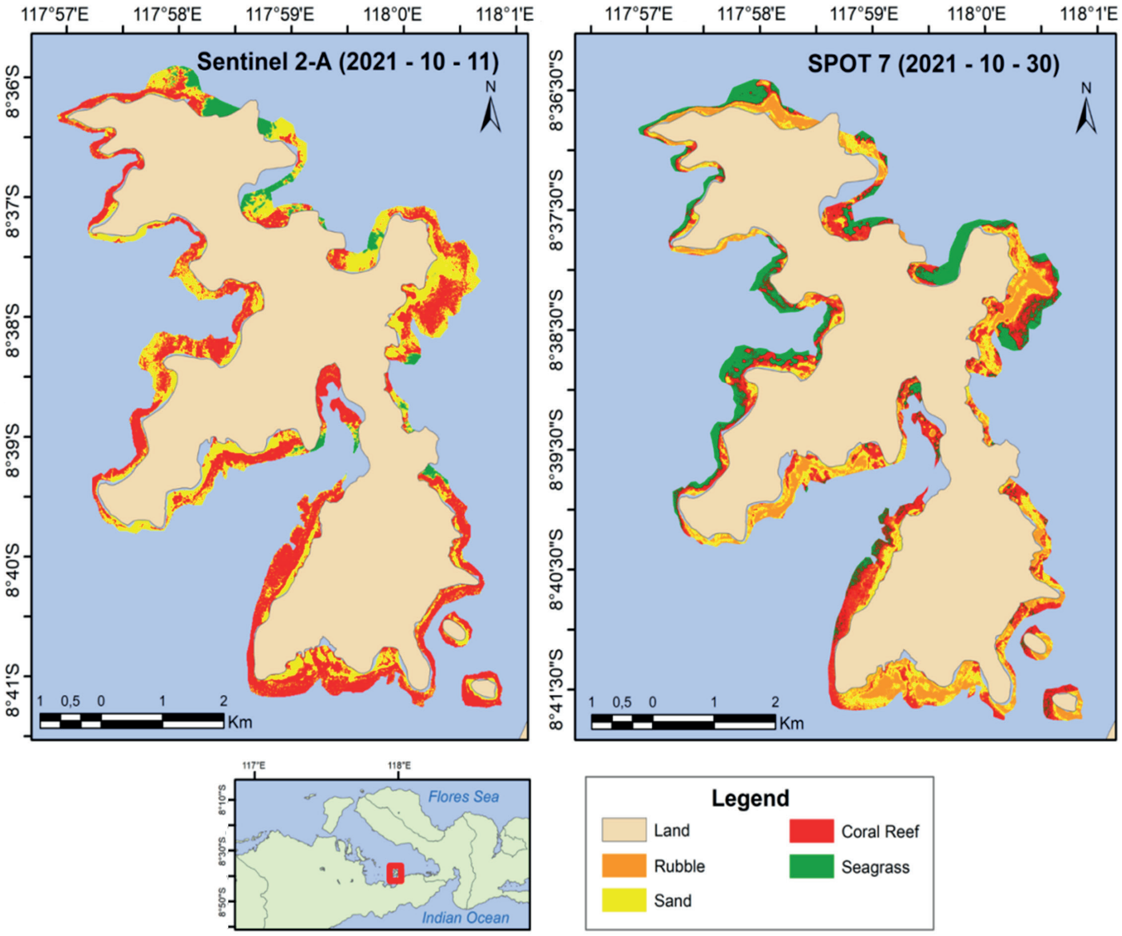


Fig. 4. Spatial distribution of coral reef habitats in Rakit Island in 2021 based on Sentinel-2A (left panel) and SPOT-7 (right panel) satellite imagery

coral degradation or disturbance, are more distinct and localized in the SPOT-7 image. This enhanced detectability is important for identifying areas impacted by natural changes or anthropogenic activities, such as coastal development, destructive fishing, or sedimentation (Putra et al. 2023). The ability to capture such subtle topographic and ecological variations highlights the value of high-resolution imagery for targeted reef monitoring and restoration.

The comparison between Sentinel-2A and SPOT-7 in this context clearly illustrates how differences in resolution and image acquisition techniques can significantly influence habitat mapping outcomes. While Sentinel-2A remains useful for regional-scale monitoring, it may lack the spatial detail required for precise habitat delineation. SPOT-7, on the other hand, offers sharper imagery that enhances accuracy in detecting ecosystem boundaries and degradation zones. These findings support the study by Cruz et al. (2024), who emphasized that combining medium- and high-resolution satellite imagery improves the accuracy and reliability of coastal habitat assessments. This approach is particularly relevant for conservation planning and monitoring ecosystem responses to climate change and human pressures. The integration of Sentinel-2A and SPOT-7 datasets provides complementary benefits, balancing regional coverage and spatial detail to enhance the effectiveness of coastal ecosystem monitoring and management.

Quantitative Comparison of Habitat Classification from Sentinel-2A and SPOT-7 Satellite Imagery

To support the visual interpretations of habitat distribution presented in the previous figures (Figs. 3 and 4), Table 5 and Table 6 provide a detailed quantitative assessment of the classified benthic habitat areas derived from Sentinel-2A Level-2A and SPOT-7 satellite imagery. These tables represent the area (in hectares/ha) and percentage coverage of four dominant coastal ecosystem classes, namely rubble, sand, coral reef, and seagrass, as mapped in two different years and locations. Table 5 shows data from Liang Island in 2016, while Table 6 shows similar classification results for Rakit Island in 2021.

By comparing the area values derived from Sentinel-2A and SPOT-7, these tables highlight the degree of agreement and difference between medium- and high-resolution satellite imagery in mapping the spatial extent of each habitat class. The data also reveal consistent patterns, with coral reef and seagrass generally occupying the largest proportion, followed by rubble and sand. These two tables allow a cross-site comparison of habitat classification accuracy between Liang Island (2016) and Rakit Island (2021), thus strengthening the evaluation of satellite image performance in varying ecological and geomorphological settings. For more details, see Tables 5 and 6 below.

These classification results from Liang Island in 2016 (Table 5) are further complemented by similar analyses conducted for Rakit Island in 2021, as shown in Table 6 below, allowing spatial comparisons between two different coral reef ecosystems within Saleh Bay that were observed in different years.

Based on Tables 5 and 6 (also Fig. 5) above, a comparison was made between habitat classifications for Liang Island (2016) and Rakit Island (2021) using Sentinel-2A Level-2A and SPOT-7 imagery. Each dataset reflects the spatial extent (in hectares/ha and percentages) of four key benthic habitat types: rubble, sand, coral reefs, and seagrass. This comparison offers insight into how spatial resolution affects habitat delineation outcomes across both locations. In the case of Liang Island, Sentinel-2A estimated rubble at 1.14

Table 5. Area of classification class in Liang Island Sentinel-2A Level-2A and SPOT-7 2016

No	Class	Level-2A Sentinel Area (ha)	Percentage Sentinel-2A Level-2A (%)	SPOT-7 area (ha)	Percentage SPOT-7 (%)
1	Rubble	1.14	24.8	1.19	25.1
2	Sand	1.12	24.3	1.16	24.5
3	Coral Reef	1.21	26.3	1.19	25.1
4	Seagrass	1.13	24.6	1.2	25.3

Table 6. Area of classification class on Rakit Island Sentinel-2A Level-2A and SPOT-7 2021

No	Class	Level-2A Sentinel Area (ha)	Percentage Sentinel-2A Level-2A (%)	SPOT-7 area (ha)	Percentage SPOT-7 (%)
1	Rubble	1.13	25.6	1.192	25.2
2	Sand	1.19	27.0	1.193	25.2
3	Coral Reef	1.18	26.8	1.197	25.3
4	Seagrass	0.91	20.6	1.15	24.3

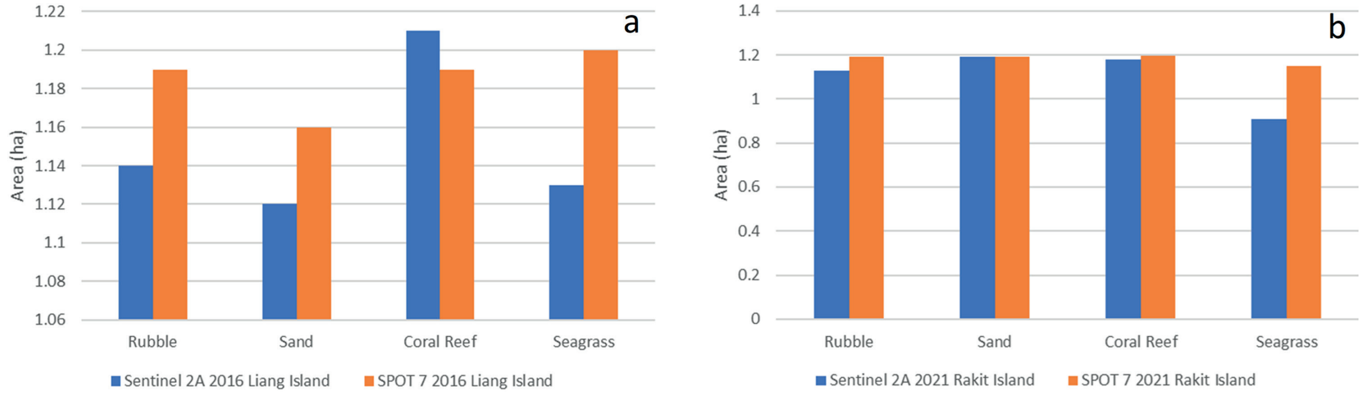


Fig. 5. Class area differences between Sentinel-2A and SPOT-7 habitat classifications for Liang Island (2016, panel a) and Rakit Island (2021, panel b)

ha (24.8%), sand at 1.12 ha (24.3%), coral reefs at 1.21 ha (26.3%), and seagrass at 1.13 ha (24.6%). Meanwhile, SPOT-7 recorded slightly higher values for rubble (1.19 ha), sand (1.16 ha), and seagrass (1.20 ha), with coral reefs slightly lower at 1.19 ha. These marginal differences suggest that SPOT-7's finer resolution enables better discrimination of habitat boundaries, especially for features that are spatially limited or irregular in shape. In the Rakit Island 2021 dataset, discrepancies became more distinct. Sentinel-2A identified seagrass at only 0.91 ha (20.6%), compared to 1.15 ha (24.3%) in the SPOT-7 image. Coral reef, sand, and rubble classes also showed minor differences, but the most pronounced gap remained in seagrass detection. This suggests that Sentinel-2A's resolution may be insufficient to accurately capture narrow, fragmented, or submerged vegetation, particularly in optically complex coastal environments.

Fig. 5 visually supports these findings through side-by-side bar graphs, reinforcing that SPOT-7 consistently maps larger areas of seagrass and coral reefs at both sites. These results point to the superior ability of high-resolution imagery to detect fine-scale spatial patterns and benthic complexity, which are often underestimated by medium-resolution sensors. This outcome aligns with the findings of Eugenio et al. (2017), who emphasized the advantages of high-resolution data for improving benthic ecosystem classification. Further supporting this, Caras et al. (2017) highlighted the importance of fine spatial resolution in capturing narrow reef features that are frequently missed by coarser sensors. Similarly, Purkis et al. (2018) noted that SPOT-7 is more effective at identifying transitional substrates and degraded reef zones, which are crucial indicators of ecological condition. The integration of multiple spatial resolutions has also been shown to enhance classification performance. For example, Hedley et al. (2018) demonstrated that combining Sentinel-2A with higher-resolution imagery improves mapping accuracy in heterogeneous benthic environments. DeVries et al. (2017) further reported that medium-resolution sensors tend to underestimate submerged vegetation in turbid or shallow waters. Meanwhile, Lyons et al. (2014) concluded that classification error rates significantly decrease when imagery resolution is below 5 meters, a condition satisfied by SPOT-7. These results affirm the value of SPOT-7 for detailed, site-specific habitat mapping, particularly in areas where precision is essential for coastal management, reef restoration, and conservation zoning. When used in tandem with broader-coverage platforms like Sentinel-2A, high-resolution imagery becomes a vital component of a multi-scale remote sensing strategy, yielding more comprehensive and reliable information for benthic habitat monitoring and spatial planning.

CONCLUSIONS

This study confirms that SPOT-7 satellite imagery offers significant advantages over Sentinel-2A in the classification and spatial delineation of coral reef habitats, particularly in capturing seagrass beds, coral structures, and rubble zones with greater clarity. The comparative analysis conducted for Liang Island (2016) and Rakit Island (2021) revealed consistent patterns in which SPOT-7 detected larger and more fragmented habitat patches, reflecting its ability to resolve fine-scale ecological features often missed by coarser-resolution sensors. Although Sentinel-2A provided broader regional coverage and slightly higher estimates for sand areas, it tended to underrepresent seagrass and coral reef extents, especially in areas with complex benthic configurations. In contrast, SPOT-7 consistently produced more accurate and spatially detailed classifications, with overall classification accuracies of 84% for Liang Island and 82% for Rakit Island, underscoring its superiority in habitat boundary detection and degradation assessment. These results highlight the essential role of high-resolution satellite imagery in supporting ecosystem-based management, particularly for the monitoring, restoration, and spatial planning of coral reef environments. The integration of high-resolution data such as SPOT-7 into routine monitoring frameworks can substantially enhance the quality of information available to decision-makers, enabling more effective conservation interventions and adaptive coastal governance. ■

REFERENCES

Agus F, Jatiwibowo R, Kamil ZA and Gifari OI. (2024). Prediksi luas vegetasi Kota Samarinda-Kalimantan Timur menggunakan metode NDVI dan klasifikasi terbimbing. Journal of Tropical Forests, 12(1), 31–39. DOI: 10.20527/jht.v12i1.19021

Agus SB, Aziizah NN, Subarno T and Sunudin A. (2018). The utilization of Spot-7 imagery for mapping seagrass distribution on intertidal zone and water depth estimation in Wawonii Island. Journal of Tropical Marine Science and Technology, 10(1), 197–207. DOI: 10.29244/jitkt.v10i1.19119

Amran MA. (2024). Dasar-dasar penginderaan jauh satelit. Nas Media Library.

Arifin T, Yulius, Ramdhan M, Johan O, Gunawan DP, Heriati A, Mustikasari E, Rahmania R, Purbani D, Dillenia I, Agus SB, Asaf R, Hatta M, Marzuki I, Akhwady R, Sufyan A and Fihrin H. (2025). Mariculture site selection based on water quality, cultivation and ecological condition. Global Journal of Environmental Science and Management, 11(1), 1–18. DOI: 10.22034/gjesm.2025.01

Atmodjo E, Lamers M and Mol APJ. (2020). Governing dynamics in marine conservation tourism in Raja Ampat, Indonesia. Tourism Planning & Development, 17(6), 655–673. DOI: 10.1080/21568316.2019.1686652

Azzahra RY, Yulius, Agus SB, Aditama FA, Arifin T, Putra A, Heriati A, Purbani D, Prihantono J, Salim HL, Ramdhan M, Rahmania R, Wahyono A, Ningsih A, Saad S and Asyiri A. (2025). Advancing mangrove carbon stock monitoring with drone and Pleiades imagery on Gilibakau Island–Saleh Bay, Indonesia. Journal of Marine and Island Cultures, 14(1), 88–101. DOI: 10.21463/jmic.2025.14.1.06

Bernardo N, Watanabe F, Rodrigues T and Alcântara E. (2017). Atmospheric correction issues for retrieving total suspended matter concentrations in inland waters using OLI/Landsat-8 image. Advances in Space Research, 59(9), 2335–2348. DOI: 10.1016/j.asr.2017.02.017

Borfecchia F, Consalvi N, Micheli C, Carli FM, Cognetti De Martiis S, Gnisci V, Piermattei V, Belmonte A, De Cecco L, Bonamano S and Marcelli M. (2018). Landsat 8 OLI satellite data for mapping of the Posidonia oceanica and benthic habitats of coastal ecosystems. International Journal of Remote Sensing, 39(20), 6715–6732. DOI: 10.1080/01431161.2018.1528020

Caras T, Hedley J and Karnieli A. (2017). Implications of sensor design for coral reef detection: Upscaling ground hyperspectral imagery in spatial and spectral scales. International Journal of Applied Earth Observation and Geoinformation, 63, 68–77. DOI: 10.1016/j.jag.2017.07.009

Cruz C, Perrin PM and Martin JR. (2024). Mapping of temperate upland habitats using high-resolution satellite imagery and machine learning. Environmental Monitoring and Assessment, 196, Article 869. DOI: 10.1007/s10661-024-12998-0

Darweesh KF, Hellal AM, Saber SA, EL-Kafrawy SB, Abdelsalam AA and EL-Naggar HA. (2021). Three decades of remote sensing habitat mapping, along the Egyptian coast of Aqaba Gulf, Red Sea. Egyptian Journal of Aquatic Biology & Fisheries, 25(6), 467 – 487

DeVries B, Huang C, Lang MW, Jones JW, Huang W, Creed IF and Carroll ML. (2017). Automated quantification of surface water inundation in wetlands using optical satellite imagery. Remote Sensing, 9(8), 807. DOI: 10.3390/rs9080807

Edrus IN, Arief S and Setyawan E. (2017). Kondisi kesehatan terumbu karang Teluk Saleh, Sumbawa: Tinjauan aspek substrat dasar terumbu dan keanekaragaman ikan karang. Indonesian Fisheries Research Journal, 16(2), 147–161. DOI: 10.15578/jppi.16.2.2010.147-161

Eugenio F, Marcello J, Martín J and Rodríguez-Esparragón D. (2017). Benthic habitat mapping using multispectral high-resolution imagery: Evaluation of shallow water atmospheric correction techniques. Sensors, 17(11), 2639. DOI: 10.3390/s17112639

Farhum MF, Jompa J, Restu M and Salman D. (2021). Management effectiveness of the Bunaken National Park, Indonesia. AACL Bioflux, 14(2), 1089-1097.

Fakan EP, Dubuc A, Hemingson CR, McCormick MI and Hoey AS. (2025). Habitat degradation has species-specific effects on the stress response of coral reef fishes. Journal of Experimental Marine Biology and Ecology, 582, 152070. DOI: 10.1016/j.jembe.2024.152070

Febriandi, Fatimah S, Triyatno, Hermon D, Putra A, Mutmainah H, Arifin T and Akhwady R. (2025). Predicting of land cover changes until 2030 and assessing sustainability status in the Mandeh region, Indonesia. Geographia Technica, 20(1), 263–280. DOI: 10.21163/GT_2025.201.18

Flament T, Trapon D, Lacour A, Dabas A, Ehlers F, Huber D and et al. (2021). Aeolus L2A aerosol optical properties product: Standard correct algorithm and Mie correct algorithm. Atmospheric Measurement Techniques, 14(12), 7851–7871. DOI: 10.5194/amt-14-7851-2021

Herdianti TP, Yulius, Agus SB, Arifin T, Putra A, Prihantono J, Heriati A, Hartati ST, Akhwady R, Suryono DD, Rahmania R, Ramdhan M, Hilmawan A, Ningsih A, Saad S, Mudjiono M and Asyiri A. (2025). Habitat assessment of whale sharks (Rhincodon typus) in Saleh Bay, Indonesia: Linking chlorophyll-a and sea surface temperature using Aqua MODIS data. Geographia Technica, 20(2), 52–64. DOI: 10.21163/GT_2025.202.04

Hedley J D, Harborne A R and Mumby P J. (2005). Technical note: Simple and robust removal of sun glint for mapping shallow-water benthos. International Journal of Remote Sensing, 26(10), 2107–2112. DOI: 10.1080/01431160500034086.

Hedley JD, Roelfsema C, Brando V, Giardino C, Kutser T, Phinn S, Mumby PJ, Barrilero O, Laporte J and Koetz B. (2018). Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8. Remote Sensing of Environment, 216, 598–614. DOI: 10.1016/j.rse.2018.07.014

Ilyas TP, Nababan B, Maduppa H and Kushardono D. (2020). Seagrass ecosystem mapping with and without water column correction in Pajenekang Island Waters, South Sulawesi. Jurnal Ilmu dan Teknologi Kelautan Tropis, 12(1), 9–23. DOI: 10.29244/jitkt.v12i1.26598

Indey JFI and Supangat S. (2024). Implementasi algoritma Naïve Bayes dalam sistem pengarsipan surat berbasis AI di GPI Papua Klasis Mimika Papua Tengah. Jurnal Ilmiah Informatika, 12(02), 102–113. DOI: 10.33884/jif.v12i02.9087

Irawan J, Sasmito B and Suprayogi A. (2017). Pemetaan sebaran terumbu karang dengan metode algoritma Lyzenga secara temporal menggunakan citra Landsat 5 7 dan 8 (Studi Kasus: Pulau Karimunjawa). Jurnal Geodesi Undip, 6(2), 56–61. DOI: 10.14710/jgundip.2017.16256

Ji X, Yang B, Wei Z, Wang M, Tang Q, Xu W, Wang Y, Zhang J and Zhang L. (2024). Benthic habitat sediments mapping in coral reef area using amalgamation of multi-source and multi-modal remote sensing data. Remote Sensing of Environment, 304, 114032. DOI: 10.1016/j.rse.2024.114032

Johan O, Yulius, Salim HL, Ardi I, Abrar M and Daulat A. (2019). The existence of ornamental coral in different live coral coverage conditions in Saleh Bay, West Nusa Tenggara. Segara Journal, 15(2), 99–108. DOI: 10.15578/segara.v15i2.6592

Kusumawati E, Susilo SB, Agus SB, Arifin T and Yulius. (2019). Analisis penentuan sebaran konsentrasi klorofil-a dan produktivitas primer di perairan Teluk Saleh menggunakan citra satelit Landsat OLI 8. Jurnal Pengelolaan Sumberdaya Alam dan Lingkungan, 9(3), 671–679. DOI: 10.29244/jpsl.9.3.671-679

Lyons MB, Murray NJ, Kennedy EV, Kovacs EM, Castro-Sanguino C, Phinn SR, Borrego Acevedo R, Ordóñez Alvarez A, Say C, Tudman P, Markey K, Roe M, Canto RF, Fox HE, Bambic B, Lieb Z, Asner GP, Martin PM, Knapp DE, Li J, Skone M, Goldenberg E, Larsen K and Roelfsema CM. (2024). New global area estimates for coral reefs from high-resolution mapping. Cell Reports Sustainability, 1(2), 100015. DOI: 10.1016/j.crsus.2024.100015

Panjaitan JP, Mahendratama G and Siregar V. (2025). Benthic habitats mapping in Pari Island, Kepulauan Seribu, Indonesia using drone and Sentinel-2B imagery with object based image analysis method. BIO Web of Conferences, 168, 05004. DOI: 10.1051/bioconf/202516805004

Poursanidis D, Topouzelis K and Chrysoulakis N. (2018). Mapping coastal marine habitats and delineating the deep limits of the Neptune's seagrass meadows using very high resolution Earth observation data. International Journal of Remote Sensing, 39(23), 8670–8687. DOI: 10.1080/01431161.2018.1490974

Pratomo DG, Cahyadi MN, Hariyanto IH, Syariz MA and Putri SA. (2024). Lyzenga algorithm for shallow water mapping using multispectral Sentinel-2 imageries in Gili Noko waters. BIO Web of Conferences, 89, 07006. DOI: 10.1051/bioconf/20248907006

Purkis SJ. (2018). Remote sensing tropical coral reefs: The view from above. Annual Review of Marine Science, 10, 149–168. DOI: 10.1146/annurev-marine-121916-063249

Putra A, Triyatno, Darwis R, Dewata I, Tanto TA, Mustapha MA, Razi P and Zainul R. (2023). Suitability of mangrove ecosystems as a protected zone based on land-use changes. Physical Oceanography, 30(6), 866–881.

Putra A, Faristyawan R, Triyatno T, Putra BG, Prasetyo A, Yulius, Ameliya Z, Arifin T, Razi P and Triarjunet R. (2026). Land cover changes (2002–2022) using Landsat imagery and pixel-based classification: Supporting coastal conservation strategies in Tanjung Mutiara. Springer Proceedings in Earth and Environmental Sciences, Part F1158, 751–763. DOI: 10.1007/978-981-95-3075-5_57

Rahman AM, Purbani D, Agus SB, Arifin T, Yulius, Putra A, Heriati A, Prihantono J, Akhwady R, Sufyan A, Mustikasari E, Purnamaningtyas SE, Rahmania R, Saad S and Tjahjo DWH. (2025). Hydrodynamic modeling of pollutant distributions from the Bera watershed and its impact to the coastal area of Saleh Bay, West Nusa Tenggara, Indonesia. Geographia Technica, 20(2), 31–51. DOI: 10.21163/GT_2025.202.03

Rahmaddeni R, Wulandari D, Renova M, Ramadhan G and Sari R. (2024). Machine learning. Padang: Serasi Media.

Simarmata N, Wikantika K, Tarigan TA, Aldyansyah M, Tohir RK and Putra A. (2024). Comparative analysis of reflectance values on Sentinel-2A image with field spectroradiometer in mangrove forest on East Coast Lampung. AIP Conference Proceedings, 3001(1), 090001. DOI: 10.1063/5.0184159

Suasti Y, Prarikeslan W, Syah N, Triyatno and Putra A. (2020). A mapping of changes in coral reefs condition based on development the marine ecotourism in the southern part coast of Padang City – Indonesia. International Journal of GEOMATE, 19(76), 157–164. DOI: 10.21660/2020.76.ICGeo4

Tanis FJ and Lyzenga DR. (1981). Development of Great Lakes algorithms for the Nimbus-G Coastal Zone Color Scanner: Phase I – Final technical report (NASA-CR-173511). NASA/Lewis Research Center and Environmental Research Institute of Michigan.

Turker M and Ozdarici A. (2011). Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: A comparison study. International Journal of Remote Sensing, 32(24), 9735–9768. DOI: 10.1080/01431161.2011.576710

Xie Y, Sha Z and Yu M. (2008). Remote sensing imagery in vegetation mapping: A review. Journal of Plant Ecology, 1(1), 9–23. DOI: 10.1093/jpe/rtm005

Yulius, Ramdhan M, Prihantono J, Priyambodo DG, Saepuloh D, Salim HL, Rizaki I and Zahara RI. (2019). Budidaya rumput laut dan pengelolaannya di pesisir Kabupaten Dompu, Provinsi Nusa Tenggara Barat berdasarkan analisa kesesuaian lahan dan daya dukung lingkungan. Jurnal Segara, 15(1), 19–30. DOI: 10.15578/segara.v15i1.7429

RESEARCH PAPER

Check for updates

CC BY

ECOSYSTEM SERVICES AND THE TEMPORAL-SPATIAL DYNAMICS OF LAND USE IN THE PEDRAS RIVER BASIN, SOUTH OF BRAZIL

Pedro Almeida[1*], Leandro Redin Vestena[1]
[1]Universidade Estadual do Centro-Oeste (UNICENTRO), Ln. Élio A. Dalla Vecchia, Guarapuava, 85040-167, Brazil
***Corresponding author:** 1912306@unicentro.edu.br

Received: February 6th 2025 / Accepted: July 11th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3879

ABSTRACT. Changes in land use in Brazil, particularly agricultural expansion, have significant impacts on ecosystems and their associated services. This study aimed to evaluate the temporal and spatial dynamics of ecosystem services (ES) in the Pedras River basin (PRB), the main water supply for Guarapuava, southern Brazil, between 2002, 2018 and 2023. The analysis was based on the Ecosystem Services Identification Matrix, integrating land use and land cover (LULC) maps and field observations of environmental and socioeconomic characteristics in the area. This approach allowed the classification and assessment of the potential supply of provisioning, regulating, and cultural ES in the basin. The findings between 2002 and 2018 revealed a general decline in most regulating and cultural services, with notable reductions in genetic resources, local climate regulation, air quality regulation services, and medicinal resources. Conversely, provisioning services related to food provision, energy provision, and soil quality regulation showed increases. Between 2018 and 2023, food provision, water quality regulation, soil quality, and pollination declined, along with decreases in cultural services. In contrast, the increase occurred in water supply, raw materials, genetic resources, ornamental resources (18.16%), and air quality and climate regulation services. Agricultural expansion resulted in a trade-off, increasing food provision while simultaneously degrading other ES. The study emphasizes that sustainable land management is necessary to preserve ecosystem functions and services, thereby preventing the degradation of their potential.

KEYWORDS: spatial analysis, sustainable management, agricultural expansion, trade-off dynamics

CITATION: Almeida P., Vestena L. R. (2026). Ecosystem Services And The Temporal-Spatial Dynamics Of Land Use In The Pedras River Basin, South Of Brazil. Geography, Environment, Sustainability, 1 (19), 17-28
https://doi.org/10.24057/2071-9388-2026-3879

ACKNOWLEDGEMENTS: The authors would like to thank the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for granting a scholarship to the authors, the Araucária Foundation for its financial support, and the Guarapuava City Hall for the data provided. The authors would also like to thank the Writing Center (CERTA – Centro de Escrita, Revisão e Tradução Acadêmica – www3.unicentro.br/centrodeescritaacademica) of the Midwestern State University of Paraná (UNICENTRO) for assistance with English language translation and developmental editing.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

The PRB is the water supply area for the municipality of Guarapuava in the state of Paraná, Brazil. In the PRB, terrestrial ecosystems are associated with the Mixed Ombrophilous Forest (MOF) and aquatic ecosystems with river courses and lakes. They are directly related to land use and occupation since, for example, the removal of vegetation cover for cultivation alters hydrological and biological processes, affecting the balance of the ecosystem. In the PRB, anthropogenic activities cause the degradation of natural habitats, loss of biodiversity, and intensification of erosion processes that affect soil fertility and deteriorate water quality.

The expansion in recent years, mainly of agricultural activities involving soybean plantation, has been requiring new land for planting, causing significant changes in land use and cover to the detriment of areas with tree and shrub vegetation, which occupied 73.37% of the areas in the PRB in 2002 (Vestena et al. 2004). The use of agrochemicals, particularly glyphosate, is associated with intensive agriculture and has been detected in the water sourced from the PRB, often exceeding the safe tolerable limit set by the European Union (Bombardi 2017; Hess 2018; Aggio et al. 2021; Piassetta et al. 2021).

Other anthropogenic actions in the basin that are harmful to the ecosystem include the disposal and release of industrial effluents associated with agricultural activities

[1]Aranha, A. and Rocha, L. (2019). "Coquetel" com 27 agrotóxicos foi achado na água de 1 em cada 4 municípios. Repórter Brasil, Agência Pública. Available at: https://reporterbrasil.org.br/2019/04/coquetel-com-27-agrotoxicos-foi-achado-naagua-de-1-em-cada-4-municipios/ [Accessed 23 Jun. 2024].

(Santos and Kobiyama 2003; Teodoro and Santos 2009), intensive logging with the removal of tree vegetation and the intensification of erosion processes (Teodoro and Santos, 2009); the presence of sediment in the water of river courses from crops, roads, and paths (Cunha 2011; Cunha et al. 2013; Cunha et al. 2014; Cunha and Thomaz 2015).

Therefore, land use is a considerable human alteration to the earth's system that modifies the structure and functioning of ecosystems (Vitousek et al. 1997). This process directly threatens ecosystems' capacity to provide ES (Brito et al. 2016).

ES are increasingly used to analyze the interactions between humans and nature, as they represent the contributions of ecosystems to human well-being, resulting from the interaction of biotic and abiotic processes (Costanza et al. 1997; MEA 2005; Benabou et al. 2022). This research adopts the classification proposed by the Millennium Ecosystem Assessment (MEA), which comprises four types of services: provisioning, regulating, cultural, and supporting. As in Githiora-Murimi et al. (2022), supporting services, considered fundamental for the supply of other service categories, were excluded from the analysis to avoid double counting.

The increasing popularity of the concept of ES has led to the development of several assessments and mappings for different contexts and scales (Martínez-Harms and Balvanera 2012). In Brazil, the dissemination of the ES concept is connected to public policy strategies outlined in the National Policy for Payment for Environmental Services (PNPSA), which regulates mechanisms that promote the adoption of conservation-oriented agricultural practices, the restoration and preservation of natural permanent preservation and natural legal reserve areas, as well as the conservation of remnants of native vegetation (Brasil 2021).

In this context, watersheds are the territorial unit established by Brazilian legislation for the planning and management of water resources (Brasil 1997). Additionally, watersheds or basins provide multiple ES by ensuring a reliable water supply. For this reason, they are frequently used as spatial units for the identification and mapping studies in ES thematic studies, as in the research of Chaves et al. (2021) in the Rio São José, Pernambuco; Hinata and Basso (2022) in the Rio Passo Fundo, Rio Grande do Sul; Sampaio and Bacani (2025) in the Rio Bebedouro, Mato Grosso do Sul; and Simões et al. (2022) in the Rio Paraíba do Sul, São Paulo.

This research aimed to evaluate the temporal and spatial dynamics of ES potential in the PRB, specifically focusing on the years 2002, 2018, and 2023, through the qualitative mapping of potential ES.

MATERIALS AND METHODS

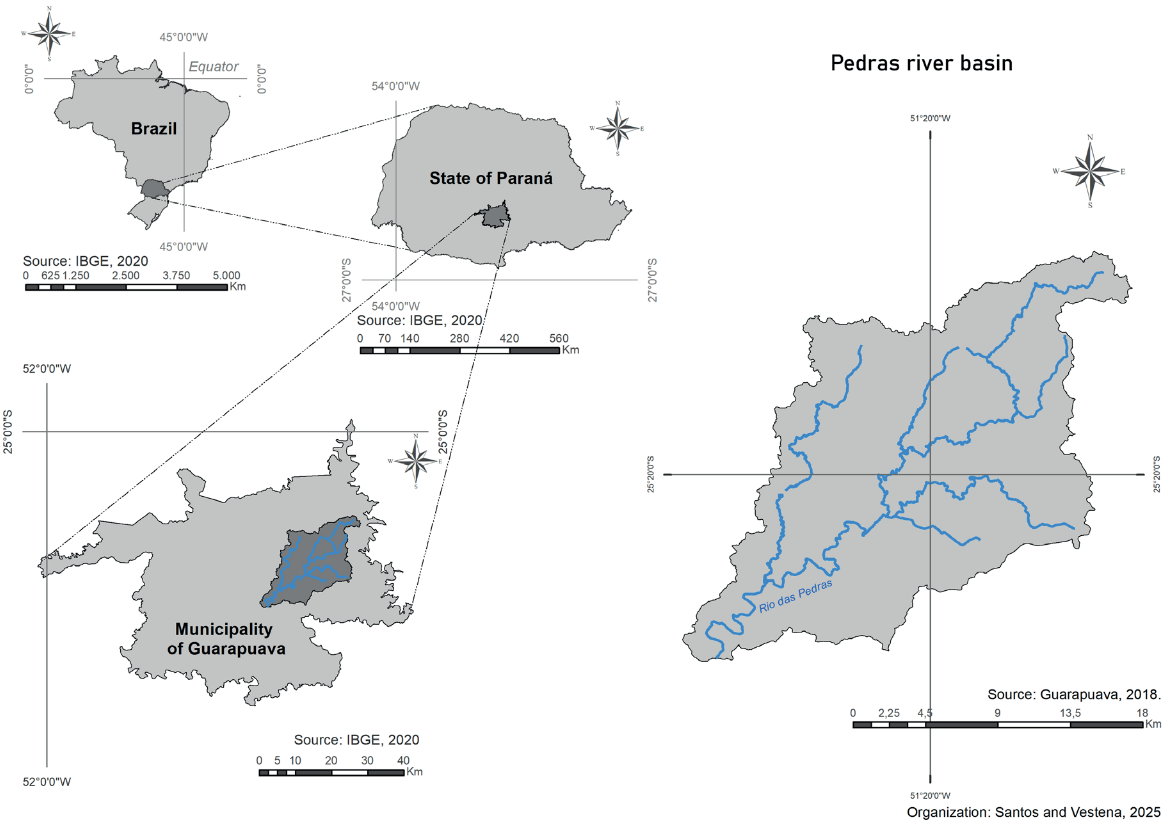
Study area

The PRB has been a water source area for Guarapuava since July 1st, 1974. The basin has an area of 330.42 km^2 and is located between the coordinates 25°17′55″ S 51°16′10″ W and 25°28′0″ S 51°26′15″ W (Fig. 1). The Pedras River is a tributary of the Jordão River, a right-bank tributary of the Iguaçu River, which flows into the Paraná River.

Methodological procedures

The methodological framework for this study was structured in steps, as illustrated in the flowchart (Fig. 2): 1) acquisition and organization of LULC mapping supplemented by field surveys; 2) identification of relevant ES in the area; 3) scoring the ES potential from each LULC type using a matrix; and 4) mapping the ES potential for three periods, 2002, 2018, and 2023.

Research in the field of ES can be either valuation studies, which assign monetary values to ES in various land cover types (Costanza et al. 1997), or socio-cultural studies with methods that can evaluate social values of ES and link them to land cover maps (Iniesta-Arandia et al. 2014).



Source: IBGE, 2020

Source: IBGE, 2020

Source: IBGE, 2020

Source: Guarapuava, 2018.

Organization: Santos and Vestena, 2025

Fig. 1. Study area: the PRB

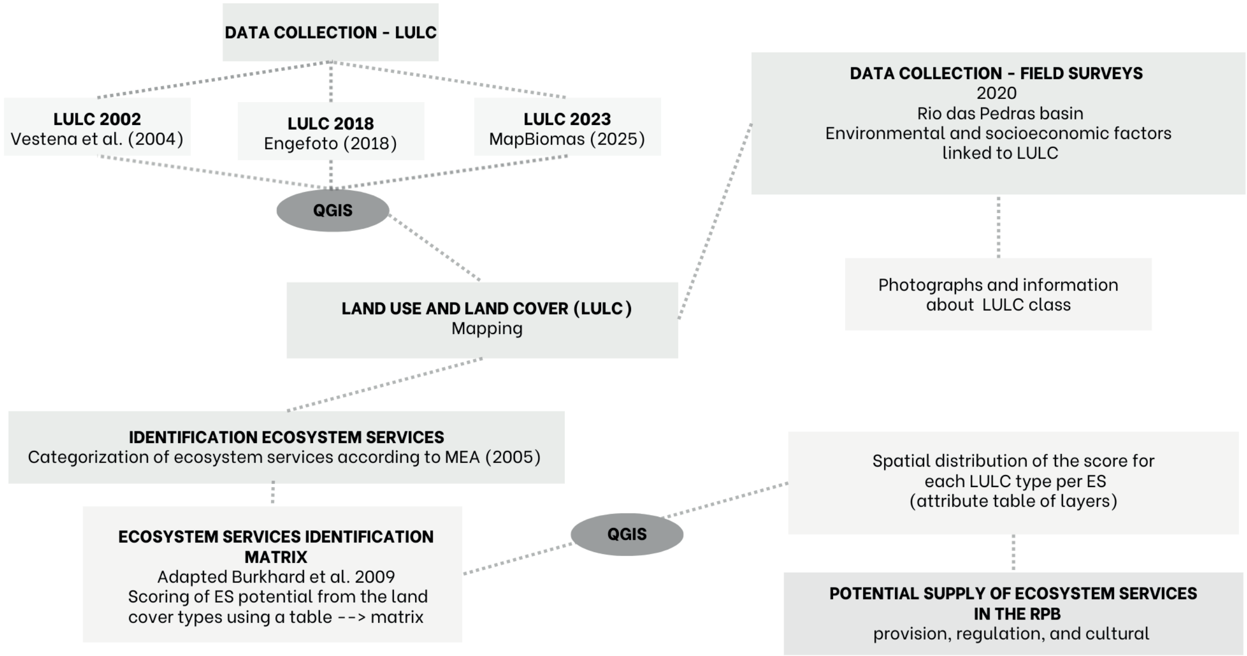


Fig. 2. Methodological flowchart

This study adopted the matrix model developed by Burkhard et al. (2009; 2017), as applied in studies such as Githiora-Murimi et al. (2022) in Kenya and Sohel et al. (2015) in Bangladesh. In this model, the LULC is used as a proxy for ES supply and scored in a table using information from multiple sources combined with a judgment of scores, and the resulting table quantifies the capacity of each class of use to provide ES (Burkhard et al., 2009).

LULC for 2002 (Fig. 3a) were mapped by Vestena et al. (2004), based on a visual interpretation of orthophotos at 1:30,000 scale from an aerial survey conducted in September 2002 (Engefoto; Guarapuava 2003). LULC mapping for 2018 (Fig. 3b) was generated through supervised automatic interpretation and classification of orthorectified images acquired between January and June 2018 by Engefoto (Engefoto; Guarapuava 2018). The classification, symbols, and colors used in the maps are in accordance with the Technical Manual of Land Use of the Brazilian Institute of Geography and Statistics (IBGE, 2013). LULC mapping for 2023 (Fig. 3c) was mapped by MapBiomas at a 1:30,000 scale from collection 9, using annual satellite images (Landsat, Sentinel, and MODIS) processed in Google Earth Engine. The method is based on pixel-by-pixel supervised classification using the Random Forest algorithm, calibrated with reference samples and time series of spectral indices. The process includes mosaicking, temporal filtering, accuracy validation, and thematic integration steps (MapBiomas 2025). In addition, we made a reclassification to merge and simplify similar classes of farming uses (temporary and perennial crops).

Field surveys conducted in 2020 aimed to identify environmental and socioeconomic characteristics associated with different LULC classes of the PRB. This information was crucial for validating LULC classification and informing the ES scoring process.

Ecosystem service scores were tailored to the specific classification detail of the 2002, 2018, and 2023 LULC datasets to mitigate discrepancies in data sources. This approach involved underestimating potential values where necessary to prevent the overrepresentation of functions, thereby respecting the internal variability of each class over time. The assessment of ES was performed based on functionally equivalent LULC groups, using LULC as proxies for ES provision, as proposed by Burkhard et al. (2009; 2017).

The qualitative mapping of LULC provides an important basis for identifying ES using the Ecosystem Services Identification Matrix. The matrix structure with an X and Y axis was constructed for each year analyzed, linking the LULC classes mapped in the area (Y axis) with the ES (X axis), which were categorized into provisioning, regulating, and cultural services, following the MEA (2005) framework.

Once the matrix was built, each ES was assigned a score representing its potential supply, classified on a scale from 0 to 5. A value on the scale of "0" represents no relevant potential for ES in a given use class, "1" very low potential, "2" low potential, "3" medium potential, "4" high potential, and "5" maximum potential for providing ES (Burkhard and Maes 2017).

Scores were assigned based on field survey observations, like in Githiora-Murimi et al. (2022), and a synthesis of information from academic literature and government data about the characteristics of the PRB, described in IBGE, 2002; Vestena et al. 2004; Donatti et al. 2004; Vestena e Thomaz 2006; IBGE 2018; Paraná 2019; Favaro et al. 2020.

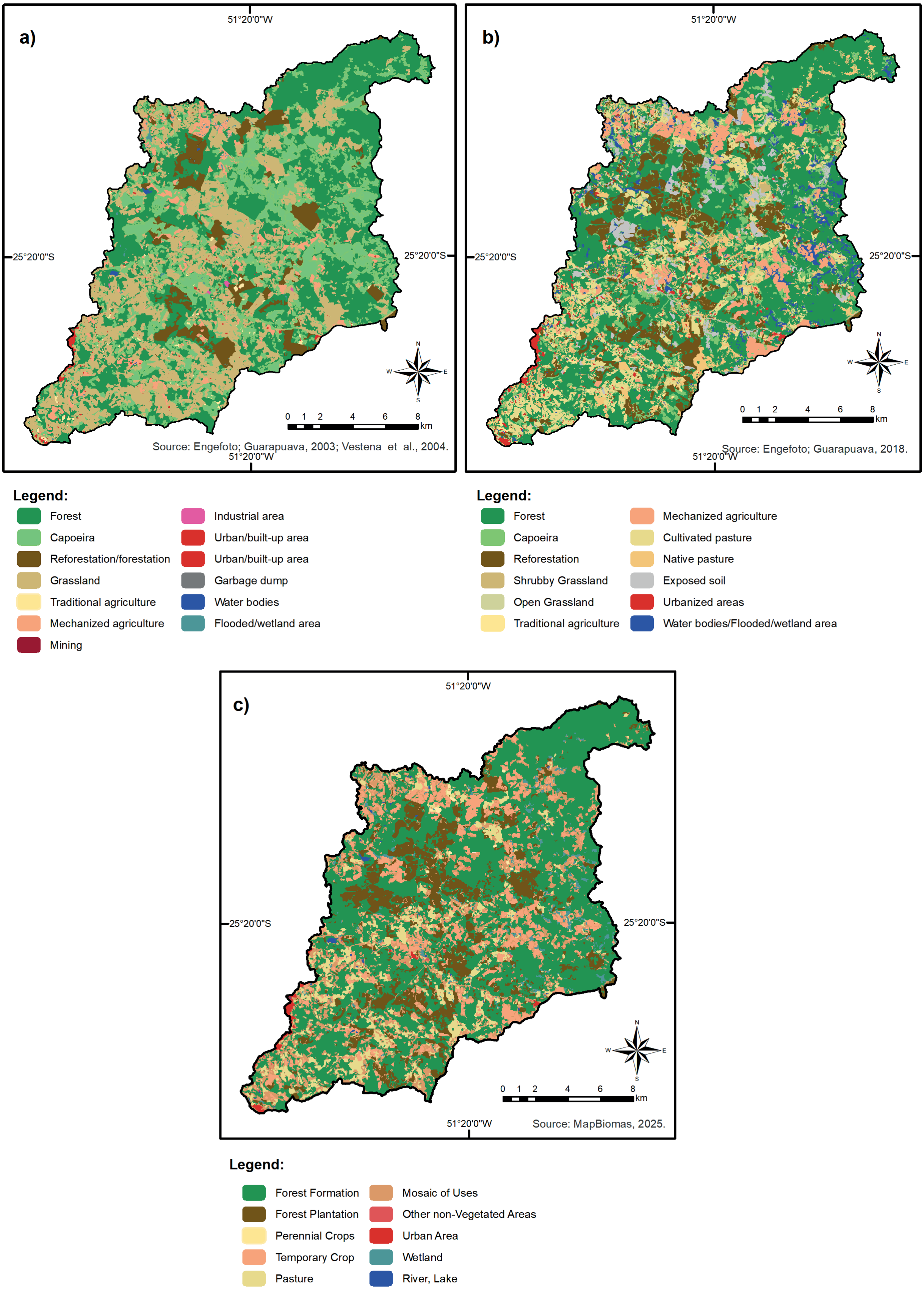
The spatial mapping of ES potential for provisioning, regulating, and cultural services was performed in QGIS. As proposed by Burkhard et al. (2012), the matrix scores were integrated into the LULC vector maps by joining them to the attribute table of each LULC polygon. For each period, 2002, 2018, and 2023, a new field was created for each individual ES, and the corresponding score (0-5) was assigned to each LULC class. This resulted in a set of maps illustrating the spatial distribution of the ES potential in the PRB for the three years.

RESULTS AND DISCUSSION

Ecosystem properties define the potential for providing ES (Bastian et al. 2012). The potential for ES can be obtained using different indicators, which can be ecological, social, and economic (Zhen and Routray 2003). These indicators were used to weight the potential values for ES.

The matrix with the estimated values of the relationship between land use, land cover, and potential ES in 2002, 2018, and 2023 is summarized in Tables 1, 2, and 3.

The scores established from the matrix were spatialized to generate summary maps for the years 2002, 2018, and 2023, representing the spatio-temporal dynamics of ES potential in the study area (Fig. 4).



Organization: Santos and Vestena, 2025

Fig. 3. LULC mapping of the PRB for (a) 2002, (b) 2018 and (c) 2023

Table 1. Matrix of ES by LULC classes (2002)

LULC*	Ecosystem Services															
	Provisioning							Regulation					Cultural			
	Food	Water Supply	Raw Materials	Genetic resources	Medicinal resources	Ornamental resources	Energy	Air quality regulation	Local climate regulation	Water quality regulation	Soil quality regulation	Pollination	Recreation and leisure	Educational value	Inspiration for art and design	Tourism
U1	4	5	4	5	5	5	4	5	5	5	5	5	4	5	5	4
U2	5	0	0	3	0	0	0	2	1	1	0	1	1	1	0	1
U3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U4	5	0	4	4	0	0	0	0	0	0	1	2	0	5	1	0
U5	0	4	0	4	4	0	0	4	4	4	4	5	3	4	3	0
U6	5	5	0	0	0	0	4	5	5	5	5	0	5	5	5	5
U7	0	0	0	0	0	5	0	0	0	0	0	0	0	5	0	0
U8	0	5	0	0	0	0	0	5	5	5	5	2	0	4	0	0
U9	5	0	0	4	4	0	0	4	4	4	4	5	4	5	0	5
U10	3	3	5	0	0	0	5	3	2	1	1	2	0	5	0	0
U11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

■ Not relevant potential ■ Low potential ■ High potential
■ Very low potential ■ Medium potential ■ Maximumpotential

*Land uses: U1 Forest; U2 Grassland; U3 Urban/built-up area; U4 Mechanized agriculture; U5 Capoeira[1]; U6 Water bodies; U7 Mining; U8 Flooded/wetland area; U9 Traditional agriculture; U10 Reforestation/forestation; U11 Industrial area; U12 Garbage dump. Organization: Authors (2021).

[1]Note: "Capoeira" is a Brazilian Portuguese term for secondary vegetation that grows after the suppression of primary forest, typically characterized by shrubs and young trees.

Table 2. Matrix of ES by LULC classes (2018)

LULC*	Ecosystem Services															
	Provisioning							Regulation					Cultural			
	Food	Water Supply	Raw Materials	Genetic resources	Medicinal resources	Ornamental resources	Energy	Air quality regulation	Local climate regulation	Water quality regulation	Soil quality regulation	Pollination	Recreation and leisure	Educational value	Inspiration for art and design	Tourism
U1	4	5	4	5	5	5	4	5	5	5	5	5	4	5	5	4
U2	0	0	0	0	0	0	0	0	0	2	1	1	0	1	0	0
U3	0	3	0	0	0	0	0	0	0	3	3	4	0	1	0	0
U4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U5	5	0	5	5	1	0	0	0	0	0	1	2	0	5	1	0
U6	5	0	0	5	5	0	5	3	3	4	3	5	4	5	2	5
U7	0	4	0	4	4	0	0	4	4	4	4	5	3	4	3	0
U8	5	0	0	0	0	0	0	0	0	1	1	1	0	4	0	0
U9	5	1	0	0	0	0	0	0	0	4	3	2	0	3	0	0
U10	3	0	5	0	0	0	5	4	4	3	3	4	0	5	1	0
U11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U12	1	5	0	0	1	0	0	5	5	5	5	3	0	4	0	0

■ Not relevant potential ■ Low potential ■ High potential
■ Very low potential ■ Medium potential ■ Maximumpotential

*Land uses: U1 Forest; U2 Open Grassland; U3 Shrubby Grassland; U4 Urbanized areas; U5 Mechanized agriculture; U6 Traditional agriculture; U7 Capoeira[1]; U8 Cultivated pasture; U9 Native pasture; U10 Reforestation; U11 Exposed soil; U12 Flooded/wetland area. Organization: Authors (2021).

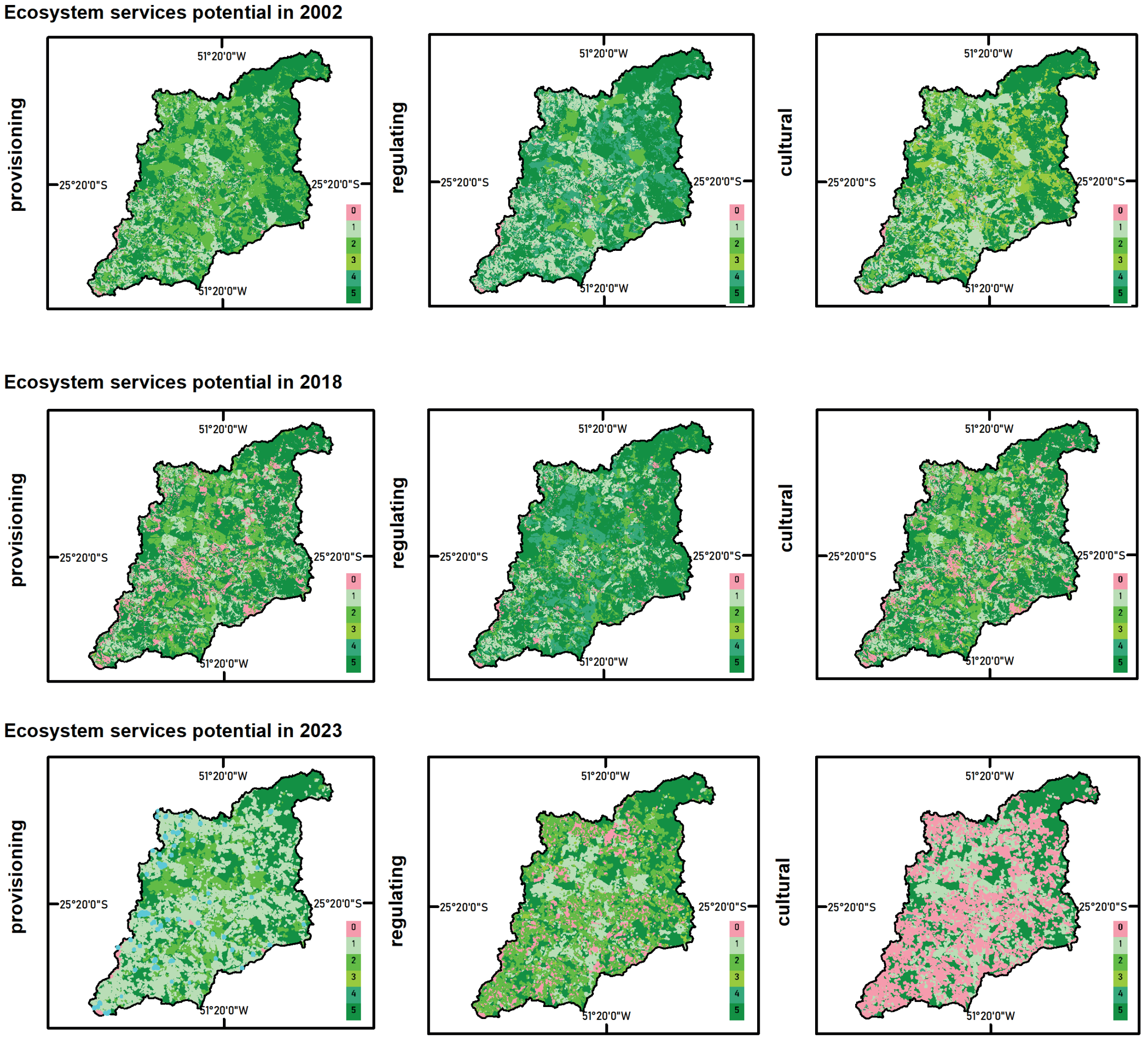
[1]Note: "Capoeira" is a Brazilian Portuguese term for secondary vegetation that grows after the suppression of primary forest, typically characterized by shrubs and young trees.

Table 3. Matrix of ES by LULC classes (2023)

LULC*	Ecosystem Services															
	Provisioning							Regulation					Cultural			
	Food	Water Supply	Raw Materials	Genetic resources	Medicinal resources	Ornamental resources	Energy	Air quality regulation	Local climate regulation	Water quality regulation	Soil quality regulation	Pollination	Recreation and leisure	Educational value	Inspiration for art and design	Tourism
U1	4	5	5	5	5	5	4	5	5	5	5	5	4	5	5	4
U2	2	2	3	0	0	1	4	2	2	0	0	0	0	4	0	0
U3	0	3	0	3	1	0	0	2	2	2	2	2	0	1	0	0
U4	0	0	3	0	0	4	0	0	0	0	0	0	0	0	0	0
U5	4	0	3	2	0	0	0	0	0	1	0	0	0	1	0	0
U6	4	0	3	2	4	0	4	2	2	2	2	3	4	4	4	4
U7	4	0	3	2	0	0	0	0	0	0	0	1	0	4	0	0
U8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U9	4	5	0	5	0	0	4	4	4	4	4	0	4	4	4	4
U10	0	4	0	0	0	0	0	4	4	4	4	0	0	4	0	0

🟥 Not relevant potential
🟩 Very low potential
🟩 Low potential
🟩 Medium potential
🟩 High potential
🟩 Maximumpotential

*Land uses: U1 Forest formation; U2 Forest plantation; U3 Mosaic of uses; U4 Other non-vegetated Areas; U5 Pasture; U6 Perennial Crop; U7 Temporary Crop; U8 Urban Area; U9 River and lakes; U10 Wetland. Organization: Authors (2025).


Ecosystem services potential in 2002

Ecosystem services potential in 2018

Ecosystem services potential in 2023

Organization: Santos and Vestena, 2025.

Fig. 4. Spatial distribution of ES potential (provisioning, regulating, and cultural) in the PRB for the years 2002, 2018 and 2023

The evolution of the dynamics of ES in the PRB in 2002 and 2018 (Table 3) showed significant decreases, especially in ES of genetic resources provision (-30.37) and medicinal resources (-3.86); the ES of air quality regulation (-22.16), local climate regulation (-22.16), water (-2.46) and pollination (-8.32); and decreases in all classes of cultural ES. It is noteworthy that there was a significant increase in the ES of soil quality regulation (14.27), and in the provision of food (6.04) and energy (5.24). Between 2018 and 2023 (Table 3), there was a decrease in the ES of food provision (-2.82), water quality regulation (-12.23), soil quality regulation (-11.40), and pollination (-13.36), as well as reductions in the cultural ES of inspiration for art and design (-17.12) and recreation and leisure (-1.02). In contrast, there was a significant increase in the ES of water supply (16.13), raw and other materials (17.67), genetic resources (24.86), ornamental resources (18.16), air quality regulation (13.82), and local climate regulation (13.82).

The increase in the potential for ES of *provision of food* in 2002 and 2018 is related to the increase in areas where animal husbandry, cultivation, and forest areas occur. Although the decrease in the potential for food provision in 2023 can be attributed to land use shifts, for instance, in several portions of the basin, traditional and diversified agricultural areas typically associated with higher food provision potential were replaced by mechanized monocultures or temporarily converted to other production systems. The production of commodities (soy, wheat and corn) in a mechanized agricultural area is part of a low-cost food system that does not promote food security and provides low monetary compensation for rural workers (Pereira et al. 2020).

It is important to emphasize that in animal husbandry and intensive agricultural production, there is a conflict situation or *trade-off* with regulatory services because in these areas there is the use of pesticides, chemical

Table 4. Dynamics of ES in the PRB (%) between 2002–2018 and 2018–2023

ES	ES class	Area (ha)* 2002	Percentage of total area of PRB 2002 (%) **	Area (ha) 2018*	Percentage of total PRB area in 2018 (%)**	Change in potential area 2002 – 2018 (%)	Area (ha)* 2023	Percentage of total PRB area in 2023 (%)**	Change in potential area 2018 – 2023 (%)
Provisioning	Food	263,58	79,7	283,53	85.81	6.04	274.19	82.99	-2.82
	Water supply	219.96	66.5	223.34	67.60	1.02	276.63	83.73	16.13
	Raw materials and other materials	203.86	61.7	215.35	65.18	3.48	273.74	82.85	17.67
	Genetic resources	295.5	89.4	195.17	59.07	-30.37	277.31	83.93	24.86
	Medicinal resources	219.47	66.4	206.71	62.56	-3.86	224.92	68.08	5.52
	Ornamental resources	157.14	47.5	162.16	49.08	1.52	222.17	67.24	18.16
	Energy	184.5	55.8	201.8	61.08	5.24	222.73	67.41	6.33
Regulation	Air quality regulation	304.29	92.1	231.08	69.94	-22.16	276.74	83.76	13.82
	Local climate regulation	304.29	92.1	231.08	69.94	-22.16	276.74	83.76	13.82
	Water quality regulation	304.29	92.1	296.16	89.64	-2.46	255.75	77.41	-12.23
	Soil quality regulation	249.01	75.3	296.16	89.64	14.27	258.51	78.24	-11.4
	Pollination	323.65	97.9	296.16	89.64	-8.32	252.03	76.28	-13.36
Cultural	Recreation and Leisure	276.14	83.5	180.76	54.71	-28.87	177.39	53.69	-1.02
	Educational values	325.07	98.3	296.16	89.64	-8.32	328.20	99.33	9.69
	Inspiration for art and design	238.56	72.20	233.95	70.81	-1.40	177.39	53.69	-17.12
	Tourism	216.83	65.63	163.02	49.34	-16.29	177.39	53.69	4.35

Note: *Area in hectares corresponding to the sum of the land use and cover classes relative to the potential for ES. **Percentage of ES concerning the total area of the PRB (330.42 ha).

fertilizers, and antibiotics, which leads to possible long-term declines in human well-being due to the alteration of the functioning or balance of the ecosystem (Foley et al. 2005). It is necessary to analyze strategies to reduce these conflicting situations in agricultural production.

The effects of intensive agriculture, which is constantly expanding, are harmful to human well-being and to the *food supply* in the future. In Brazil, it is already possible to see some of the effects of climate change on agriculture, such as reduced productivity due mainly to rising temperatures and extreme weather events such as precipitation and droughts, which are occurring with increasing intensity and frequency (Jägermeyr et al. 2021; IPCC 2021).

Traditional and subsistence agriculture, as well as organic food production practices, have decreased in the PRB. This decline is driven by public policies that mandate pesticide use as a condition for accessing agricultural credit (Favaro et al. 2020). In addition, squatters and owners of small properties face problems such as land concentration and conflicts over land, possession, use, and ownership through resistance and confrontation (CPT 2018).

Despite this, traditional agriculture has the greatest potential for *food supply*, as it is responsible for approximately 1/5 of the planet's food production. Traditional farmers interact directly with the benefits generated by the ecosystem and are aware of the importance of biodiversity, the preservation of springs and hillsides, Permanent Preservation Areas, and Legal Reserves, as they are more directly connected to the land and the causes and consequences of actions in ecosystems.

Agrosystems in general, when combined with agroecology, can provide sustainable food production, allowing for a synergistic relationship between other ES. Sustainable land management practices can help reconcile agricultural production with the maintenance of ES.

In the Brazilian context, sustainable land management practices have emerged, for example, integrated Crop–Livestock–Forestry (ICLF), no-till farming, and crop rotation. In the integrated ICLF systems, the productive arrangements combine agriculture, livestock, and forestry through intercropping, rotation, or succession. These systems optimize land use, enhance soil and water conservation, promote biodiversity, and contribute to carbon sequestration, being widely implemented in degraded and non-degraded pasturelands (Balbino et al. 2012; Skorupa and Manzatto 2019; Leite et al. 2023; Marchão et al. 2024). The no-till farming practice is based on permanent soil cover, absence of plowing, and diversification of cropping systems, including cover crops. It reduces erosion, improves soil quality, and increases resilience to climatic variability (Mello and Van Raij 2006), is integrated with the crop rotation, enhances physical, chemical, and biological soil conditions, diversifies production, and reduces both climatic and market risks (Franchini et al. 2011).

In forested areas, the potential food supply does not affect the other ESs due to the synergy between them, with food being extracted in a controlled manner, as is the case with the harvesting, sale, transportation, and storage of ripe pine nuts, which takes place with the permission of the IAT (Instituto Água e Terra – is the Water and Land Institute of Paraná State).[2]

The increase in areas with vegetation cover, which favors the *water supply*, directly helps regulate the river regime and water quality. However, the importance of sustainable practices in cultivated areas stands out, as they have vegetation cover and, even temporarily, can cause trade-offs with the ES.

The supply of *raw materials and other materials* has increased due to the expansion of reforestation/afforestation and forest areas, which largely provide this type of ES. As well as the potential for *energy supply*, which is primarily associated with vegetation. However, in these areas, there is a conflicting situation or trade-off, since intensive logging can affect the ES that regulates soil, air, and water quality due to the impacts of removing and transporting materials. Even if the structure of the forest is not altered by the exploitation of raw materials such as wood, there is degradation in the forest productivity, biomass, and biodiversity, as well as the introduction of pests and pathogens (Folley et al. 2005).

Despite the expansion of forest areas, the increase in 2023 in the ES of the *supply of genetic* and medicinal resources is linked to the total vegetated area and to the diversification of land-cover types that contribute to these functions. Early-successional mosaics, such as capoeira, can temporarily elevate the availability of genetic resources. In addition, the persistence of small patches of traditional agriculture and secondary vegetation maintains species used for medicinal purposes that tend to disappear under intensive mechanized agriculture. Thus, even with reductions in some capoeira areas, the structural diversification of the landscape and the presence of regenerating vegetation contributed to the increase in the potential supply of genetic and medicinal resources.

The *soil quality regulation* ES increased between 2002 and 2018 but declined in 2023. This shift is linked to the expansion of mechanized agriculture and pasturelands, introducing trade-off dynamics. While vegetation in these areas offers some protection against precipitation-driven erosion, other practices, such as intensive machinery use and animal trampling, accelerate soil disaggregation and erosion processes, ultimately reducing the system's regulating capacity.

The *local climate regulation* service had a decrease related to agricultural expansion in the PRB, reiterating that these activities are highly degrading and lead to a decrease in synergy between ES, mainly due to soil and water pollution, loss of biodiversity, soil erosion, decrease in carbon stocks, deforestation, and forest fragmentation (Vestena and Thomaz 2006; Pessoa et al. 2020; Polidoro et al. 2020; Kumi et al. 2021). Despite the overall degradation pressures, the *climate regulation service* shows an increase in 2023 due to the role of regenerating vegetation and soil-based carbon recovery processes. Even though capoeira areas were reduced compared to previous years, the remaining secondary vegetation patches, and the increase in forest formations continued to accumulate biomass and organic matter.

Vegetation and soil are responsible for most of the carbon storage and organic matter stocks, which contribute to climate regulation (Machado 2005; Uzeda 2020). The degrading activities' expansion reduces vegetation cover and increases impervious areas, affecting the regional climate through changes in net radiation and energy flow in the soil (which make up the energy balance in the interaction between soil and atmosphere) and the water balance (Foley et al. 2005).

In the years mapped, capoeira areas were reduced to a large extent and were replaced by mechanized agriculture. This fact makes it important to protect secondary

---

[2] IAT libera colheita e venda do pinhão em todo o Paraná, mas com exigência de maturação. Agencia Estadual de notícias. Available at: https://www.aen.pr.gov.br/Noticia/IAT-libera-colheita-e-venda-do-pinhao-em-todo-o-Parana-mas-com-exigencia-de-maturacao [Accessed 10 Apr. 2024].

vegetation, as it is resilient, easily regenerates, mitigates climate change, and conserves biodiversity. In some cases, it sequesters around eleven times more carbon than vegetation in advanced stages of regeneration and preservation and provides other ES such as the provision of water, energy, and raw materials. It can also contribute to restoring ecosystems and improving environmental conditions (Poorter et al. 2016; 2021).

Air quality regulation has also seen a significant decrease due to changes in land use related to anthropogenic activities, which, through various sources of pollution such as dust particles and other particulate matter, nitrous oxide emission from the use of agrochemicals (inorganic fertilizers), methane emissions from livestock, pollution from pesticide spraying, the burning of biomass, and the emission of polluting gases from vehicles, alter atmospheric conditions and consequently air quality (Folley et al. 2005; Miller Junior 2011). The 2023 mapping shows an increase in *air quality regulation* in ES, driven by localized gains in vegetation cover, particularly the expansion of forest formations and young secondary vegetation. Although these improvements do not fully counterbalance pollution sources, they generate a measurable enhancement in air quality regulation relative to previous years.

Water quality regulation is related to forest ecosystems, such as natural and planted forests and riparian forests (Rares and Brandimarte 2014). The reduction of this regulatory ES necessitates attention to preserving areas with vegetation cover, which must be distributed along the entire length of the watershed to maintain its functions, primarily in regulating biogeochemical cycles, water and nutrient flow, reducing erosion, and purifying water (Tundisi and Matsumura-Tundisi 2010; Adas et al. 2020). Water quality is degraded depending on the dynamics of land use. Furthermore, the expansion of intensive agriculture tends to increase erosion and sediment load and release nutrients and pesticides into groundwater, streams, and rivers (Folley et al. 2005). It is noteworthy that the water that supplies the municipality of Guarapuava comes from the PRB and has been contaminated by pesticides (Aggio et al. 2021; Piasseta et al. 2021).

The decline observed in *water quality* regulation in 2023 is associated with a different pattern of land use and vegetation structure. Although the area of forest formations and mosaic land-use classes increased, these mosaics consist of agro-pastoral and peri-urban mixtures where pastures, croplands, and fallow fields are interspersed with secondary vegetation at early successional stages. These transitional environments lack the structural complexity, continuous canopy cover, and well-developed root systems characteristic of mature forests, which are essential for filtering sediments, retaining nutrients, stabilizing soils, and regulating biogeochemical processes (Cui et al. 2020).

The *soil quality regulation* ES increased between 2002 and 2018 concerning the other regulating ES because of the increase in areas with vegetation cover. However, the 2023 mapping shows that the soil regulation has decreased, so the apparent increase in vegetated area does not translate into improved regulating functions. Instead, the replacement or fragmentation of mature forests exacerbated by agricultural pressures in surrounding transitional zones reduces the landscape's capacity to control erosion (Cui et al. 2020). In addition, the ongoing expansion of the PRB and the degrading potential of agroforestry or agricultural activities that do not incorporate soil conservation practices further intensify this decline (Miller Junior 2011).

The results indicated a decrease in *pollination potential*, mainly related to the replacement of native vegetation in areas by intense agriculture, which leads to a decline in *natural habitats* for pollinators, the increased use of pesticides, a decrease in floral resources, an increase in monocultures, and the intensive use of soil and water (Imperatriz-Fonseca 2004).

Pollinators are essential components for the functioning of ecosystems, and consequently, human food depends on pollination, so it is important to adopt the controlled use of agrochemicals and pesticides, land preparation practices that protect the nests of bees that occur in the soil, the maintenance of native vegetation edges, and living fences, and the reduced use of herbicides because some species of ruderal plants also provide resources for pollinators (Imperatriz-Fonseca 2004; FAO 2004).

The increase in educational values and tourism is attributed to a qualitative, rather than quantitative, shift in land-use practices within the basin. Although the areas historically associated with these cultural activities, such as traditional agriculture, declined, this loss was counterbalanced by the growth of environmental education initiatives, outreach programs, and revitalization projects. Consequently, the cultural tourism ES was either maintained or, in certain areas, expanded, supported by the diversification of rural recreational practices.

CONCLUSIONS

This study demonstrates that LULC changes in the PRB between 2002, 2018 and 2023 are directly associated with shifts in ES potential. While the expansion of anthropogenic activities, including mechanized agriculture and increased provisioning services (such as food and energy), these gains occurred alongside critical trade-offs, such as intensified soil erosion, forest fragmentation, and biodiversity loss. These dynamics in the PRB indicate that increases in vegetated areas are insufficient to ensure the integrity of ecosystem functions.

The matrix-based approach used in this study relies on LULC classes as proxies for ES potential. To account for discrepancies between LULC mappings, ES scores were assigned relative to each specific year, considering their respective scales and levels of generalization. Consequently, the results represent relative potentials that, while year-specific, allow for an evaluation of temporal and spatial trends. Thus, the observed changes reflect both land-use dynamics and the inherent specificities of the datasets used.

Our findings emphasize the need to implement integrated strategies and multi-level governance to promote sustainable land management. Future research should prioritize harmonized LULC classifications, integrate quantitative biophysical indicators with socio-economic data, and incorporate participatory approaches to better capture ES demand and access. We recommend developing monitoring protocols integrated into constantly updated databases to facilitate ES valuation. Such an approach would enable long-term management decisions on land use and ecosystems to safeguard human well-being for current and future generations. ■

REFERENCES

Adas, M. A. A., Hardt, E., Miraglia, S. G. E. and Semensatto, D. (2020). Reforest or perish: ecosystem services provided by riparian vegetation to improve water quality in an urban reservoir (São Paulo, Brazil). Sustentabilidade em Debate, 11 (1), Avaliable at: https://www.cabidigitallibrary.org/doi/full/10.5555/20203272295 [Accessed 07 Jan. 2026].

Aggio, C. M., Leite, A., Malaquias, T. S. M. and Haddad, M. C. F. L. (2021). Caracterização das notificações de intoxicações por agrotóxicos na 5º Regional de Saúde do Paraná (PR). Revista Eletrônica Acervo Saúde, v. 13, n. 2, p. e5816-e5816, 2021. Available at: https://acervomais.com.br/index.php/saude/article/view/5816 [Accessed 23 Jun. 2024].

Alvarenga, R. C., Silva, V. P., Gontijo-Neto, M. M., Viana, M. C. M., and Vilela, L. (2010). Sistema integração lavoura-pecuária-floresta: condicionamento do solo e intensificação da produção de lavouras. Informe Agropecuário, v. 31, p. 59-67. Avaliable at: https://www.alice.cnptia.embrapa.br/handle/doc/868431 [Accessed 07 Jan. 2026].

Andrade, D. C. and Romeiro, A. R. (2009). Serviços ecossistêmicos e sua importância para o sistema econômico e o bem-estar humano. Texto para discussão. IE/UNICAMP, v. 155, p. 1- 43. Avaliable at: https://repositorio.unicamp.br/acervo/detalhe/1211804 [Accessed 07 Jan. 2026].

Balbino, L. C., Cordeiro, A. M. C., Oliveira, P., Kluthcouski, J., Galerani, P. R., and Vilela, L. (2012). Agricultura sustentável por meio da integração lavoura-pecuária-floresta (iLPF). Informativo Agronômico, v. 138, p. 1-18. Avaliable at: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1111127/agricultura-sustentavel-por-meio-da-integracao-lavoura-pecuaria-floresta-ilpf [Accessed 07 Jan. 2026].

Bastian, O., Haase, D. and Grunewald, K. (2012). Ecosystem properties, potentials and services: the EPPS conceptual framework and an urban application example. Ecological Indicators, v. 21, p. 7-16. Avaliable at: https://doi.org/10.1016/j.ecolind.2011.03.014 [Accessed 07 Jan. 2026].

Benabou A., Moukrim S., Laaribya S., Aafi A., Chkhichekh A., Maadidi T.E., Aboudi A.E. Mapping Ecosystem Services of Forest Stands: Case Study of Maamora, Morocco. Geography, Environment, Sustainability. 2022;15(1):141-149. Avaliable at: https://doi.org/10.24057/2071-9388-2021-047 [Accessed 07 Jan. 2026].

Bombardi, L. M. (2017). Geografia do uso de Agrotóxicos no Brasil e Conexões com a União Europeia. São Paulo: FFLCH/ USP. 292 p.

Brasil. Lei nº 9.433, de 8 de janeiro de 1997. (1997). Brasília: Diário Oficial da União. Available at: http://www.planalto.gov.br/ccivil_03/leis/l9433.htm [Accessed 10 Sep. 2025].

Brasil. Lei nº 14.119, de 13 de janeiro de 2021. (2021). Brasília: Diário Oficial da União. Available at: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14119.htm [Accessed 10 Sep. 2025].

Brito, E. M., Mello, K., Costa. D. R., Faria, L. C. and Valente, R. A. (2016). Geotecnologias aplicadas ao estudo do conflito de uso do solo visando o planejamento ambiental. Irriga, v. 21, n. 3, p. 577-590. Avaliable at: https://doi.org/10.15809/irriga.2016v21n3p577-590 [Accessed 07 Jan. 2026].

Burkhard, B. and Maes, J. (2017). Mapping ecosystem services. Advanced books, v. 1, p. e12837.

Burkhard, B., Kroll, F., Nedkov, S. and Müller, F. (2012). Mapping ecosystem service supply, demand and budgets. Ecological indicators, v. 21, p. 17-29. Avaliable at: https://doi.org/10.1016/j.ecolind.2011.06.019 [Accessed 07 Jan. 2026].

Chaves, A. M. S., Vieira, A. G. T., França, E. M. S. De, Santos, E. J. Dos, Teixeira, G. Dos S. S., Silva, J. I. dos, and SOUZA, R. M. e. (2021). Análise dos serviços ecossistêmicos na paisagem semiárida da Bacia do Riacho São José, Pernambuco. Geosaberes, v. 12, p. 139-158. DOI: 10.26895/geosaberes.v12i0.1147.

CPT (Comissão Pastoral da Terra) (2018). Conflitos no campo Brasil (2018). Available at: https://www.cptnacional.org.br/publicacoes-2/destaque/4687-conflitos-no-campo-brasil-2018 [Accessed 15 Aug. 2024].

Costanza, R.; D'Arge, R.; Groot, R.; Farber, S.; Grasso, M.; Hannon, B.; Limburg, K.; Naeem, S.; O'Neill, R. V.; Raulelo, J.; Raskin, R. G.; Sutton, P.; Belt, M. (1997). The value of the world's ecosystem services and natural capital. Nature, v. 387, n. 6630, p. 253-260. Available at: https://www.nature.com/articles/387253a0 [Accessed 20 Jun. 2024].

Cui, N., Zhang, X., Cai, M., Zhou, L., Chen, G., and Zou, G. (2020). Roles of vegetation in nutrient removal and structuring microbial communities in different types of agricultural drainage ditches for treating farmland runoff. Ecological Engineering, 155, 105941. Avaliable at: https://doi.org/ 10.1016/j.ecoleng.2020.105941 [Accessed 07 Jan. 2026].

Cunha, M. C., Santos, V. S. and Cruz, A. A. (2014). Levantamento das medidas de manutenção aplicado nas estradas rurais na bacia do Rio das Pedras, Guarapuava-PR com avaliação qualitativa do estado de conservação. Caderno de Geografia, v. 24, n. 2, p. 124-138. Avaliable at: https://doi.org/10.5752/P.2318-2962.2014v24n42p124 [Accessed 07 Jan. 2026].

Cunha, M. C., Thomaz, E. L. and Vestena, L. R. (2013). Medidas de controle de erosão em estradas rurais na bacia do rio das Pedras, Guarapuava-PR. Sociedade e Natureza, v. 25, n. 1, p. 107-118. Avaliable at: https://doi.org/10.1590/S1982-45132013000100009 [Accessed 07 Jan. 2026].

Cunha, M. C. and Thomaz, E. L. (2015). É possível reduzir a turbidez da água em bacia rural por Meio de implantação de caixas de infiltração?. Revista Brasileira de Geomorfologia, v. 16, n. 4. Avaliable at: https://doi.org/10.20502/rbg.v16i4.725 [Accessed 07 Jan. 2026].

Cunha, M. C. (2011). Caracterização das estradas rurais não pavimentadas como elementos presentes na paisagem da bacia hidrográfica do rio das pedras, Guarapuava-PR. Revista de Geografia, Meio Ambiente e Ensino, v. 1, n. 2. Avaliable at: https://periodicos.unespar.edu.br/geomae/article/view/5754 [Accessed 07 Jan. 2026].

De Groot, R. S. (2002). A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecological economics, v. 41, n. 3, p. 393-408. Avaliable at: https://doi.org/10.1016/S0921-8009(02)00089-7 [Accessed 07 Jan. 2026].

Donatti, L., Rebeca, R., Zalenski, T., Felski, G., Wolff, L. L., And Viana, D. (2004). Percepção e educação ambiental: Percepção ambiental pela população ribeirinha da ictiofauna existente na bacia hidrográfica do rio das Pedras. In: BATTISTELI, M. et al. (Eds.), Proteção e manejo da bacia do rio das Pedras: relato de experiências. Guarapuava, PR. Ed. Gráfica BandD.

FAO (Food and Agriculture Organization of the United Nations). Conservation and management of pollinators for sustainable agriculture – the international response. In: Freitas, B.M.; Pereira, J.O.P. (eds.) Solitary bees: conservation, rearing and management for pollination. Imprensa Universitária. Fortaleza, 2004.

Favaro, J. L., Ikuta, F. K., And Bertotti, L. G. Agricultura familiar no território Paraná-Centro 2020: desafios para o fortalecimento da agricultura camponesa e agroecológica. In: Favaro, Jorge Luis et al. (Org.). Experiências e reflexões extensionistas: Núcleo Multidisciplinar de Estudo em Agroecologia e Produção Orgânica da Unicentro. CandA Alfa Comunicação. Goiânia, 2020.

Franchini, J. C. et al. (2011). Importância da rotação de culturas para a produção agrícola sustentável no Paraná. Londrina: Embrapa Soja.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N. and Snyder, P. K. (2005). Global consequences of land use. Science, 309(5734), 570-574. Avaliable at: https://doi.org/10.1126/science.1111772 [Accessed 07 Jan. 2026].

Githiora-Murimi, Y. W., Owuor, M. A., Abila, R., Olago, D., and Oriaso, S. (2022). Integrating stakeholder preferences into ecosystem services mapping in Yala wetland, Kenya. Ecosystems and People, v. 18, n. 1, p. 146–163. Avaliable at: https://doi.org/10.1080/26395916.2022.2039774. [Accessed 07 Jan. 2026].

Guarapuava. Municipal Government of Guarapuava. Aerial photographs obtained by ENGEFOTO of the Municipality of Guarapuava/PR. 2003. Scale: 1:30,000.

Guarapuava. Municipal Government of Guarapuava. Aerial photographs obtained by ENGEFOTO of the Municipality of Guarapuava/PR. 2020. Scale: 1:30,000.

Hess, S. C. (2018). Ensaios sobre poluição e doenças no Brasil. Outras Expressões.

Hinata, S. Da S., And Basso, L. A. (2022). Mapeamento do uso e cobertura do solo como subsídio à avaliação de serviços ecossistêmicos na sub-bacia hidrográfica do Arroio Passo Fundo, Guaíba-RS. Estudos Geográficos (UNESP), v. 43, n. 2.

IBGE (Instituto Brasileiro de Geografia e Estatística). Manual técnico de uso da terra. Rio de Janeiro, 1999.

IBGE (Instituto Brasileiro de Geografia e Estatística). (2002). PAM - Produção Agrícola Municipal 2002. Rio de Janeiro: IBGE. Retrieved from: https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhesandid=766. Accessed on: Sep. 10, 2025.

IBGE (Instituto Brasileiro de Geografia e Estatística). (2018). PAM - Produção Agrícola Municipal. Retrieved from: https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9117-producao-agricola-municipal-culturas-temporarias-permanentes.html?edicao=25369andt=conceitos-e-metodos. Accessed on: Sep. 10, 2025.

Iniesta-Arandia, I. Et Al. (2014). Socio-cultural valuation of ecosystem services: uncovering the links between values, drivers of change, and human well-being. Ecological economics, v. 108, p. 36-48. Avaliable at: https://doi.org/10.1016/j.ecolecon.2014.09.028 [Accessed 07 Jan. 2026].

Imperatriz-Fonseca, V. L. and Nunes-Silva, P. (2010). As abelhas, os serviços ecossistêmicos e o Código Florestal Brasileiro. Biota Neotropica, 10, 59-62. Avaliable at: https://doi.org/10.1590/S1676-06032010000400008 [Accessed 07 Jan. 2026].

Imperatriz-Fonseca, V. L. (2004). Serviços aos ecossistemas, com ênfase nos polinizadores e polinização. São Paulo: USP. Avaliable at: http://www.biotaneotropica.org.br/v10n4/en/abstract?article+bn00910042010 [Accessed 07 Jan. 2026].

IPCC. Climate Change 2022: Impacts, Adaptation and Vulnerability, 2021. Available at: https://www.ipcc.ch/library/ [Accessed 15 Jan. 2024].

Jägermeyr, J. et al. (2021) Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. Nature Food, v.2 n.11 p. 873-885. Available at: https://www.nature.com/articles/s43016-021-00400-y [Accessed 15 Oct. 2024].

Kumi, S., Addo-Fordjour, P., Fei-Baffoe, B., Belford, E. J. D., and Ameyaw, Y. (2021). Land use land cover dynamics and fragmentation-induced changes in woody plant community structure in a mining landscape, Ghana. Trees, Forests and People v. 4, p. 100070. Avaliable at: https://doi.org/10.1016/j.tfp.2021.100070 [Accessed 07 Jan. 2026].

Leite, F. F. G. D. Et Al. (2023). Greenhouse gas emissions and carbon sequestration associated with Integrated Crop–Livestock–Forestry (ICLF) systems. Environmental Reviews, v. 31, n. 4, p. 589-604. Avaliable at: https://doi.org/10.1139/er-2022-0095 [Accessed 07 Jan. 2026].

Machado, P. A. (2005). Soil carbon and the mitigation of global climate change. Química Nova, v. 28, n. 2, p. 329-334. Avaliable at: https://doi.org/10.1590/S0100-40422005000200026 [Accessed 07 Jan. 2026].

MapBiomas. (2025). Visão geral da metodologia de cobertura e uso da terra (Coleção 9.0). Projeto MapBiomas. Available at: https://brasil.mapbiomas.org/visao-geral-da-metodologia/ [Accessed 15 Nov. 2025].

Marchão, R. L. Et Al. (2024). Integrated crop–livestock–forestry systems for improved soil health, environmental benefits, and sustainable production. In: Soil Health Series: Volume 3 Soil Health and Sustainable Agriculture in Brazil (p. 19-61).

Martínez-Harms, M.-J., And Balvanera, P. (2012). Methods for assessing ecosystem services: a review. Biological Conservation, v. 154, p. 319-333.

MEA (Millennium Ecosystem Assessment). Ecosystems and human well-being: synthesis. Island, Washington, DC, 2005.

Metzger, M.J., Rounsevell, M.D.A., Acosta-Michlik, L., Leemans, R. and Schröter, D. (2006). The vulnerability of ecosystem services to land use change. Agriculture, ecosystems and environment, n. 114, p. 69-85, 2006. Available at: https://doi.org/10.1016/j.agee.2005.11.025 [Accessed 07 Jan. 2026].

Mello, I., And Van Raij, B. (2006). No-till for sustainable agriculture in Brazil. Proceedings of the World Association of Soil and Water Conservation, v. 1, p. 49-57.

Mello, K. De Et Al. (2020). Multiscale land use impacts on water quality: Assessment, planning, and future perspectives in Brazil. Journal of environmental Management, v. 270, p. 110879. Available at: https://doi.org/10.1016/j.jenvman.2020.110879 [Accessed 07 Jan. 2026].

Miller Junior, G. T. Ciência Ambiental. São Paulo: Cengage Learning, 2011.

Muñoz, A. M. M. and De Freitas, S. R. (2017). Importância dos Serviços Ecossistêmicos nas Cidades: Revisão das Publicações de 2003 a 2015. Revista de Gestão Ambiental e Sustentabilidade, v. 6, n. 2, p. 89-104. Available at: https://doi.org/10.5585/geas.v6i2.853 [Accessed 07 Jan. 2026].

Odum, E.P. Ecologia. Rio de Janeiro: Ed. Guanabara. 1988.

Paraná. (2019). Valor Bruto da Produção de 2018. Retrieved from: https://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/documento/2019-10/graficosmunicipaisvbp2018r_0.pdf. Accessed on: Sep. 10, 2025.

Pereira, E. M. et al. (2020). Agroecologia e saúde: luta por alimentos saudáveis no Brasil. In: Favaro, J. L. et al. (Org.). Experiências e reflexões extensionistas: Núcleo Multidisciplinar de Estudo em Agroecologia e Produção Orgânica da Unicentro. CandA Alfa Comunicação. Goiânia.

Pessoa, M. M. de L., Carvalho, D. C. de, Santos, J. M. dos, and Magalhães, L. M. S. (2020). Landscape dynamics and its impacts in the Atlantic Forest area of Southeastern Brazil. Revista Brasileira de Ciências Agrárias v. 15, n. 1, p. 1-8. Available at: https://doi.org/10.5039/agraria.v15i1a6971 [Accessed 07 Jan. 2026].

Piassetta, R. R. L., Souza N. J., Mikos, A. P. and Auer, C. G. (2021). Restrictive legislation regarding the use of pesticides in Paraná cities. BIOFIX Scientific Journal, v. 6, n. 1, p. 75-83, 2021. Available at: https://doi.org/10.5380/biofix.v6i1.77378 [Accessed 07 Jan. 2026].

Polidoro, J. C., Freitas, P. L. de, Hernani, L. C., Anjos, L. H. C. dos, Rodrigues, R. de A. R., Cesário, F. V., Andrade, A. G. de, and Ribeiro, J. L. (2021). Potential impact of plans and policies based on the principles of conservation agriculture on the control of soil erosion in Brazil. Land Degradation and Development, 1–12. Available at: https://doi.org/10.1002/ldr.3876 [Accessed 07 Jan. 2026].

Poorter, L. et al. (2016). Biomass resilience of Neotropical secondary forests. Nature, v. 530, n. 7589, p. 211-214. Available at: https://www.nature.com/articles/nature16512 [Accessed 10 Jun. 2024].

Poorter, L. et al. (2021). Multidimensional tropical forest recovery. Science, v. 374, n. 6573, p. 1370-1376. Available at: https://doi.org/10.1126/science.abh3629 [Accessed 07 Jan. 2026].

Rares and Brandimarte 2014 - Rares, C. D. S., and Brandimarte, A. L. (2014). The challenge of aquatic environments conservation and continuity of environmental services in urban green areas: the case of Cantareira State Park. Ambiente and Sociedade, v. 17, p. 111-128. Available at: https://doi.org/10.1590/S1414-753X2014000200008 [Accessed 07 Jan. 2026].

Sampaio, B. D. S., And Bacani, V. M. (2025). Avaliação dos serviços ecossistêmicos e pagamento por serviços ambientais: aplicação na bacia hidrográfica do Córrego Bebedouro, Mato Grosso do Sul. Geografia (Rio Claro-SP), v. 50, n. 1, p. 124-150.

Santos, A. F. and Kobiyama, M. (2003). Contribuição potencial de cargas poluentes na bacia do Rio das Pedras, no município de Guarapuava-PR. RECEN-Revista Ciências Exatas e Naturais, v. 5, n. 1, p. 33-46. Available at: https://revistas.unicentro.br/index.php/RECEN/article/download/431/586 [Accessed 20 Jun. 2024].

Santos, C. R. Et Al. (2024). Análise dos serviços ecossistêmicos nos diferentes biomas brasileiros: uma revisão sistemática. Boletim Goiano de Geografia, v. 43, n. 1. Available at: https://doi.org/10.5216/bgg.v43i01.71185[Accessed 07 Jan. 2026].

Simões, G. Dos S., Fiore, F. A., And Silva, L. C. (2022). Mapeamento dos Serviços Ecossistêmicos providos pela Área de Proteção Ambiental do Rio Paraíba do Sul. Ambiente e Sociedade, v. 25, p. e01882. Available at: http://dx.doi.org/10.1590/1809-4422asoc20210188r2vu2022L2AO [Accessed 07 Jan. 2026].

Skorupa, L. A., And Manzatto, C. V. (2019). Avaliação da adoção de sistemas de integração lavoura-pecuária-floresta (ILPF) no Brasil. In: Sistemas de integração lavoura-pecuária-floresta no Brasil: Estratégias regionais de transferência de tecnologia, avaliação da adoção e de impactos (p. 340-79). Brasília, DF: Embrapa.

Sohel, M. S. I., Mukul, S. A., And Burkhard, B. (2015). Landscape's capacities to supply ecosystem services in Bangladesh: a mapping assessment for Lawachara National Park. Ecosystem Services, v. 12, p. 128-135. Available at: http://doi.org/10.1016/j.ecoser.2014.11.015 [Accessed 07 Jan. 2026].

Teodoro, P. F. and Santos, A. F. (2009). Qualidade da água da bacia do Rio das Pedras–Guarapuava (PR), baseado nos parâmetros que definem o Índice de Qualidade da Água (IQA). Guairacá-Revista de Filosofia, v. 25, n. 1. Available at: https://revistas.unicentro.br/index.php/RECEN/article/download/431/586 [Accessed 20 Jun. 2024].

Tundisi, J. G. and Matsumura-Tundisi, T. (2010). Impactos potenciais das alterações do Código Florestal nos recursos hídricos. Biota Neotropica, v. 10, n. 4, p. 67-76. Available at: http://doi.org/ 10.1590/S1676-06032010000400010 [Accessed 07 Jan. 2026].

Uzeda, M. C. (2020). Estratégias de diversificação da paisagem agrícola como ferramentas para manutenção dos estoques de carbono do solo. Anais da Semana Científica Johanna Döbereiner.

Vestena, L. R. et al. (2004). Uso da terra da bacia hidrográfica do Rio das Pedras. In: Battistelli et al. Proteção e Manejo da Bacia do Rio das Pedras: Relato de Experiências. Editora BandD Ltda. Guarapuava.

Vestena, L. R. and Thomaz, E. L. (2006). Avaliação de conflitos entre áreas de preservação permanente associadas aos cursos fluviais e uso da terra na Bacia do Rio das Pedras, Guarapuava-PR. Ambiência, v. 2, n. 1, p. 73-85. Available at: https://revistas.unicentro.br/index.php/ambiencia/article/view/362 [Accessed 25 May. 2024].

Vitousek, P. M., Mooney, H. A., Lubchenco, J. and Melillo, Jerry M. (1997). Human domination of Earth's ecosystems. Science, v. 277, n. 5325, p. 494-499. Available at: http://doi.org/ 10.1126/science.277.5325.494 [Accessed 07 Jan. 2026].

Zhen, L. and Routray, J. K. (2003). Operational indicators for measuring agricultural sustainability in developing countries. Environmental Management, New York, v. 32, n. 1, p. 34-46 Available at: http://doi.org/10.1007/s00267-003-2881-1 [Accessed 07 Jan. 2026].

Check for updates

cc BY

RESEARCH PAPER

IDENTIFICATION OF BIOLOGICAL AEROSOLS BY ITS FLUORESCENCE SPECTRA AND COMPARISON WITH VOLUMETRIC DATA

Egor A. Illarionov[1,2*], Svetlana V. Polevova[3], Oleg S. Panin[1], Anastasiya A. Konstatinidi[1], Elena E. Severova[3*]
[1]Moscow State University, Department of Mechanics and Mathematics, Leninskiye Gory 1, Moscow, 119991, Russia
[2]Moscow Center of Fundamental and Applied Mathematics, Leninskiye Gory 1, Moscow, 119991, Russia
[3]Moscow State University, Department of Biology, Leninskiye Gory 1, Moscow, 119991, Russia
***Corresponding author:** egor.illarionov@math.msu.ru, elena.severova@mail.ru

Received: May 20th 2025 / Accepted: January 12th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4053

ABSTRACT. An automatic bioaerosol classification system is an attractive alternative to the standard visual identification and counting of pollen in the standard volumetric method of aerobiological monitoring. While various physical principles can be used for automatic measurement of the parameters of particles present in the air, the key problem becomes the development of a classification model based on these data. In particular, practical application of the models becomes challenging due to the large variability of particles present in real air compared to laboratory experiments in which models are usually trained. Instead of training the model on data obtained in laboratory conditions, we applied a clustering algorithm to fluorescence spectra collected during daily measurements in the 2024 season with the Rapid-E+ automatic detector installed at the monitoring station at Moscow State University. Comparison of the temporal distribution of particles in each cluster with the seasonal dynamics of eleven pollen types obtained from standard aerobiological monitoring with a volumetric trap at the same station at Moscow State University shows that some clusters (i.e., fluorescence spectra of specific shape and amplitude) demonstrate temporal patterns similar to pollen seasons. However, the fluorescence spectra alone are not sufficient for differentiation of individual pollen types, and they can only provide detection of larger groups of bioaerosols. Interestingly, the detected larger groups show more diverse seasonal patterns than those observed by volumetric monitoring at the station at Moscow State University. This result demonstrates that automatic detectors can provide more useful information on the content and seasonal distribution of bioaerosols compared to standard volumetric methods.

KEYWORDS: bioaerosol classification, cluster analysis, automatic detector, pollen

CITATION: Illarionov E. A., Polevova S. V., Panin O. S., Konstatinidi A. A., Severova E. E. (2026). Identification Of Biological Aerosols By Its Fluorescence Spectra And Comparison With Volumetric Data. Geography, Environment, Sustainability, 1 (19), 29-35
https://doi.org/10.24057/2071-9388-2026-4053

ACKNOWLEDGEMENTS: The study was conducted under the state assignment of Lomonosov Moscow State University. EI acknowledges the Lomonosov-2 supercomputer center at MSU for providing computational resources and support of the Moscow Center of Fundamental and Applied Mathematics of Lomonosov Moscow State University under agreement No. 075-15-2025-345.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Plant pollen and spores are the main components of biological aerosols and, simultaneously, one of the leading causes of allergic diseases (pollinosis) (Valenta et al. 1992; D'Amato 2000; Sofiev, Bergman 2012; Fröhlich-Nowoisky et al. 2016; D'Amato et al. 2017). Allergies reduce the quality of life for approximately 25–30% of the global population (Dykewicz & Hamilos 2010, Akdis et al. 2015). An essential element of the complex of anti-allergic measures is aerobiological monitoring, which makes it possible to track and predict the dynamics of the concentration of the main allergens in the atmosphere to adjust the therapy and lifestyle of patients with pollinosis.

Currently, the "gold standard" of pollen monitoring is the volumetric pollen trap used for detection and counting of pollen (Hirst 1952, Galán et al. 2014, Sikoparija et al. 2017, Buters et al. 2018), but these devices are not able to analyze the composition of aerosols in real time. Volumetric samples are identified by a microscopic analysis, which is time-consuming, demands specialized expertise, and involves considerable uncertainties (Šaulienė et al. 2019, Oteros et al. 2017). These problems can be solved with automated aerosol composition analyzers (see An et al. 2024, Kabir et al. 2020, Maya-Manzano et al. 2020, and Huffman et al. 2019, for review of recent progress in different detection techniques), one of which (Rapid E+, Plair, Switzerland) was purchased under the Moscow State University development program and installed in early 2022. In this instrument, the technology for measuring the characteristics of particles in the air stream is based on a combination of light scattering for determination of particle morphology, size, and velocity, and fluorescence spectroscopy for analysis of fluorescence spectra.

Recently, Erb et al. (2024), Brdar et al. (2023), Boldeanu et al. (2022), Šikoparija et al. (2022), Crouzy et al. (2020), Šaulienė et al. (2019) demonstrated that machine learning models trained on fluorescence spectra combined with other measured pollen parameters can differentiate pollen types. However, these results were obtained in laboratory conditions when the collected pollen of different taxa was injected into the detector. The authors of these models note that adaptation of the models to detection of the learned pollen types in real outdoor conditions meets significant difficulties. The key problem is that a model trained to distinguish pollen of a specific type, such as type A, from another specific type, such as type B, becomes ambiguous when applied to real-world data that includes type A, type B, and a various other pollen types. A practically usable model should be able to distinguish pollen of type A from all other particles present in the air (referred to background particles). However, preparing a corresponding dataset with examples of all other possible particles is unfeasible. As a workaround, for example, Matavulj et al. (2025) proposed to use data collected during the no-pollen season as examples of background particles. However, this approach still cannot capture the full variability of background particles in the pollen season.

Another problem noted in Matavulj et al. (2022, 2025) is that the measurements between different detectors (of the same type) are weakly consistent. Thus, models trained on data from one detector cannot be directly applied to data from another without additional adaptation.

Instead of direct training of classification models, Swanson & Huffman (2020) and Daunys et al. (2021) proposed usage of clustering techniques that help to avoid the problem of collection of the training dataset but require identification of the obtained clusters with specific pollen types.

In this work, we apply a clustering algorithm to fluorescence spectra measured by the automatic atmospheric particle detector Rapid-E+ (which is different from the fluorescence detector used in Swanson & Huffman, 2020) during the 2024 season, and we compare the temporal distribution of particles in each cluster with the temporal distribution of pollen types obtained from daily monitoring with a standard volumetric trap (which extends the approach proposed in Daunys et al., 2021).

MATERIALS AND METHODS

Sampling Campaign

Data for this study were collected from March through October 2024 at the monitoring station at Moscow State University (Moscow, Russia) using two different sampling campaigns conducted in parallel. In the first campaign, data was collected using the automatic aerosol detector Rapid E+ installed on the roof of one of the buildings on Moscow State University's territory. This device measures various particle parameters based on the physical effects of light scattering and fluorescence, among which we consider only particle size and fluorescence spectrum. It is important to note that by design the automatic detector cannot analyze all particles present in a given volume of flowed air, but only a certain number of them. Moreover, the device measures particles of all types uniformly and cannot prioritize the measurement of one type of particle (e.g., pollen) and ignore other types. To filter out only particles that could potentially include pollen, we apply threshold filtering and select particles with a size larger than 2μm and a mean fluorescence intensity above 4000 (this value is also recommended by the manufacturer for pollen analysis). Fig. 1 shows the daily number of particles meeting these criteria during the 2024 season. Due to technical reasons, the device was turned off in June and early July, as well as in mid-August (these periods correspond to zero values in Fig. 1).

In the second campaign, parallel to the automatic detector, data was collected using a Hirst-type volumetric pollen and spore trap (Lanzoni VSSP 2010, Italy) located at the same place as the automated detector. We focused on eleven pollen types *(Acer, Alnus, Betula, Corylus, Pinus, Poaceae, Populus, Quercus, Salix, Ulmus, Urtica)*. Fig. 2 shows the number of these pollen types and the total amount of detected pollen.

In addition to the above daily measurements, samples of pure pollen from *Betula, Quercus,* and *Phleum* were collected from the plants during their pollen periods. Using these examples of pollen, we conducted a laboratory experiment in which the automatic aerosol detector Rapid E+ was placed indoors, and these particles and mixtures of these particles were injected into the device. In more detail, we injected first clean pollen of *Betula*, then *Quercus*, then the mixture of *Betula* and *Quercus* with mass ratio 1:1, then the mixture of *Betula* and *Quercus* with mass ratio 1:3, then clean pollen of *Phleum*, then the mixture of *Betula* and *Phleum* with mass ratio 1:1, then the mixture of *Betula, Quercus*, and *Phleum* with mass ratio 1:3:2 correspondingly. Fig. 3 shows the total number of particles detected by the Rapid-E+ per minute, and vertical lines indicate the time moments when the next portions of pollen were injected.

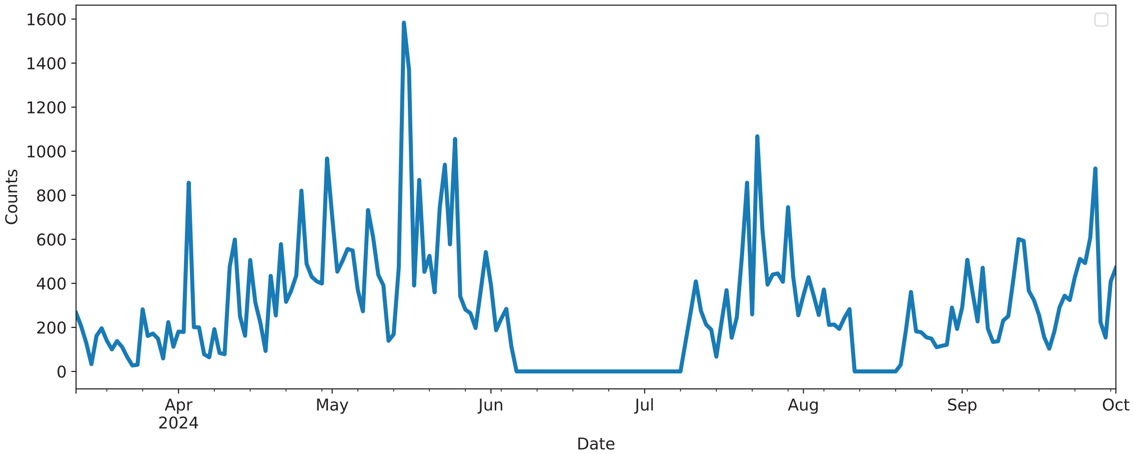


Fig. 1. Number of particles (with size > 2 μm and fluorescence threshold > 4000) detected by the automatic detector at the monitoring station at Moscow State University during the 2024 season (zero values mean that the device was turned off)

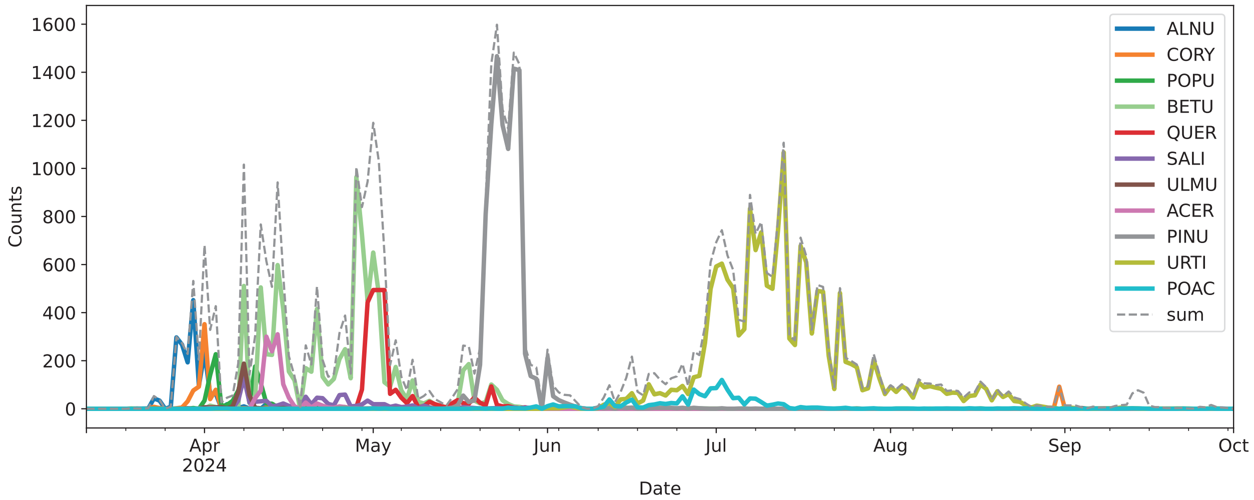


Fig. 2. Daily counts of different pollen types obtained using the Hirst volumetric method at the monitoring station at Moscow State University during the 2024 season (ALNU - *Alnus*, CORY - *Corylus*, POPU - *Populus*, BETU - *Betula*, QUER - *Quercus*, SALI - *Salix*, ULMU - *Ulmus*, ACER - *Acer*, PINU - *Pinus*, URTI - *Urtica*, POAC - *Poaceae*)

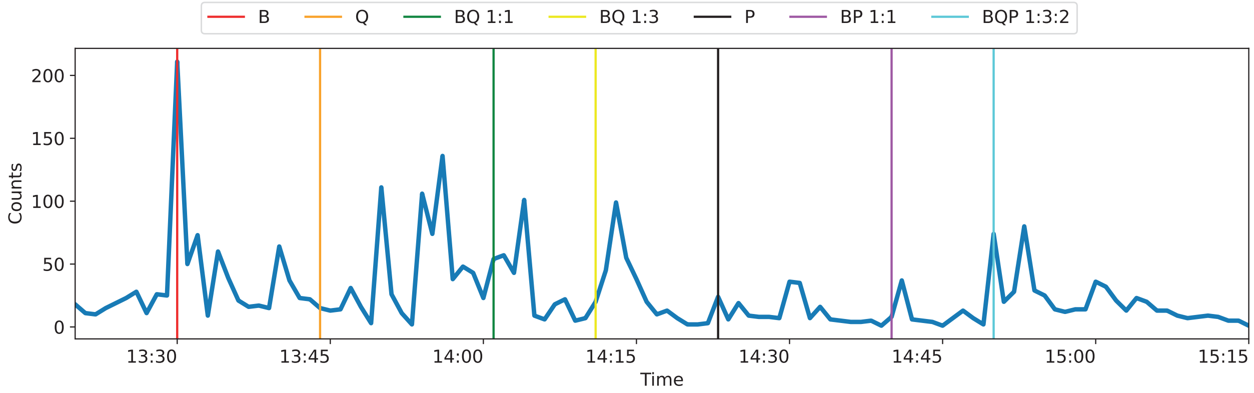


Fig. 3. Particle counts (blue line) during the laboratory experiment. Vertical lines indicate time moments when the new portion of pollen was injected (B is for *Betula*, Q is for *Quercus*, P is for *Phleum*, BQ 1:1 and the similar notations mean the mixture of B and Q with mass ratio 1:1)

Methods

We aim to isolate specific groups of fluorescence spectra within the entire dataset of measured spectra and associate them with particular pollen types using daily monitoring data. The fluorescence spectrum of each particle is given by 16 values representing fluorescence intensity measured at 16 different wavelengths. Particles of different types yield different shapes and amplitudes of the fluorescence spectra. Three different samples shown in Fig. 4 illustrate the variability of the spectra. We observe in Fig. 4 that the spectrum of particle 1 is middle-peaked and strongly non-symmetric; the spectrum of particle 2 is middle-peaked but rather symmetric; and the peak of the spectrum of particle 3 is shifted to the right. Also, amplitudes of the spectra of particles 1 and 3 are much higher than for particle 2. The observed difference in shape and amplitudes motivates us to apply a clustering method for the isolation of different types of spectra.

To divide all measured spectra into subgroups (clusters) containing spectra of similar shape and amplitude, we apply the *K*-means clustering algorithm (see, e.g., Bishop, 2007). Specifically, each measured spectrum is considered a point in a 16-dimensional space. The goal of the K-means algorithm is to place K new points, considered as cluster centers, in this space in such a way as to minimize the sum of squares of the distance from each point to the nearest center. The number of clusters, *K*, is set by the user, and the distance is the Euclidean distance. The algorithm finds the optimal location of cluster centers iteratively, starting from some arbitrary positions. Once the method has converged, each point is assigned to the cluster with the nearest center. The number of clusters, *K*, is adjusted manually to obtain representative results. It turned out to be useful to isolate a larger number of clusters (compared to the potential number of pollen types) and then merge some clusters that have similar spectrum shapes (but possibly different amplitudes) and similar seasonal patterns.

The final step is to associate each cluster with a specific pollen type. Similar to Fig. 2, we plot daily counts of particles from a given cluster and compare this plot with corresponding plots arising from daily monitoring of known pollen types. If a similar seasonal pattern is found with any known pollen type, then the cluster is identified with that pollen type.

RESULTS

Laboratory experiment

We start with a demonstration of clusters obtained in the laboratory experiment described above (see Fig. 3). To identify clusters corresponding to Betula, Quercus, and Phleum, the K-means clustering algorithm was applied with the number of clusters K=10, and the obtained clusters are shown in Fig. 5. The red line in Fig. 5 shows the averaged spectra or cluster center in terms of the K-means algorithm. Some clusters have similar shapes but different amplitudes, for example, clusters 1 and 10 or clusters 2 and 4. One can assume that they represent the same pollen type.

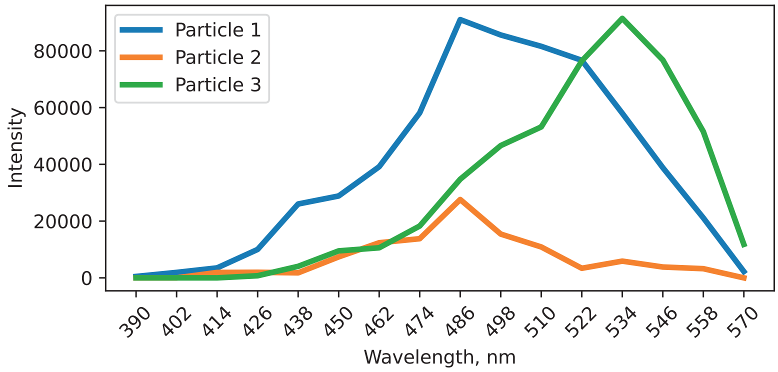


Fig. 4. Examples of fluorescence spectra for three different particles

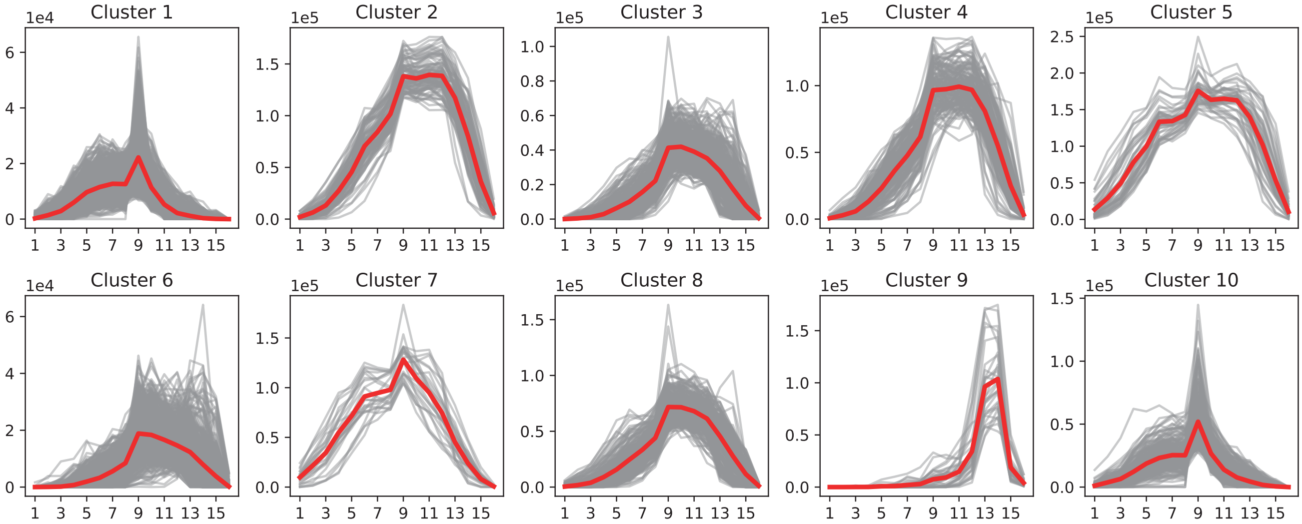


Fig. 5. Clusters identified by the *K*-means algorithm in the laboratory experiment. Gray lines show spectrum lines of the individual particles in each cluster. The red line is the average spectrum line in each cluster

Now consider the temporal distribution of particles in each cluster. Fig. 6 shows that particles in clusters 1 and 10 first appear when we inject *Betula* pollen into the detector (this moment is marked with a red vertical line) and then appear only in those periods when mixtures containing pollen of *Betula* were injected. Thus, one can identify the spectral profiles in clusters 1 and 10 as corresponding to *Betula*. Using the same reasoning, spectral profiles in clusters 2, 4, and 5 are identified with *Quercus* pollen. Surprisingly, particles from these clusters do not appear when the mixtures of *Betula* and *Quercus* are injected; however, they are observed when the mixture of all three pollen types is injected. The reason is not clear to us, and the only thing that we can note is that the mass content of *Quercus* pollen in the latter case was several times larger than in the previous mixtures. Considering the remaining clusters, one can note that cluster 7 has the first pronounced peak when *Phleum* pollen is injected and thus can be identified as this plant. In contrast, clusters 3, 6, and 9 appear at all times and even before the experiments started and can be attributed to background particles.

This laboratory experiment demonstrates that the proposed approach can be used to identify pollen types using spectrum profiles. The next step is to apply this method to the data from daily outdoor monitoring in the 2024 season.

Daily monitoring data

For the outdoor spectra measurements, we applied clustering with a large number of clusters (K=70) and obtained two groups of clusters. Particles in one group do not demonstrate any localized (in time) seasonal pattern. Their temporal distribution is similar to the distribution of the total number of particles and is therefore non-informative. Particles in the second group have localized (in time) patterns and thus can be (potentially) identified with some specific pollen. Additionally, clusters whose temporal distribution had a similar pattern were merged into new superclusters. In fact, superclusters include spectra of similar shape but with different amplitudes. The final plots contain the temporal distribution of the merged clusters. Fig. 7 shows the temporal distributions of the particles in (some) superclusters, while Fig. 8 shows centers of the clusters that contribute to each supercluster.

Comparison of Fig. 7 with Fig. 2 does not reveal such clear correspondence as in the laboratory experiment. We can only propose some tentative interpretations. For example, cluster 6, spanning from mid-May to early June, looks consistent with the *Pinus* pollen. Cluster 3, mostly concentrated in the time interval from mid-April to mid-May, could be associated with *Betula pollen*. However, the peak in early April could indicate that *Acer* or *Corylus pollen* is mixed in this group. Cluster 2 is close in time with *Acer*. Cluster 5 is more prevalent during the summer and could contain grass pollen. Interestingly, the prominent peak in cluster 1 does not exhibit a correspondence with typical pollen in Fig. 2. In contrast, cluster 8 is present uniformly throughout all seasons, and we consider it background particles.

DISCUSSION

In this study, we examined the extent to which fluorescence spectra, measured by the Rapid E+ detector, can be linked to specific pollen types. We applied the *K*-means clustering algorithm to isolate groups of spectra with specific profiles and amplitudes within the whole set of

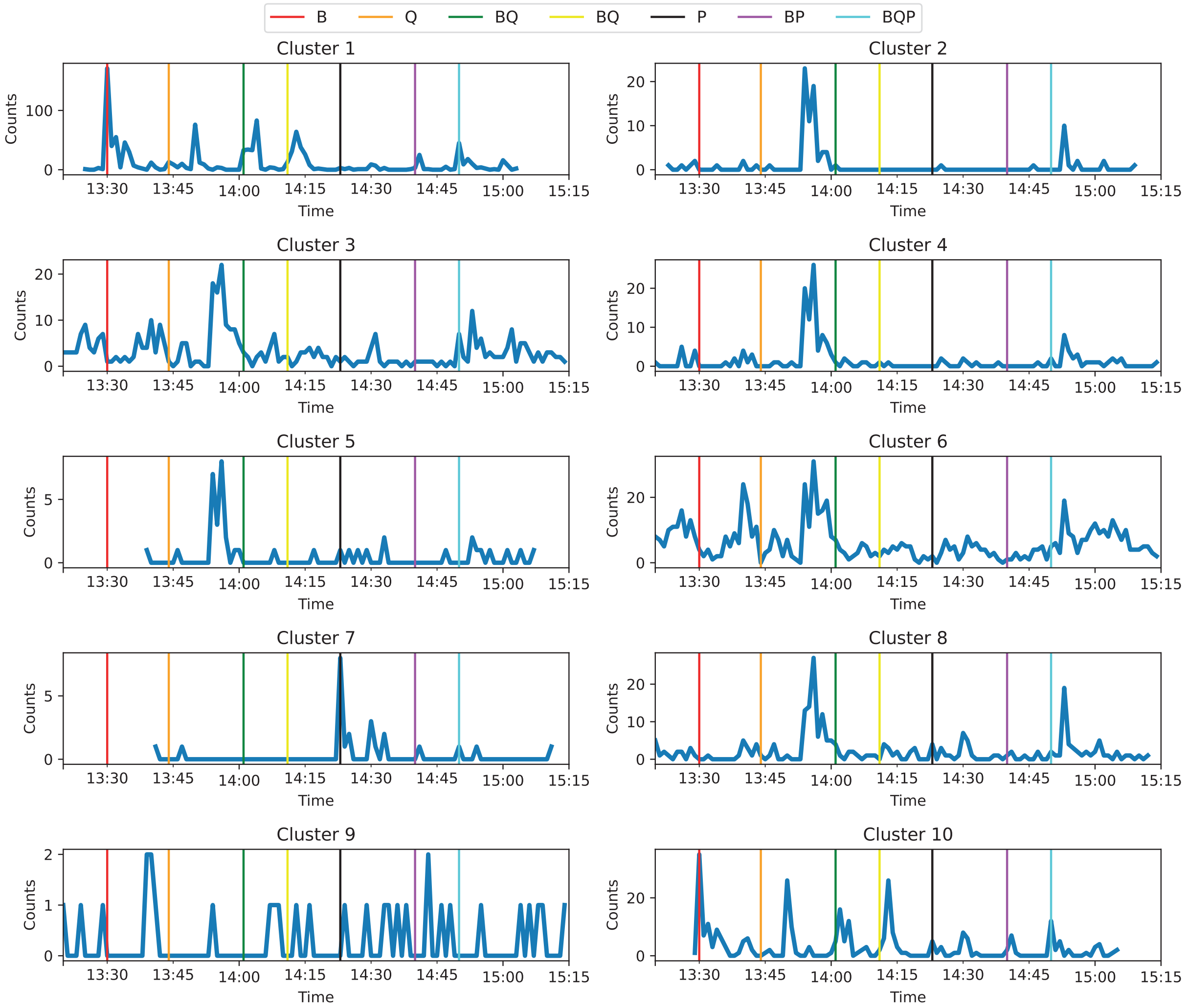


Fig. 6. Temporal distribution of the clusters identified by the *K*-means algorithm in the laboratory experiment

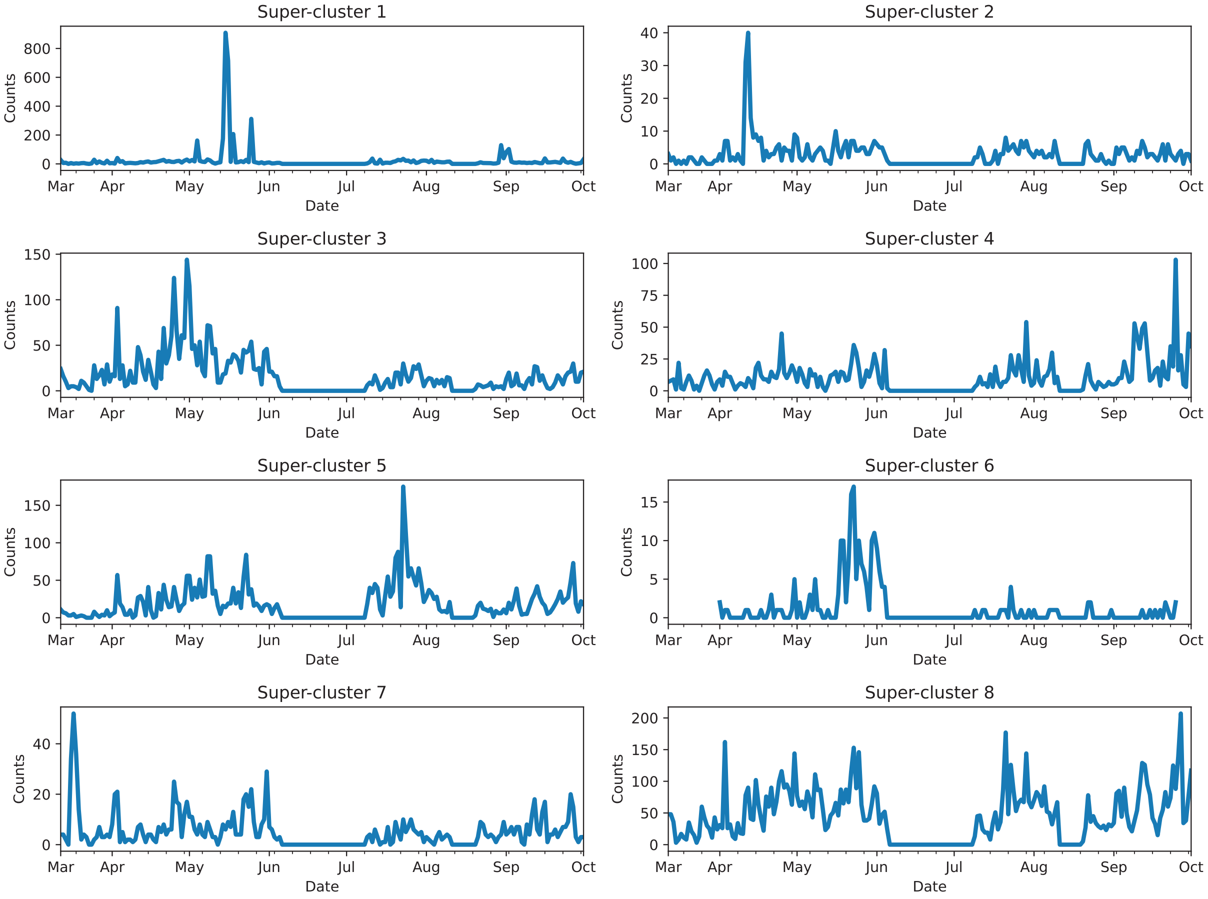


Fig. 7. Temporal distributions of particles in superclusters during the season 2024

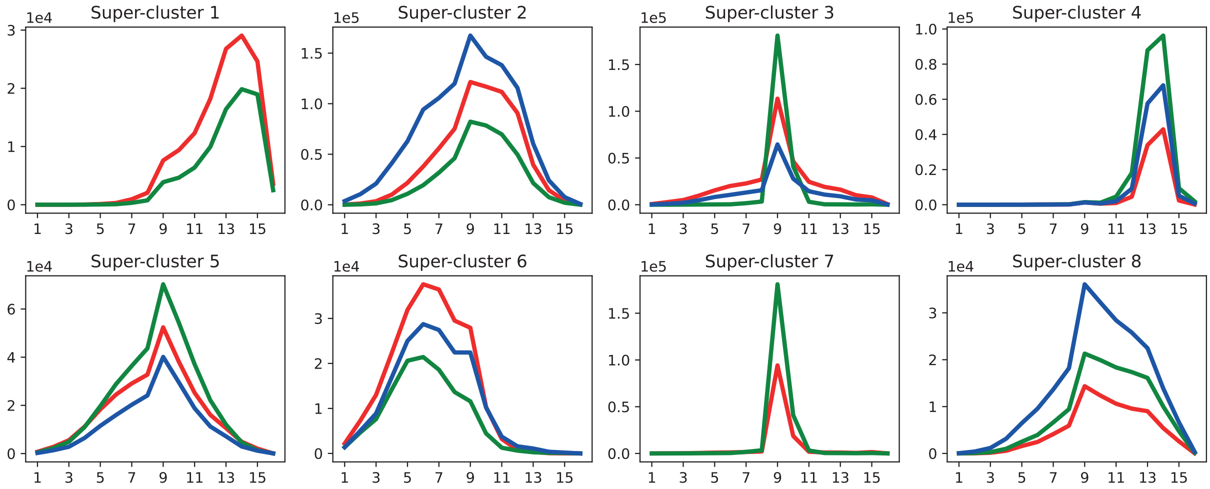


Fig. 8. Centers of the clusters that were merged to the superclusters. Each colored line represents the center of a separate cluster. The number of lines in each plot indicates the number of clusters merged into the super-cluster

measured spectra and analyzed the temporal distribution of the particles of the obtained groups (clusters). In laboratory conditions, when only a few pollen types are investigated and pollen is injected into the detector in portions in a clear form or as a known mixture, the proposed method yields quite straightforward associations. We observe that different pollen types have different spectrum profiles. In contrast, applying this approach to outdoor observation and comparing it with the volumetric data does not reveal as direct an association as one would like to see. We discuss some possible reasons below.

First, the clustering procedure is ambiguous. It depends on how the similarity (distance) between any two objects (see Blanco-Mallo, 2023, for further discussion) is measured, on the particular algorithm of clustering (*K*-means is probably the simplest but not unique one), and on the specific parameters of the particular algorithm (in the case of K-means, it depends on the number of clusters, initial position of cluster centers, and a number of more finer details).

In this study, we used the simplest Euclidean distance metric (sum of the squared differences between the corresponding components of the two vectors). Other distance metrics require modification of the *K*-means clustering algorithm itself (see, e.g., Kaufman & Rousseeuw, 1990). Actually, various clustering algorithms and various metrics were investigated, but qualitatively, the results appear similar.

The clustering results also depend on data preprocessing. For example, data normalization could help join spectra with similar shapes but different amplitudes into the same cluster. However, in fact this process often leads to joining in the same cluster spectra with different shapes, since it becomes more difficult to distinguish data by shape alone, rather than by shape and amplitude. For simplicity, we did not use any data preprocessing.

Instead, it turned out to be more practical to isolate a large number of clusters (much larger than the potential number of expected pollen types) and then join clusters that yield similar temporal distribution into superclusters (see, e.g., Xu et al., 2016, for possible estimates of the optimal number of clusters). In this research, we created superclusters manually based on visual analysis, but this step can also be automated using, say, clustering the time series.

Another thing to discuss is the correspondence of the clusters obtained in the laboratory experiment and during outdoor monitoring. Pollen used in the laboratory experiment is dry, while in outdoor monitoring, a mixture of dry and hydrated pollen may be observed. We do not know how the degree of hydration affects the fluorescence spectrum, so we leave it out of this research.

It should also be noted that outdoor measurements contain spectra of a much broader set of particle types than is reported in standard monitoring. In particular, we observe prominent clusters that are very localized in time but do not correspond to any of the pollen types reported in standard monitoring.

Finally, we used only a small part of the measurements provided by the automatic detector. It looks more than reasonable that finer separation of pollen types requires a combination of measured parameters. Furthermore, the known mass ratios of the pollen types used in laboratory experiments could be used in the future to estimate model accuracy in more detail.

CONCLUSIONS

In this work, we proposed a method for identification of pollen types using their fluorescence spectra measured by the automatic atmospheric particle detector Rapid-E+. The method is based on clustering of all fluorescence spectra measured by the detector and comparison of temporal patterns of occurrence of particles in each cluster with temporal patterns of occurrence of specific pollen types measured by the standard volumetric method. In a set of laboratory and outdoor experiments, it was demonstrated that fluorescence spectra are clustered into a set of groups that yield different distributions in time and that at least some of these groups are associated with specific pollen types. The obtained association can be used in the future to create an automatic pollen identification system. However, fluorescence data alone are insufficient for fine differentiation of all pollen types, and in fact, much larger groups representing mixtures of different pollen types are identified. At the same time, among these larger groups we observe those that are not identified by volumetric monitoring, and understanding the source of these particles is a matter of further research. ■

REFERENCES

Akdis C. A., Hellings P. W., Agache I. (Eds.) (2015). European Academy of Allergy and Clinical Immunology. Global atlas of allergic rhinitis and chronic rhinosinusitis, EAACI, Zürich.

An T., Liang Z., Chen Z., Li G. (2024). Recent progress in online detection methods of bioaerosols. Fundamental Research, 4 (3), DOI: 10.1016/j.fmre.2023.05.012.

Bishop, C. M. (2007). Pattern Recognition and Machine Learning (Information Science and Statistics), Springer.

Blanco-Mallo E., Morán-Fernández L., Remeseiro B., Bolón-Canedo V. (2023), Do all roads lead to Rome? Studying distance measures in the context of machine learning. Pattern Recognition, 141, DOI: 10.1016/j.patcog.2023.109646.

Boldeanu M., Burileanu C., Cucu H., Luminita Marmureanu L. (2022): Pollen classification using classical ML algorithms on fluorescence and scattering data.. U.P.B. Sci. Bull., Series C, 84 (4).

Brdar S., Panić M., Matavulj P. et al. (2023). Explainable AI for unveiling deep learning pollen classification model based on fusion of scattered light patterns and fluorescence spectroscopy. Sci Rep 13, 3205, DOI: 10.1038/s41598-023-30064-6.

Buters J. T., Antunes C., Galveias A. et al. (2018). Pollen and spore monitoring in the world. Clinical and translational allergy, 8, 1-5. DOI:10.1186/s13601-018-0197-8

Crouzy B., Stella M., Konzelmann T., Calpini B., Clot B. (2016). All-optical automatic pollen identification: Towards an operational system. Atmospheric Environment, 140, DOI: 10.1016/j.atmosenv.2016.05.062.

D'Amato G. (2000). Urban air pollution and plant-derived respiratory allergy. Clin. Exp. Allergy, 30, 628–636. DOI: 10.1046/j.1365-2222.2000.00798.x.

D'Amato,G., Vitale C., Sanduzzi A., Molino A., Vatrella,A., D'Amato M. (2017). Allergenic pollen and pollen allergy in Europe. Allergy and allergen immunotherapy, 287-306. DOI:10.1111/j.1398-9995.2007.01393.x.

Daunys G., Šukienė L., Vaitkevičius L. et al. (2021). Clustering approach for the analysis of the fluorescent bioaerosol collected by an automatic detector. PLoS One, 16 (3), DOI: 10.1371/journal.pone.0247284.

Dykewicz M. S., Hamilos D. L. (2010). Rhinitis and sinusitis. Journal of Allergy and Clinical Immunology, 125 (2), 103-115, DOI: 10.1016/j.jaci.2009.12.989.

Erb, S., Graf, E., Zeder, Y., Lionetti, S., Berne, A., Clot, B., Lieberherr, G., Tummon, F., Wullschleger, P., and Crouzy, B. (2024). Real-time pollen identification using holographic imaging and fluorescence measurements. Atmos. Meas. Tech., 17, 441–451, DOI: 10.5194/amt-17-441-2024, 2024.

Fröhlich-Nowoisky J., Kampf C. J., Weber B. et al. (2016). Bioaerosols in the Earth system: Climate, health, and ecosystem interactions. Atmospheric Research, 182, 346-376, DOI: 10.1016/j.atmosres.2016.07.018

Galán C., Smith M., Thibaudon M. et al. (2014). Pollen monitoring: Minimum requirements and reproducibility of analysis. Aerobiologia, 30 (4), 385–395. DOI: 10.1007/s10453-014-9335-5

Hirst J. M. (1952). An automatic volumetric spore trap. The Annals of Applied Biology, 39 (2), 257–265.

Huffman J. A., Perring A. E., Savage N. J. et al. (2019). Real-time sensing of bioaerosols: Review and current perspectives. Aerosol Science and Technology, 54 (5), DOI: 10.1080/02786826.2019.1664724.

Kabir E., Azzouz A., Raza N. et al. (2020). Recent Advances in Monitoring, Sampling, and Sensing Techniques for Bioaerosols in the Atmosphere. ACS Sensors, 5 (5), 1254-1267, DOI: 10.1021/acssensors.9b02585.

Kaufman, L. and Rousseeuw, P.J. (1990). Partitioning Around Medoids (Program PAM). In Finding Groups in Data (eds L. Kaufman and P.J. Rousseeuw). DOI: 10.1002/9780470316801.ch2

Matavulj P., Cristofori A., Cristofolini F., Gottardini E., Brdar S., Sikoparija B. (2022). Integration of reference data from different Rapid-E devices supports automatic pollen detection in more locations. Science of The Total Environment, 851 (2), DOI: 10.1016/j.scitotenv.2022.158234.

Matavulj P., Jelic S., Severdija D., et al. (2025). Domain adaptation for improving automatic airborne pollen classification with expert-verified measurements. Appl Intell 55, 430, DOI: 10.1007/s10489-024-06021-9.

Maya-Manzano J. M., Smith M., Markey E., Hourihane Clancy J., Sodeau J., O´Connor D. J. (2020). Recent developments in monitoring and modelling airborne pollen, a review. Grana, 60 (1), 1–19. DOI: 10.1080/00173134.2020.1769176.

Oteros J., Buters J., Laven G., Roseler S. et al. (2017). Errors in determining the flow rate of Hirst-type pollen traps. Aerobiologia, 33, 201–210. DOI: 10.1007/s10453-016-9467-x, 2017.

Šaulienė I., Šukiene L., Daunys G. et al. (2019). Automatic pollen recognition with the Rapid-E particle counter: the first-level procedure, experience and next steps. Atmospheric Measurement Techniques, 12, DOI: 10.5194/amt-12-3435-2019.

Sikoparija B., Galán C., Smith M. et al. (2017). Pollen-monitoring: between analyst proficiency testing. Aerobiologia, 33, 191–199. DOI:10.1007/s10453-016-9461-3

Šikoparija B., Matavulj P., Mimić G., Smith M., Grewling L., Podraščanin Z. (2022). Real-time automatic detection of starch particles in ambient air. Agricultural and Forest Meteorology, 323, DOI: 10.1016/j.agrformet.2022.109034.

Sofiev M., Bergman K.-C. (Eds.) (2012). Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts. Springer.

Swanson, B. E., Huffman, J. A. (2020). Pollen clustering strategies using a newly developed single-particle fluorescence spectrometer. Aerosol Science and Technology, 54 (4), 426–445, DOI: 10.1080/02786826.2019.1711357.

Valenta R., Vrtala S., Ebner C., Kraft D., Scheiner O. (1992). Diagnosis of grass pollen allergy with recombinant timothy grass (Phleum pratense) pollen allergens. Int. Arch. Allergy Immunol., 97, 287–294, DOI: 10.1159/000236135.

Xu S., Qiao X., Zhang Y., Xue C., Li L. (2016). Reviews on Determining the Number of Clusters. Applied Mathematics & Information Sciences, 10, 1493-1512, DOI: 10.18576/amis/100428.

RESEARCH PAPER

SOIL ORGANIC MATTER MINERALIZATION AND TRANSFORMATION IN INNER SHELF OASES OF EAST ANTARCTICA: LABORATORY ASSESSMENTS AND ROLE OF ENVIRONMENTAL DRIVERS

Ivan I. Alekseev[1,3]*, Elena N. Grek[2]
[1]Arctic and Antarctic Research Institute, Bering str. 38A, Saint Petersburg, 199397, Russia
[2]State Hydrological Institute, 2nd line of Vasilevsky island 23, Saint Petersburg, 199004, Russia
[3]Karelian Research Centre of the Russian Academy of Science, Pushkinskaya str. 11, Petrozavodsk, 185000, Russia
*Corresponding author: alekseevivan95@gmail.com

Received: June 25th 2024 / Accepted: December 5th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4107

ABSTRACT. Soils of ice-free regions (oases) in East Antarctica have been rarely investigated via the prism of combined detailed biological and chemical methods. The main purpose of this work is to assess the influence of environmental factors, including microrelief and local features of the accumulation of biogenic material on the mineralization and humification of the organic matter in topsoils from Bunger Hills and Schirmacher Oasis. We used multiple techniques to analyze soil basal respiration and mineralization rates, extract humic acids, and investigate their molecular structure by ^{13}C-NMR spectroscopy.

The results reveal that soil formed under moss cushions exhibits the highest levels of organic carbon (up to 2.43%), alongside elevated basal respiration rates, which reached up to 2.979±0.015 mg g^{-1} day^{-1}in moist conditions. Soil pH ranged from slightly acidic to alkaline, influenced by salt accumulation, which adversely affects plant communities and limits biomass production. Using ^{13}C-NMR spectroscopy, although in a limited number of samples, we identified a predominance of aliphatic structures in humic acids with carboxylic functional groups, indicating significant vegetation influence on organic matter complexity. The sp^2/sp^3 carbon ratios (0.724 for Bunger Hills and 0.408 for Schirmacher Oasis) indicate a balance between potentially decomposable aliphatic and stable aromatic structures. Mineralization rates were significantly higher in soils with greater moisture retention, with cumulative CO_2 release reaching up to 150 mg CO_2/kg^{-1} soil day^{-1} over a 30-day period. Overall, our work revealed complex relationships between the environmental conditions and soil characteristics that significantly influence biological activity, carbon storage, and organic matter structure.

KEYWORDS: carbon sequestration, Antarctic, soil organic matter, humic acids, ^{13}C-NMR

CITATION: Alekseev I. I., Grek E. N. (2026). Soil Organic Matter Mineralization And Transformation In Inner Shelf Oases Of East Antarctica: Laboratory Assessments And Role Of Environmental Drivers. Geography, Environment, Sustainability, 1 (19), 36-50
https://doi.org/10.24057/2071-9388-2026-4107

ACKNOWLEDGEMENTS: The authors would like to acknowledge the Russian Antarctic Expedition, Antonina Chetverova (Arctic and Antarctic Research Institute) and colleagues for their crucial help in fieldwork.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Ice-free areas in Antarctica, known as oases, nunataks, and dry valleys, cover approximately 55,000 km^2 and serve as critical sites for studying soil formation processes (Bockheim et al. 2008; Convey et al. 2014). Despite its remoteness and harsh climate, Antarctic soils are vital for understanding broader ecological dynamics and the impacts of climate change (Hodgson et al. 2014). Historically, Antarctic soils have been underinvestigated compared to other global regions. Pioneering studies by Ugolini et al. (1982) and Tedrow and Ugolini (1966) laid the groundwork for our understanding of these unique ecosystems. More recent investigations by Michel et al. (2014), Bockheim et al. (2015), Mergelov et al. (2015), and Lupachev et al. (2020) have further illuminated the diversity and complexity of Antarctic soils. However, significant gaps remain in our knowledge, particularly regarding the intricate relationships between soil composition, microbial communities, and environmental factors (Cary et al. 2010; Elberling et al. 2013). Due to their geographically unique positions, the lowland Schirmacher Oasis and Bunger Hills are located along the coast in the ablation zone. The land areas and internal marine bays (epishelf water bodies) that make up these oases are separated from the open ocean by shelf glaciers that can be tens to hundreds of kilometers wide, and are bounded on the opposite side by outlet glaciers or the continental ice sheet. Due to the properties of the underlying surfaces in the oases (soil, rocks, water bodies), air temperatures throughout the year are on average 1-2 °C higher than the the surrounding snow and ice areas; soil surfaces in the oases can warm to +41 °C, while water surfaces can reach +17 °C (Gore and Leishman 2020). The oases also exhibit much more extreme temperature values than the surrounding glaciers (Kruchinin and Simonov 1967; Rusin 1958; Rusin 1961; Simonov 1971). The high frequency of strong winds (up to hurricane strength), low humidity, and limited precipitation (less than 200 mm per year) combined with high evaporation rates (over 500 mm) are also distinctive features of these lowland shelf oases.

In environments characterized by minimal decomposition and humification of organic residues, coupled with a limited duration of biological activity, soil formation processes are inherently slow. East Antarctic oases are characterized by the complete absence of typical soil-forming agents and humus sources such as vascular plants. Instead, these regions are dominated by mosses, lichens, and cyanobacteria, which create specialized microhabitats that support the development of unique microbial communities. Within these extreme conditions, soil microorganisms assume a central role in soil genesis, owing to their remarkable adaptations that enable survival and function under harsh environmental stresses (Nikitin et al. 2017; Savaglia et al. 2024). Microbial assemblages are fundamental to the functioning of Antarctic soil ecosystems. They are responsible not only for the decomposition of organic matter but also for facilitating nutrient cycling and contributing to the formation and stabilization of soil structure (Cary et al. 2010; Elberling et al. 2013). Recent investigations have emphasized the critical importance of understanding microbial community dynamics in relation to climate change, particularly concerning permafrost thawing and the consequent release of greenhouse gases (Schuur et al. 2015; Turetsky et al. 2020). The sensitivity of microbial activity to temperature variations is especially noteworthy; studies have demonstrated that microbial metabolic rates tend to increase with rising temperatures. Nonetheless, this response varies significantly among different microbial taxa and communities (Rinnan et al. 2009). Microorganisms adapted to cold environments generally exhibit lower temperature sensitivity compared to those from warmer regions (Nikitin et al. 2017) . This differential responsiveness underscores the potential vulnerability of Antarctic microbial ecosystems as global temperatures continue to rise.

To better understand the effects of permafrost degradation and the potential release of greenhouse gases, precise knowledge of the spatial distribution of soil organic matter is needed, both in terms of quantity and quality (e.g. biodegradation, chemical composition, and degree of humification). Different authors used various indicators to identify the rate of humification in polar soils and the stability of the process, including the level of aromaticity. The rate of humification in soils can be inferred from ^{13}C-NMR by observing changes in the relative proportions of different carbon structures, as more decomposed humic substances, indicating higher humification, typically show increased aromaticity and decreasing alkyl and O-alkyl carbon content. Therefore, a higher proportion of aromatic carbon and a lower proportion of alkyl and O-alkyl carbon in ^{13}C-NMR spectra of soil organic matter suggest a higher rate and degree of humification (Vasilevich et al. 2018; Knicker 2006). The structural and molecular composition of soil organic matter in Antarctica has been studied previously by ^{13}C-NMR spectroscopy (Alekseev and Abakumov 2024; Beyer et al. 1997, Abakumov and Alekseev 2018).

The assessment of carbon dioxide (CO_2) emissions in laboratory settings provides researchers with the ability to control various environmental factors that influence these emissions, such as temperature (Carvalho et al. 2010). However, it is important to acknowledge that this approach has inherent limitations, including potential inaccuracies and significant time requirements. Laboratory conditions frequently alter the natural soil structure, significantly impacting microbial processes (Oertel et al. 2016). Typically, laboratory assessments focus primarily on respiration associated with the mineralization of soil organic matter (Carvalho et al. 2010; Oertel et al. 2016; Thomazini et al. 2016). Consequently, findings derived from laboratory studies cannot be directly correlated with in situ measurements, despite occasional similarities in observed trends. In contrast, *in situ* investigations encompass both autotrophic and heterotrophic (microbial) components contributing to ecosystem respiration (Oertel et al. 2016; Thomazini et al. 2016). Therefore, while laboratory experiments enable researchers to effectively evaluate the temperature sensitivity of soil organic matter mineralization under warming scenarios, they also facilitate estimates of CO_2 production potentials (Pires et al. 2017).

In the context of East Antarctica, it is evident that soil biological activity is generally low and heavily influenced by microclimatic conditions. Factors such as soil temperature and moisture levels critically affect primary productivity and carbon sequestration (Thomazini et al. 2016). The primary sources of soil organic carbon in ice-free regions predominantly include plants – such as mosses, lichens, and algae – as well as marine animals and birds (Michel et al. 2006; Zhu et al. 2014). Furthermore, bacterial mats occur in areas where shallow ponds previously existed, resulting in accumulations of minimally decomposed cyanobacteria that contribute insignificantly to terrestrial organic carbon reservoirs. Notably, the organic contributions from algae and lichens are substantially lower than those from bryophytes, which serve as the principal source of organic material in this region (Claridge et al. 2000). Additionally, carbon pools associated with Antarctic ice sheets or snow are rare due to the near absence of organic life forms (Claridge et al. 2000). Nonetheless, a minor influx of microbial organic material can enhance carbon storage capabilities within these soils (Claridge et al. 2000). Plants play a crucial role in facilitating carbon uptake through photosynthesis, while birds contribute locally through significant guano deposits (Michel et al. 2006; Thomazini et al. 2016). Moreover, substantial inputs of organic matter—particularly from guano—are recognized as critical processes driving soil formation and increasing organic carbon stocks within these ecosystems (Simas et al. 2007). The immobilization of carbon within soils through active incorporation by biomass and microorganisms not only fosters carbon sequestration but also enhances water retention capacity while influencing nutrient bioavailability and heavy metal dynamics (Schaefer et al. 2008; Thomazini et al. 2016).

Field studies combined with controlled laboratory experiments are commonly used to assess losses associated with organic carbon emissions (Oertel et al. 2016; Thomazini et al. 2016). In particular, potential mineralization losses from soils can be effectively evaluated through meticulously controlled laboratory settings. These methodologies are essential for analyzing the apparent stability of soil organic matter alongside its mineralization losses while predicting potential stability and losses within polar environments. However, it is crucial to recognize that data obtained from controlled laboratory experiments cannot be directly compared with those derived from field studies – such as those utilizing closed chamber methodologies. Nevertheless, data collected under standardized conditions can provide valuable insights for subsequent simulation modeling regarding SOM transformation. Such studies are essential for evaluating the stability of organic matter and the potential for mineralization processes. It is important to emphasize that results from these laboratory experiments cannot be directly equated with those obtained from field studies, including methods like the closed chamber approach (Anderson 1982). However, the findings outlined in this section were derived from consistent laboratory conditions, which may enable their use in future modeling efforts related to the transformation of soil organic matter (Schmidt et al. 2011).

Previous research revealed that increased temperature sensitivity of microbial communities is linked with elevated mean annual soil temperature, thus proving the idea that microbial communities from colder regions are less temperature sensitive compared to warmer analogues (Rinnan et al. 2009).

Considering the relatively little attention given to the investigation of mineralization and microbial communities in soils of East Antarctica (especially in remote inner shelf oasis), our research is contributing crucial data on laboratory assessment of soil respiration rates and stabilization of organic matter in various soil types of rarely investigated areas of Bunger Hills and Schirmacher Oasis. In turn, this data is important for modelling of greenhouse gas emissions, carbon turnover, and soil organic matter stability under changing environmental conditions of Antarctica. The following objectives were chosen:

– Detailed studying of physical-chemical parameters of soils investigated across landscapes of ice-free areas of Bunger Hills and Schirmacher Oasis;

– Investigate the molecular composition of humic acids isolated from various soil types investigated in Bunger Hills and Schirmacher Oasis using 13C-NMR spectroscopy.

– Assess the soil organic matter mineralization rates in various soil types of Bunger Hills and Schirmacher Oasis using a laboratory incubation experiment.

MATERIALS AND METHODS

Soil samples for this study have been collected during the fieldwork of the 69th Russian Antarctic Expedition (February–March 2024) in two ice-free areas of East Antarctica – Bunger Hills and Schirmacher Oasis (Fig. 1). The main climatic parameters of the studied areas are provided in Table 1. In the field, soils were sampled from 20×20 cm soil pits from different depths. The samples were stored in double sterile plastic bags, labeled, and transported to the laboratory (Arctic and Antarctic Research Institute, Saint Petersburg, Russia).

Regional setting

Bunger Hills

Bunger Hills region is characterized by a complex geomorphological landscape comprising numerous islands and inlets situated on the inner continental shelf, overlain by morainic deposits. Geologically, the area is predominantly composed of metamorphic and igneous rocks, including gneiss, granite, and migmatite, reflecting its intricate tectonic history (Gore and Leishman 2020). A distinctive feature of this region is the pervasive presence of surface salts originating from marine aerosols, notably halite (NaCl) and thenardite (Na_2SO_4), which are observed on both lithic and soil surfaces (Gore and Leishman 2020). These salts are mobilized and redistributed during the austral summer months through aeolian processes; salt-laden aerosols are transported west-northwestward by prevailing winds emanating from adjacent marine inlets

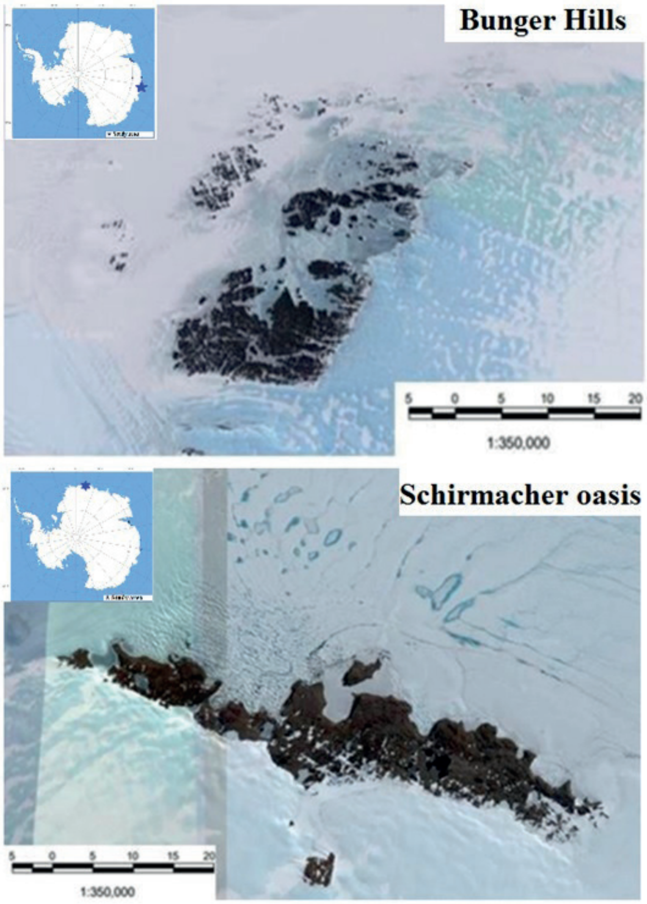


Fig. 1. Location and maps of Bunger Hills and Schirmacher Oasis

and hypersaline lakes. Wind-driven salt deposition results in the accumulation of evaporitic minerals on terrestrial surfaces located downwind of saline sources. Biotic colonization within the Bunger Hills is limited due to extreme environmental conditions; nonetheless, biological communities are represented primarily by lichens and mosses. The character of mosses' and lichens' distribution in Bunger Hills depends mostly on environmental conditions, particularly soil salinity and the availability of habitat and water (Leishman et al. 2020). Lichens are mostly represented by *Buellia frigida, Umbllicaria decussata, U. aprina, Xanthoria candelaria, X. elegans, Physcia caesia, Candelariella antarctica, Pseudephebe minuscula, Pertusaria globulifera, Buellia ligioides, Lecidea lapicida, Rhizoplaca melanophthalma, Lecanora expectans, L. polytropa, Rhizocarpon flavum, Rinodina petermannii, Usnea acromelaena, U. antarctica, Acarospora gwynnii,* and *A. Petalina* (Andreev, 1990). Mosses often described growing on soils and are represented by Bryum algens *Card., Bryum argenteum Hedw., Ceratodon purpureus (Hedw.) Brid., Schistidium antarcticum (Card.) Sav.-Lyub. et Z. Smirn.*

These organisms are predominantly confined to meltwater channels – often associated with periglacial cracking phenomena – and to south-facing rock slopes in proximity to the Apfel Glacier (Leishman et al. 2020). These microhabitats provide critical refugia for extremophile communities adapted to survive under harsh polar conditions. The region also hosts a substantial lacustrine system, including Algae Lake, which ranks among the largest freshwater lakes in Antarctica and holds ecological significance for regional hydrology and biogeochemical cycles (Klokov and Verkulich 1994). Furthermore, the area supports an extensive terrestrial drainage network, recognized as the third-longest in Antarctica, facilitating surface runoff and subsurface flow pathways that influence landscape evolution and sediment transport processes (Gibson et al. 2002). The lichenometric method, based on the size of Buellia frigida, delineated that deglaciation was occurring from the center of the oasis in all directions towards its modern edges (Bolshiyanov and Verkulich 1992).

Schirmacher Oasis

The Schirmacher Oasis is an ice-free area on the border between the mainland and the ice shelf. The oasis is about 17 km long; it is stretched as a narrow three-kilometer strip in the direction from west-northwest to east-southeast. The oasis is separated from the Lazarev Sea (80 km), to the north by the Nivlisen ice shelf. On the south side, the oasis is bordered by the continental ice sheet, with the Wohlthat mountain range located on it (100 km). The most intense cyclonic winds from the east and southeast are dominant, which is accompanied by significant clouds, blizzards, snowfalls, and gale-force winds. Katabatic winds often blow, and the wind from the southeast direction causes a sharp decline in air temperature and wind speed, combined with clear weather and a decrease in air humidity. The oasis is composed mainly of Precambrian age strata consisting of gneisses and crystalline schists. The oasis is characterized by a rather weak development of cellular forms of weathering and desquamation tiles on rock surfaces, as well as fairly recent traces of glacial impact, which indicates a relatively recent deglaciation, according to various estimates from 7 to 10 thousand years ago (Verkulich 2007). The vegetation of the oasis is exceptionally poor and is represented by some rare patches of lichens and mosses on rock substrates and fine earth. This decline is due to the harsh thermal regime, acute lack of liquid precipitation, low humidity, and frequent and strong winds, which create extremely unfavorable conditions for plant development (Kurbatova and Ochyra 2012; Singh et al. 2008). The frequency of species occurrence and changes in species composition varied across different locations. The lichen species growing on soil-moraine and moss habitats in Schirmacher Oasis include *Acarospora williamsii Filson, Buellia grimmiae Filson, Caloplaca citrina (Hoffm.) Th. Fr., Candelariella flava (C. W. Dodge & Baker) Castello & Nimis, Lepraria cacuminum (A. Massal.) Lohtander, Lecanora expectans Darb., Lecanora geophila (Th. Fr.) Poelt, Lecidella siplei (CW Dodge & GE Baker) May. Inoue, Lecidella sp. B, Physcia caesia (Hoffm.) Furnr. Rinodina olivaceobrunea CW Dodge & GE Baker.* Mosses are most abundant along soil habitats near water bodies and meltwater streams and include *Bryum argenteum Hedw. var. muticum Brid., Bryum archangelicum Bruch & Schimp, Bryum pseudotriquetrum (Hedw.) P. Gaertn., B. Mey. & Scherb, Ceratodon purpureus (Hedw.) Brid, Orthogrimmia sessitana (De Not.) Ochyra & Zarnowiec, Syntrichia sarconeurum Ochyra & RH Zander.* Recent observations by ornithologists showed only Snow Petrels *(Pagodroma nivea)* and Wilson's Storm Petrels *(Oceanites oceanicus)* were reported breeding in the past (no current observation) with only South Polar skua *(Stercorarius maccormicki)* confirmed breeding in recent years (Ryan 2024). And, despite ornithogenic effect in Schirmacher Oasis being believed to be minor, new data obtained during current fieldwork revealed the appearance of breeding colony of Adelie penguins *Pygoscelis adeliae* (20 breeding pairs were found in December 2024–January 2025) (Galustov et al. 2025).

Laboratory analyses and procedures

The soil samples were air-dried at room temperature and subsequently passed through a 2 mm plastic sieve prior to chemical analysis. To determine the pH, a pH meter was used (pH-150 M), utilizing a soil-to-solution ratio of 1:2.5. The soil organic carbon, nitrogen, and hydrogen content was quantified using a Vario EL III analyzer (Elementar, Germany). For the assessment of particle size distribution, the pipette-sedimentation method described by Kachinsky (1958) was used.

Soil mineralization rates were obtained through a controlled laboratory incubation experiment, which was conducted in five replicates for each soil horizon. The CO_2

Table 1. Climatic parameters of the study areas (data from ROSHYDROMET, http://www.wmo.int/pages/prog/www/Antarctica)

MAAT, °C	MAST, °C	MAGT, °C	Mean annual wind speed, m s^{-1}	Annual precipitation, mm (in liquid equivalent)
Bunger Hills				
-9	-8,8	-7,8	6,8	204
Schirmacher Oasis				
-10,5	-10,0	-10,0	9,7	264,5

respiration rate emissions were measured following the methodology outlined by Anderson (1982). Specifically, plastic cylinders with dimensions of 10 cm in diameter and 20 cm in height were employed for this purpose. During the controlled laboratory incubation, 10 g of mineral soil and 5 g of organic material were placed within the cylinders. Subsequently, these cylinders containing soil samples – maintaining a moisture content of approximately 60% of their water holding capacity – were positioned in beakers filled with water at the bottom to simulate field conditions typical of paddy environments. Throughout the incubation period, the water content (%) was monitored to ensure consistency.

Following this initial phase, aliquots of the soil samples were transferred to sealed plastic bottles containing 1 M NaOH and maintained at a temperature of 20°C. The amount of CO_2 trapped by the alkaline solution was quantified through titration after a period of 7 days of incubation. The overall incubation duration lasted for 112 days, during which duplicate measurements were conducted for 10% of the samples to ensure reliability.

We followed the standard methods set out by the International Humic Substances Society (IHSS 1981) to get humic acids (HAs) from topsoil horizons. We used Swift (1996) to get humic substances and the residue. Initially, HAs from the studied soils were air-dried before recording their ^{13}C-NMR spectra using a JNM-ECA 400 NMR spectrometer (JEOL, Tokyo, Japan). This analysis employed the Cross-Polarization Magic Angle Spinning procedure at a spinning speed of 6 kHz with a contact time of 5 ms and a recycle delay of 5 s. Adamantane (29.46 ppm) served as the reference material for chemical shifts. Data corrections were applied for water and ash content; additionally, oxygen content was calculated based on the difference between sample mass and gravimetric concentrations of carbon (C), nitrogen (N), hydrogen (H), and ash. Leonardite HA standard (1S104H) and Elliot soil HA standard (1S102H) were utilized as reference materials during this analysis. The acquired ^{13}C-NMR spectra underwent transformation into quantitative forms using Fourier transformation techniques. Fragment matching was performed according to literature data (Kechaykina et al. 2011; Chefetz et al. 2002), based on specific chemical shift ranges as detailed in Table 2. Data processing was conducted using MestreNova software (Mestrelab Research S.L., Spain).

The sp^2/sp^3 carbon ratio was calculated according to Eq. (1) (Mao et al. 2000):

$$sp^2/sp^3 = (\text{area } (108\text{–}220 \text{ ppm}))/(\text{area } (0\text{–}108 \text{ ppm})) \quad (1)$$

where sp^2 is sp^2 hybridized carbons; sp^3 is sp^3 hybridized carbons.

RESULTS

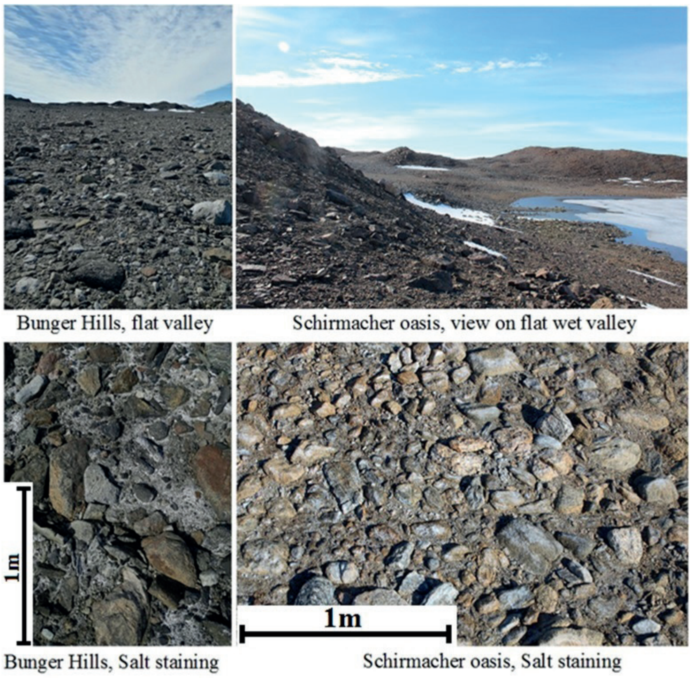
Soil morphology

Both in Bunger Hills and Schirmacher Oasis, an extensive part of the ice-free area is occupied by inter-ridge valleys (Fig.2), where cryoturbated soils with high gravel content (less often sandy) predominate – more often Leptosols are identified (Table 2). These soils are formed in the complete absence of higher plants, so organogenic horizons are formed mainly of microbial and cryptogamic autotrophs. In some cases, organogenic horizons are not formed at all, leading to negligible TOC content (Mergelov et al. 2018). These soils are especially widespread in parts of the oases where surface accumulation of soluble salts is observed. Wind erosion is actively involved in the formation of this type of soil, which contributes to the corrasion of upper horizons (Tedrow and Ugolini 1966). In addition, most of the landscapes are occupied by bedrock and aggregations of coarse-grained material, on which only epilithic and endolithic soil-like bodies are developed, which were previously described (Mergelov et al. 2018; Friedmann et al. 1982).

In inter-ridge valleys, where Leptosols are mainly found, the main factor determining the development of biotic complexes on the soil surface is the diversity of micro- and mesorelief forms - wind shelters can exist on the sides of valleys and in rock cavities, which contributes to the formation of perennial snowfields and the redistribution of moisture (Dolgikh et al. 2015; Zazovskaya et al. 2017). Therefore, differences in the thickness and composition of organogenic horizons are determined primarily by the distribution of melting waters over meso- and microrelief (Fig. 3). These can be surface algobacterial complexes of subaqueous soils developing along the shores of lakes with a pulsating water regime, surface or subaerial horizons dominated by mosses, or hidden under gravel pavements - hypolythic algobacterial horizons. At the same time, an interesting feature of amphibian soils under algal-bacterial mats is the development of a saline-type profile (dark gray horizons of sulfide salinity, glaucous peeled and ferruginous-metamorphic horizons).

Table 2. Structural composition of HAs based on chemical shifts on ^{13}C-NMR spectra (Kogel-Knabner 1997)

Chemical Shift (ppm)	Type of Structural Fragment
0-10	Methyl ($–CH_3$) and methylene ($–CH_2$) groups
10-27	H- and C-substituted aliphatic fragments
27-50	CH2-alkyl structures
50-70	Methoxyl ($–OCH_3$) fragments, amino acid groups, and ether groups
70-100	O- and N-substituted aliphatic fragments
100-108	Anomeric carbon in sugars
108-135	Protonated aromatic carbon
135-150	Alkylaromatic groups
150-170	Aromatic carbon of phenols and esters
170-190	Carboxyl (–COOH) and carbonyl (C=O) groups
190-220	Carbonyls in conjugated systems



Bunger Hills, flat valley

Schirmacher oasis, view on flat wet valley

Bunger Hills, Salt staining

Schirmacher oasis, Salt staining

Fig. 2. Typical flat valleys and salt staining manifestations in Bunger Hills and Schirmacher Oasis, East Antarctica

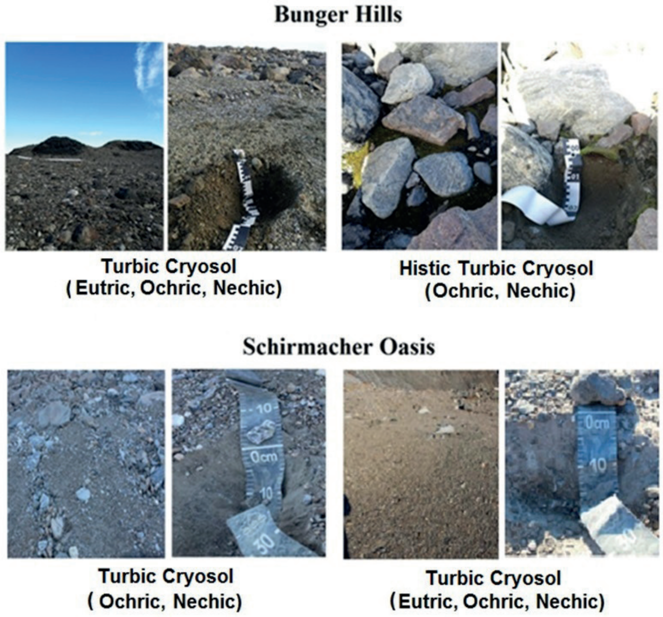


Fig. 3. Studied soil profiles and respective landscapes of Bunger Hills: slope of a valley composed of boulder debris in a sandy matrix (left), wind shelter near the melting snow patch (right); Schirmacher Oasis – patterned ground in the dry valley (left); dry flank of a valley with gravel pavement and salt staining (right)

In Bunger Hills, due to the formation of large flat surfaces of inter-ridge valleys as well as strong katabatic winds, fresh snow is actively blown out (Gore and Leishman 2020). In addition, moraines with high gravel content are very widespread, which leads to greater moisture deficit in the upper soil horizons as well as smaller spatial coverage of organogenic soils compared to other ice-free areas of Antarctica (even compared to East Antarctica oases). Another characteristic feature of Bunger Hills is very limited ornithogenic transfer, which is due to the absence of penguin rookeries and very rare nesting of flying birds. In addition, carbonate-chloride-sulfate and carbonate salts on the surface are extensively represented in the oasis, which also complicates the development of soils (Gore and Leishman 2020). Soils with the most developed vertical profiles are formed in wind shelters and locally in places of ephemeral melting of snow patches with thicker moss-lichen cushions (Fig. 4). However, this soil variety occupies less than 0.1% of the area. An interesting feature of the area is the predominance of loamy parent material, which determines a weak development of hypolythic organogenic horizons together with the occurrence of cryometamorphic soils (characterized by a specific granular structure in the sub-surface horizon).

Soil physico-chemical characteristics

Schirmacher Oasis

The highest levels of organic carbon and nitrogen accumulation, along with the greatest rates of basal respiration, were observed in Histic Turbic Cryosol that develops beneath a moss cushion approximately 3 cm thick, in in a micro-depression where water likely accumulates from melting snowfields during the summer months (Table 3). It is important to note that such conditions are atypical for the harsh climatic environment characteristic of this offshore oasis.

Fig. 4. Moss cushions in Bunger Hills (wind shelter in south-western part) and in Schirmacher Oasis (wet valley in south-eastern part)

Table 3. Soil physico-chemical characteristics in Bunger Hills and Schirmacher Oasis

Soil ID and name	Soil depth, cm	pHH_2O	pHKCl	TOC, %	N, %	H, %	C:N	Basal respiration, mg g $^{-1}$ day^{-1}	Fine earth (<2 mm), %
Bunger Hills									
Bunger 1 Turbic Cryosol	0-2	8.12 ± 0.32	–	0.77 ± 0.10	0.073 ± 0.005	0.21 ± 0.04	10.45 ± 1.27	0.009 ± 0.001	18.69 ± 1.13
Bunger 1 Turbic Cryosol	2-20	7.56 ± 0.21	–	0.86 ± 0.16	0.062 ± 0.005	0.12 ± 0.02	12.26 ± 2.09	0.012 ± 0.001	21.71 ± 1.87
Bunger 2 (wind shelter, moss) Histic Turbic Cryosol	0=1	5.32 ± 0.19	4.87±0.12	3.43 ± 0.43	0.270 ± 0.090	0.49 ± 0.10	14.30 ± 1.23	0.095 ± 0.010	25.71 ± 2.87
Bunger 2 (BC) Histic Turbic Cryosol	1-23	5.76 ± 0.23	5.34 ± 0.12	0.75 ± 0.11	0.071 ± 0.005	0.15 ± 0.03	10.32 ± 1.57	0.009 ± 0.001	21.49 ± 1.14
Bunger 3 (salt staining) Turbic Cryosol	0-1	8.87 ± 0.23	–	0.70 ± 0.14	0.068 ± 0.005	0.17 ± 0.06	11.08 ± 2.12	0.007 ± 0.001	21.11 ± 1.47
Bunger 3 Turbic Cryosol	1-16	7.42 ± 0.11	–	0.70 ± 0.12	0.065 ± 0.005	0.12 ± 0.02	10.56 ± 1.37	0.006 ± 0.001	23.71 ± 1.87
Bunger 6 (moss) Histic Turbic Cryosol	0-2	5.12 ± 0.15	–	2.43 ± 0.23	0.210 ± 0.090	0.31 ± 0.10	11.30 ± 1.23	0.050 ± 0.009	32.71 ± 1.07
Bunger 6 (BC) Histic Turbic Cryosol	2-21	6.12 ± 0.23	–	0.89 ± 0.09	0.091 ± 0.005	0.12 ± 0.03	8.32 ± 1.09	0.021 ± 0.001	22.71 ± 2.17
Schirmacher Oasis									
Novo-3 (moss) Histic Turbic Cryosol	0-3	6.09 ± 0.12	5.87 ± 0.14	2.43 ± 0.12	0.170 ± 0.080	0.49 ± 0.98	14.30 ± 0.51	2.979 ± 0.015	28.71 ± 1.67
Novo-3 (BC) Histic Turbic Cryosol	3-6	5.66 ± 0.15	5.12 ± 0.15	0.50 ± 0.07	0.075 ± 0.010	0.17 ± 0.03	6.63 ± 0.10	0.149 ± 0.010	23.22 ± 1.62
Novo-7 (moss) Histic Turbic Cryosol	0-2	7.70 ± 0.15	–	0.06 ± 0.01	0.019 ± 0.007	0.14 ± 0.02	2.73 ± 0.05	0.195 ± 0.010	31.71 ± 1.85
Novo-7 (BC) Histic Turbic Cryosol	2-8	6.25 ± 0.16	5.96 ± 0.10	0.05 ± 0.01	0.023 ± 0.005	0.16 ± 0.03	2.30 ± 0.12	0.138 ± 0.010	24.22 ± 1.12

*values show mean ± standard deviation

In both the Schirmacher Oasis and the Bunger Hills, instances of surface salt accumulation have been frequently documented, which adversely impacts the development of plant communities, particularly mosses and lichens. Consequently, the spatial extent of areas exhibiting increased biomass production is extremely limited within the studied region of the Schirmacher Oasis. The pH levels of the examined soils range from slightly acidic to acidic in horizons located beneath moss cushions within wind shelters, transitioning to alkaline conditions in areas with active salt accumulation. Elevated soil respiration rates typically corresponded with increased concentrations of organic carbon and nitrogen, as well as a modest rise in fine earth content. The density of the solid phase of the soils increases from surface horizons – where minimum density is characteristic for peaty layers—to central horizons, a trend associated with varying degrees of fine earth presence and its weathering processes. Areas subjected to anthropogenic influence exhibit higher bulk density values.

Bunger Hills

Soil pH in the Bunger Hills exhibits variability, ranging from slightly acidic to acidic in horizons located beneath moss cushions within wind shelters, while alkaline conditions are observed in areas characterized by active salt accumulation. Salt accumulation is influenced by the active weathering of massive crystalline and friable rock formations. Research indicates that the oasis is delineated by a "salt line", which partitions it into two distinct regions, thereby influencing the intensity of rock weathering as well as the distribution and growth of salt-sensitive flora, including mosses and lichens (Gore and Leishman 2020).

Total carbon content within these soils varies considerably, with values ranging from 0.70% to 2.43%. Soils located in wet valleys and wind shelters that support developed moss covers exhibit higher carbon concentrations compared to other soil-like bodies. A notable characteristic of the Bunger Hills is the creation of conditions conducive to a significant deficit of liquid moisture during summer months, even in the upper soil horizons. Consequently, the extent of moistened areas – primarily resulting from snowmelt or ephemeral stream flows – is minimal relative to other previously studied ice-free regions. This phenomenon is attributed to the widespread presence of moraines with high gravel content in valleys and troughs.

Soil basal respirometry

The oases of East Antarctica are generally characterized by much more severe climatic conditions (a short period of biological activity, harmful ultraviolet radiation for macroscopic life forms, strong winds, large temperature fluctuations), as well as a more homogeneous composition of organic residues with a simultaneous decrease in the total organic carbon content. On average, soil carbon dioxide emissions on King George Island and Ardley Island are 0.076-0.122 mg g^{-1} day $^{-1}$, which is higher than the average values described for the oases of East Antarctica (Fig. 5). The highest rates of soil respiration were described for soils developing under Deschampsia Antarctica – up to 0.210 mg g^{-1} day $^{-1}$, which was described earlier (Abakumov and Mukhametova 2014; Thomazini et al. 2016).

Our results showed that the soils of wet valleys, as well as the soils of wind shelters, are the most important participants in the biogeochemical carbon cycle in the coastal and offshore oases of East Antarctica (especially in the almost absence of direct and indirect ornithogenic input of organic matter, which is observed in Bunger Hills). Soil respiration in the oases of East Antarctica is characterized by the absence of root respiration, as there are no representatives of higher vegetation. Thus, soil respiration is mainly characterized by the dominance of microbial respiration (both heterotrophic and autotrophic). In the case of the development of moss cushions in conditions of wind shelters and moist valleys, respiration due to moss rhizoids is also observed in the soils. Generally, in the studied oases, soil respiration rates have been found with the highest values of basal respiration being typical for soils of wind shelters (up to 0.125 mg g^{-1} day^{-1}) (Table 3). CO_2 emissions are lower in the surface horizons of soils in moist valleys with less dense moss cover (0.068–0.098 mg g^{-1} day^{-1}). "Amphibian soils" developing on the shores of lakes with a pulsating water regime and

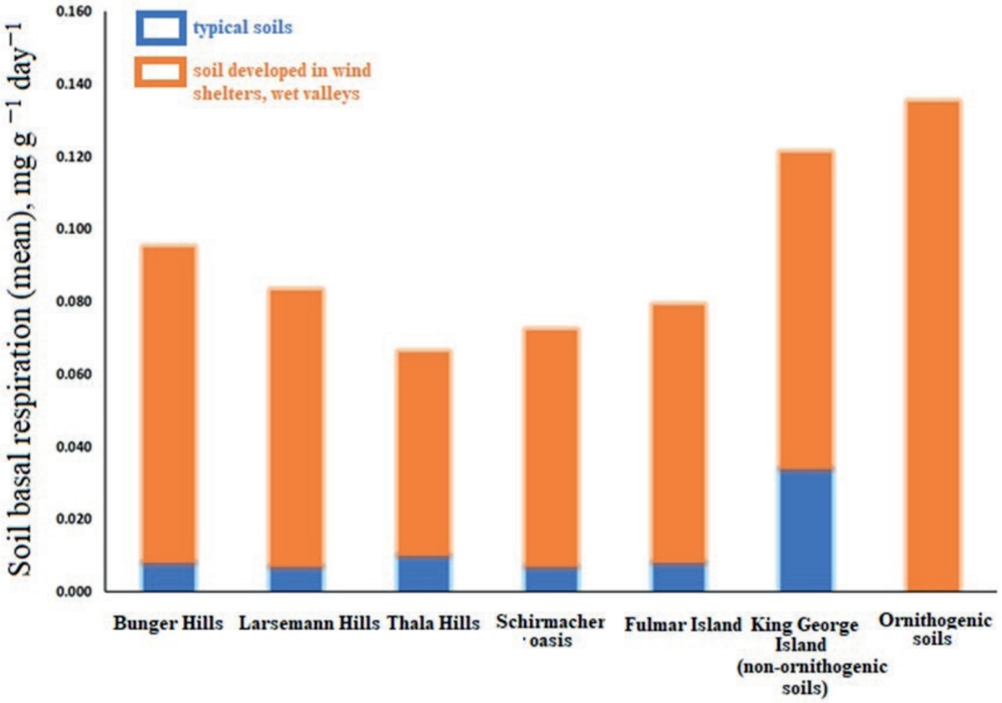


Fig. 5. Soil basal respiration rates in different ice-free areas of East and Maritime Antarctica (own data, unpublished)

surface algobacterial mats are characterized by carbon dioxide emission levels of 0.042–0.102 mg g^{-1} day^{-1}. Finally, the lowest CO_2 emissions are typical for "ahumic" soils without macroscopic organogenic horizons (0.005–0.010 mg g^{-1} day^{-1}). Soil moisture from melting snowfields has a critical effect on carbon dioxide emission levels during periods of low liquid precipitation in the oases of East Antarctica. Previously, authors showed that a decrease in soil moisture from 40 to 10% leads to a reduction in carbon dioxide emissions by more than three times (Mergelov et al. 2015, Mergelov et al. 2018). At the same time, seasonal and interannual dynamics in CO_2 emission levels depend more on the level of moisture in a particular summer season than on the duration of exposure of vegetation without snow in summer (in days), when biological activity is suppressed in the case of a low level of meltwater intake from summer snowfalls.

Although soil respiration levels in wind-sheltered soils within East Antarctic oases are comparable to those observed in Maritimr Antarctic soils – and align with previously reported levels in the Russian Arctic (Alekseev et al. 2022) – the spatial extent of such sites in East Antarctica is limited. Average soil CO_2 emissions in the studied oases of Bunger Hills and Schirmacher Oasis exceed those documented for the Dry Valleys (Archer et al. 2018), yet remain lower than emissions observed on King George Island and Ardley Island. According to Dennis et al. (2013), climate warming is expected to impact soil microbial communities differently in subantarctic versus coastal Antarctic environments. Studies examining latitudinal gradients in Antarctic soil biological parameters have demonstrated that microbial temperature sensitivity increases with higher mean annual soil temperatures. Consequently, bacterial communities in colder regions, such as East Antarctic oases, tend to be less responsive to temperature changes than those in warmer, more maritime regions of Antarctica (Rinnan et al. 2009).

These findings suggest a pronounced gradient in microbial activity from the coastal and inland oases of East Antarctica toward the less extreme environments of Maritime Antarctica. This gradient is influenced by variations in soil temperature regimes, enzymatic activity, and pools of organic carbon, nitrogen, and phosphorus. It is worth highlighting that there remains limited research focused on soil respiration and biological activity across the diverse eco-climatic regions of Antarctica. Earlier, Russian researchers (Lupachev et al. 2020) proposed potential ecological and climatic shifts that could influence soil formation in East Antarctic oases. They suggest that increases in mean annual air temperatures and changes in hydrological conditions – such as prolonged soil biological activity periods and enhanced moisture availability – may lead to greater accumulation of organic matter. This, in turn, could promote the stabilization of surface cover by mosses and lichens and facilitate the prolonged preservation of organic material in situ.

Estimation of CO_2 emission rate and mineralization of organic matter in Antarctic soils

Cryogenic soils show negligible organic decomposition, driven by persistent low temperatures that severely restrict microbial mineralization, oxygen deficiency that inhibits oxidase activity, and complex substrates with very low nitrogen content that further hinder microbial metabolism (Freeman et al. 2001; Moore and Basiliko 2006).

Despite the fact that the number of studies on the assessment of carbon pools in cryogenic soils has increased over the past decades (Michel et al. 2014; Mergelov et al. 2015; Schaefer et al. 2008), the issues of quantifying soil carbon reserves and mineralization potential are still problematic and controversial. Despite the fact that the period of biological activity and positive temperatures in the oases of East Antarctica is short, there is still a certain variety of soil processes that are associated with the transformation of organic matter and the activity of biota. In particular, the question of the rate of mineralization and intensity in the case of different soils and different sources of humus formation in both Western and Eastern Antarctica was discussed earlier (Abakumov 2010). It can be said that mineralization and humification processes can occur in all Antarctic soils where organic matter is present. However, it is worth noting that, unlike most soils, the humification process in Antarctic soils is accompanied by the formation of an increased amount of fulvic acids with a low proportion of humic substances and the dominance of aliphatic structures, which will be shown below using the example of ^{13}C-NMR spectra (Alekseev and Abakumov 2024). The rate of CO_2 emission from the surface horizons of the soils of East Antarctica oases was different (Table 4). The cumulative production of C-CO_2 by the surface horizons of soils in wind shelters and wet valleys averaged 282.97 mg of CO_2/kg^{-1} soil day^{-1}, while soils developing without local mitigation of soil formation conditions – 69.92 CO_2/kg^{-1} soil day^{-1}. As shown in Table 4, soil CO_2 emissions displayed two pronounced peaks: an initial one within the first 1–3 weeks, followed by a second, more substantial peak at 13–14 weeks. This dual-peak pattern is a universal trait observed across nearly all studied soils.

During soil incubation, the amount of C-CO_2 released reflects both microbial vitality and the organic matter's vulnerability to breakdown. The data show a rapid and continuous decline in CO_2 emissions, especially in the

Table 4. Average CO_2 emission rates, mg CO_2/kg^{-1} soil day^{-1}

Soil ID and horizon depth, cm	1-3 week	4-7 week	8-11 week	12-15 week	16-19 week
Bunger 1 0–2 cm	163.2 ± 12.1	142.0 ± 7.2	123.2 ± 9.7	92.3 ± 7.6	81.2 ± 4.5
Bunger 1 2–20 cm	191.2 ± 10.5	161.2 ± 12.1	142.1 ± 8.8	92.3 ± 6.5	71.2 ±5.4
Bunger 3 0–1 cm (salt staining)	221.2 ±17.6	181.2 ± 14.1	152.1 ± 10.4	122.3 ± 9.2	101.2 ± 10.4
Bunger 3 1–16 cm	181.2 ± 6.1	161.2 ± 7.1	142.1 ± 10.5	94.3 ± 8.1	71.2 ± 8.2
Bunger 6 0–2 cm (moss)	753.1 ± 5.1	641.2 ± 12.4	574.2 ± 15.6	531.2 ± 20.1	421.2 ± 10.4
Bunger 6 2–21 cm (BC)	563.2 ± 5.7	454.2 ± 15.4	453.1 ± 12.3	374.0 ± 9.7	333.1 ± 12.4
Novo 7 0–2 cm	563.1 ± 10.4	452.1 ± 14.1	341.2 ± 13.1	191.1 ± 7.6	181.2 ± 12.7
Novo 3 0–3 cm (moss)	567.6 ± 12.4	364.2 ± 19.2	174.3 ± 8.7	202.1 ± 12.5	91.2 ± 8.5

surface organic layers, signaling a swift shift toward stabilization that makes organic matter highly resistant to decay from biological or environmental forces. In contrast, destabilization reverses this trend, stripping organic components of their resistance and making them easily accessible for microbial consumption (Six et al. 2006). Ornithogenic soils are characterized by organic matter stability, while in deeper mineral horizons, mineralization rates are lower – likely due to changes in humic acid composition, such as altered ratios of carboxyl groups and aliphatic structures driven by ornithogenic influence. These profound structural modifications play a decisive role in controlling decomposition rates, demonstrating that ornithogenic factors fundamentally reshape the organic matter system and promote its stabilization (Ejarque and Abakumov 2016).

^{13}C-NMR spectroscopy of humic substances

The application of ^{13}C-NMR spectroscopy in studying Antarctic Cryosols provides critical insights into carbon storage mechanisms, nutrient cycling processes, historical climate records, and potential feedback loops that could influence future climate scenarios. It is crucial to note that the limited amount of samples analyzed in this section primarily stems from the challenges encountered in obtaining sufficient quantities of humic acids from each soil sample, which is largely attributed to the low levels of organic matter present in soils of Leptosols in East Antarctica. Moreover, there is a significant deficiency of published studies addressing the molecular composition of soil organic matter (SOM) in East Antarctica. That is why it is a vital imperative to undertake extensive instrumental investigations in these remote offshore environments to deepen our understanding of organic matter composition and the mechanisms underlying its stabilization. We have analyzed two topsoil samples from Bunger Hills (Turbic Cryosol, Bunger 1 0–2 cm) and Schirmacher Oasis (Histic Turbic Cryosol, Novo 7 0–2 cm).

The findings regarding the elemental composition of the isolated humic acids (HAs) from the topsoil horizons are presented in Table 5. Notably, the low carbon content observed in HAs can be attributed to a predominance of aliphatic structures, with aromatic components being present only in trace amounts. This suggests that the HAs in this region are characterized by a higher proportion of non-aromatic carbon compounds. The C/N ratio, which ranges from 10.06 to 10.73 in the studied samples, further supports the notion of low carbon availability in peripheral compounds, coupled with a relatively high nitrogen content.

The ^{13}C-NMR spectral shapes of humic acids extracted from investigated topsoil horizons are shown in Fig. 6. Both samples exhibit a predominance of aliphatic structures, which have been documented in previous studies (Beyer et al. 1997, Alekseev and Abakumov 2024; Zech et al. 1997). A particularly notable feature observed in the ^{13}C-NMR spectrum from the Turbic Cryosol in Bunger Hills is the presence of a prominent peak at a chemical shift of 173.89 ppm. This peak is characteristic of carbonyl carbons (C=O) associated with functional groups such as carboxylic acids, esters, and amides. The intensity and position of this peak suggest significant organic matter transformation processes influenced by the accumulation of moss-lichen vegetation, which contributes to the development of a more complex peripheral aliphatic component within humic acids. In addition to this key peak, several other significant chemical shifts were observed in the spectra. For instance, peaks around 30 ppm correspond to methylene carbon atoms (–CH$_2$–), which are often associated with lipid components within organic matter (Alekseev and Abakumov 2024; Keeler et al. 2006). The presence of these lipids indicates microbial activity and organic matter decomposition processes that are critical for nutrient cycling in these cold environments. The methylene carbon signals can also reflect the degree of saturation and structural complexity of lipids present, providing insights into microbial community composition and metabolic pathways.

In general, most of the peaks on the ^{13}C-NMR spectra of Turbic Cryosol from Bunger Hills were in the region of alkyl-C (0–60 ppm). Previous research has established that the decomposition of soil organic matter is generally accompanied by a shift in carbon functional groups, notably an increase in alkyl carbon content and a concomitant decrease in O-alkyl carbon fractions (Shen et al. 2018). This transformation reflects the preferential breakdown of labile, easily degradable components such as polysaccharides, which are highly accessible to microbial enzymes. As microbial activity progresses, these labile compounds are rapidly mineralized, leading to a relative enrichment of more recalcitrant alkyl structures that are resistant to further decomposition. This pattern underscores the dynamic nature of organic matter turnover and the selective degradation of specific molecular pools during soil organic matter decomposition processes.

The most intense peak on a ^{13}C-NMR spectra of HA isolated from Histic Turbic Cryosol from wind shelter in Schirmacher Oasis is found at 173.18 ppm. This is most likely due to carboxyl carbon on aromatic rings or in esterified fatty acids (Li et al. 2003; Preston and Schnitzer 1987). The peak at 30.39 ppm in the alkyl-C region is assigned as -CH2- methylene carbons. At 22.73 ppm, a prominent peak of CH$_3$ was also observed, which could be due to the presence of aliphatic chains of various components such as lipids/waxes (Simpson et al. 2011). In soil from Schirmacher Oasis, another (less prominent) peak was observed at 71.77 ppm, which is attributed by some authors to quaternary C bonded to oxygen (O) (Nebbioso et al. 2015). However, as discussed by Gamage et al. (2024) the peak at 70 ppm could have originated from the modified and relatively hydrophobic carbohydrate structures. ^{13}C-NMR spectra of humic acids from both Schirmacher Oasis and Bunger

Table 5. The elemental composition of humic acids from studied topsoil horizons

Soil ID	C, %	H, %	N, %	O, %	C/N	H/C	O/C	Ash, %
Turbic Cryosol, Bunger 1 0–2 cm	49.01	5.01	4.87	41.11	10.06	0.10	0.83	4.35
Histic Turbic Cryosol, Novo 7 0–2 cm	52.03	4.92	4.85	38.2	10.72	0.09	0.73	4.20
CV, %	4,23	1,28	0,29	5,19	0,80	7,44	9,07	2,48
Leonardite HA standard (1S104H)	63.81	3.7	1.23	31.27	51.88	0.006	0.49	2.6
Elliot soil HA standard (1S102H)	58.13	3.68	4.14	34.08	14.04	0.06	0.59	0.9

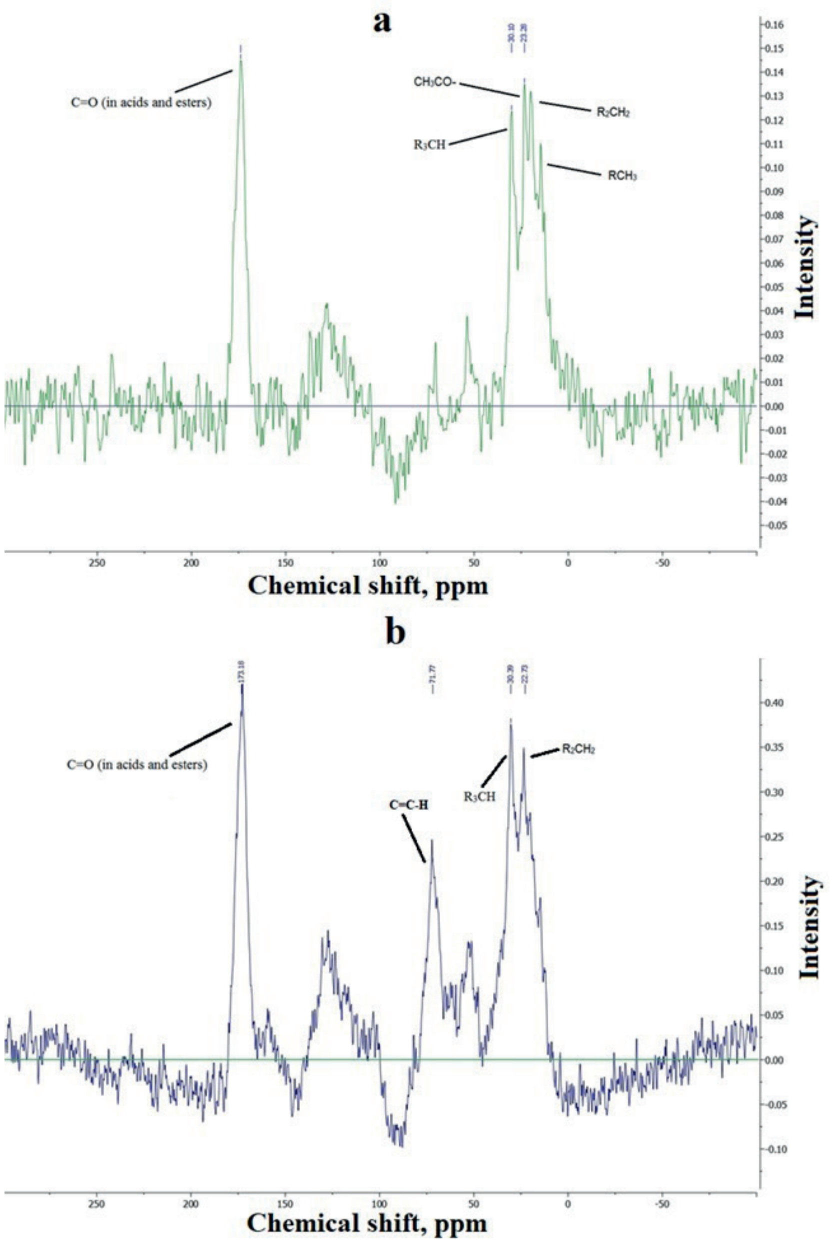


Fig. 6. ^{13}C-NMR spectra of Has isolated from studied topsoils in Bunger Hills (Bunger 2, 0–1 cm, a) and Schirmacher Oasis (Novo 7, 0–2 cm, b)

Hills showed quite low signs in the area of 110–140 ppm, which indicate weak contributions of phenol as it occurs in the lignin or aromatic amino acids group (Knicker et al. 2006).

Overall, the hydrophobic nature of aliphatic components described makes humic substances resistant to microbial degradation (Mylotte et al. 2016). Previous findings by Gamage et al. (2024) revealed that this can be explained by highly decomposed aliphatic/alicyclic molecules. At both ^{13}C-NMR spectra of humic acids, we have found significant contributions from non-protonated O-alkyl-C and non-polar alkyl protons, which could be attributed to their origin from the carboxyl-rich alicyclic CRAM structures (Cao et al. 2016; Garage et al. 2024). We can also prove this finding, since we also found a notable presence of carboxylic acid, indicated by a peak at 170 ppm characterizing CRAM-like structures (Hertkorn et al. 2016).

Analysis of both ^{13}C-NMR spectra from Bunger Hills revealed that SOM is more degraded (or humified) compared to that from Schirmacher Oasis, where some plant material (moss) inputs were detected. This is also supported by a lower contribution of O-alkyl carbon on a spectrum (a) and a reduced O-alkyl C/alkyl C ratio, which shows a lesser extent of decomposition. The same were previously observed by Pradel et al. (2023). However, a higher intensity of peaks was noted in the carboxylic region between 220 and 160 ppm in the topsoil sample from Schirmacher Oasis. Our findings align with previous studies, confirming that aliphatic carbon predominates over aromatic carbon while exhibiting lower levels of carbonyl carbon.

Moreover, smaller peaks observed between 50–60 ppm can be attributed to carbon atoms involved in ether linkages or aliphatic carbon structures, suggesting contributions from plant-derived materials and microbial metabolites. These shifts reflect the structural diversity within soil organic matter and its dynamic interactions with soil minerals. The presence of ether linkages is particularly important, as they can enhance the stability and recalcitrance of organic matter against microbial degradation. Peaks in the range of 100–110 ppm may indicate aromatic carbon structures, which are typically resistant to decomposition and play a crucial role in long-term carbon storage within soil profiles. Aromatic compounds contribute to soil stability through their ability to form strong interactions with soil minerals, thereby influencing overall carbon dynamics and nutrient retention.

Almost no peaks were observed around 60–80 ppm, which usually represent carbons attached to hydroxyl groups (–OH) or those involved in hydrogen bonding within humic substances. These functional groups are essential for understanding how organic matter interacts with soil minerals and affects nutrient availability. Hydroxyl groups can also participate in complexation reactions with metal ions, further influencing nutrient cycling processes.
The comprehensive analysis of these chemical shifts not only enhances our understanding of the biochemical composition and functional roles of organic matter in Antarctic Cryosols but also provides insights into how these ecosystems respond to environmental changes. By elucidating the structural characteristics and dynamics of soil organic matter through ^{13}C-NMR spectroscopy, we can better assess potential impacts on carbon cycling and storage under changing climatic conditions. Furthermore, specific attention should be given to interpreting minor peaks that may appear within the spectra. For example, signals between 0–10 ppm could indicate methyl ($-CH_3$) groups associated with microbial byproducts or fresh plant material inputs. These low-field signals can provide valuable information regarding recent biological activity and organic matter turnover rates.

The ratio of sp^2 to sp^3 carbon in Antarctic soils serves as a crucial indicator for understanding the composition, stability, and dynamics of soil organic in these distinct and harsh ecosystems. Antarctic environments are characterized by low temperatures, limited biological activity, and specific vegetation that shapes the organic material. In humic acids extracted from soils of Bunger Hills and Schirmacher Oasis, we found sp2/sp3 ratios of 0.42/0.58=0.724 and 0.29/0.71=0.408, respectively (Table 6). These results show that while there is significant aliphatic content (indicating potential for decomposition), there is also a notable presence of aromatic structures that contribute to soil stability over time. Previous research has shown that certain Antarctic soils possess significant amounts of stable organic carbon, indicated by elevated sp^2 levels (Beyer et al. 1997). This finding implies that these soils are capable of storing carbon effectively over extended periods, even in the face of limited biological activity. Our findings partly confirmed previous results (Beyer 1997; Pradel et al. 2023) showing the prevalence of aliphatic structure over aromatic in studied Antarctic soils; however they do not match perfectly with previously reported low levels of carbonyl groups. As mentioned above, in humic acids from topsoil of Bunger Hills Turbic Cryosol, where fresh C input was observed (probably from moss cushion), we could identify a low ratio of organic mineralization and a higher degree of humification.

Although there is a scarcity of literature regarding the molecular composition of soil organic matter (SOM) in Antarctic soils, our results are consistent with earlier studies from various regions of the continent, which showed that aliphatic carbon compounds are more prevalent than aromatic ones in typical organo-mineral soils (Calace et al. 1995). Additionally, recent research has revealed comparable traits in soils from different offshore oases. ^{13}C-NMR spectroscopy serves as an invaluable tool for characterizing soil organic matter composition at a molecular level. The interpretation of chemical shifts allows researchers to infer structural features that govern biochemical processes within soils, ultimately contributing to our understanding of ecosystem functioning and resilience in polar regions facing rapid environmental change. By integrating ^{13}C-NMR data with other analytical techniques such as elemental analysis or isotopic studies, we can develop a more comprehensive picture of how Antarctic Cryosols contribute to global carbon cycles.

CONCLUSIONS

The research conducted in Bunger Hills and Schirmacher Oasis provides valuable insights into the unique soil characteristics, organic matter composition, and carbon dynamics within two remote inner shelf oases poorly investigated in the context of complex soil chemical and microbiological studies. Our work revealed a complex interplay between environmental conditions and soil characteristics that significantly influence biological activity, carbon storage, and the molecular composition of organic matter. Soils of East Antarctic oases have been investigated only sporadically, despite their fundamental importance for revealing information on soil genesis as well as ensuring the effective realization of environmental protection measures and saving unique ecosystems under the Antarctic Treaty System.

In both studied oases, the soils exhibited a range of physical-chemical properties shaped by factors such as moisture availability, salt accumulation, and vegetation cover. Notably, Turbic Histic Cryosol found beneath moss cushions demonstrated the highest levels of organic carbon and nitrogen accumulation. This is in line with previous soil research done in East Antarctica oases and suggests that microhabitats with moisture retention play a crucial role in supporting greater biological activity. We found pH values varying from slightly acidic to acidic conditions (under moss cushions), while alkaline conditions prevailed in areas affected by salt accumulation. Such variability highlights the significant impact of environmental factors on soil chemistry and biology. Higher TOC content was found in wet valleys. Soil respiration rates were generally lower in East Antarctic oases compared to other regions like King George Island, indicating limited biological activity due to the harsh climatic conditions prevalent in these areas. Our results demonstrated that soil moisture gradient (taking into account studied soil in wind shelters and dry flanks of the valleys) could play a critical factor for increasing carbon dioxide emission levels in harsh environmental conditions of East Antarctic oases. The general trend showed the highest values of basal respiration in wind shelters, followed by wet valleys, subaqueous soils on the lakes shoreline, and the lowest respiration rates in ahumic soils.

The molecular composition of humic substances was analyzed on two samples using ^{13}C-NMR spectroscopy, which revealed a predominance of aliphatic structures in humic acids extracted from studied samples from both oases. Although our preliminary results of ^{13}C-NMR

Table 6. The sp^2/sp^3 ratios of studied ^{13}C-NMR spectra of humic substances isolated from Antarctic Cryosols

Sample	sp^2/sp^3 ratio
Turbic Cryosol, Bunger 1 0–2 cm	0.724
Histic Turbic Cryosol, Novo 7 0–2 cm	0.408
Leonardite	2.11

spectroscopy showed predominantly common features of molecular composition of HAs with those previously described for other ice-free areas of Antarctica, other findings depicted lower levels of carbonyl carbon groups. This finding suggests ongoing microbial activity and organic matter decomposition processes that are critical for nutrient cycling in these cold environments. Significant peaks associated with carboxylic acids indicated that organic matter transformation processes are influenced by vegetation cover, particularly moss-lichen communities. We have observed a low ratio of organic mineralization and a higher degree of humification in topsoil from Bunger Hills (Turbic Cryosol), where fresh carbon input affected the spectral shape of ^{13}C-NMR.

The stability of soil organic matter is further underscored by the sp^2/sp^3 carbon ratios observed in humic acids from both Bunger Hills and Schirmacher Oasis. These ratios indicate a balance between aliphatic content – suggesting potential for decomposition – and aromatic structures that contribute to long-term stability. Such findings align with previous research indicating that Antarctic soils can effectively store stable organic carbon despite limited biological activity. By elucidating the physical-chemical properties, molecular composition, and mineralization rates of soil organic matter, this study contributes to a deeper understanding of how Antarctic soils respond to environmental changes and their implications for greenhouse gas emissions. As climate change continues to impact these fragile ecosystems, future studies integrating various analytical techniques will be essential for enhancing our knowledge of their dynamics and resilience in an era of rapid environmental transformation.

Recent findings of our research project in Schirmacher Oasis (unexplored findings on influence of penguin colonies on soil formation in Schirmacher Oasis, Galustov et al. 2025), even more so, show the necessity of more detailed research in inner shelf oases to confirm or expand the ideas of soil formation, mineralization, and stability of organic matter previously accepted.

FUNDING

The research was supported by the Russian Science Foundation (project № 24-27-00361 «Soils in Eastern Antarctica oases: biogeochemistry, stability of organic matter and environmental risks»). ■

REFERENCES

Abakumov E. and Alekseev I. (2018). Stability of soil organic matter in Cryosols of the maritime Antarctic: Insights from 13 C NMR and electron spin resonance spectroscopy. Solid Earth, 9(6):1329-1339. https://doi.org/10.5194/se-9-1329-2018

Abakumov E. and Mukhametova N. (2014). Microbial biomass and basal respiration of selected Sub-Antarctic and Antarctic soils in the areas of some Russian polar stations. Solid Earth, 5(2):705-712. https://doi.org/10.5194/se-5-705-2014

Abakumov E.V. (2010). The sources and composition of humus in some soils of West Antarctica. Eurasian Soil Sci, 43:499-508. https://doi.org/10.1134/S1064229310040050

Alekseev I., Abakumov E. (2022). Soil organic carbon stocks and stability of organic matter in permafrost-affected soils of Yamal region, Russian Arctic. Geoderma Regional, 28: e00454. https://doi.org/10.1016/j.geodrs.2022.e00454

Alekseev I., Abakumov E. (2024). Soil organic matter and biogenic-abiogenic interactions in soils of Larsemann Hills and Bunger Hills, East Antarctica. Polar Sci, 40:101040. https://doi.org/10.1016/j.polar.2024.101040

Anderson J.P. (1982). Soil respiration. Methods of soil analysis: part 2 chemical and microbiological properties 9:831-871.

Archer S.D., Lee K.C., Caruso T., Maki T., Lee C.K., Cowan D.A., Pointing S.B. (2018). Microbial dispersal limitation to isolated soil habitats in the McMurdo Dry Valleys of Antarctica. bioRxiv, 493411.

Beyer L., Knicker H., Blume H.P., Bölter M., Vogt B., Schneider D. (1997). Soil organic matter of suggested spodic horizons in relic ornithogenic soils of coastal continental Antarctica (Casey Station, Wilkes Land) in comparison with that of spodic soil horizons in Germany. Soil Sci, 162(7):518-527.

Bockheim J.G., Lupachev A.V., Blume H.P., Bölter M., Simas F.N.B., McLeod, M. (2015). Distribution of soil taxa in Antarctica: A preliminary analysis. Geoderma, 245–246, 104–111. https://doi.org/10.1016/j.geoderma.2015.01.017

Bockheim J.G., McLeod M. (2008). Soil distribution in the McMurdo dry valleys, Antarctica. Geoderma, 144(1-2):43-49.

Bolshiyanov D.Yu., Verkulich S.R. (1992). New data on the environmental history of Bunger Hills (East Antarctica). Antarctica: Doklady Komissii, 30: 58-64 (in Russian).

Calace N., Campanella L., Paolis F.D., Petronio B.M. (1995). Characterization of humic acids isolated from Antarctic soils. Int J Environ Anal Chem, 60(1):71-78.

Cao X., Aiken G.R., Spencer R.G., Butler K., Mao J., Schmidt-Rohr K. (2016). Novel insights from NMR spectroscopy into seasonal changes in the composition of dissolved organic matter exported to the Bering Sea by the Yukon River. Geochim Cosmochim Acta, 181:72-88.

Cary S.C., McDonald I.R., Barrett J.E., Cowan D.A. (2010). On the rocks: the microbiology of Antarctic Dry Valley soils. Nat Rev Microbiol, 8(2):129-138.

Chefetz B., Salloum M.J., Deshmukh A.P., Hatcher P.G. (2002). Structural components of humic acids as determined by chemical modifications and carbon-13 NMR, pyrolysis- and thermochemolysis-gas chromatography/mass spectrometry. Soil Sci Soc Am J, 66(4):1159-1171.

Claridge G.G.C., Campbell I.B., Sheppard D.S. (2000). Carbon pools in Antarctica and their significance for global climate change. In: Smith J (ed) Global climate change and cold regions ecosystems CRC Press:71-90.

Convey P. et al (2014). The spatial structure of Antarctic biodiversity. Ecol Monogr, 84(2):203-244.

de Souza Carvalho J.V. et al. (2010). Impact of expected global warming on C mineralization in maritime Antarctic soils: results of laboratory experiments. Antarct Sci, 22(5):485-493.

Dennis PG et al (2013) Warming constrains bacterial community responses to nutrient inputs in a southern but not northern maritime Antarctic soil. Soil Biol Biochem 57:248-255.

Ejarque E., Abakumov E. (2016). Stability and biodegradability of organic matter from Arctic soils of Western Siberia: insights from 13 C-NMR spectroscopy and elemental analysis. Solid Earth, 7(1):153-165.

Elberling B.O. et al. (2013). Long-term CO2 production following permafrost thaw. Nat Clim Change, 3(10):890-894.

Freeman C. et al. (2001). Export of organic carbon from peat soils. Nature, 412(6849):785.

Friedmann E.I. (1982). Endolithic microorganisms in the Antarctic cold desert. Science, 215(4536):1045-1053.

Galustov K., Zagumyonnyi D., Chupin I., Ryan P. (2025) Inland breeding by Adelie penguins Pygoscelis adeliae at Schirmacher Oasis, East Antarctica. Polar Biol, 48, 99. https://doi.org/10.1007/s00300-025-03421-6.

Gamage J. et al. (2024). Chemical composition of soil humin in an organic soil profile. Applied Geochemistry, 165:105954.

Gore D.B., Leishman M.R. (2020). Salt, sediments and weathering environments in Bunger Hills. Antarctic Science, 32(2),138–152.

Hertkorn N. , Harir M., Cawley K.M., Schmitt-Kopplin P., Jaffé R. (2016). Molecular characterization of dissolved organic matter from subtropical wetlands: a comparative study through the analysis of optical properties, NMR and FTICR/MS. Biogeosciences,13(8),2257–2277.

Hodgson D.A., Graham A.G., Roberts S.J., Bentley M.J., Cofaigh C.Ó., Verleyen E., Smith J.A. (2014). Terrestrial and submarine evidence for the extent and timing of the Last Glacial Maximum and the onset of deglaciation on the maritime-Antarctic and sub-Antarctic islands. Quaternary Science Reviews, 100 ,137–158.

IHSS - International Humic Substances Society (1981). Outline of Extraction Procedures .Distributed to the Members of International Humic Substances Society .International Humic Substances Society.Saint Paul.

Kachinsky N.A. (1958). Mechanical and micro - aggregate composition of soil, methods of its study. Academy of science USSR,Moscow .193p.[in Russian].

Kechaykina I.O., Rumin A.G., Chukov S.N. (2011). The post-aggrogenic transformation of the organic matter sod-podzolic soils. Pochvovedenie 10 :1178–1193.

Keeler C., Kelly E.F., Maciel G.E. (2006). Chemical – structural information from solid - state13C NMR studies a suite humic materials from lower montane forest soil Colorado USA. Geoderma, 130(1 -2),124–140.

Klokov V.D., Verkulich S.R. (1994). Peculiarities hydrological regime water bodies Bunger Hills (East Antarctica). Informational Bulletin Russian Antarctic Expedition, 118:60–68 (in Russian).

Knicker, H., Almendros, G., González-Vila, F.J., González-Pérez, J.A., Polvillo, O. (2006) Characteristic alterations of quantity and quality of soil organic matter caused by forest fires in continental Mediterranean ecosystems: a solid-state 13C NMR study. Eur J Soil Sci, 57:558-569. https://doi.org/10.1111/j.1365-2389.2006.00836.x

Kögel-Knabner, I. (1997) 13C and 15N NMR spectroscopy as a tool in soil organic matter studies. Geoderma, 80(3-4), 243-270.

Kruchinin Yu.A., Simonov I. M. (1967). New type of lakes in Antarctica. Informational Bulletin of Soviet Antarctic Expedition, 66:12-17 (in Russian).

Kurbatova L.E., Ochyra R. (2012). Two noteworthy additions to the moss flora of the Schirmacher Oasis in continental Antarctica. Cryptogamie, Bryologie, 33(2), 159-167.

Leishman M.R., Gibson J.A., Gore D.B. (2020). Spatial distribution of birds and terrestrial plants in Bunger Hills. Antarctic Science, 32(2), 153-166.

Li L., Huang W., Peng P.A., Sheng G., Fu J. (2003) Chemical and molecular heterogeneity of humic acids repetitively extracted from a peat. Soil Sci Soc Am J, 67:740-746. https://doi.org/10.2136/sssaj2003.7400

Lupachev A.V., Gubin S.V., Abakumov E.V. (2020). Levels of biogenic-abiogenic interaction and structural organization of soils and soil-like bodies in Antarctica. In: Processes and Phenomena on the Boundary Between Biogenic and Abiogenic Nature. Springer International Publishing, pp 481-500.

Lupachev A., Abakumov E., Goryachkin S., Veremeeva A. (2020). Soil cover of the Fildes Peninsula (King-George Island, West Antarctica). Catena, 193, 2020 (104613). https://doi.org/10.1016/j.catena.2020.104613.

Mao J.D., Hu W.G., Schmidt-Rohr K., Davies G, Ghabbour E.A, Xing B. (2000). Quantitative characterization of humic substances by solid-state carbon-13 nuclear magnetic resonance. Soil Sci Soc Am J, 64:873 -884. https://doi.org/10.2136/sssaj2000.643873x

Mergelov N.S., Konyushkov D.E., Lupachev A.V., Goryachkin S.V. (2015). Soils of Macrobertson Land: The soils of Antarctica. In: The Soils of Antarctica, pp 65-86.

Mergelov N., Mueller C.W., Prater I., Shorkunov I., Dolgikh A., Zazovskaya E. et al (2018). Alteration of rocks by endolithic organisms is one of the pathways for the beginning of soils on Earth. Sci Rep, 8:3367. https://doi.org/10.1038/s41598-018-21724-y

Michel R.F.M, Schaefer C.E.G.R., Dias L.E., Simas F.NB., de Melo Benites V., de Sá Mendonça E. (2006). Ornithogenic Gelisols (Cryosols) from maritime Antarctica: pedogenesis, vegetation and carbon studies. Soil Sci Soc Am J, 70:1370-1376. https://doi.org/10.2136/sssaj2005.0222

Michel R.F.M., Schaefer C.E.G.R., López-Martínez J., Simas F.N.B., Haus N.W., Serrano E., Bockheim J.G. (2014) Soils and landforms from Fildes Peninsula and Ardley Island, Maritime Antarctica. Geomorphology, 225, 76–86. https://doi.org/10.1016/j.geomorph.2014.03.041

Moore T. and Basiliko N. (2006). Decomposition in boreal peatlands. In: Boreal Peatland Ecosystems. Springer Berlin Heidelberg, pp 125-143.

Mylotte R. et al. (2016). Insights into the composition of recalcitrant organic matter from estuarine sediments using NMR spectroscopy. Org Geochem, 98:155-165. https://doi.org/10.1016/j.orggeochem.2016.05.002

Nebbioso A. et al. (2015). Unveiling the molecular composition of the unextractable soil organic fraction (humin) by humeomics. Biol Fertil Soils, 51:443-451.

Nikitin D.A. et al. (2017). Microbial biomass and biological activity of soils and soil-like bodies in coastal oases of Antarctica. Eurasian Soil Sci, 50:1086-1097.

Oertel C. et al. (2016). Greenhouse gas emissions from soils—A review.* Geochemistry, 76(3):327-352.

Pires C.V. et al. (2017). Soil organic carbon and nitrogen pools drive soil C-CO2 emissions from selected soils in Maritime Antarctica. Sci Total Environ, 596:124-135.

Pradel P. et al. (2023). Microbial response to warming and cellulose addition in a maritime Antarctic soil. Permafrost Periglacial Process, 34(3):370-383.

Preston C.M. and Schnitzer M. (1987). 13C NMR of humic substances: pH and solvent effects. J Soil Sci,38(4):667-678.

Rinnan R. and Bååth E. (2009). Differential utilization of carbon substrates by bacteria and fungi in tundra soil. Appl Environ Microbiol, 75(11):3611-3620.

Rusin N.P. (1958). Meteorological regime of Antarctic oases and its influence on the environment. Meteorology Hydrology, 3:14-19.

Rusin N.P. (1961). Meteorological and radiation regime of Antarctica. Gidrometeoizdat, Leningrad (in Russian).

Ryan P. (2024). Notes on the birds of Schirmacher oasis, Dronning Maud Land, East Antarctica. Mar Ornithol, 52:197–202.

Savaglia V. et al. (2024). Geology defines microbiome structure and composition in nunataks and valleys of the Sør Rondane Mountains,East Antarctica. Front Microbiol, 15:1316633.

Schaefer C.E.G. et al. (2008). Micromorphology and microchemistry of selected Cryosols from maritime Antarctica. Geoderma, 144(1–2):104–115.

Schmidt M.W. et al. (2011). Persistence of soil organic matter as an ecosystem property. Nature, 478(7367):49–56.

Schuur E.A. et al. (2015). Climate change and the permafrost carbon feedback. Nature, 520(7546):171–179.

Sheng D. et al. (2018). Increased chemical stability but decreased physical protection of soil organic carbon in response to nutrient amendment in a Tibetan alpine meadow. Soil Biol Biochem, 126:11–21.

Simas F.N. et al. (2007). Ornithogenic cryosols from Maritime Antarctica: Phosphatization as a soil forming process. Geoderma, 138(3–4):191–203.

Simonov I.M. (1971). Oases of East Antarctica. Gidrometeoizdat, Leningrad (in Russian).

Simpson A.J. et al. (2011). NMR spectroscopy in environmental research: From molecular interactions to global processes. Prog Nucl Magn Reson Spectrosc, 58(3–4):97–175.

Singh S.M., Sharma J., Ravindra R., Singh P. (2008). Plant community and nutrient status of the soils of Schirmacher Oasis, East Antarctica. Advances in Polar Science, 19(1), 63-76.

Six J. et al. (2006). Bacterial and fungal contributions to carbon sequestration in agroecosystems. Soil Sci Soc Am J, 70(2):555–569.

Swift R.S. (1996). Organic matter characterization. Methods of soil analysis: Part 3 chemical methods, 5:1011–1069.

Tedrow J.C.F., Ugolini F.C. (1966). Antarctic soils. In: Antarctic soils and soil forming processes, 8, pp 161-177.

Thomazini A., Francelino M.R., Pereira A.B., Schünemann A.L., Mendonça E.S., Almeida P.H.A., Schaefer C.E.G.R. (2016). Geospatial variability of soil CO2–C exchange in the main terrestrial ecosystems of Keller Peninsula, Maritime Antarctica. Sci Total Environ, 562:802-811. https://doi.org/10.1016/j.scitotenv.2016.04.086

Turetsky M.R., Abbott B.W., Jones M.C., Anthony K.W., Olefeldt D., Schuur E.A., McGuire A.D. (2020). Carbon release through abrupt permafrost thaw. Nat Geosci, 13(2):138-143. https://doi.org/10.1038/s41561-020-0530-7

Ugolini F.C., Jackson M.L. (1982). Weathering and mineral synthesis in Antarctic soils. Soil Sci Soc Am J, 9(5):228-229.

Vasilevich R., Lodygin E., Beznosikov V., Abakumov E. (2018) Molecular composition of raw peat and humic substances from permafrost peat soils of European Northeast Russia as climate change markers. Sci Total Environ, 615:1229-1238. https://doi.org/10.1016/j.scitotenv.2017.09.197

Verkulich S.R., Puschina Z.V., Sokratova I.N. (2007). Changes in sea level and glacioisostasy in coastal Antarctica. Materials of Glaciological Survey, 102:161-167 (in Russian).

Zech W., Senesi N., Guggenberger G., Kaiser K., Lehmann J., Miano T.M., Schroth G. (1997). Factors controlling humification and mineralization of soil organic matter in the tropics. Geoderma, 79(1-4):117-161. https://doi.org/10.1016/S0016-7061(97)00025-X

Zhu R., Bao T., Wang Q., Xu H., Liu Y. (2014). Summertime CO2 fluxes and ecosystem respiration from marine animal colony tundra in maritime Antarctica. Atmos Environ, 98:190-201. https://doi.org/10.1016/j.atmosenv.2014.08.022

Check for updates

CC BY

RESEARCH PAPER

METHODOLOGICAL APPROACHES TO CALCULATING MAXIMUM RIVER DISCHARGE IN THE UPPER IRTYSH BASIN

B. I. Gartsman[*1,2], **V. M. Moreido**[1], **A. V. Pavlenko**[3], **T. S. Gubareva**[1,2]
[1]Institute of Water Problems of the Russian Academy of Sciences, Moscow, Russia
[2]Institute of Natural-Technical Systems, Sevastopol, Russia
[3]East Kazakhstan University named after Sarsen Amanzholov, Ust-Kamenogorsk, Kazakhstan
*Corresponding author: gartsman@inbox.ru

Received: May 8th 2025 / Accepted: July 22nd 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4030

ABSTRACT. This study addresses flood estimation challenges in the Upper Irtysh River basin through comprehensive stochastic hydrological analysis. We evaluate the adequacy of various engineering methods for calculating peak discharges, with each computational approach based on probabilistic models combining: (1) theoretical probability distributions and (2) parameter estimation techniques for limited observational data. Our methodology employs an extensive range of three-parameter probability laws and frequency curve parameterization methods.

The research protocol involved: (i) rigorous stationarity testing of maximum annual discharge time series (for the period 1951-2019), and (ii) development of probabilistic frequency curves. Since conventional stochastic modelling requires stationary series, we developed methodological tools for detecting non-stationarity (particularly linear trends) and adjusting affected series through statistical normalization.

Key findings reveal that a part of studied rivers exhibit statistically significant ($p<0.05$) non-stationarity in annual peak flows observed as a linear trend. For such rivers, the time series were adjusted to stationary conditions. For all time-series - including the adjusted datasets - we constructed the complete set of considered probability models. From these, optimal models were selected representing different computational approaches: (1) the standard framework recommended by current regulatory documents, and (2) alternative schemes derived through comprehensive synthesis of published research.

Through application of multiple model quality criteria, it has been established that alternative computational schemes yield evidently better results compared to the standard methodology. The analysis further demonstrates that current non-stationarity in time series does not yet substantially affect the magnitude of the most critical design parameter - the 1% exceedance probability discharge. Future regional research should focus on: (1) identifying causes of non-stationarity in annual peak flow series, and (2) developing optimized computational frameworks for non-stationary conditions

KEYWORDS: flood risk assessment, tributaries of the Irtysh River, peak discharge, tributary systems, frequency analysis, probabilistic hydrology

CITATION: B. I. Gartsman, V.M. Moreido, A. V. Pavlenko, T. S. Gubareva (2026). Methodological Approaches To Calculating Maximum River Discharge In The Upper Irtysh Basin. Geography, Environment, Sustainability, 1 (19), 51-61
https://doi.org/10.24057/2071-9388-2026-4030

ACKNOWLEDGEMENTS: This work has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant Number BR24992899) within the framework of the project «Development of a system for forecasting catastrophic floods in the East Kazakhstan region using remote sensing data, GIS technologies, and machine learning».

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

This study investigates flood hazards in the East Kazakhstan Region (EKR), the easternmost administrative division of Kazakhstan covering 97,800 km^2 (Fig. 1). The region shares borders with Russia's Altai Territory and Altai Republic to the north, China to the east and southeast, and Kazakhstan's Abay Region to the west. Ust-Kamenogorsk serves as the regional administrative centre.

The East Kazakhstan region's river network, as documented in the Republic of Kazakhstan's water and energy cadastre, comprises over 800 rivers exceeding 10 km in length, including 48 rivers longer than 50 km and 20 rivers surpassing 100 km. All waterways in the region constitute tributaries of varying orders within the Irtysh River system. Based on hydrological regime characteristics (Fig. 1), these tributaries can be classified into three distinct groups:

— Group 1: right-bank tributaries of the South-Western Altai (e.g., Bukhtarma, Uba, Ulba, Kurshim, Qalzhyr, Naryn): Characterized by perennial flow and high discharge capacity;

— Group 2: left-bank tributaries of the Kalbinsky Range (e.g., Ulken-Boken, Qaiyndy, Ablaketka, Ulanka, Dresvyanka, Kyzylsu): Exhibiting reduced but generally sustained flow;

— Group 3: southern Zaisan Lake basin rivers (e.g., Qandysu, Uydene, Kenderlyk): Typically, ephemeral systems that frequently terminate in alluvial sand deposits or experience complete desiccation.

The territory of EKR is characterized by several types of floods of different origins, including spring freshets, rain-induced floods, ice-jam and debris-jam floods, and wind-driven surges. The main channel of the Irtysh River is controlled by Bukhtarma Reservoir and rarely inundates coastal areas. Flooding during freshets and flash floods is typical for all tributaries of the Irtysh, particularly in sections with more uniform riverbeds, whereas ice and debris jams occur in mountainous areas where the channel narrows. Inundation of coastal areas due to wind-driven surges is primarily observed in Lake Zaisan and the Bukhtarma Reservoir.

An analysis of materials from the Ministry of Emergency Situations of the Republic of Kazakhstan (MES RK) has identified a number of key areas regularly at risk of flooding. The highest concentration of such areas is observed along rivers of the first group, particularly at the confluences of tributaries of the Irtysh River, including the Ulba River near Ridder, as well as the Krasnoyarka, Glubochanka, and other rivers. Several flood-prone zones are also located along the Bukhtarma River and its tributaries. Certain settlements in the lower reaches of the Kalzhir, Kurchum, and Naryn rivers are also susceptible to flooding. The most flood-prone areas of the first group of rivers include Chapaevo village on the Krestovka River (Altai District), Ust-Talovka settlement on the Talovka River, Ubinka village on the Oba River (Shemonaikha District), Karatogai village on the Kalgutty River (Kurchum District).

Among the most significant floods was the spring freshet that occurred in March 2018 near Ust-Kamenogorsk on the Ulba River. According to media reports, more than ten villages were affected, with over 700 residents losing their homes due to the inundation of more than 480 houses (Vostochny Potop 2018). The direct damage to the region's infrastructure was estimated at 3.2 billion tenge (Year 2018 Is Not Over Yet 2018). The disaster was triggered by an abrupt and unusually rapid temperature rise, combined with heavy precipitation and frozen ground, which prevented water absorption and intensified surface runoff.

The coastal areas of second-group rivers, due to their low water flows, are significantly less prone to flooding. However, isolated flooding events have occasionally been recorded along the Ulken-Naryn, Kayyndy, Lailinka, Ulanka, and Tainty rivers. Among the most problematic areas are Samarskoye village on the Lailinka River, Mirolyubovka village on the Kayyndy River (Samar District), Ulanskoye village and Zhanuzak village on the Ulanka River, Asubulak village on the Ungyrdy River, Besterek village on the Kolbala Stream (Ulan District).

The tributaries of the third group are generally characterized by rapid water level rises during snowmelt floods, which often leads to inundation, particularly in their lowland sections. Certain areas along these rivers experience recurrent flooding, including The Kandysu River and rivers near Zaisan City, Kensai, Zharsu, Bakasu, and Saryzhira villages on the Uidene River, Tugyl settlement on the Kabyrgatal River (Zaisan District).

In Soviet and subsequently Russian engineering practice, the development of technological and informational tools for flood risk management has been based on approaches established in the mid-20th century

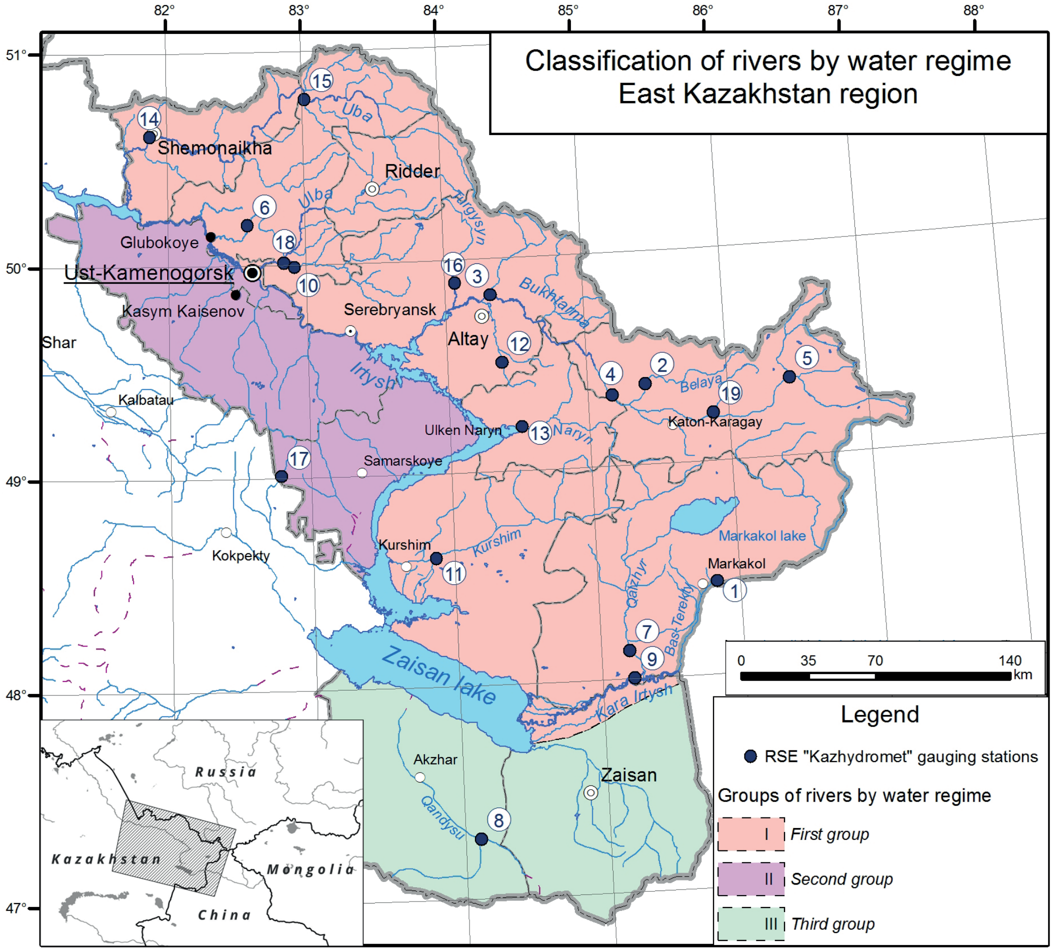


Fig. 1. Map of the grouping of the East Kazakhstan region of the Republic of Kazakhstan by water regime, indicating the location of hydrological observation network (the gauging station numbers correspond to the numbers in Table 1)

(Rozhdestvensky, Chebotarev 1974). These approaches primarily focus on determining design hydrological characteristics (water discharge or levels) with low probability (i.e., rare recurrence intervals) using available observational data series of river flow.

The objective of this study was to conduct a statistical analysis of flood hazards based on actual maximum flow data from Irtysh River tributaries. The analysis employed various methods of stochastic hydrology recommended by both current regulatory documents in force in Russia and Kazakhstan and authoritative literary sources.

The article evaluates the adequacy and effectiveness of various engineering calculation methods for determining maximum water discharges with specified exceedance probabilities, as applied to rivers in the study region. Each computational approach is based on an appropriate probabilistic model incorporating: (1) a theoretical probability distribution for the studied variable, and (2) a parameter estimation method for limited observational datasets. The research employs a comprehensive range of stochastic hydrology methods for developing flow frequency curves, including:

— The standard calculation framework recommended by current regulatory documents in Russia (() SP 529.1325800.2023 (2023)) and Kazakhstan (MSP 3.04-101-2005 (2006));

— An alternative computational scheme developed by Yu.B. Vinogradov (1988);

— Several probability distributions validated through international flood frequency analysis practice (Gubareva 2010, 2011);

— The state-of-the-art L-moments method for distribution parameter estimation, predominantly used in international studies (Hosking & Wallis 1997; Gubareva & Gartsman 2010).

The majority of EKR lies within the mountain system of the Southwestern Altai, except for its southern portion which partially encompasses the Saur-Tarbagatai mountain range. Progressing westward and southwestward, the mountains gradually transition into the more subdued topography of the Kazakh Uplands (Kazakhskiy Melkosopochnik). A prominent river valley, formed by the Irtysh River and its numerous tributaries, cuts through the mountainous terrain from the southwest to the northeast. This valley includes the intermontane basin of Lake Zaisan. Overall, the predominantly mountainous landscape of the region exhibits a wide range of elevations, from 200 to 4,500 meters above sea level, with a general slope trending northwestward and westward (Belyanin et al. 2013; Egorina, Zinchenko, Zinchenko 2000; Egorina et al. 2015).

The climate of the region, situated in the central part of the Eurasian continent, is classified as harsh continental and further complicated by mountainous terrain, following the principles of altitudinal zonation. These characteristics significantly influence the distribution of most meteorological parameters. Winter in the region is cold and prolonged, with mean January temperatures ranging from -12°C to -17°C in lowland areas to -23°C to -27°C in high-altitude zones. Absolute minima in some years can drop to -51°C to -54°C. Summer is hot, with July averages between 15°C and 24°C, while absolute maxima reach 35°C to 45°C. The number of days with temperatures above 0°C varies from fewer than 200 in mountainous areas to 230 in the southern lowlands of EKR. Precipitation is highly unevenly distributed, ranging from 400–650 mm in mountainous regions to less than 200 mm in the Zaisan Depression. Mountainous zones typically experience sufficient or excessive moisture, whereas lowland areas face moisture deficits. Mean annual wind speeds across the oblast generally range from 2–5 m/s, though in some areas, they can exceed 15 m/s (Egorina & Popova 1989; Egorina et al. 2015).

The mountainous landscapes of EKR exhibit distinct altitudinal zonation with four characteristic elevation belts. The lower belt, extending to 500-600 meters above sea level, encompasses plains and foothills. In the northwestern foothills, chernozem soils support feather grass-forb steppe communities, while the left bank of the Irtysh River valley features feather grass-fescue vegetation on dark kastanozems (dark chestnut soils). The Zaisan Depression displays unique arid-environment vegetation including wormwood-fescue communities on light kastanozems (light chestnut soils) and wormwood-anabasis associations on brown soils, with widespread occurrence of solonchaks, solonetz soils, and dune sands. At middle elevations (up to 1900-2000 meters), the forest belt dominates with mixed woody vegetation growing on brown forest soils. Higher elevations (up to 2800-3000 meters) are occupied by the subalpine-alpine belt characterized by meadow communities developing on mountain meadow soils. The uppermost elevations form the nival belt, where mountain peaks contain permanent snowfields, glaciers, and exposed bedrock surfaces. This vertical zonation reflects the transition from steppe ecosystems through forested middle elevations to alpine and ultimately glacial environments (Belyanin et al. 2013; Egorina et al., 2015).

The primary watercourse in EKR is the Irtysh River segment flowing from the border with China to the administrative boundary between EKR and Abai Region. This approximately 800 km long stretch incorporates several major hydrological features: the Kara Ertis (Black Irtysh) river section, the through-flow Zaisan Lake, and three major reservoirs (Bukhtarma, Ust-Kamenogorsk, and Shulba). Within Kazakhstan, the Irtysh River basin covers approximately 545,000 km^2. The river's flow regime in EKR exhibits mixed feeding sources, predominantly snowmelt and glacial meltwater. Discharge patterns are influenced by tributary inflows and water withdrawals from China, where Kara Ertis waters are extensively used for agricultural irrigation. Three major hydroelectric dams along the main channel create a fully regulated flow regime that virtually eliminates flood risks in most reaches. Exceptions occur in braided river sections and during wind-driven surge events in lacustrine portions of the channel. The Irtysh maintains year-round flow without stable ice cover due to its broad channel and sustained velocities (1.2-1.4 m/s) even in downstream reaches. The spring freshet period (March-June) is significantly attenuated by dam operations. Long-term monitoring (1961-2022) at the Ust-Kamenogorsk gauging station records mean annual discharge of 559 m^3/s at corresponding water level of 287.25 m (Belyanin et al. 2013; Egorina et al. 2015; Pavlenko et al. 2024).

MATERIALS AND METHODS

The study utilized maximum annual discharge records from gauging stations spanning their entire observation periods through 2022. The primary data sources included annual publications and reference materials from the State Water Cadastre maintained by RSE "Kazhydromet" (State Water Cadastre of RSE Kazhydromet 2025). The analysis focused on gauging stations with the most extensive observation histories, though nearly all available series contained significant gaps in their records. Selected stations provided continuous maximum annual discharge data series of at least 30 years, with the shortest complete series covering 33 years and the longest extending to 82 years. Approximately 85% of the analyzed data came from rivers of the first group. The complete list of monitoring stations and their respective observation periods is presented in Table 1.

The research methodology included, first, testing the observation time-series for stationarity and, second, constructing a probabilistic model—a frequency curve of maximum annual water discharges—to obtain design values with low exceedance probabilities required for engineering design and flood prevention planning. The current methodological framework for probabilistic modelling in engineering hydrology is based on the assumption of stationarity in observation series and can only be correctly applied to such series. At present, clearly evident climate changes, as well as significant anthropogenic transformations of river catchments, have led to non-stationarity in hydrological series, including maximum flow data. This necessitates the development of a methodological approach to identify various forms of non-stationarity (threshold changes in mean and variance parameters, linear trends, cyclical variations, etc.) using statistical criteria.

In this study, the maximum discharge series were tested for the presence of a linear trend — that is, a monotonic change in mean values over time — using two statistical criteria described below. For series where no significant linear trend was detected, a probabilistic model was developed by selecting the optimal combination of an analytical distribution curve and a parameter estimation method that provided the closest approximation to the empirical distribution curve constructed from the sample of measured values.

For time series demonstrating significant trends, there exists no generally accepted methodological framework for constructing probabilistic models. Therefore, estimates of design discharges with low exceedance probabilities can only be obtained through special studies extending beyond the analysis of sample data. In this research, we limited ourselves to approximate estimates of design discharges by eliminating the identified trends — that is, by adjusting the sample to stationary conditions corresponding to the maximum mean value.

For series with downward trends, adjustment was made to the first year of observations, while for series with upward trends — to the last year. Comparing the calculation results for the adjusted series with the results of conventional calculations performed on the same series without accounting for the presence of trends allowed for a quantitative assessment of their influence on the design values.

For time series with downward trends, adjusting to the beginning of the series allows for calculations with a safety margin, which is generally beneficial when estimating maximum discharges. However, additional analysis is required to prevent excessive safety margins that could lead to unjustified costs for flood protection measures. In the case of upward trends, adjusting the sample to the end of the observation period still does not provide design values of the required reliability, since the trend will continue into the future to unknown extents. This situation demands

Table 1. List of gauging stations and used observation periods

No.	River	Basin area (km^2)	River group	Observation periods (years)	Total years
1	Bas-Terekty – Moiyldy	184	1	1962-64, 1966-86, 1988-91, 2003-22	33
2	Belaya – Beloe	945	1	1954-62, 1964, 1966-98, 2005-22	48
3	Bukhtyrma – Berel	1,850	1	1958-97, 2005-22	58
4	Bukhtyrma – Lesnaya Pristan	10,700	1	1954-2022	69
5	Bukhtyrma – Pechi	6,860	1	1940-44, 1947-98, 2000-22	80
6	Glubochanka – Belokamenka	47	1	1978-98, 2003-22	41
7	Kalzhir – Kalzhyr	3,150	1	1940-46, 1949-52, 1955-64, 1966-96, 1998-2000, 2013-22	57
8	Kandysu – Saryolen	2,610	3	1973-94, 2012-22	33
9	Kara Ertis – Boran	55,900	1	1940-2000, 2002-22	82
10	Kishi Ulbi – Gornaya Ulbinka	2,170	1	1953, 1955-64, 1966-87, 1989-91	36
11	Kurchim – Voznesenka	5,840	1	1940-45, 1948-52, 1954-97, 1999-2022	72
12	L. Berezovka – Sredigornoe	251	1	1948-57, 1959-2022	66
13	Naryn – Ulken Naryn	1,960	1	1955-91, 1993-2022	67
14	Oba – Shemonaikha	8,470	1	1958-64, 1966-2021	63
15	Oba – Karakozha	2,768	1	1959-64, 1967-98, 2006-13, 2020-22	44
16	Turgysyn – Kutikha	1,200	1	1949-57, 1959-93, 2008-22	56
17	Ulken Boken – Djumba	758	2	1957-2000, 2002-22	66
18	Ulbi – Ulbi-Perevalochnaya	4,900	1	1940-2001, 2003-22	82
19	Chernovaya – Chernovoe	488	1	1955-69, 1971-77, 1979-98	36

particular caution in engineering decision-making. Thus, the results presented in this article for non-stationary series represent preliminary estimates that require special case-by-case verification studies in each specific instance.

Analysis of Linear Trends. To assess the linear trend for each original data series, an equation of the following form is constructed:

$$\overline{Q}_i = q \times i + p \qquad (1)$$

where i is the year number in the multi-year series, counted from the beginning of the series; $\overline{Q}_i$ is the moving mathematical expectation of discharge for the i-th year; q is the linear trend coefficient, calculated based on the correlation coefficient between the values of the characteristic and their chronological sequence numbers; p is the regression intercept (constant term).

Significance of the Linear Trend is assessed using a specialized modification of Student's t-test (Handbook of Hydrology…, 1993):

$$T_c = \left| \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \right| > T_{1-a/2,\, n-2} \qquad (2)$$

where: T_c is the test statistic, $T_{1-a/2,\, n-2}$ is the quantile of Student's t-distribution with (n-2) degrees of freedom of probability *(1-a/2), a* - is the significance level of the estimate, r is the correlation coefficient characterizing the trend, n is the length of the observation series (sample size). To assess the significance of the trend, the root mean square error of the correlation coefficient may also be used (Rozhdestvensky, Chebotarev 1974), as defined by the formula:

$$\sigma_r = \frac{1-r^2}{\sqrt{n-1}} \qquad (3)$$

The trend is considered statistically significant if the correlation coefficient is at least twice as large as the root mean square error. The adjustment of individual series values to achieve stationary conditions is performed using the formula:

$$Q_i^* = Q_i + q \times \left(n^* - i + 1\right) \qquad (4)$$

where: Q_i and Q_i^* are the original and adjusted (stationary) values of the characteristic in the i-th year of the multi-year series, n^* is the duration of the observation period from the first to the last year. It should be emphasized that the duration of the observation period from the first to the last year exceeds the length of the observation series if there are missing years in the record. In this case, the year index i corresponds strictly to the sequential numbering of years in the continuous period from the first observation year to the last.

Probability Distribution Functions and Parameter Estimation Methods. The guidelines of the current Russian regulatory document SP 529.1325800.2023 (2023), which align with the applicable Kazakh standards MSP 3.04-101-2005 (2006), prescribe the use of the following distributions in hydrological calculations: primarily curves derived from the gamma distribution, including the Pearson Type III (binomial) distribution and the three-parameter Kritsky-Menkel gamma distribution. The method of moments is recommended as the primary approach for estimating the parameters of analytical curves based on sample data. For the Pearson Type III distribution, an additional graphical-analytical method is suggested, while the Kritsky-Menkel distribution calls for the approximate maximum likelihood method. Notably, the regulations do not prohibit the use of alternative calculation techniques, provided they are properly justified. Hereafter, we will refer to the normative-recommended computational framework as the *baseline approach*.

An alternative framework by Yu.B. Vinogradov (Vinogradov 1988) employs a family of functionally normal curves and nonparametric methods for estimating their parameters from samples. In the present study, the three-parameter lognormal distribution and the C3 distribution (described via transformation) were utilized:

$$z = 0.5\,(x^a + 1)\, ln\; x \qquad (5)$$

here z is the normally distributed random variable, x – the initial variable, a – the transformation parameter.

The study employed direct numerical fitting of analytical curves to sample points (calibration) using various convergence measures. For sample points, unbiased estimates of the empirical distribution function coordinates were adopted. The parameters of the analytical distribution function were computed based on minimizing the convergence measure.

The total relative divergence between the empirical and analytical curves in terms of probability was used as the convergence measure:

$$\omega = \sum_{i=1}^{n} \left(\frac{\left|p^* - p^{**}\right|}{p_b^* - p_a^*} \right)_i \qquad (6)$$

where $p^* = m/n$ is the empirical probability of the order statistic members; p^{**} is the analytical probability; p_a^* and p_b^* are the confidence interval bounds for probability p given m and n at a specified significance level; m and n represent the rank of the i-th value and the total sample size, respectively (Vinogradov 1988, p. 251). The minimization of the ω metric ensures probability-based convergence between the analytical and empirical curves and characterizes the reliability of the adopted solution.

An additional convergence measure based on absolute magnitude was employed, defined as the root mean square deviation between the ordinates of the empirical and analytical curves. This metric characterizes the precision of the adopted solution.

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(k_p^* - k_p^{**}\right)^2} \qquad (7)$$

where: k_p^* and k_p^{**} are equally probable quantiles of the empirical and analytical distributions, expressed as modular coefficients. Thus, Vinogradov's alternative approach enables the construction of four distribution curve variants:

— Lognormal with ω-based approximation
— Lognormal with s-based approximation
— C3 with ω-based approximation
— C3 with s-based approximation

The metrics ω (reliability) and s (precision) subsequently serve as criteria for comparative evaluation of all computed probability curves. A key adequacy criterion for the analytical probability curve is its containment within the confidence interval bounds (p_a, p_b) of the empirical curve.

Additionally, among the widely used and internationally recommended probability distributions for peak flow calculations, the following were selected:

— Generalized Extreme Value (GEV) distribution
— Generalized Logistic (GLO) distribution
— Generalized Pareto (GPO) distribution
— Two-parameter Gumbel (GM2) distribution

The selection of these distributions is supported by established research (Pisarenko et al., 2002; Naydenov, 2004; Gubareva, 2010), which demonstrates that heavy-tailed distributions are most appropriate for probabilistic modeling of extreme hydrological characteristics, particularly in the upper tail section. The application of these distributions also leads to alternative computational approaches (Gartsman, Gubareva, Kichigina, 2020).

The L-moments method, proposed relatively recently as an alternative approach for characterizing probability distribution shapes (Hosking & Wallis, 1997), offers several key advantages over conventional methods and has gained widespread international adoption. Theoretically, L-moments represent a modification of the probability-weighted moments (PWMs) originally introduced by Greenwood et al. (1979).

The procedure begins by calculating unbiased sample estimates of probability-weighted moments (PWMs) from an ascendingly ordered sample $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$ of size n that can be described as follows:

$$b_0 = n^{-1} \sum_{j=1}^{n} x_{j \div n} \tag{8}$$

The generalized expression for b_r is

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_{j \div n} \tag{9}$$

The sample L-moments are derived as

$$\begin{gathered} l_1 = b_0,\ l_2 = 2b_1 - b_0, \\ l_3 = 6b_2 - 6b_1 + b_0, \\ l_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \end{gathered} \tag{10}$$

Or in general

$$l_{r+1} = \sum_{k=0}^{r} p^{*}_{r,k} b_k;\ r - 0.1,\ \ldots,\ n-1 \tag{11}$$

where $p^{*}_{r,k}$ coefficients are derived as

$$p^{*}_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2 (r-k)!} \tag{12}$$

The sample L-moments of r-th order are derived as

$$t = \frac{l_2}{l_1},\ t_r = \frac{l_r}{l_2},\ r = 3,4 \ldots \tag{13}$$

where t – sample L-coefficient of variation, t_3 – sample L-coefficient of skewness, t_4 – sample L-kurtosis. In the study by (Gubareva & Gartsman 2010), algorithms are provided for the mutual computation of L-moments and parameters for several three-parameter distribution laws. These distributions are characterized by location (shift), shape, and scale parameters.

Table 2 presents the analytical probability distributions and parameter estimation methods employed in this study to develop probability curves for peak water discharge values. Each computational approach combines a specific probability distribution with a corresponding parameter fitting technique, forming a unique variant for analysis.

The performance of these variants is evaluated through multiple criteria. The metrics ω (reliability) and s (precision) are computed for each variant to enable quantitative comparison. Additionally, the analysis examines how closely each fitted probability curve remains within the 90% confidence interval boundaries *(p_d, p_b)* of the empirical probability curve across all probability points. This provides a measure of statistical consistency between the analytical and observed data. Complementing these quantitative assessments, the study incorporates expert evaluation of the hydrological plausibility of the peak discharge estimates obtained from each variant. This qualitative judgment considers whether the results align with physical expectations and regional hydrological characteristics.

Results of Extreme Flood Probability Assessment

Table 3 presents the evaluation of linear trend significance based on the Student's t-test and correlation coefficient error across all analyzed observation series (refer to **Table 1**). The analysis reveals statistically significant trends in nearly one-third of the 19 examined time series. Specifically, five series exhibit downward trends indicating decreasing flood magnitudes, while one series demonstrates an upward trend suggesting increasing flood magnitudes. Importantly, all detected trends show consistent significance when assessed through both applied statistical criteria. The identified trends may stem from diverse underlying factors, including climatic influences such as shifting precipitation regimes or anthropogenic impacts like land-use modifications and water management practices. Given this complexity, a targeted follow-up study is recommended to elucidate the precise drivers behind these observed hydrological changes.

Figure 2 presents the development of probabilistic flood frequency models for the Kalzhyr River at Kalzhyr village (catchment area 434 km², 65-year observation period), which exhibits a statistically significant downward trend. The modeling approach involved an exhaustive evaluation of all possible combinations of theoretical distribution laws and parameter estimation methods described in Table 2, with final model selection based on both quantitative

Table 2. Methodical tools for constructing of probability curves

Analytical Distribution Laws	Parameter Estimation Methods (Approximation)
Pearson Type III (PIII) Three-parameter gamma distribution by Kritsky-Menkel (KM3) Three-parameter lognormal distribution (LN3) Functional-normal distribution C3 by Vinogradov (VC3) Generalized Extreme Value distribution (GEV) Generalized Logistic distribution (GLO) Generalized Pareto distribution (GPD) Two-parameter Gumbel distribution (GM2)	Method of Moments (Mom) Approximate Maximum Likelihood Method (MLh) Graphical-Analytical Method (GA) L-Moments Method (LMo) Nonparametric Calibration Method: - By reliability (Pl) - By accuracy (Ac)

Table 3. Testing the series of observed annual maximum discharges for stationarity

No.	River	Basin area (km^2)	Sample length	Significance test	
				R error	t-test
1	Bas-Terekty – Moiyldy	184	48	-	-
2	Belaya – Beloe	945	61	-	-
3	Bukhtyrma – Berel	1,850	58	-	-
4	Bukhtyrma – Lesnaya Pristan	10,700	69	-	-
5	Bukhtyrma – Pechi	6,860	80	-	-
6	Glubochanka – Belokamenka	47	41	-	-
7	Kalzhir – Kalzhyr	3,150	65	<0.05 (-)	<0.01 (-)
8	Kandysu – Saryolen	2,610	33	<0.05 (-)	<0.05 (-)
9	Kara Ertis – Boran	55,900	82	-	-
10	Kishi Ulbi – Gornaya Ulbinka	2,170	36	-	-
11	Kurchim – Voznesenka	5,840	79	-	-
12	L. Berezovka – Sredigornoe	251	74	-	-
13	Naryn – Ulken Naryn	1,960	67	<0.05 (+)	<0.05 (+)
14	Oba – Shemonaikha	8,470	63	<0.05 (-)	<0.01 (-)
15	Oba – Karakozha	2,768	49	<0.05 (-)	<0.05 (-)
16	Turgysyn – Kutikha	1,200	59	<0.05 (-)	<0.05 (-)
17	Ulken Boken – Djumba	758	65	-	-
18	Ulbi – Ulbi-Perevalochnaya	4,900	82	-	-
19	Chernovaya – Chernovoe	488	42	-	-

goodness-of-fit criteria and qualitative expert judgment. The model selection process was conducted in two distinct phases. The initial phase considered only those methods explicitly recommended by current regulatory guidelines, while the subsequent phase expanded the evaluation to include alternative computational approaches not covered by standard protocols. Given the presence of a statistically significant trend in the series, the analysis was performed separately for both the original observed data series (Fig. 2a) and a detrended series normalized to initial conditions (Fig. 2b). Model verification employed specialized normal probability paper, which transforms the cumulative normal distribution function into a linear plot, with discharge values plotted on a logarithmic scale. This visualization technique allows for immediate assessment of model performance by examining how closely the theoretical distribution curves align with the empirical data points and remain within the 5-95% confidence intervals across the entire probability range.

Appendix A presents a comprehensive comparison of optimal flood frequency models for annual maximum discharge series at 19 gauging with records exceeding 30 years (see **Table 1**). The table systematically organizes selection results for both standard methodology and alternative approaches, enabling direct comparison of their performance characteristics. For each station, the analysis provides parallel sets of parameters for the optimal standard methodology solution and the best-performing alternative approach. The tabulated parameters include key metrics for model evaluation: the observed maximum discharge (Q_{max}) as a reference value, the selected probability distribution type, the parameter estimation method, quantitative reliability (ω) and precision (s) estimates, and the calculated 1% exceedance discharge $(Q_{1\%})$. These metrics collectively allow for assessment of both statistical adequacy and engineering safety implications, with particular attention to differences in extreme quantile estimation. For the six stations exhibiting statistically significant trends (previously identified in **Table 3**), **Appendix A** presents comparative results for both the original observed series and the detrended, normalized series.

Thus, **Appendix A** presents data from 25 individual calculations, each performed using both standard and alternative methodological approaches. The analysis reveals several key findings regarding the performance of different probability distributions and parameter estimation techniques. When applying the standard methodology, the Pearson Type III distribution demonstrated superior performance in 17 out of 25 cases, while the Kritsky-Menkel distribution proved optimal in the remaining 8 cases. The standard approach did not recommend any other probability distributions for these datasets. Regarding parameter estimation techniques, the method of moments and the graphical-analytical method each provided the best solution in 12 cases, whereas the maximum likelihood method yielded optimal results in only 1 out of 25 instances. An important engineering safety consideration emerges from the observation that in 20 out of 25 cases, the calculated 1% exceedance discharge $(Q_{1\%})$ exceeds the observed maximum discharge (Q_{max}), which

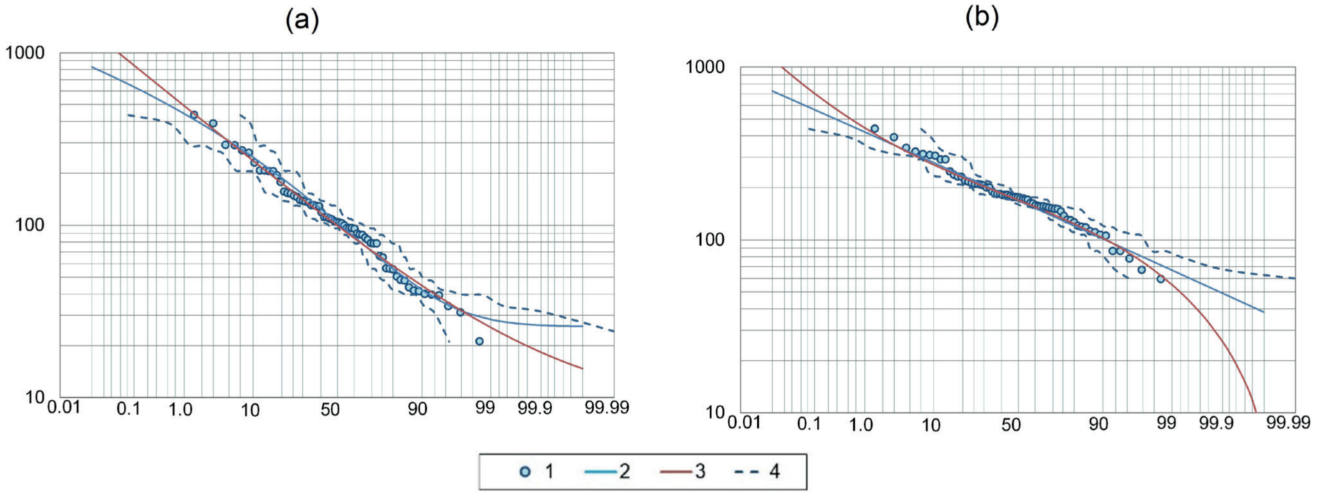


Fig. 2. Empirical and analytical exceedance probability curves for annual maximum discharges at the Kalzhyr River gauging station (catchment area = 434 km², record length = 65 years). Legend: 1 – Empirical exceedance probability curve, 2 – Optimal analytical curve from standard methodology, 3 – Optimal alternative analytical curve, 4 – 5-95% confidence interval bounds for empirical curve; (a) Analysis of original discharge series (trend not accounted for): best fit within the standard methodology: the Pearson Type III (PIII) distribution with graphical-analytical (GA) parameter estimation; best fit among alternative approaches: the three-parameter lognormal distribution (LN3) with accuracy-based calibration (Ac); (b) Analysis of trend-corrected discharge series: best fit within the standard methodology: the Kritsky-Menkel (KM3) distribution using method of moments (Mom) parameterization; best fit among alternative approaches: the generalized logistic distribution (GLO) with L-moments (Lmo) estimation

may be interpreted as an expert criterion for sufficient safety margin in engineering design.

For non-stationary series subjected to trend correction, the Q1% values exhibited an equal probability of either increasing or decreasing (3 cases each), despite the consistent increase in mean values following detrending procedures. In their original form, all non-stationary series were best described by the Pearson Type III distribution. However, after trend adjustment, half of these cases showed improved fit with the Kritsky-Menkel distribution, indicating the significant influence of trend treatment on distribution selection.

The alternative calculation approaches demonstrate substantially greater diversity in both probability distributions and parameter estimation methods, revealing noteworthy patterns in their application outcomes. Among the 25 cases analyzed, the three-parameter lognormal distribution emerged as optimal in 7 instances, followed by the generalized logistic distribution (6 cases), Pearson Type III distribution with L-moments parameterization (5 cases), Vinogradov's C3 distribution and generalized extreme value distribution (3 cases each), and the two-parameter Gumbel distribution (1 case). Parameter estimation methods exhibited more consistent performance characteristics. From all available alternatives, either the L-moments method (15 cases) or the accuracy-based calibration (s-method, 10 cases) consistently provided optimal solutions. The latter approach, as previously noted, involves direct fitting of analytical curves to empirical data points by minimizing the mean squared differences in discharge values.

For the six non-stationary time series analyzed with both original and trend-corrected approaches, the distribution selection demonstrates specific patterns. In four cases, either the same distribution type was maintained or replaced with a functionally similar alternative (e.g., different parameterizations of gamma-type distributions). However, two cases exhibited more substantial changes - transitioning from Pearson Type III to Generalized Extreme Value distribution in one instance, and from three-parameter lognormal to Generalized Logistic distribution in another.

The parameter estimation methods showed greater stability during trend correction procedures, remaining unchanged in five out of six cases. Notably, the alternative calculation approaches produce more conservative engineering estimates. The computed 1% exceedance discharge ($Q_{1\%}$) exceeds the observed maximum discharge (Q_{max}) in 22 of 25 cases (88%), with these exceedances being more pronounced than those obtained through standard regulatory methods. Furthermore, when applying alternative methods to detrended series, $Q_{1\%}$ values more frequently increased than decreased (4 cases versus 2), indicating an inherent tendency toward greater safety margins in the alternative framework.

The quantitative evaluation metrics (reliability *ω* and precision *s*) provide compelling evidence for the superiority of alternative approaches. In 20 of 25 cases (80%), these criteria unequivocally indicate better performance characteristics for probability models developed using alternative methodologies compared to standard regulatory solutions.

CONCLUSION

The results of maximum flow analysis in the study region generally correspond to previously established global patterns in flood frequency distributions (Gubareva 2011). The application of a comprehensive suite of stochastic hydrology tools for probabilistic estimation of extreme discharges in rivers of the East Kazakhstan region leads to several fundamental conclusions.

The analysis demonstrates the clear superiority of alternative computational approaches over the standard SP 33-101-2003 methodology. This conclusion is supported by three key factors:

— the alternative schemes employ heavy-tailed probability distributions that more accurately characterize extreme flood behavior, as documented in hydrological literature (Naidenov, 2004; Gubareva, 2010, 2011);

— the alternative schemes consistently achieve superior performance metrics, showing significantly better values for both reliability (ω) and precision (s) indicators;

— the alternative approaches systematically produce higher estimates for the 1% exceedance discharge (Q1%), thereby providing increased safety margins for flood protection infrastructure design.

The prevalence of different parameter estimation methods among the selected optimal models serves as indirect evidence supporting the greater adequacy of alternative approaches. Within the 25 optimal solutions obtained using the standard methodology, the approximate maximum likelihood method appears only once, despite its theoretical superiority as the most statistically efficient estimation technique (Gubareva, Gartsman 2010). This apparent contradiction can be explained by noting that the method's advantages are strictly contingent upon correct distributional assumptions. The observed results therefore suggest potential inadequacies in the standard framework's prescribed probability distributions for modeling extreme flood characteristics. In contrast, alternative methodologies demonstrate fundamentally different patterns. The L-moments method, which shows comparable precision and robustness to maximum likelihood estimation according to established research (Hosking & Wallis, 1997; Gubareva & Gartsman, 2010), appears in over half of the optimal alternative solutions. This striking methodological consistency strongly indicates that the alternative probability distributions better correspond to the actual statistical properties of maximum discharge series. The robust performance of L-moments estimation in this context provides compelling evidence for the theoretical soundness of the alternative distributional models, as the method's statistical properties are known to be particularly sensitive to misspecification of the underlying probability distribution.

The second key finding of the regional analysis reveals the relatively limited impact of non-stationarity – at least in the form of downward trends – on the estimation of rare flood discharge quantiles. This conclusion emerges from a detailed examination of trend-adjusted series, where maximum discharge records were normalized to the highest moving average observed during the monitoring period. While this adjustment consistently produces mean values significantly higher than in the original series, the resulting changes in 1% exceedance discharge (Q1%) estimates remain within a modest range of -10% to +11% deviation from original values. This observed variation in $Q_{1\%}$ estimates due to trend correction falls well within the combined error envelope encompassing both measurement inaccuracies and computational uncertainties inherent in flood frequency analysis. Constructing the correct and adequate flood frequency curves based on relevant probability distributions may aide to future regional hydrological model development, which can be calibrated using signature measures based on flow duration curves (Gartsman, Solomatine, Gubareva, 2024).

The observed non-stationarity in river flow regimes across Eastern Kazakhstan represents a complex phenomenon requiring systematic investigation. Current evidence points to climate change as the primary driver, characterized by a well-documented rise in mean annual temperatures throughout the region. This warming trend has fundamentally altered the hydrological cycle through several interconnected mechanisms. A well-established redistribution of precipitation patterns has occurred, with meteorological records from RSE "Kazhydromet" and supporting studies documenting decreased summer rainfall alongside increased winter precipitation. This seasonal shift in moisture availability has significantly modified the hydrological behavior of regional rivers (Stambekov, Turulina, 2016, Salnikov et al., 2014). Increased aridity has led to more frequent forest fires, resulting in substantial reductions in forest cover – a critical natural regulator of both surface and subsurface flow regimes (Lebed', Eserkepova, Suleimenov, 2020). While illegal logging activities represent an additional stressor, their contribution to overall forest loss appears secondary compared to climate-driven impacts. These transformations manifest most distinctly in Group I river basins, including the Oba, Turgysyn, Naryn, and Kalzhyr watersheds, where hydrological changes have been particularly pronounced.

Human economic activities represent a significant factor influencing alterations in hydrological regimes, particularly through agricultural expansion and water infrastructure development (Milanović Pešić 2024). In the Tarbagatai district, for instance, water from the Kandysu River is diverted through an extensive network of irrigation canals to support growing cropland areas. Under increasingly arid climatic conditions, agricultural water withdrawals have risen substantially, creating measurable impacts on river discharge patterns. The observed non-stationarity in flow records for the Turgysyn River may be directly attributed to the construction and commissioning of the Turgysyn Hydroelectric Power Station in 2021. Such hydraulic engineering projects fundamentally modify natural flow regimes through flow regulation, sediment trapping, and alteration of seasonal discharge patterns. ■

REFERENCES

Belyanin V.I., Manucharyan L.S., Novikov Yu.A., Chursin A.S., Pavlenko A.V. Operational Resources of Groundwater in Eastern Kazakhstan. Ust-Kamenogorsk, 2013. 57 p. (in Russian)

Eastern Flood (Electronic resource) / Vremya. URL: https://time.kz/articles/territory/2018/03/26/vostochnij-potop (accessed: 26.03.2018) (in Russian)

Egorina A.V., Popova K.I., Dyukarev A.D., Kondratyev V.P., Chursin A.S. Climate of Southwestern Altai / Ed. by A.V. Egorina. Ust-Kamenogorsk: Semey-Pechat, 2015. 315 p. (in Russian)

Egorina A.V., Zinchenko Yu.K., Zinchenko E.S. Physical Geography of East Kazakhstan Region. Ust-Kamenogorsk, 2000. 124 p. (in Russian)

Gartsman B.I., Gubareva T.S., Kichigina N.V. Possibilities of Estimating Maximum Discharge Probability in Angara River Tributaries (Case Study of the Iya River) // Hydrosphere. Hazardous Processes and Phenomena. 2020. Vol. 2. No. 4. Pp. 347-364. DOI: 10.34753/HS.2020.2.4.347 (in Russian)

Gartsman B.I., Solomatine D.P., Gubareva T.S. A method of multi-site calibration of distributed hydrological models based on the Nash-Sutcliffe efficiency. GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY. 2024;17(4):76-87. https://doi.org/10.24057/2071-9388-2024-3564

Greenwood J.A., Landwehr J.M., Matalas N.C, Wallis J.R. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form // Water Res. Research. 1979. V. 15. № 5. P. 1049–1054. DOI: 10.1029/WR015i005p01049

Gubareva T.S. Distribution Types for Extreme Flood Estimation // Geoecology. 2010. No. 5. Pp. 446-457. (in Russian)

Gubareva T.S. Spatial Patterns of Flood Distribution in Temperate Zone Rivers of the Northern Hemisphere // Proceedings of the Russian Academy of Sciences. Geographical Series. 2011. No. 2. Pp. 65-77. (in Russian)

Gubareva T.S., Gartsman B.I. Estimation of Extreme Hydrological Values Distribution Parameters Using L-Moments Method // Water Resources. 2010. Vol. 37. No. 4. Pp. 398-407. (in Russian)

Handbook of hydrology // Ed. Maidment D.R. McGRAW-HILL, INC. 1993.

Hosking J.R.M., Wallis J.R. Regional Frequency Analysis. An Approach Based on L-Moments. Cambridge: Cambridge Univers. Press, 1997. 224 p. DOI: https://doi.org/10.1017/CBO9780511529443

Lebed L.V., Yeserkepova I.B., Suleimenov N.K. Forest and Steppe Fires in Kazakhstan as Sources of Gas and Particulate Emissions into the Atmosphere // Hydrometeorology and Ecology. 2020. No. 1 (96). Pp. 175-186. (in Russian)

Milanović Pešić A., Jakovljević D., Ćulafić G., Milivojević M. Water regime variability of selected rivers on the Balkan peninsula: a comparative study of central Serbia and northern region of Montenegro. GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY. 2024;17(4):35-43. https://doi.org/10.24057/2071-9388-2024-3404

MSP 3.04-101-2005. Determination of Main Design Hydrological Characteristics. Moscow, 2006. 95 p. URL: https://online.zakon.kz/Document/?doc_id=30099725 (in Russian)

Naidenov V.I. Nonlinear Dynamics of Land Surface Waters. Moscow: Nauka, 2004. 318 p. (in Russian)

Pisarenko V.F., Bolgov M.V., Osipova N.V., Rukavishnikova T.A. Application of Extreme Value Theory in Probability Distribution Approximation of Maximum Water Discharges // Water Resources. 2002. Vol. 29. No. 6. Pp. 645-657. DOI: 10.7868/S0002333714030077 (in Russian)

Rozhdestvensky A.V., Chebotarev A.I. Statistical Methods in Hydrology. Leningrad: Gidrometeoizdat, 1974. 424 p. (in Russian)

Salnikov V.G., Turulina G.K., Polyakova S.E., Petrova E.E. Climate Change and Its Regional Manifestations in Kazakhstan // Hydrometeorology and Ecology. 2014. No. 2 (73). Pp. 17-31. (in Russian)

SP 529.1325800.2023. Determination of Main Design Hydrological Characteristics. Moscow, 2023. 110 p. (in Russian)

Stambekov M.D., Turulina G.K. Temperature Regime Features in Western and Eastern Kazakhstan During Spring in Recent Decades // Hydrometeorology and Ecology. 2016. No. 4 (83). Pp. 7-23. (in Russian)

State Water Cadastre of RSE Kazhydromet (Electronic resource). URL: https://www.kazhydromet.kz/ru/gidrologiya/gosudarstvennyy-vodnyy-kadastr-poverhnostnye-vody (accessed: 01.02.2025) (in Russian)

Vinogradov Yu.B. Mathematical Modeling of Runoff Formation Processes. Moscow: Gidrometeoizdat, 1988. 365 p. (in Russian)

Year 2018 Is Not Over Yet: Enormous Damage Already Caused to the State by Natural Disasters (Electronic resource) / Caravan.kz. URL: https://www.caravan.kz/news/2018-god-eshhe-ne-zakonchilsya-kakojj-kolossalnyjj-ushherb-uzhe-poneslo-gosudarstvo-izza-prirodnykh-katastrof-446862/ (accessed: 23.05.2018) (in Russian)

Appendix

Appendix A. Optimal design curves for annual maximum instantaneous peak discharges of rivers in the East Kazakhstan Region

№	River, gauge	Q_{max}, m^3/s	Standard methodology					Alternative approach				
			Distr. type	Param. estimat.	*ω*	*s*	$Q_{1\%}$	Distr. type	Param. estimat.	ω	s	Q1%
1	Bas-Terekty – Moiyldy	69.7	PIII	GA	11.9	1.46	58.6	VC3	Ac	8.03	0.57	84.7
2	Belaya – Beloe	305	KM3	Mom	11.3	0.12	316	GM2	LMo	11.1	0.11	322
3	Bukhtyrma – Berel	444	PIII	Mom	6.39	0.07	476	PIII	LMo	6.39	0.06	482
4	Bukhtyrma – Lesnaya Pristan	2740	KM3	Mom	9.67	0.07	2641	GLO	LMo	5.66	0.06	2738
5	Bukhtyrma – Pechi	1340	KM3	Mom	14.8	0.28	1324	GLO	LMo	10.5	0.27	1321
6	Glubochanka – Belokamenka	11.7	KM3	Mom	11.2	0.36	11.9	LN3	Ac	10.4	0.34	12.6
7	Kalzhir – Kalzhyr	434	PIII	GA	10.9	0.52	445	LN3	Ac	10.1	0.31	496
	Same – detrended		KM3	Mom	12.5	0.25	428	GLO	LMo	8.77	0.24	447
8	Kandysu – Saryolen	19.7	PIII	GA	4.72	0.02	22.7	GLO	LMo	3.20	0.03	21.7
	Same – detrended		KM3	Mom	3.50	0.02	21.4	GLO	LMo	3.28	0.02	21.8
9	Kara Ertis – Boran	2330	PIII	GA	14.1	0.13	2468	GLO	LMo	12.1	0.09	2485
10	Kishi Ulbi – Gornaya Ulbinka	1060	PIII	Mom	7.34	0.19	1223	PIII	LMo	7.37	0.18	1283
11	Kurchim – Voznesenka	1050	KM3	MLh	7.30	0.06	1074	PIII	LMo	7.22	0.05	1071
12	L. Berezovka – Sredigornoe	27.1	PIII	GA	11.1	0.52	33.4	PIII	LMo	13.8	0.38	31.0
13	Naryn – Ulken Naryn	166	PIII	Mom	22.7	0.85	155	VC3	Ac	7.99	0.26	191
	Same – detrended		PIII	GA	26.9	0.51	168	LN3	Ac	13.6	0.17	187
14	Oba – Shemonaikha	3050	PIII	Mom	6.27	0.04	3178	PIII	LMo	5.79	0.03	3254
	Same – detrended		PIII	GA	5.76	0.03	3428	GEV	LMo	6.05	0.03	3366
15	Oba – Karakozha	2580	PIII	Mom	8.40	0.34	2639	LN3	Ac	9.29	0.24	2909
	Same – detrended		PIII	GA	7.74	0.15	2926	LN3	Ac	8.17	0.13	2954
16	Turgysyn – Kutikha	733	PIII	GA	7.90	0.11	833	GEV	LMo	7.98	0.11	773
	Same – detrended		KM3	Mom	6.94	0.08	826	GEV	LMo	7.02	0.07	845
17	Ulken Boken – Djumba	428	PIII	GA	10.5	0.47	460	LN3	Ac	8.78	0.36	498
18	Ulbi – Ulbi-Perevalochnaya	2710	PIII	GA	39.0	1.18	2717	VC3	Ac	24.6	0.63	2762
19	Chernovaya – Chernovoe	117	PIII	GA	10.6	0.29	121	LN3	Ac	8.21	0.20	129

Check for updates

CC BY

RESEARCH PAPER

THE INFLUENCE OF METEOROLOGICAL FACTORS ON SULFUR DIOXIDE (SO_2) AND NITROGEN DIOXIDE (NO_2) AND PREDICTION MODEL FOR RAINWATER ACIDITY BASED ON THEIR CONCENTRATIONS IN JAKARTA CITY

Marwah Noer[1,2], **Eko Kusratmoko**[2], **Asep Karsidi**[2]
[1]Geography Department, University of Indonesia, Margonda Raya Street, 16424, Depok, Indonesia
[2]Ministry of Agrarian Affairs and Spatial Planning/National Land Agency, Sisingamangaraja Street No.2, 12014, Jakarta, Indonesia
*Corresponding author: Marwah Noer

Received: November 13rd 2024 / Accepted: November 12th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3724

ABSTRACT. In the last few decades, rapid industrial growth and increasing urban traffic density have caused air quality problems, including in Jakarta. One indicator of air pollution is acid rain. Acid rain occurs due to pollutants in the form of SO_2 and NO_2 reacting with water (H_2O). The impact of acid rain, among other things, can damage soil fertility, affect the quality of human life, and damage objects and infrastructure. This research aims to analyze the distribution of SO_2 and NO_2 spatially and temporally and create a rainwater acidity model based on the distribution of SO_2 and NO_2 in Jakarta. The distribution of SO_2 and NO_2 was obtained using remote sensing techniques using Sentinel 5P Satellite imagery. Processing is carried out using GEE. From the results of the bivariate analysis, it is known that the spatial distribution of SO_2 is influenced by rainfall and is not influenced by wind speed. Meanwhile, the distribution of NO_2 is significantly influenced by rainfall and wind speed. Temporally, the distribution of SO_2 in 2023 has the highest value in June, and the distribution of NO_2 has the highest value in August. The prediction model for rainwater acidity levels was obtained based on the distribution of SO_2 and NO_2 in 2023 in Jakarta. The results of multiple linear regression show a correlation between rainwater acidity and the distribution of SO_2 and NO_2. The correlation coefficient is (-) 0.7305, which means the correlation is in the strong category. The negative correlation explains that the higher the levels of SO_2 and NO_2, the more acidic the rainwater will be. A value of 13% was obtained in the MAPE calculation, which means the prediction model is included in the good category and can be used to predict rainwater acidity in Jakarta.

KEYWORDS: air pollution, rain acidity, SO_2, NO_2, Sentinel 5P, bivariate analysis

CITATION: M. Noer, E. Kusratmoko, A. Karsidi. (2026). The Influence of Meteorological Factors on Sulfur Dioxide (SO_2) and Nitrogen Dioxide (NO_2) and Prediction Model For Rainwater Acidity Based On Their Concentrations in Jakarta City. Geography, Environment, Sustainability, 1 (19), 62-71
https://doi.org/10.24057/2071-9388-2026-3724

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Air pollution is a condition of the atmosphere in which the concentration of a particular substance has adverse effects on human health and the environment, including global warming, acid rain, and ozone layer depletion [1],[2]. The World Health Organization (WHO) states that air pollution kills millions of people worldwide every year [3]. Research results show that most of the world's population faces this problem, especially in big cities [4]. In Southeast Asia, air pollution is the biggest cause of non-communicable diseases [5]. In recent decades, rapid industrial growth and increasing urban traffic density have become serious air quality problems [6]. The decline in air quality is one of the consequences of the growth and development of a city [7]. Air mixed with hazardous components at certain levels can be categorized as air pollution. In 2020, the government of the Republic of Indonesia determined that air quality is measured from the concentration of seven parameters, namely carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), PM 10, PM 2.5, sulfur dioxide (SO_2), and hydrocarbons (HC) [5].

As a pollutant, SO_2 has a significant impact on the environment and global climate [8]. SO_2 is a toxic gas that is colorless and has an odor if its concentration is more than 5 ppm [9]. The effect of SO_2 on humans is to cause respiratory disorders such as asthma and throat irritation. Even high levels of SO_2 inhaled by humans for a long time can cause lung cancer [10]. NO_2 is a reaction between nitrogen oxide and oxygen. NO_2 is a grayish-red substance that can cause eye irritation and even with high exposure can cause lung disease in humans [11]. In clean air, NO_2 cannot be found [12], therefore, if NO_2 is found in the air at a certain level, it can be ascertained that the area is polluted. In urban areas, the concentration of NO is usually 10-100 times greater than in rural areas [13]. SO_2 and NO_2

gases in the air are the cause of acid rain. Rainwater plays a role in washing pollutants in the air and carrying them from the atmosphere to the soil or water below [14], [15]. Acid rain can harm the environment and also the health of living things [16]. One indicator of air pollution is acid rain [17]. Acid rain occurs when pollutants in the form of SO_2 and NO_2 react with water (H_2O). Acid rain can affect soil fertility, residents' quality of life, and damage objects or infrastructure made of iron, limestone, concrete, and marble [18]. Rainwater that has a pH level of less than 5.6 can be categorized as acid rain [19].

The air quality of Jakarta City has become a serious topic of discussion among the national and international public. Throughout 2023, Jakarta's air quality even showed a fairly serious condition because it was categorized as unhealthy. Jakarta City has several times become the city with the worst air quality in the world in August and September 2023. The Ministry of Environment and Forestry (KLHK) revealed that several factors can cause air pollution in Jakarta, both natural and unnatural. Natural factors can be seasons, wind direction, and speed, and even the city landscape can also cause high pollution in Jakarta. When air pollution reaches a high level, it is typically during the dry season in Jakarta, when rainfall is minimal and air temperatures exceed 35 °C. This natural factor is a factor that is difficult for humans to control. Unnatural factors come from activities carried out by humans in an area such as the transportation sector, industry, and household activities.

Air is a very important element for living things [20], so air quality is something that every country needs to pay attention to following the Sustainable Development Goals (SDGs), which call for reducing deaths and diseases due to air pollution (point 3.9.1), ensuring access to clean energy at home (point 7.1.2), reducing the impact of urban environments by improving air quality (point 11.6.2). A spatial-temporal study is needed on the distribution of pollutants in Jakarta City and their relationship to meteorological factors and rainwater acidity. Therefore, the purpose of this study is to analyze the spatial and temporal distribution of SO_2 and NO_2 in 2023 using remote sensing data, and to develop a prediction model for rainwater acidity levels based on the distribution of SO_2 and NO_2 in 2023 within the study area.

MATERIAL AND METHOD

Study Area

The research was conducted in Jakarta City, which consists of North Jakarta, Central Jakarta, East Jakarta, West Jakarta, and South Jakarta. Jakarta has an area of 661.23 km^2 (Decree of the Minister of Home Affairs Number 050-145 of 2022 concerning the Granting and Updating of Codes, Data on Government Administrative Areas and Islands in 2021), consisting of 44 sub-districts and 267 urban villages. Geographically, Jakarta Province is between 106°22'42" East Longitude to 106°58'18" East Longitude and 5°19'12" South Latitude to 6°23'54" South Latitude. Jakarta has a tropical climate and has two seasons, namely the dry season and the rainy season. The peak of the rainy season usually occurs in January and February with an average temperature of 27 °C. In comparison, the peak of the dry season occurs in August-October with temperatures that can reach 40 °C. In 2023, the highest rainfall experienced in Jakarta was in February, 461.58 mm, and the lowest rainfall was in September, 3.44 mm. Based on the data, the population of DK Jakarta in 2023 reached 10,672,100 people [22]. This number has increased from the previous year, which was recorded at 10,640,007 people. The population of Jakarta City is equivalent to 3.87% of the total population in Indonesia, making Jakarta the province with the highest population density in Indonesia.

Methodology

This study used survey data, remote sensing data or satellite imagery, and observation data from agencies. The data obtained from the direct survey were pH data or rainwater acidity taken daily on March 1-15 when it rained. The remote sensing data or satellite imagery used was Sentinel 5P imagery. Its main purpose is to conduct atmospheric measurements with high spatial-temporal resolution, which are used for monitoring air quality, ozone, and ultraviolet radiation, and climate forecasting [23]. Sentinel 5P imagery was processed using Google Earth Engine (GEE) to see the distribution of SO_2 and NO_2 by month throughout 2023. Observation data on SO_2, NO_2, and rainwater pH from the Meteorological, Climatological, and Geophysical Agency (BMKG) are also needed to predict the estimated model for rainwater acidity in the study area. Wind speed data were obtained from windrose at each air observation station owned by the DKI Jakarta Environmental Service (DLH), and rainfall data were obtained from the BMKG.

The software used in this study was GEE, ArcGIS, and Microsoft Excel. GEE was used to analyze the distribution of SO_2 and NO_2 pollutants. GEE is a cloud-based geospatial processing platform for large-scale environmental monitoring and analysis [24]. GEE can be used for free and provides access to remote sensing imagery that is ready to use for various studies. The remote sensing image data used in this study was Sentinel 5P, which can present the distribution of SO_2 and NO_2. This distribution was then analyzed to determine whether it affected rainfall and wind speed. ArcGIS acts as a layout in presenting maps so that they are more informative and easier to understand. Microsoft Excel was used for calculating simple linear regression, multiple linear regression, and rainwater acidity prediction models. Simple linear regression was carried out to analyze the distribution of SO_2 and NO_2 with meteorological conditions, namely rainfall and wind speed. Multiple linear regression was carried out on SO_2, NO_2, and rainwater pH so that the effect of SO_2 and NO_2 on rainwater pH or rainwater acidity at each sample point could

Table 1. Data Sources and Collection Technique

Data Type	Component	Source	Data Collection Techniques
Primary Data	Rainwater Acidity	Direct Survey	Measured Using pH Meter
		BMKG	Request Data From The Agency
	SO_2 and NO_2	BMKG	Request Data From The Agency
Secondary Data	SO_2 and NO_2	Sentinel 5P Satellite Imagery	Google Earth Engine (GEE)
	Wind Speed	DLH DKI Jakarta	Literature Study From Web
	Precipitation	BMKG	Request Data From The Agency

be determined. The rainwater acidity prediction model can be obtained using multiple linear regression by calculating the coefficients generated from the previous calculations. The resulting model is then applied to the Sentinel 5P satellite imagery to produce a rainwater acidity prediction map.

The correlation value between SO_2, NO_2, and pH in this study is explained by the Pearson correlation coefficient value and determined by the equation:

$$r = \frac{\sum (xi - \overline{x})(yi - \overline{y})}{\sqrt{\sum (xi - \overline{x})^2 \sum (yi - \overline{y})}} \tag{1}$$

r is the correlation coefficient, while *x* is the value of the variable *x* at each point minus the average $(xi - \overline{x})$, and y is the value of the variable y at each point minus the average $(yi - \overline{y})$. The closer the correlation coefficient is to 1, the stronger the correlation. Correlation coefficients are categorized as <0.20, meaning very low; 0.21–0.40, meaning low; 0.41–0.60, meaning moderate; 0.61–0.80, meaning strong; and >0.81, meaning very strong. [25].

Model evaluation is carried out using Mean Absolute Percentage Error (MAPE). MAPE can provide an overview of how big the error or prediction error is by comparing it with the actual value and calculations using the model. MAPE is a measure that can be used to determine the percentage of deviation from the prediction results. Mathematically, the MAPE formula is as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y - y'|}{y} \times 100\,\% \tag{2}$$

where: *y* =actual value, *i* =time period index, *y´* =prediction value, *n* =amount of data

If the MAPE calculation results are included in the category of reasonable to very good, it can be interpreted that the model has predictive capabilities that can be used to predict several future periods [26]. The lower the MAPE value, the better the forecasting model. There are categories to classify the resulting MAPE values [26], the classification of MAPE values is < 10 % means highly accurate forecasting, 10 – 20 % means good forecasting, 20 – 50 % means reasonable forecasting, and > 50 % means weak and inaccurate forecasting .

RESULT AND DISCUSSION

Distribution of Sulfur Dioxide (SO_2) and Nitrogen Dioxide (NO_2) in 2023

The distribution of SO_2 in this research area was processed using Google Earth Engine (GEE). The distribution of SO_2 was obtained every month in 2023. The distribution of SO2 is the average SO_2 level every month throughout 2023 (Figure 1). The SO_2 level was then divided into 5 categories so that each month's spatial pattern was known.

Similar to SO_2, the distribution of NO_2 was also obtained by processing using GEE. NO_2 levels are

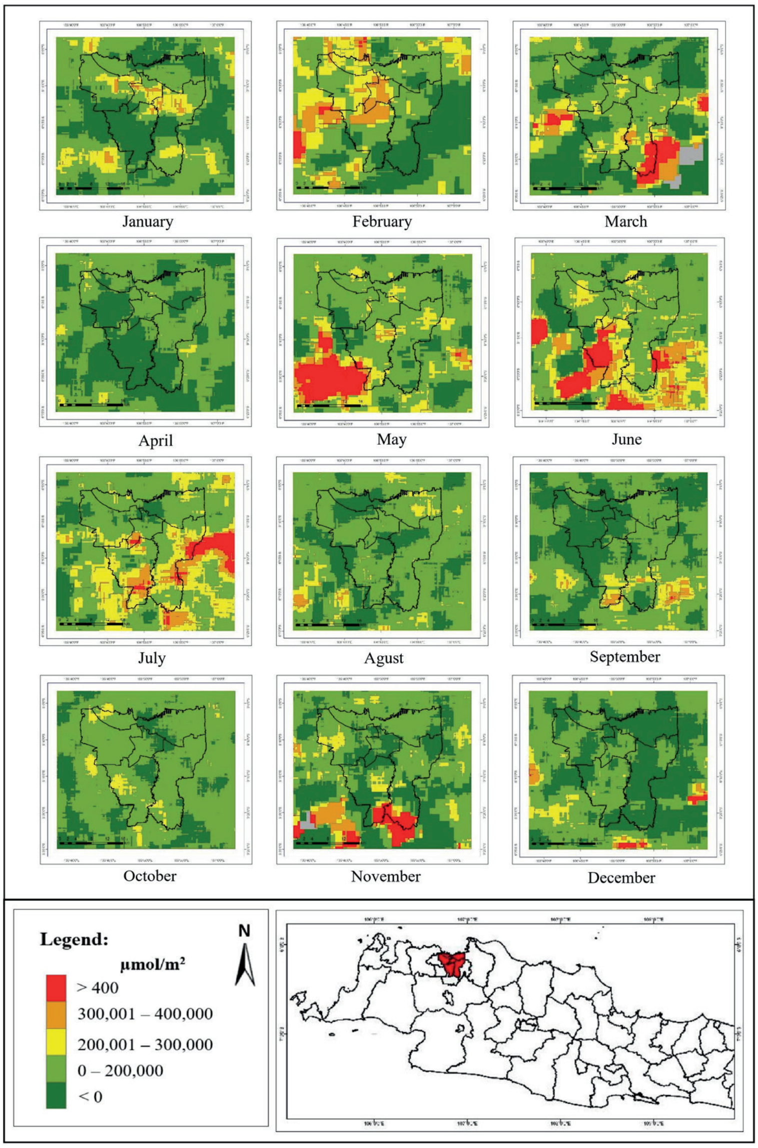


Fig. 1. Monthly SO_2 Distribution in Jakarta City Year 2023

presented in monthly ranges during 2023. The distribution of NO_2 levels in the study area has a very different pattern when compared to the distribution of SO_2 levels. In the distribution of NO_2, a fairly clear gradation of values is seen from high to low levels. In January-April, NO_2 levels appear to be dominated by very low levels (Figure 2).

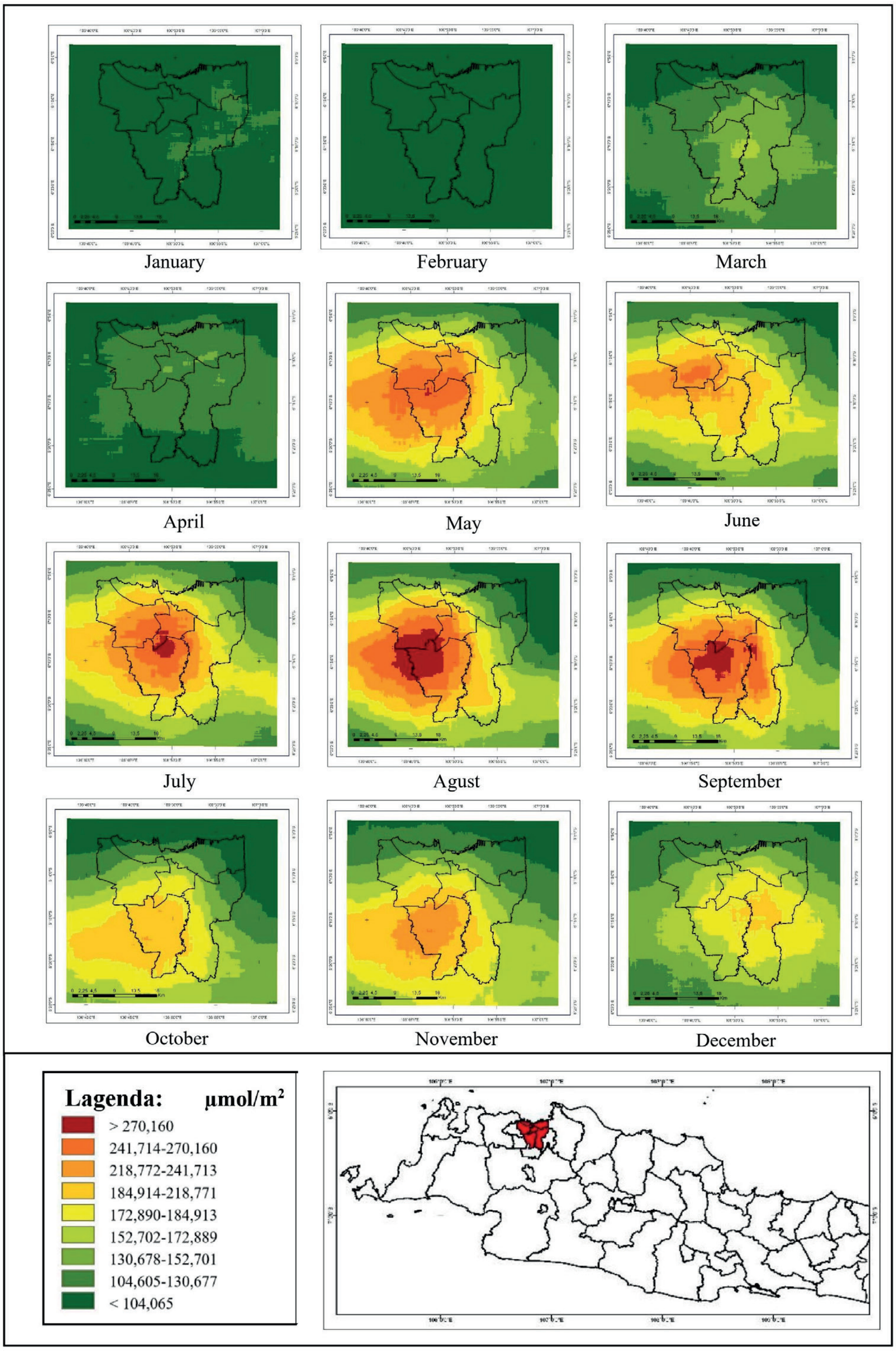
Correlation of SO_2 and NO_2 Distribution with Meteorological Factors

Correlation with Precipitation

The distribution of pollutants in the air can be influenced by several meteorological factors [27], [28], one of which is precipitation. Precipitation can act as a cleaner of pollutants in the air. Jakarta has a tropical climate because it is located close to the equator. The seasons in Jakarta are the rainy season and the dry season. The rainy season usually falls in December-March. While the dry season falls in June-September.

In 2023, Jakarta experienced a drier dry season than the last three years. This was caused by the El Nino event in Indonesia. The severe drought occurred not only in Jakarta but also in other provinces such as Aceh, Bali, West Nusa Tenggara, and Lampung. Jakarta experienced an increase in scorching temperatures, reaching almost 400^0. The dry season conditions were also said to be the cause of the poor air quality in Jakarta, which was the worst in recent years. Low precipitation caused the haze of pollution in the air to persist for a long time. In August 2023, DLH DKI Jakarta and BMKG collaborated to utilize artificial rain technology to reduce the impact of pollution in Jakarta. These efforts succeeded in reducing air pollution, although not significantly.

Data from BMKG shows that precipitation in September is the lowest throughout the year. August


January February March

April May June

July Agust September

October November December

Fig. 2. Monthly NO_2 Distribution in Jakarta City Year 2023

and October are very dry months due to minimal rainfall. Meanwhile, February and March are the months with the highest precipitation in 2023. This precipitation data is then compared with SO$_2$ and NO$_2$ levels obtained from Sentinel 5P satellite processing. SO$_2$ and NO$_2$ level data are first averaged from eight BMKG observation stations to produce an average SO$_2$ and NO$_2$ level for each month in 2023 (Table 2). The locations of the eight BMKG observation stations can be seen in figure 3.

From figure 4, it is known that NO$_2$ levels are influenced by the amount of precipitation. The lowest NO$_2$ levels were obtained in February, when the rainfall that fell was the highest throughout the year. On the other hand, in August-October, rainfall was very low, at that time NO$_2$ levels increased. This proves that the higher the precipitation, the lower the NO$_2$ levels. Regression calculations were carried out to calculate the strength of the correlation between rainfall and NO$_2$, and the results of the correlation coefficient were obtained as much as (-) 0,725, which means that the correlation that occurred was included in the strong category. The calculation of the determination coefficient showed a value of 0.526, which means that precipitation affects NO$_2$ levels by 52%, the rest is influenced by other factors.

The SO$_2$ level has a different pattern when compared to NO$_2$, in figure 4, it can be seen that the SO$_2$ level is not inversely proportional to the high precipitation. In February when the rainfall is very high, the SO$_2$ level also shows a high level. The SO$_2$ level decreased in August-November, where in that month the rainfall also decreased. The correlation between precipitation and SO$_2$ levels tends to show a positive correlation. Regression calculations were carried out to see the correlation between the two and the results obtained a correlation coefficient of (+) 0,460 which means that the correlation is included in the moderate category. The resulting determination coefficient is 0.211 which shows that rainfall affects SO$_2$ levels by only 21%.

Table 2. Monthly Precipitation and Pollutant Levels in 2023

Month	Jakarta's Monthly Precipitation (mm)	SO$_2$ (μmol/m^2)	NO$_2$ (μmol/m^2)
January	162,54	171,17	97,00
February	461,58	254,83	84,63
March	283,72	231,67	127,88
April	111,29	136,00	117,88
May	136,20	146,86	213,63
June	111,83	142,75	197,88
July	47,58	287,25	223,00
Agust	4,10	68,00	220,38
September	3,44	126,25	214,00
October	9,17	81,88	171,38
November	180,67	22,00	177,63
December	95,93	32,33	168,38

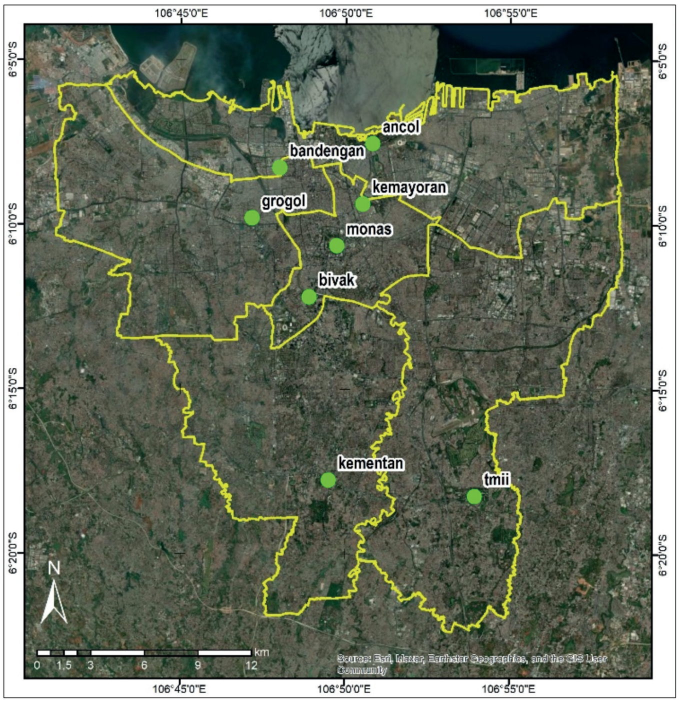


Fig. 3. BMKG Observation Station Location

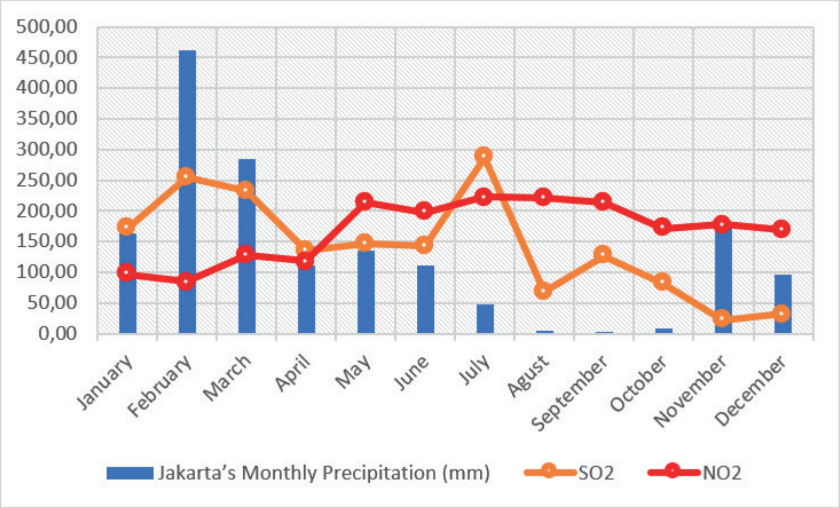


Fig. 4. Time Series of SO_2 and NO_2 Distribution with Precipitation in Jakarta

Correlation with Wind Speed

The correlation between wind speed and SO_2 and NO_2 pollutants was also analyzed from Sentinel 5P image processing data. Monthly wind speed data was obtained by calculating the average wind speed each month at five observation stations.

From the data above, a regression analysis was carried out between wind speed and NO_2, as well as between wind speed and SO_2. The results obtained explain that wind speed has a weak correlation with SO_2. This is proven by the results of calculating the correlation coefficient which is only (+) 0,2418. A P value of more than 0,05 confirms that wind speed does not have a significant influence on SO_2 distribution. This is different from the correlation between wind speed and NO_2 distribution. The results of the regression calculations show that wind speed and NO_2 distribution have a very strong correlation, which can be proven by the correlation coefficient which shows the number (-) 0,8022. The P value even shows that wind speed significantly influences the distribution of NO_2 because it is less than 0,05. The correlation between wind speed and NO_2 is negative, which means the higher the wind speed, the lower the NO_2 levels in the air.

Rainwater Acidity Level Prediction Model Based on SO_2 and NO_2 Distribution

One of the aims of this research is to create a prediction model for rainwater acidity based on the distribution of SO_2 and NO_2 in the research area. To ensure that the prediction model has a good level of confidence, it is necessary to calculate the correlation between the dependent variable (rainwater acidity) and the independent variables (SO_2 and NO_2). The correlation level calculation is carried out using multiple linear regression so that the correlation coefficient and determination coefficient of these variables can be known. Data on the acidity level (pH) of rainwater, SO_2 and NO_2 in the prediction calculations were obtained from BMKG (Table 4). All data is an average for each month in 2023.

The results of the regression calculations show that there is a correlation between the acidity (pH) of rainwater and the SO_2 and NO_2 values. The correlation coefficient is (-) 0,7305, which means the correlation is in the strong category. The coefficient of determination resulting from the calculation is 0,5337, which also means that the influence of SO_2 and NO_2 on acidity is strong. The value of 0,5337 explains that the influence of SO_2 and NO_2 on pH is 53,37%, while the remaining 46,63% is influenced by other factors. The coefficients of variables X1 and X2 are both negative, which indicates that the correlation

Table 3. Monthly Wind Speed and Pollutant Levels in 2023

Month	Jakarta's Monthly Wind Speed (m/s)	SO_2 ($\mu mol/m^2$)	NO_2 ($\mu mol/m^2$)
January	0,8592	171,17	97,00
February	0,8512	254,83	84,63
March	0,6204	231,67	127,88
April	0,6772	136,00	117,88
May	0,5322	146,86	213,63
June	0,5264	142,75	197,88
July	0,5436	287,25	223,00
Agust	0,6052	68,00	220,38
September	0,6662	126,25	214,00
October	0,6736	81,88	171,38
November	0,6038	22,00	177,63
December	0,5772	32,33	168,38

Table 4. Monthly Average Rainwater Acidity (pH), SO_2, and NO_2 in Jakarta City

Month	pH	SO_2 (ppm)	NO_2 (ppm)
January	4,78	0,0064	0,0162
February	5,09	0,0064	0,0174
March	4,99	0,0048	0,0202
April	4,87	0,0052	0,0150
May	5,11	0,0056	0,0220
June	4,59	0,0090	0,0222
July	4,88	0,0056	0,0210
Agust	4,45	0,0070	0,0214
September	-	0,0052	0,0178
October	-	0,0050	0,0174
November	4,66	0,0054	0,0260
December	4,36	0,0072	0,0280

Source: (BMKG, 2024)

is negative [30]. The greater the SO_2 and NO_2 levels, the greater the acidity of rainwater.

The rainwater acidity level prediction model is obtained using the formula:

$$y = a + x1coefficient + x2coefficient \qquad (3)$$

where: y =*Rainwater Acidity (pH)*, $x1 = SO_2$, $x2 = NO_2$

In this research, the rainwater acidity level prediction model was obtained:

y=5,994005+(-98,9218*x1)+(-28,4983*x2)

To find out whether the rainwater acidity prediction model is valid or can be accounted for, the prediction model was tested with rainwater acidity data taken directly by the author in March 2024. After that, the Mean Absolute Percentage Error (MAPE) value was calculated to see the accuracy of the prediction number for rainwater acidity level (pH).

Rainwater samples are collected directly from coordinate points adjacent to BMKG's observation station. The rainwater sample is then measured using a pH meter so that the acidity is known. Direct measurements in the field were carried out every time it rained from 1 to 15 March. The measurement locations were carried out at eight observation station points belonging to BMKG. The acidity (pH) data obtained is then calculated as an average (Table 5).

The results of the MAPE calculation are obtained by subtracting the actual and predicted pH values, and then the difference is calculated in absolute terms. The results are divided by the actual pH value at each station. Then the total value is added up and divided by 8 or as many air quality monitoring stations. To get the MAPE, the final result is multiplied by 100 to show the percentage number. The smaller the percentage, the more the predicted value is considered feasible to use. In the calculation results of this study, the MAPE value was obtained at 13%, which means that the rainwater acidity prediction model is included in the good forecasting category (Table 6).

Although the forecasting results from this study showed positive results, several limitations need to be addressed for future improvements. For example, extending the pH measurement timeframe. Other meteorological factors, such as wind direction, could also be added to future research to obtain more representative results. The resulting model can be applied using SO_2 and NO_2 distribution images (Sentinel 5P), so that a monthly rainwater acidity prediction map can be produced. Rainwater acidity predictions are generated using the raster calculator feature in ArcMap. The predicted values of monthly rainwater acidity throughout 2023 and January to May 2024 can be seen in figure 5.

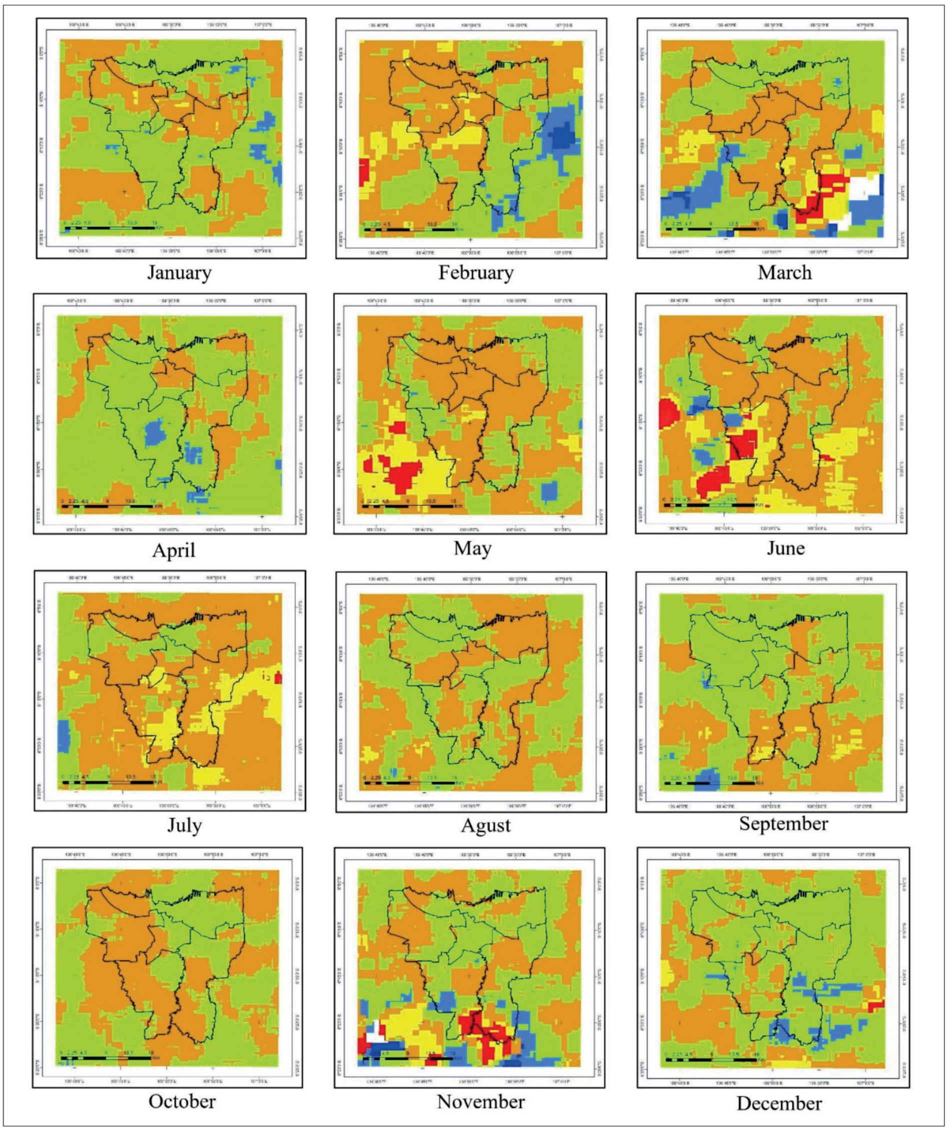
Table 5. Rainwater Acidity (pH) measurement Survey Results

BMKG Observation Station	MARCH 2024															Average
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Kemayoran			4,8	4,9					5,2	5,0	5,2		5,5	4,7	5,6	5,11
Ancol		4,0	4,9	4,9					5,4	5,5	4,9		4,8	4,4	5,0	4,87
Bandengan		7,0	6,6						6,3	7,0	5,9		6,3	6,0	6,8	6,49
Bivak			5,5	6,6	6,4				6,2	5,5	5,4		5,5	5,6	5,9	5,84
Grogol	5,4		5,5		6,0			5,4	5,6	5,5	5,7		5,2	5,0	5,4	5,47
Kementan	7,1	7,2	7,6		7,5			7,2	7,2	7,1	7,4	7,1	7,1	7,2	7,1	7,23
Monas			5,6	5,7					5,6	5,4	5,4		5,2	5,5	5,8	5,53
TMII	7,0	7,4		7,2	7,4	7,4		7,3	7,2	7,5	7,4	7,1		7,4	7,5	7,32

Table 6. MAPE Calculation of Rainwater Acidity Prediction Model

BMKG Observation Station	SO_2	NO_2	Actual pH (Survey Results)	Prediction pH	Error	Absolute Error	Absolut (Eror/ Actual pH)
Kemayoran	0,004	0,0178	5,11	5,09	0,02	0,21	0,004196
Ancol	0,005	0,0186	4,87	4,97	-0,10	0,10	0,021095
Bandengan	0,003	0,0176	6,56	5,20	1,36	1,36	0,207632
Bivak	0,002	0,0184	5,84	5,27	0,57	0,57	0,097982
Grogol	0,002	0,0202	5,47	5,22	0,25	0,24	0,045613
Kementan	0,002	0,0172	7,23	5,31	1,93	1,92	0,266453
Monas	0,007	0,0156	5,53	4,86	0,67	0,66	0,120909
TMII	0,002	0,0196	7,32	5,24	2,08	2,08	0,284156
			Total				1,048036
MAPE			= (1,048036/8)*100				13.1 %

(a)



January February March

April May June

July Agust September

October November December

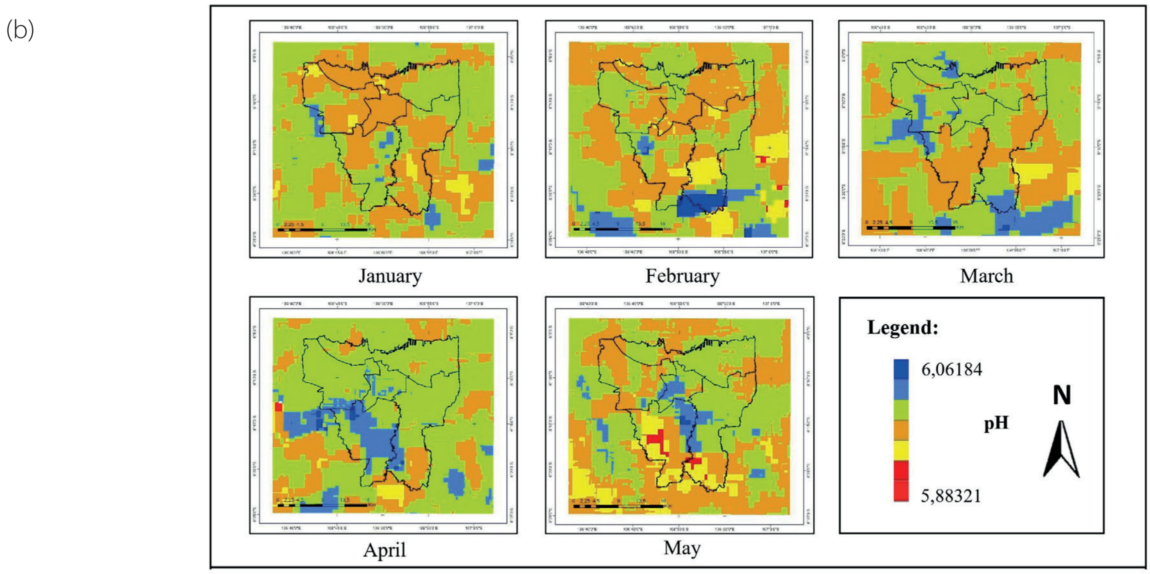


Fig. 5. Rainwater Acidity (pH) Prediction Based on the Distribution of SO_2 and NO_2 in Jakarta City (a) 2023 (b) 2024

CONCLUSION

The spatial-temporal distribution of SO_2 and NO_2 in 2023 produced by Sentinel 5P imagery has different patterns. The spatial distribution of SO_2 has a scattered pattern and is statistically influenced by rainfall with a moderate but insignificant positive correlation category. The spatial distribution of SO_2 is not influenced by wind speed, this is evident from statistical calculations which show that the distribution of SO_2 has a weak and insignificant positive correlation with wind speed. Meanwhile, the NO_2 distribution pattern is more clustered and has a gradation of values that are sequential from the highest to the lowest. The distribution of NO_2 is influenced by rainfall and also wind speed. The distribution of NO_2 has a strong and significant negative correlation with rainfall and a very strong negative correlation with wind speed. Temporally, the distribution of SO_2 in 2023 has the highest value in June, and the distribution of NO_2 has the highest value in August.

The prediction model for the acidity level of rainwater based on the distribution of SO_2 and NO_2 in 2023 in Jakarta was obtained from monthly data throughout 2023. The results of multiple linear regression show that there is a correlation between rainwater acidity and the distribution of SO_2 and NO_2. The correlation coefficient is (-) 0.7305, which means that the correlation is in the strong category. The correlation is negative, which explains that the higher the levels of SO_2 and NO_2, the more acidic the rainwater will be. To test whether the resulting model is suitable for use or not, a MAPE calculation is carried out. MAPE is obtained by comparing the actual acidity taken directly from the direct survey in March 2024 with the predicted acidity produced by the model. The results of the calculations from this study obtained a MAPE value of 13%, which means that the prediction model is included in the good forecasting category and can be used to predict rainwater acidity in Jakarta. ■

REFERENCES

T. Hong, X. Huang, X. Zhang, and X. Deng, “Correlation modelling between land surface temperatures and urban carbon emissions using multi-source remote sensing data: A case study,” Physics and Chemistry of the Earth, vol. 132, Dec. 2023, doi: 10.1016/j.pce.2023.103489.

S. S. Al-Alola, I. I. Alkadi, H. M. Alogayell, S. A. Mohamed, and I. Y. Ismail, “Air quality estimation using remote sensing and GIS-spatial technologies along Al-Shamal train pathway, Al-Qurayyat City in Saudi Arabia,” Environmental and Sustainability Indicators, vol. 15, Sep. 2022, doi: 10.1016/j.indic.2022.100184.

M. Roser, “Data Review: how many people die from air pollution?,” Our World in Data, 2021.

J. Rentschler and N. Leonova, “Global air pollution exposure and poverty,” Nat Commun, vol. 14, no. 1, Dec. 2023, doi: 10.1038/s41467-023-39797-4.

T. Handhayani, “An integrated analysis of air pollution and meteorological conditions in Jakarta,” Sci Rep, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-32817-9.

F. Mohammadi, H. Teiri, Y. Hajizadeh, A. Abdolahnejad, and A. Ebrahimi, “Prediction of atmospheric PM2.5 level by machine learning techniques in Isfahan, Iran,” Sci Rep, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-52617-z.

M. Astriyani, “Analisis Klasifikasi Data Kualitas Udara DKI Jakarta Menggunakan Algoritma C.45,” Jurnal Sistem & Teknologi Informasi Komunikasi, 2022.

Y. Huanhuan, C. Liangfu, S. Lin, T. Jinhua, and Y. Chao, “SO2 columns over China: Temporal and spatial variations using OMI and GOME-2 observations,” IOP Conf Ser Earth Environ Sci, vol. 17, no. 1, 2014, doi: 10.1088/1755-1315/17/1/012027.

G. A. K. S. Maharini, “STUDI REDUKSI SULFUR DIOKSIDA (SO2) UDARA AMBIEN OLEH RUANG TERBUKA (RTH) UNTUK WILAYAH PERMUKIMAN DAN TRANSPORTASI DI KOTA SURABAYA,” Institut Teknologi Sepuluh Nopember, Surabaya, 2017.

J. Lee et al., “Mortality from Lung Cancer in Workers Exposed to Sulfur Dioxide in the Pulp and Paper Industry,” Environmental Health Perspectives • VOLUME, vol. 110, no. 10, 2002, [Online]. Available: http://ehpnet1.niehs.nih.gov/docs/2002/110p991-995lee/abstract.html

B. A. Dewapandhu and A. Pribadi, “Analisis Penyebaran Gas Nitrogen Dioksida (NO2) di Jalan Raya Dramaga – Ciampea Kabupaten Bogor dengan Menggunakan Model Caline-4,” Jurnal Teknik Sipil dan Lingkungan, vol. 8, no. 1, pp. 67–76, Apr. 2023, doi: 10.29244/jsil.8.1.67-76.

N. R. Maherdyta, A. Syafitri, F. Septywantoro, P. A. Kejora, S. D. Gulo, and D. Sulistiyorini, “Analisis Risiko Kesehatan Lingkungan Paparan Gas Nitrogen Dioksida (NO2) dan Sulfur Diokida (SO2) pada Masyarakat di Wilayah Yogyakarta,” Jurnal Sanitasi Lingkungan, vol. 2, no. 1, pp. 51–59, May 2022, doi: 10.36086/jsl.v2i1.1040.

A. R. Pujaardana, "Studi Pemanfaatan Nitrogen Dioksida (NO2) Dari Satelit GOME 2 METOP-A Untuk Pembuatan Model NO2 Ambien dan Penggunaan Lahan," Institut Teknologi Sepuluh Nopember, Surabaya, 2016.

Z. Wang, F. Zheng, W. Zhang, and S. Wang, "Analysis of SO2 Pollution Changes of Beijing-Tianjin-Hebei Region over China Based on OMI Observations from 2006 to 2017," Advances in Meteorology, vol. 2018, 2018, doi: 10.1155/2018/8746068.

A. Kurniawan, "Estimasi Tingkat Keasaman Air Hujan (pH) Akibat Absorpsi Gas NO2 dan SO2," Jurnal Ilmu Lingkungan, vol. 17, no. 3, p. 399, Dec. 2019, doi: 10.14710/jil.17.3.399-407.

L. S. Supriatin, W. E. Cahyono, and S. Syafrizon, "The Effect of Rainwater Quality on The Methane Concentration," JKPK (Jurnal Kimia dan Pendidikan Kimia), vol. 2, no. 2, p. 103, Sep. 2017, doi: 10.20961/jkpk.v2i2.11972.

R. Katulistiyani and A. Ihwan, "ANALISIS TERJADINYA HUJAN ASAM DI KOTA PONTIANAK AKIBAT EMISI GAS DARI INDUSTRI DAN KENDARAAN," PRISMA FISIKA, vol. III, no. 01, pp. 15–20, 2015.

]P. Brimblecombe, The Effects of Air Pollution on the Built Environment, vol. 2. London: Imperial College Press, 2002.

WMO, "World Meteorology Organization." Accessed: Mar. 02, 2024. [Online]. Available: wmo.int

A. P. Desvina, "PERAMALAN PENCEMARAN UDARA OLEH SULFUR DIOKSIDA (SO2) DI PEKANBARU DENGAN MODEL AR(3)," Jurnal Sains Teknologi dan Industri, vol. 9, no. 2, 2012.

Kepmendagri Nomor 050-145, Keputusan Menteri Dalam Negeri Nomor 050-145 Tahun 2022 Tentang Pemberian dan Pemutahiran Kode, data Wilayah Administrasi Pemerintahan dan Pulau Tahun 2021. 2022. [Online]. Available: www.wikidata.org

BPS, "Badan Pusat Statistik." Accessed: Mar. 02, 2024. [Online]. Available: jakarta.bps.go.id/

sentinel.copernicus.eu, "sentinel.copernicus.eu." Accessed: Mar. 22, 2024. [Online]. Available: sentinels.copernicus.eu

H. Tamiminia, B. Salehi, M. Mahdianpari, L. Quackenbush, S. Adeli, and B. Brisco, "Google Earth Engine for geo-big data applications: A meta-analysis and systematic review," Jun. 01, 2020, Elsevier B.V. doi: 10.1016/j.isprsjprs.2020.04.001.

J. Cohen, "Set Correlation and Contingency Tables," 1988, [Online]. Available: http://www.copyright.com/

I. Nabillah and I. Ranggadara, "Mean Absolute Percentage Error untuk Evaluasi Hasil Prediksi Komoditas Laut," JOINS (Journal of Information System), vol. 5, no. 2, pp. 250–255, Nov. 2020, doi: 10.33633/joins.v5i2.3900.

Y. Liu, Y. Zhou, and J. Lu, "Exploring the relationship between air pollution and meteorological conditions in China under environmental governance," Sci Rep, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-71338-7.

B. Wieczorek, "Air Pollution Patterns Mapping of SO2, NO2, and CO Derived from TROPOMI over Central-East Europe," Remote Sens (Basel), vol. 15, no. 6, Mar. 2023, doi: 10.3390/rs15061565.

bmkg.go.id, "bmkg.go.id." Accessed: Mar. 29, 2024. [Online]. Available: bmkg.go.id

G. K. Uyanık and N. Güler, "A Study on Multiple Linear Regression Analysis," Procedia Soc Behav Sci, vol. 106, pp. 234–240, Dec. 2013, doi: 10.1016/j.sbspro.2013.12.027.

Check for updates

CC BY

RESEARCH PAPER

EXPLORING THE ROLE OF LAND USE SETTINGS ON LANDSLIDE ELLIPTICITY IN VOLCANIC FOOT SLOPE OF THE SOUTHERN SUMBING MOUNTAIN, CENTRAL-JAVA INDONESIA

Rina Purwaningsih[1,4], Junun Sartohadi[2,4*], Guruh Samodra[3], Christopher Gomez[5]
[1]Department of Environmental Science, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[2]Department of Soil Science, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[3]Department of Environmental Geography, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[4]Research Center of Land Resource Management, Universitas Gadjah Mada, Yogyakarta, 55281, Indonesia
[5]Faculty of Maritime Science, 5-1-1 Fukaeminami-machi Higashinada-ku, 658-0022 Kobe, Japan
***Corresponding author:** junun@ugm.ac.id

Received: March 20th 2025 / Accepted: January 27th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3962

ABSTRACT. This study investigates the effects of landslide processes on land morphology and their management strategies. Land use types determine the size and morphology of the landslide, which is critical for assessment and planning. LiDAR imaging helps to locate the various features of a landslide and human intervention. The research measures the ellipticity of landslides for a 3.2 km^2 area using LiDAR-derived data and ellipse numerical modeling. Landslide boundaries were delineated through automated and manual techniques, considering topographical index, visual interpretation, and expert knowledge. Convex hulls were used to approximate landslide shapes, which were then used to calculate the various ellipticity indices (e_L). Bare land, mixed gardens, rain-fed rice fields, and intermittent rivers have a higher e_L. Local roads, houses, and irrigated rice fields indicated receding e_L values due to the increased level of land management. Changes in eL are affected by the intensity of land use and disturbances related to the wetting of landslide slip areas. Higher ellipticity indicates dynamics in the landslide activity. Therefore, land use planning should be performed considering both natural and non-natural elements in the area.

KEYWORDS: landslide boundary shape, ellipticity index, land use settings

CITATION: Purwaningsih R., Sartohadi J., Samodra G., Gomez C. (2026). Exploring The Role Of Land Use Settings On Landslide Ellipticity In Volcanic Foot Slope Of The Southern Sumbing Mountain, Central-Java Indonesia. Geography, Environment, Sustainability, 1 (19), 72-85
https://doi.org/10.24057/2071-9388-2026-3962

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Landslides are ubiquitous geomorphic processes shaping the Earth's surface, continuously modifying landscapes through erosion and deposition (Jin et al. 2021; Pudasaini and Krautblatter 2021). These mass movements follow fundamental geomorphological principles where processes observed in the past continue into the present and future, albeit with decreasing intensity over time (Hooke, 2020). The principle of dynamic equilibrium in geomorphology suggests that landscape features evolve through a balance between erosive and depositional forces (Thorn and Welford 1994). In landslide-prone areas, this balance is particularly evident in how slope morphology changes over time. Rainwater plays a crucial role in this process, creating distinct patterns: diffusive spread in residual zones leading to convex slopes, and concentrated erosive flow in upper and middle slopes resulting in concave formations (Chen et al. 2014).

Recent studies have revealed that landslides typically create elliptical patterns in two-dimensional representations (Liu et al. 2023; Taylor et al. 2018). This geometric characteristic has become an important marker for understanding landslide behavior and evolution (Yin et al. 2014). Deviations from this elliptical pattern often indicate external influences, such as flow paths, channel morphology, or the merger of multiple smaller landslides (Martel 2004). The symmetry of these patterns can be classified into two main categories: symmetrical forms occurring on flat surfaces and asymmetrical forms on curved surfaces, each providing insight into the underlying landslide mechanics (Ogita et al. 2017).

However, human activities increasingly influence these natural patterns. Consequently, anthropogenic modifications of land cover significantly and increasingly affect landslide dynamics and susceptibility (Tarolli and Sofia 2016). Different land use types - from the broad spectrum of forest vegetation to urban development - create varying conditions for slope stability. Natural vegetation can help prevent soil erosion through root systems and reduced surface runoff, while urban expansion and agricultural development can increase landslide risks by altering drainage patterns and soil structure

(Chen et al. 2019). The relationship between land use changes and landslide morphology remains particularly important in regions experiencing rapid development. Engineering practices, construction activities, and changes in vegetation cover can substantially alter both the likelihood and characteristics of landslides. These modifications often occur without adequate risk assessment or mitigation measures, leading to increased vulnerability in previously stable areas.

Recent empirical observations suggest that while new landslides typically create elliptical scars, not all landslide boundaries maintain this shape over time (Fang et al. 2024; Li et al. 2022). The transformation of these boundaries may be influenced by both natural processes and land management practices. Understanding these changes is crucial for effective land use planning and risk mitigation in landslide-prone areas. This study aims to quantify these relationships by analyzing the ellipticity index of past and current landslide sequences in a volcanic-structural transition landscape. Specifically, we investigate how land use management influences landslide morphology, seeking to distinguish between changes caused by natural landslide processes and those resulting from anthropogenic modifications. This understanding is essential for developing more effective land management strategies in landslide-susceptible regions.

MATERIALS AND METHODS

Study Area

This research was conducted at the base of Mount Sumbing's volcanic slopes in Central Java, Indonesia (Fig. 1a). The study site exhibited distinctive pedogeological characteristics, as shown in Fig. 1b, with tertiary volcanic subsurface materials influencing quaternary volcanic surface materials. Bedrock is dominated by andesitic volcanic breccias and reworked volcanic sediments that have undergone intense hydrothermal alteration, producing clay-rich mineral assemblages dominated by kaolinite, halloysite, illite, and abundant quartz (Noviyanto et al. 2020; Purwaningsih et al. 2025). These processes have resulted in steep slopes mantled by a thick debris and soil cover, generally exceeding 80–120 cm, formed from volcanic ash deposition and prolonged weathering (Pratiwi et al. 2019). These unique pedogeological features have produced steep slopes mantled by substantial soil depth composed of clay-rich, deeply weathered materials.

The thick soil in the research area, formed through volcanic ash deposition, was also influenced by Tertiary hydrothermal alteration processes (Noviyanto et al. 2020; Pulungan and Sartohadi 2018; Samodra et al. 2020). The hydrothermal alteration of Tertiary breccia rocks produced weathered material with remarkably high clay content (Pirajno 2009). When exposed to water, this clay-rich weathered material became highly saturated, becoming the primary factor contributing to landslides (Wida et al. 2019).

Landslides occurred in areas with thick, uniform soil material, creating distinctive surface features at the research site. The concave and convex surface features observed in the hillshade image of the study site (Fig. 1c) suggest evidence of past landslide activity. This soil type was characterized by its consistent texture and composition and was prone to waterlogging (Wida et al. 2019). Water saturation reduced frictional resistance along the surface, increasing the likelihood of collapse. Water infiltration increased pore water pressure, reducing soil shear strength and facilitating slippage. The combination of uniform soil structure and high-water content created conditions where soil masses underwent rotational movement around a well-defined axis, resulting in rotational landslides. In two dimensions, these rotational landslides appeared as elliptical shapes.

The site's land use history and evolution were carefully examined, as human activities can significantly affect slope stability, influencing landslide susceptibility and associated hazards (Fell et al. 2008). This anthropogenic transformation is common across various global regions, creating localized impacts known as "local geomorphic change" (Bruschi et al. 2013). Incorporating geomorphological information into land use planning is essential for risk zoning at both intermediate and advanced levels. Advanced hazard and risk zoning requires a comprehensive approach, including detailed mapping of geomorphological features at larger scales for site-specific assessments. For such detailed analyses, the recommended mapping scale exceeds 1:5000 (Fell et al. 2008).

This site was selected as a particular research location because it represents a unique geomorphological setting where multiple geological and climatic factors converge to influence slope instability. The combination of undulating-steep topography deeply weathered volcanic soils, and a history of hydrothermal alteration provides ideal conditions to study the triggering mechanisms and morphological expression of rotational landslides. Moreover, the area is representative of many densely populated volcanic landscapes in Indonesia, making it a strategic location for developing site-specific hazard assessment frameworks with broader applicability to similar environments.

Acquisition of Landslide Boundary Data

A Da-Jiang Innovations (DJI) Matrice 300 Aircraft, mounted with a DJI Zenmuse L1 laser sensor and a camera, recorded the landslide and its boundaries with high accuracy over the 3.2 km^2 area. DJI Zenmuse L1 has a vertical accuracy of ±3 cm; however, the actual measurement accuracy of the LiDAR data in the study area was 0.35 meters, as determined through Global Navigation Satellite System (GNSS) measurements at six control points (Fig. 2). Besides that, the aerial photos have a resolution of 5 cm. Ground control points (GCP) and benchmarks were determined via GNSS to enable precise data acquisition. Furthermore, point cloud filtering and intensity modeling were undertaken during post-processing, resulting in about 26 points/m^2 of ground point density. A Digital Terrain Model (DTM) with 0.5-meter resolution was generated for the study area, and landslide identification was used to assess the chronology of the first occurrence.

Land use and land cover (LULC) data were derived from orthophotos acquired through aerial photography in April 2023, conducted concurrently with DTM–LiDAR data collection. The simultaneous acquisition enabled manual delineation of nine LULC categories using on-screen digitizing based on standard visual interpretation elements, including color, texture, shape, size, pattern, shadow, and contextual association. The interpreted LULC map was subsequently verified through field validation.

DTM products derived from LiDAR are visualized using hillshade and Topographic Position Index (TPI) features (Fig. 3). Hillshade provides a visual representation of terrain morphology, emphasizing differences between convex and concave surfaces such as ridgelines, valleys, steep slopes, and terraces (Trantham 2023). The light–dark patterns in hillshade imagery often reveal evidence of mass movements (landslides), including scarps, translational zones, and depositional areas. TPI enhances local morphological interpretation, where positive values indicate ridges and negative values represent depressions

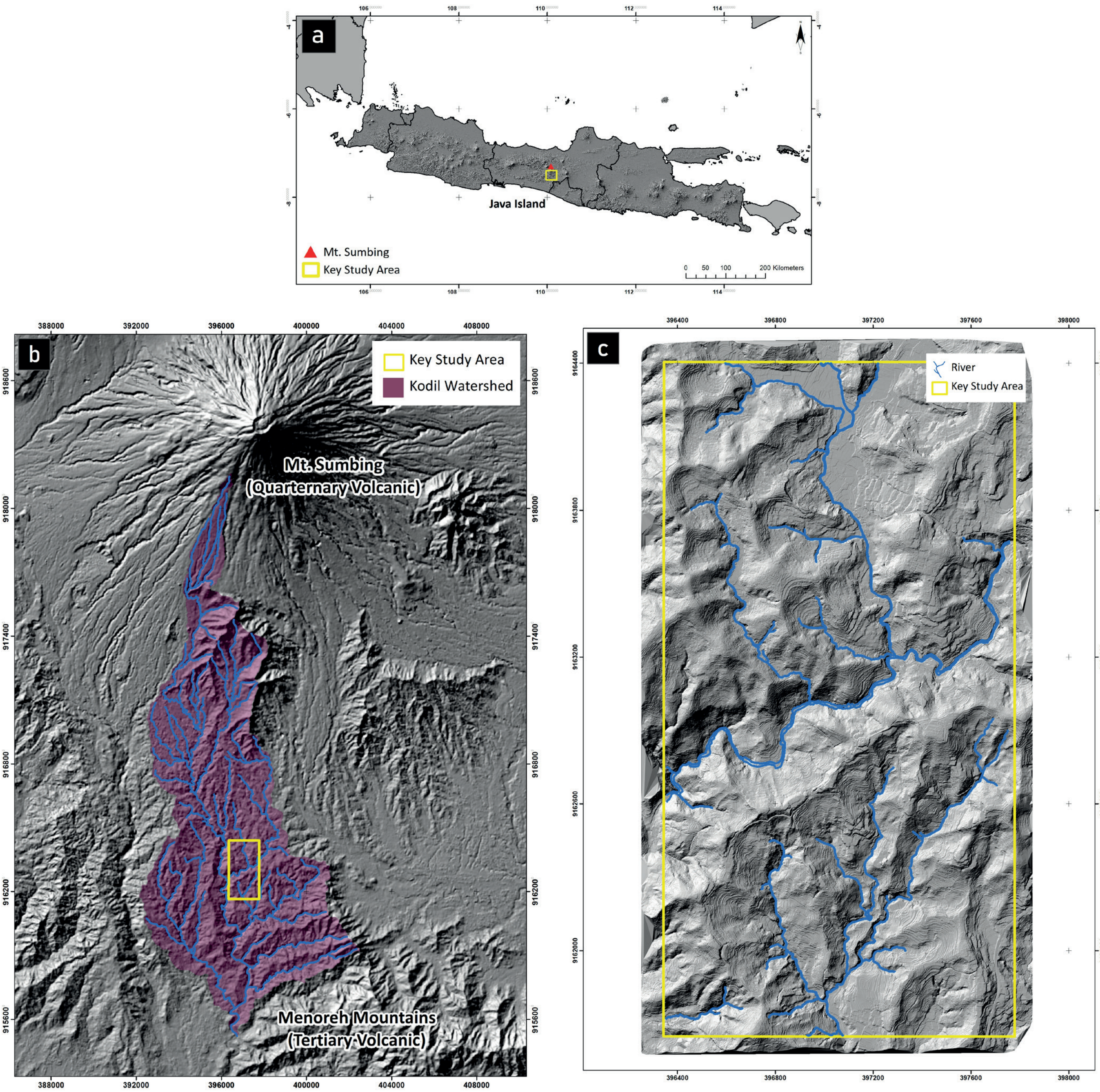


Fig. 1. Location of the Study (a) Java Island, (b) study location located between Mt. Sumbing and Menoreh Mountains, (c) Key research locations

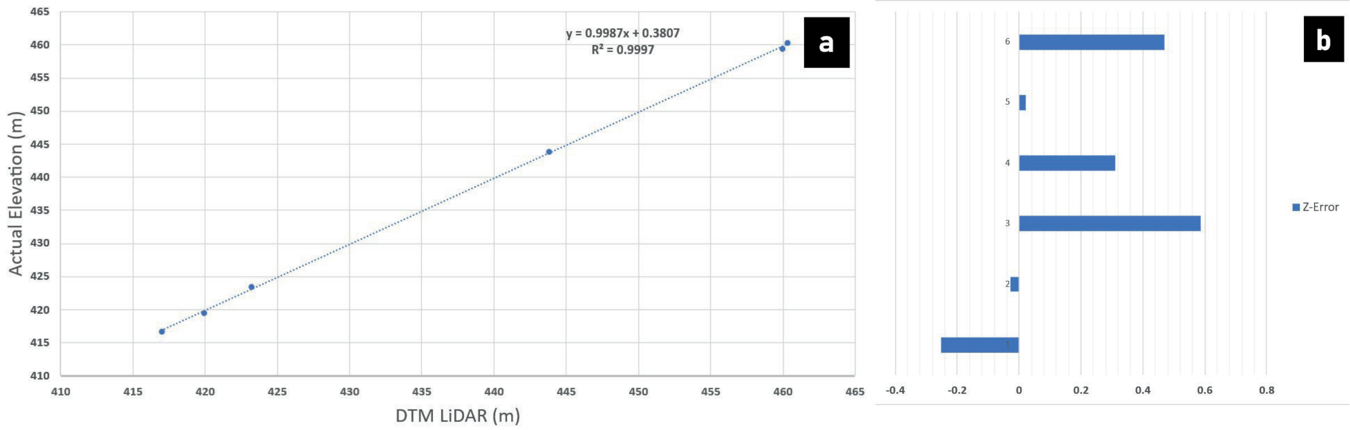


Fig. 2. (a) Linear fitting function between GNSS checkpoint elevations and DTM elevations obtained with 6 GCPs; (b) the error value represents the difference between actual and LiDAR-derived elevations

(Otgonbayar et al. 2023). In object-based analysis or terrain segmentation, hillshade and TPI support the delineation of landslide boundaries through visual interpretation.

Landslide 2D Geometrical Features and Ellipticity Identification

The elliptical descriptor developed by Schmittbuhl et al. (2003) and Tort (2003) allowed the measurement of shape, regardless of its complexity, and facilitated the conversion of morphological variations into basic geometric concepts. This procedure was challenging to perform using traditional Fourier coefficients. The Elliptical Descriptor methodology was also employed to effectively analyze the morphology of lakes, providing a robust means of describing variations and patterns between their outlines (Dhingra et al. 2019). Their synthetic morphology example confirmed that methods normalized for size and rotation were effective in grouping comparable outlines.

In this research, each delineated landslide polygon underwent ellipse numerical simulations using the methodology proposed by Taylor et al. (2018) to determine the dimensional characteristics of the landslide shapes. Computations were performed on each polygon corresponding to a landslide to determine its area *(A_L)*, perimeter *(P_L)*, central coordinates (x, y), and primary angle (Fig. 4).

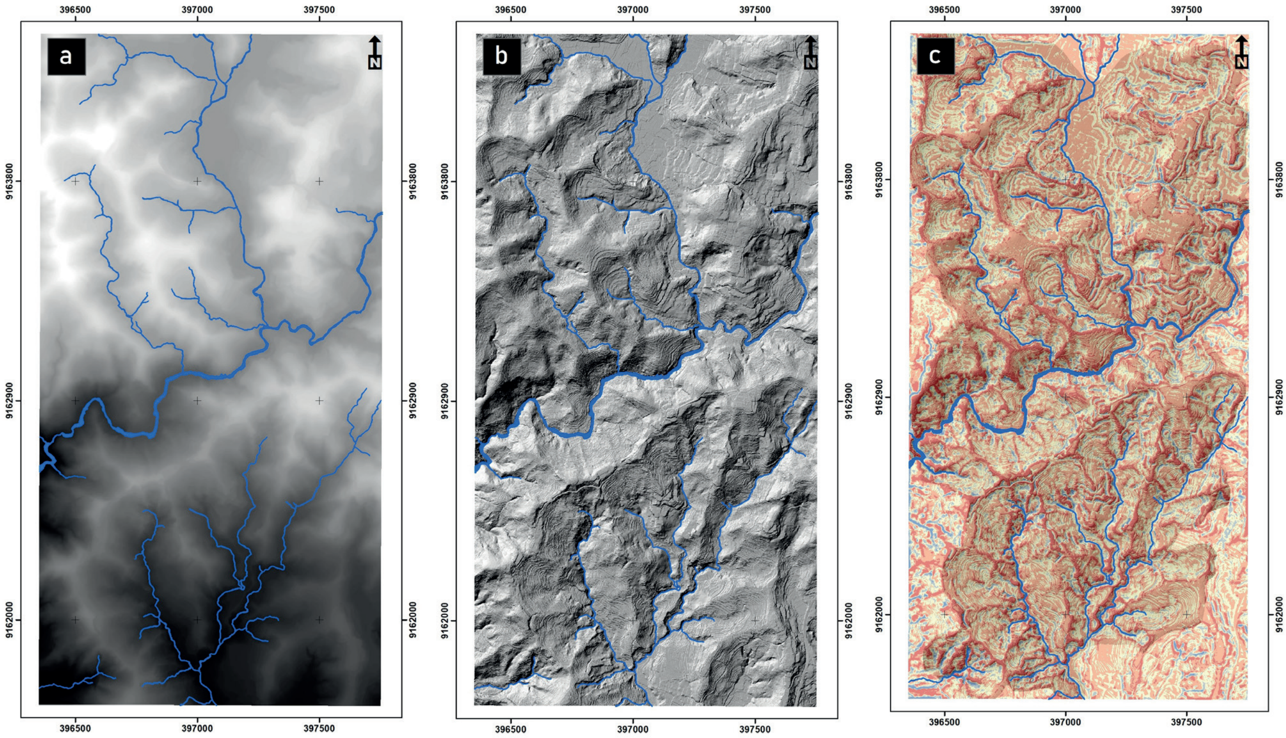


Fig. 3. DTM LiDAR (a) derivatives in the form of Hillshade (b) and Topographic Position Index (TPI) with a 0.5 × 0.5 size of the neighborhood (c)

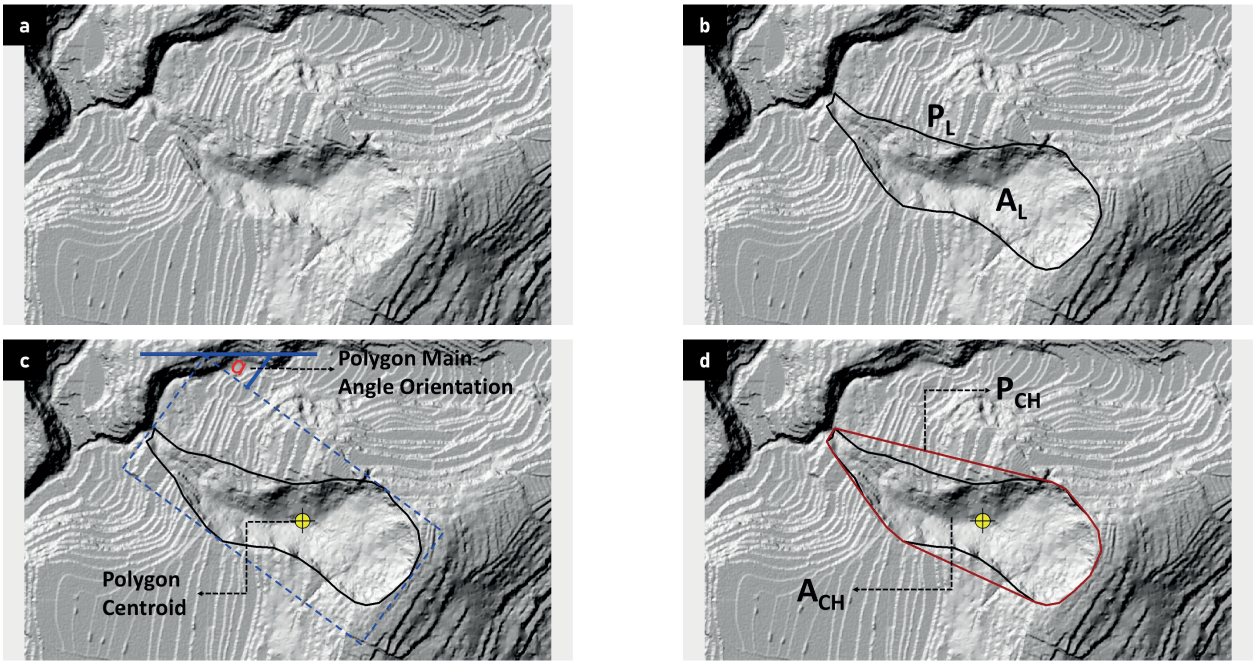


Fig. 4. Derived LiDAR data is used to demarcate the landslide boundary and determine the geometric elements required in the polygon ellipse calculation: (a) landslide feature, (b) area and perimeter of landslide boundary, (c) polygon centroid and its main angle orientation, and (d) area of convex hull and perimeter of convex hull

The landslide boundaries identified in this study were delineated based on surface morphological characteristics. However, many of these areas have since experienced changes in land use and land cover, resulting in the loss of visible landslide features in the field. The landslide delineation resulted in a polygon that was then subjected to a 2D convex hull operation. This operation found the minimum bounding geometry, which was the smallest convex polygon that fit around the input feature. For each resulting polygon obtained from the convex hull analysis, the area (A_{CH}), perimeter (P_{CH}), central coordinates (x, y), and primary angle (α) were also computed.

Mathematical computations were employed to ascertain the length-to-width ratio *(AE)*. The length-to-width ratio and ellipse dimensions were subsequently applied to the convex hull of the polygon (Eq. 1).

$$\Lambda_E = \frac{L_E}{W_E} = \frac{\sqrt[2]{\dfrac{A \times \Lambda_E}{\pi}}}{\sqrt[2]{\dfrac{A}{\pi \times \Lambda_E}}} = \frac{4\pi A_E}{P_E^2 - \sqrt{P_E^4 - 16\pi^2 A_E^2}} \quad (1)$$

Λ_E - length-to-width ratio of ellipse
L_E - length of ellipse (m)
W_E - width of ellipse (m)
P_E - perimeter of ellipse (m)

The ellipse was graphically represented using a polygon center point (x, y) endowed with length (*L* scaled *CH*) and width (*W* scaled *CH*) attributes (Eq. 2). The graphical depiction of the polygon ellipse was aligned with the orientation of the landslide. The ellipticity index (e_L) in Eq. 3 was derived from the disparity between the landslide polygon and the ellipse polygon (Lombardo, 2014). An inverse relationship was observed between the size of the area outside the polygon (A_L-$A_{L\cap E}$) and the ellipticity index; specifically, a larger area corresponded to a smaller ellipticity index and vice versa. The ultimate ellipse (Fig. 5) was determined by its area (A_E), perimeter (P_E), length (L_E), width (W_E), semi-major axis (a), and semi-minor axis (b).

$$\begin{aligned} L_M\left[\left(A_{CH}, P_{CH}\right) \to A_L\right] &= L_M\left[A_{CH}, P_{CH}\right]\sqrt{A_L/A_{CH}} \\ W_M\left[\left(A_{CH}, P_{CH}\right) \to A_L\right] &= W_M\left[A_{CH}, P_{CH}\right]\sqrt{A_L/A_{CH}} \end{aligned} \quad (2)$$

L_M - length of landslide for elliptic approximation based on convex hull (m)
W_M - width of landslide for elliptic approximation based on convex hull (m)
A_{CH} - area of landslide convex hull (m^2)
P_{CH} - perimeter of landslide convex hull (m)
A_L - area of landslide (m^2)

$$e_L = 1 - 2\frac{A_L - A_{L\cap E}}{A_L} \quad (3)$$

e_L - ellipticity indexx
A_L - area of landslide (m^2)
$A_{L\cap E}$ - area of intersection (∩) between original landslide shape and elliptical approximation (m^2)

In the previous method, the arithmetic rotation was used to automatically run polygonal main angle calculations, determining the orientation of a landslide. However, this approach produced a principal angle that did not align with the actual landslide angle. The principal angle in the context of landslides pertains to the spatial arrangement of length and width within a landslide, indicating the direction of the landslide's length from its summit to its base. Slope angles were utilized to assess locations that might become unstable (Donnarumma et al. 2013). It was imperative to rectify the principal angle of the landslide to ascertain the original ellipticity value. This correction was implemented through the application of rotation tools. The numerical representation of the landslide's principal angle indicated the rotational angle value measured from a reference point in the northward direction. The landslide direction is quantified in degrees of rotational angle measured clockwise from the north.

$$\tan\alpha = \frac{H}{L_{CL}}, \quad \alpha = \tan^{-1}\frac{H}{L_{CL}}, \quad L = \sqrt{L_{CL}^2 + H^2} \quad (4)$$

Each landslide is characterized by a principal slope angle formed between the failure surface and the horizontal plane. The landslide polygons, delineated from LiDAR-derived DEM data, provide measurements of the landslide's horizontal length *(Lc)* and vertical height *(H)*, as described in Fig. 6. These parameters are used to calculate the slope angle of the landslide (α) and the landslide's oblique length *(L)*, as defined in Eq. 4.

RESULTS

2D Geometrical Features of Landslide Boundary

Landslide geometry is controlled by the friction angle of the slope's remaining material and its original slope angle. The geometry of a landslide is defined by the ratio of its thickness to its length. We used LiDAR data to observe distinctive features of landslide geometry, enabling the creation of precise landslide boundary areas. Each landslide boundary unit possessed geometric attributes that accurately described its topographic shape,

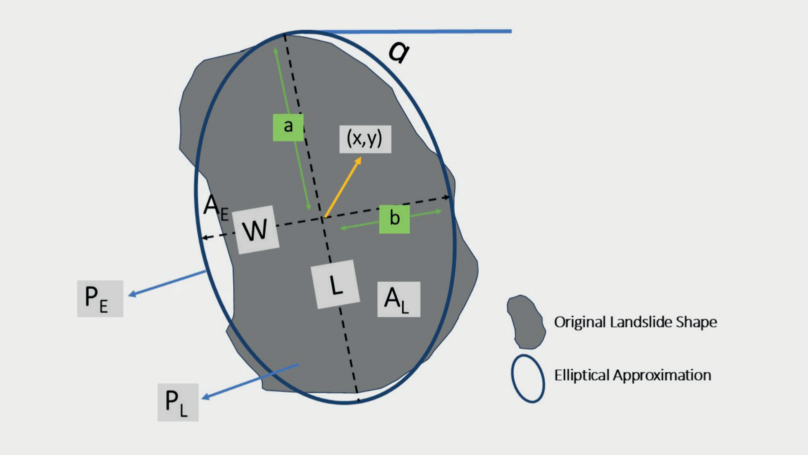


Fig. 5. Plotting Ellipse from LE, WE, centroid (x, y), and orientation (α) (Taylor et al., 2018)

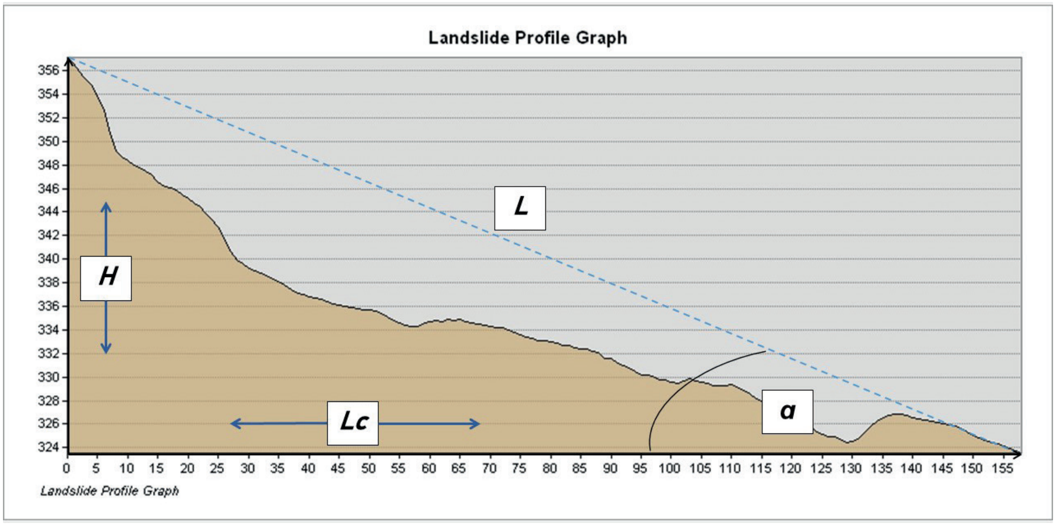


Fig. 6. How to obtain the main slope value from a 2D landslide illustration (example from landslide ID 11)

representing various landslide conditions observed in the field (Fig. 9a). The erosion-deposition process on landslide bodies creates surface shapes with both concave and convex components. These features are illustrated in 2D planes in Fig. 6 and quantified based on area, perimeter, height, width, length, and slope (Table 1).

Landslides affected 72% of the study area, with 259 landslides successfully delineated. The landslide areas ranged from 70 m^2 to 99,627 m^2. We observed a pattern in the configuration of landslide boundary polygons, where small boundary areas typically followed larger ones. Each landslide boundary unit's perimeter showed a strong positive linear relationship (0.82) with its area. The landslide elevation range *(H)*, also referred to as landslide thickness, varied between 1 and 115 m, showing a weak positive linear relationship (0.34) with area. The average landslide length at the centerline (L_{CL}) was shorter than the total length *(L)* because L included the landslide runout distance. Landslide thickness was proportional to length, with their ratio determining the overall landslide geometry. Landslide boundary unit widths ranged from 6.18 to 503.51 m. The landslide slope angle (α) varied with the landslide area, with steeper angles typically occurring in smaller landslides.

Landslide Ellipticity

The ellipticity index in the study area ranged from -0.74 to 0.96 (Fig. 7). By using the geometrical interval classification technique, the ellipticity index (e_L) was categorized into two classes: -0.74 to 0.63 (non-ellipse), and 0.63 to 0.96 (ellipse). The classification technique employed in this study aimed to represent continuous data visually, and the advantage of the geometric interval's classification is its ability to effectively handle data that do not follow a normal distribution. The study site exhibited a prevalence of high eL values, with 137 ellipse polygons and 122 non-ellipse polygons. We observed that the average landslide area decreased as the eL values increased, indicating a transition towards a more elliptical shape. This pattern indicates that smaller landslides tend to adopt more defined elliptical forms, while larger ones display more variability in

Table 1. Statistic parameters for 2D landslide geometry

Statistic Parameters	A_L	P_L	H	W	L_{CL}	L	α
Mean	8839.88	392.72	26.12	74.45	94.38	98.64	18.47
Standard Error	1010.21	26.31	1.08	4.39	5.56	5.61	0.47
Median	2680.32	227.61	20.00	47.66	62.00	65.07	17.65
Mode	#N/A	#N/A	14.00	35.02	42.00	55.71	16.70
Standard Deviation	16257.81	423.48	17.31	70.58	89.45	90.33	7.52
Sample Variance	264316351.59	179334.71	299.72	4981.75	8000.48	8158.99	56.61
Kurtosis	13.71	6.60	2.15	8.65	6.45	6.07	0.35
Skewness	3.50	2.42	1.22	2.46	2.35	2.28	0.50
Range	99556.90	2600.04	114.00	497.33	530.00	533.12	41.09
Minimum	69.80	36.83	1.00	6.18	10.00	11.40	2.64
Maximum	99626.70	2636.87	115.00	503.51	540.00	544.52	43.73
Sum	2289530.00	101713.62	6764.00	19283.43	24444.00	25547.93	4784.30
Count	259	259	259	259	259	259	259

A_L = landslide area (m^2), P_L = landslide perimeter (m), H = landslide elevation range (m), W = landslide width (m), L_{CL} = landslide length centerline (m), L = landslide total length (m), α = landslide slope angle (°)

shape. Such distinctions carry important implications for landslide risk assessment, as the morphology of a landslide may reflect its activity status. Landslides with well-preserved elliptical shapes, which have not experienced significant morphological modification, may indicate a higher potential for reactivation. In contrast, landslides whose shapes have evolved through multiple modification processes are likely associated with greater slope stability. Understanding this relationship between shape, size, and stability can support more accurate predictive modeling and guide targeted land use planning and mitigation efforts in susceptible regions.

Based on the results of aerial imagery captured in 2024, followed by ground-truthing through a field survey, 48 landslides are identified as falling within the active state. These recently occurring landslides are categorized as both ellipse and non-ellipse. Active landslides are particularly observed on depositional slopes, which are closely associated with loose soil material and its high-water retention capacity. Depositional slopes often consist of unconsolidated sediments, which are more prone to failure compared to lithified material. The loose nature of these sediments makes them susceptible to landslides, especially when combined with other triggering factors such as rivers. As a result, the dynamic flow of river water continuously influences newly formed landslides, causing modifications to the contours of the landslide boundary so that the shape can be an ellipse or non-ellipse.

Landslide direction pertains to the specific orientation or azimuth in which a slope undergoes a landslide event (Massey et al. 2013). In the investigated area, landslides predominantly manifest in the south, southeast and southwest directions (Fig. 8). How landslides move is closely connected to how the regional geological structures are shaped in the study area through local lineaments. The river flow pattern, which consists of intermittent rivers that typically flow from north to south before merging into perennial rivers that stream from northeast to southwest, illustrates the control of the geological structure (Patanduk 2020).

Landslide Ellipticity	Number of Landslide
-0.74 – 0.63 (non-ellipse)	122 (47%)
0.63 – 0.96 (ellipse)	137 (53%)

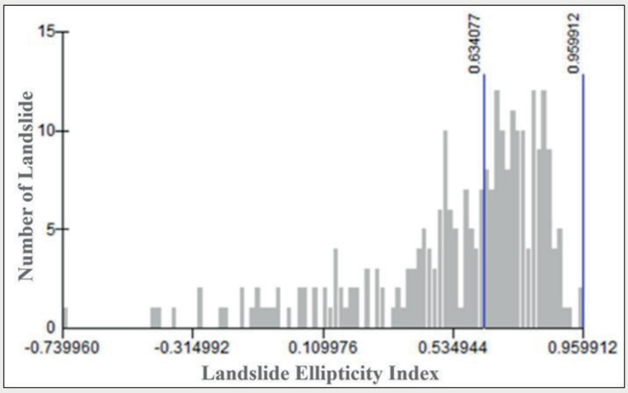


Fig. 7. Division of ellipticity index classes using the geometrical interval classification technique

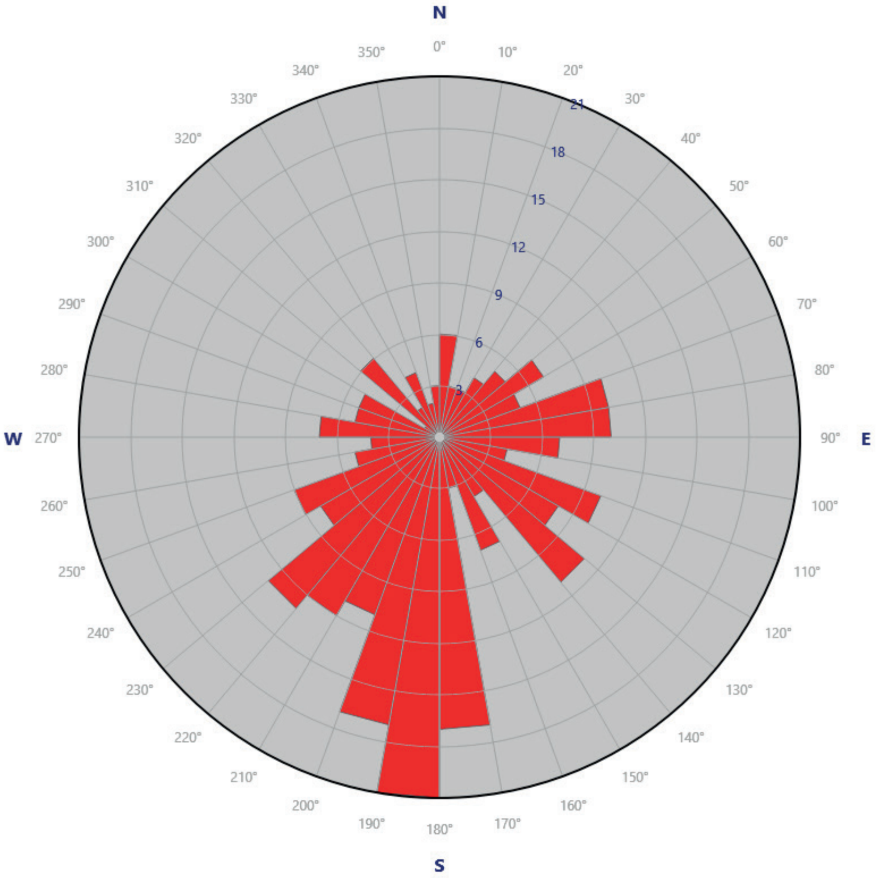


Fig. 8. The orientation of 259 landslide units within the study area is presented in a rose diagram, revealing a dominant directional trend toward the southeast, south, and southwest

Land Use and Land Cover

Land use types were interpreted from aerial photography and were classified into nine types (Fig. 9c). As presented in Table 2, the dominant land use was mixed gardens (73%). Following closely was the allocation of 18% to rice fields, houses 3%, dry field agriculture 3%, with the remaining 3% encompassing bare land, roads, and water bodies such as rivers. Rice fields exhibited a widespread distribution across valleys and inclined terrain, while dry field agriculture was dispersed proximately to these paddy fields. Houses tended to show a distribution pattern that was dominated by occupying high positions along the hillsides. The presence of bare land exhibited a random arrangement interspersed among mixed gardens and was notably associated with areas susceptible to new landslide occurrences. Intermittent rivers indicated a directional alignment from north to south, while the perennial river followed a northeast-southwest line. In the irrigated rice field land use unit, there are irrigation canals located on the slopes, interconnected with adjacent groups of irrigated rice fields. Small-sized ponds were also found that were intended for cultivating freshwater fish, most of which were located near irrigated rice fields.

DISCUSSION

Relation between Landslide Ellipticity and 2D Geometrical Features

Large dormant landslides are often accompanied by smaller, active landslides in the present day. There is a tendency for high ellipticity to occur in small landslides. This observation supports the theory that current geomorphological processes continue but with diminishing intensity over time (Guzzetti et al. 2012). The increasing steepness of slope angles following successive landslides demonstrates the dynamic equilibrium effect, where land surfaces tend toward flatness over time (Roering and Hales 2022). However, recent surface processes typically produce significantly steeper slope angles. The descriptive presentation in Fig. 10 shows that elliptical

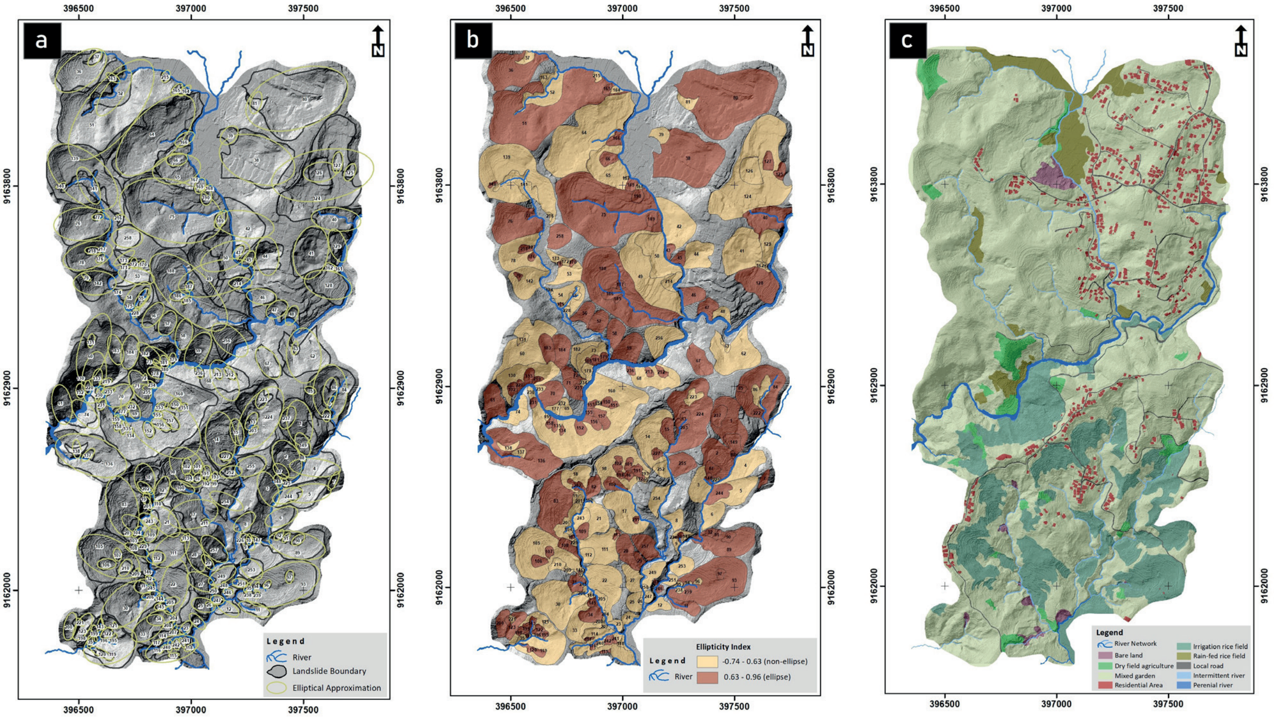


Fig. 9. (a) Mapped landslide boundary and its elliptical approximation, (b) Landslide ellipticity index, (c) Land use land cover in the key research area

Table 2. Land use land cover in the study area

No	Land Use Land Cover	Area within the study area (m^2)	%	Area within landslide boundary (m^2)	%
1	Bare land	25,750.73	0.81	24,706.74	1.08
2	Dry field agriculture	86,996.34	2.75	65,441.32	2.86
3	Houses	107,974.16	3.41	61,359.48	2.68
4	Intermittent river	15,158.46	0.48	4,972.18	0.22
5	Irrigated rice field	434,197.44	13.71	351,835.22	15.37
6	Local road	39,270.29	1.24	19,914.60	0.87
7	Mixed garden	2,317,471.15	73.18	1,729,790.30	75.55
8	Perennial river	22,314.18	0.70	2,862.75	0.13
9	Rain-fed rice field	117,504.15	3.71	28,647.43	1.25
Total		3,166,636.90	100	2,289,530.02	100

landslides are small and have relatively minor geometric parameters, particularly in the area value, perimeter, width, and elevation range. The Ellipse category is found with an average landslide size of 8175 m^2, an average perimeter of 338 m, an average width of 60 m, and an elevation range of 31 m. Meanwhile, the non-ellipse class occurred in landslides with a size of 9587 m^2, an average perimeter of 455 m, an average width of 91 m, and an elevation range of 30 m. Meanwhile, the *L* size in the ellipse category is 102 m, and in the non-ellipse category, it is 98 m. The average size of L_{CL} in the ellipse category is 117 m, and in the non-ellipse category, it is 108 m.

The geometry of the landslide, which is reflected in the ellipticity index at the research location, has varying slope angles. Landslides with steep slopes can occur in ellipse or non-ellipse conditions, and vice versa, landslides with small slopes can also occur in both. The average landslide angle at the research location, whether elliptical or non-elliptical, is 18.4 degrees. Slope size is one of the most important factors influencing slope instability as a cause of landslides. The magnitude of the landslide slope angle is inversely proportional to the landslide volume, indicating that there is a difference in the control of the slope angle on the size of the landslide for a certain material strength (Fan 2010; Li and Kong 2010; Wu et al. 2024). This condition explains how the landslide slope is affected by local situations such as slope height, soil material, landslide slip surface, and its spatial arrangement.

Relations between Landslide Ellipticity and Land Use Land Cover

Land use and land cover are some of the factors that determine the occurrence of landslides (Bartelletti et al. 2017; Cevasco et al. 2014), and the land use changes caused by human activities have a significant impact on the onset of landslide events (Meneses et al. 2019). The resulting landslide's surface shape varies depending on its LULC. This study's analysis considers nine classes of LULC at the research location. The assessment of the level of ellipticity of landslide scars in LULC units is also deepened by the trend of geometrical feature conditions. The condition of the tendency for landslides to be long-short, large-small, wide-narrow and the size of the slope angle can be seen in Fig. 12.

The boxplot in Fig. 11 illustrates the distribution of ellipticity index values for each LULC. We chose the boxplot visualization to precisely compare the distribution of each set of ellipticity values, particularly within the LULC group, and understand their range. Perennial rivers have the highest median value, 0.61, while house units have the lowest, 0.455. The spread of the distribution describes the difference between quartile 3 (upper whisker) and quartile 1 (lower whisker). The widest distribution of eL values is on local roads with a value of 0.6, followed by houses with a value of 0.585 and mixed gardens with a value of 0.545. The perennial river has the smallest distribution; the result can be seen from the thin bar diagram.

From minimum to maximum, the range (largest minus smallest) shows the width of all data. The range can be seen from the length of the lower and upper fences. The largest range occurs in mixed garden units with the lowest e_L data, namely -0.74, and the highest e_L data is 0.95. The smallest range is owned by rain-fed rice fields with the lowest e_L data, namely -0.11, and the highest e_L data, 0.96. The shape of the e_L distribution based on land use units illustrates whether the e_L value is normally distributed or not. A normal distribution can be seen if the middle 50% of the data is evenly divided, with 25% of the data between quartile 1 and quartile 2 equal to 25% of the data between quartile 2 and quartile 3. We found no normal distribution of 9-unit land use based on the eL value. The shape of the distribution of 9-unit land use is skewed left, which shows the dominance of the data

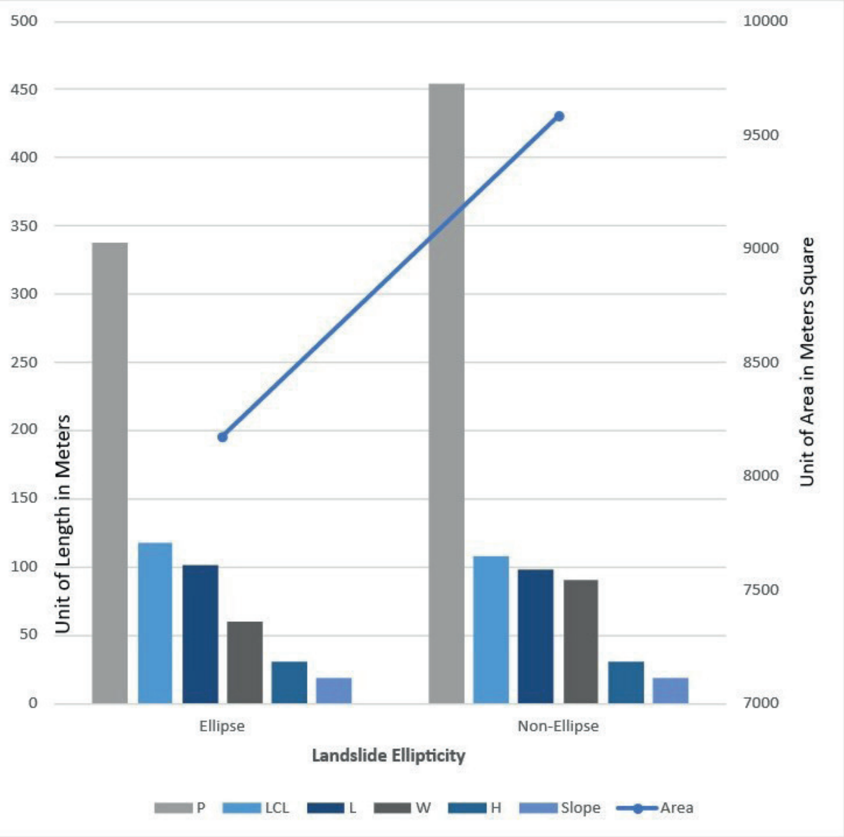


Fig. 10. The relationship between landslide ellipticity and the average value of 2D geometrical features (*P*: Landslide perimeter in meters, L_{CL}: Landslide center length in meters, *L*: Landslide total length in meters, *W*: Landslide width in meters, *H*: Landslide height in meters, Slope was in degrees, and Area was in meter square)

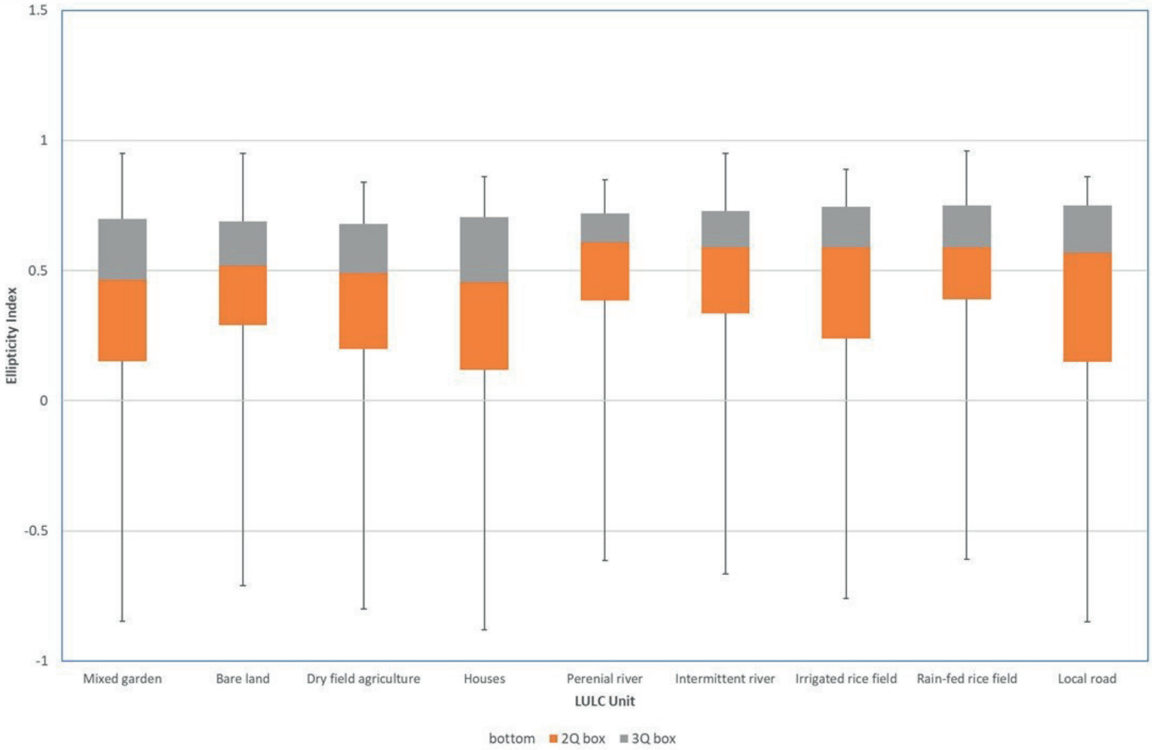


Fig. 11. Boxplot of the landslide ellipticity index distribution for each LULC unit

distribution below the median value (quartile 3). From Fig. 8, it can be concluded that LULC units with high ellipticity include bare land, mixed gardens, rain-fed rice fields, and intermittent rivers, while LULC with low ellipticity occurs in local roads, houses, irrigated rice fields, and mixed gardens.

As shown in Fig. 11, sequentially, the dominance of the highest e_L is bare land, mixed gardens, rain-fed rice fields, and intermittent rivers. High-ellipticity landslides are active landslides that have just occurred, so they still have elliptical boundaries. The recent landslide that occurred in less than 1.5 years means that the land cover above the landslide is currently bare land. Geometrically, landslides on bare land are small, short, and narrow, and they have steep slopes (Fig.12). These landslides are located along intermittent rivers on a spatial scale. Intermittent rivers flow only during certain times of the year when there is sufficient rainfall.

The mixed garden at the study location is dominated by woody plants, which vary in terms of tree height and tree types with low-to-high density levels. High e_L was found in landslides in low-density mixed-garden areas. Landslides found in mixed garden units are characterized by small, short, narrow geometry with slope sizes tending to be large (Fig. 12). This situation strengthens the evidence that many new and active landslides have been found in landscapes with mixed garden cover. The presence of various plant species and their alterations in the mixed garden area are expected to impact the soil's shear strength, leading to slope instability. In certain instances, these rapid changes can potentially initiate landslides (Davies 2015). Detailed observation of aerial photos of the active landslide polygons reveals a land cover pattern in the form of mixed gardens with sparse density and coconut trees at the top of the slope.

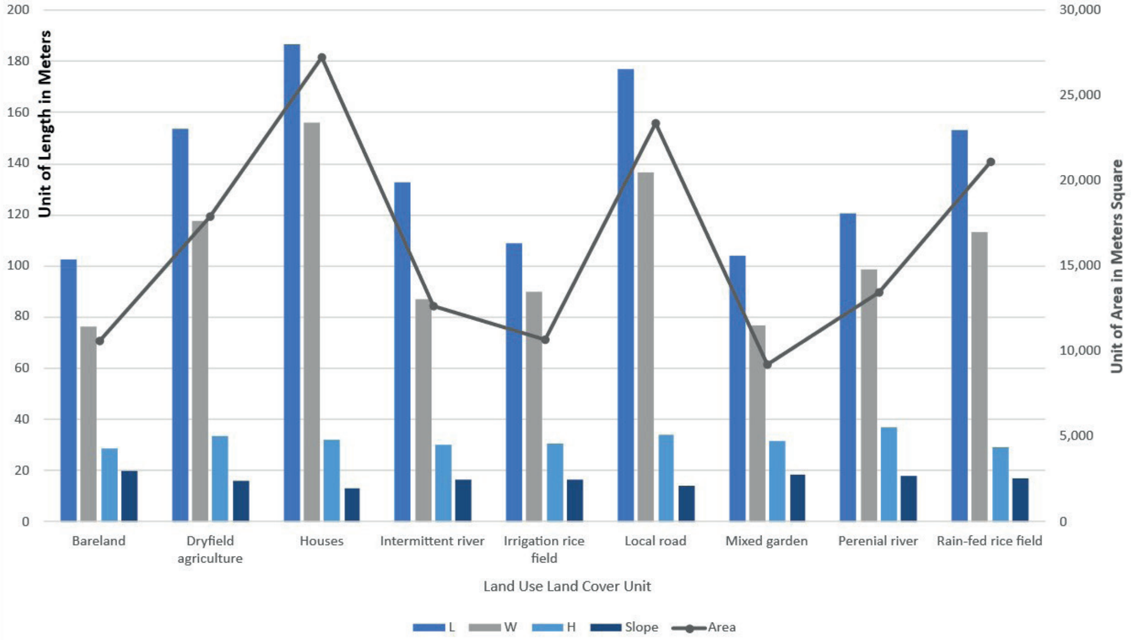


Fig. 12. Trends in 2D geometrical landslide features on LULC units, considering landslide total length in meters (*L*), landslide width in meters (*W*), landslide height in meters (*H*), landslide slope in degrees, and landslide area in square meters

The rain-fed rice fields scattered between the foot slopes of the quarter-tertiary volcanic slopes are only planted once a year on average. Rain-fed rice fields will only be cultivated intensively when the rainy season arrives. Intercrops in the growing season after rice are usually secondary crops, such as corn and horticulture. The landslides found on this land use have a larger geometric character, are a little longer, and have a narrow geometry with slope sizes tending to be large.

Specifically, the active landslides occur on depositional slopes intimately connected to water and river dynamics. The phenomenon of riverbank undercutting arises from the rapid flow of water at bends or specific segments of the river (Figs. 13a, c, e). Elevated water pressure induces erosion along slopes and riverbanks. The progressive undercutting at the base generates instability in the overlying slope, leading to subsequent collapse. As a result, the dynamic flow of river water continuously influences newly formed landslides, causing modifications to the contours of the landslide boundary, whether ellipse or non-ellipse. From the Fig. 11 graph, the landslides along the intermittent river are wide, so some of the final landslides along the river channel have a non-ellipse shape.

The distribution of low ellipticity values is owned by local roads, houses, irrigated rice fields, and mixed garden units. Landslide boundary units with irregular or non-ellipse shapes are greatly influenced by the intensity of land management (Kimura 2024; Segoni et al. 2024). The predominance of low ellipticity occurs in areas of former landslides crossed by local roads, caused by cut and fill slopes during the construction process (Figs. 13f, g, h). Slope modification has a high influence on changing the features of past landslides into non-elliptical ones (Sartohadi et al. 2024). If viewed on a site-specific basis, slope modification can increase slope instability (Jaboyedoff et al. 2016). The landslide geometry in the local road unit is larger, longer, and wider with a smaller slope (Fig. 11). Road construction and maintenance have direct or indirect impacts on slopes through slope cuts or changes in surface water runoff (Vuillez et al. 2018). The accumulation of water runoff that leads to landslide slip areas has the potential to trigger landslides (Ran et al. 2018). Additional effects on roads encompass the establishment of logging pathways which typically adhere to distinct standards in contrast to the official road regulations (Jaafari et al. 2015), which generally follow different regulations compared to official road regulations. With a minimum standard size and quality, the existence of roads in the village is also the main route used for transporting building materials and harvesting products with quite a large tonnage. Vehicle traffic with heavy loads has an impact on road stability which can trigger landslide reactivation.

Most of the houses are in flat areas, either at the top of slopes or valley plains. The current flat condition is likely to occur due to two things, occurring naturally through the erosion-deposition process during land sliding or the process of human civilization when constructing the settlement through slope modification. The 2D feature values in landslides in houses have large sizes, widths, and slopes tend to be smaller (Fig. 12). Houses and local roads are positioned at the top of slope areas, which could be the cause of landslides. Rainwater falls on residential roofs and local roads at the start of the rainy season, significantly contributing to surface flow, which then triggers the process of wetting the landslide slip surface. This concentrated surface flow acts as a source of energy that causes gully erosion and landslides (Shen et al. 2022).

Irrigated rice fields are agricultural land uses with the highest management intensity. An irrigation network follows the distribution of irrigated rice fields in the study location, which also features several fish ponds. The irrigated rice fields at the study location have terraces developed in a regular pattern following the slope pattern. A flat terrace surface will hinder surface flow and transport of soil material, but on the other hand, it will allow greater water infiltration. This event has an important role in wetting the landslide slip surface, especially the subsurface soil material, which is more impermeable. The results of a literature review (Garcia-Chevesich et al. 2021; Segoni et al. 2024) show that landslides caused by irrigation are an increasing risk in regions where there are agriculture and certain climatic and geological conditions. The characteristics of former landslides at this stage include slopes of less than eight degrees and intensive management for agriculture with more regular terracing applications. Between the terraces, there are small areas containing a group of vegetation as mixed gardens. Woody vegetation is also found extending at the foot of the slopes along the river. Land cover, especially vegetation conditions, tends to be uniform throughout the area.

Landslides with low ellipticity also occur in mixed gardens. A non-ellipse mixed garden has a different geometric character from an ellipse mixed garden, where the landslide size is larger and the slope angle tends to be smaller. The diversity of vegetation in a low ellipticity mixed garden consists of multilayer crops. Multilayer vegetation tends to stabilize slopes by anchoring soil and reducing surface runoff (Fattet et al. 2011). However, mixed gardens with low ellipticity are located close to house complexes and road networks so the intensity of maintenance, management, and harvesting of garden products is higher. High-pressure management in mixed gardens has a big role in changing the shape of landslide scars to become irregular.

The results of landslide boundary shape identification and each of its geometric characteristics serve as a significant standard for land management methods. The

Fig. 13. Active landslides were found during the field survey. (a – e) active landslides in mixed gardens adjacent to rice fields and river, (f – h) active landslides over residential areas

landslide boundary configuration is characterized by an elliptical shape, indicating recent landslide events, and post-landslide land use. The occurrence of new landslides raises concerns about the stability of the land, which is yet uncertain. Therefore, it is crucial to consider more controlled land use in areas that have recently experienced landslides. Utilizing unstable terrain without any limitations carries a substantial danger of incurring losses. After a landslide occurs, it is wise to adopt agroforestry techniques for dry field agriculture in the early stages (Garrity 2012) which provides the possibility of increasing the shear strength of the soil. Feasible cultivation of crops for dryland agriculture can be achieved by considering the stability of previous landslide-prone slopes and effectively addressing the risks of erosion and reactivation of landslides (Purwaningsih et al. 2020).

The landslide boundary is shaped in an irregular or non-elliptical form, representing an inactive landslide, with more stable slopes. Land located in places with low ellipticity can be effectively utilized for a diverse range of uses. The landslide boundary located adjacent to the ellipse in inactive landslides exhibits a range of slope conditions, ranging from flat to extremely steep. An important thing that must be considered in land use management is how to minimize the wetting of landslide slip areas when it rains, which can produce runoff that ultimately triggers landslides. Land use planning prioritizes the assessment of local landslide slopes, particularly focusing on the crown, scarp, sides, landslide body, and landslide foot. Every component of the landslide possesses a distinct ecological role, necessitating appropriate adjustments in land utilization. The combination of the preexisting landslide boundary shape and its land use serves as the primary input for generating suggestions for the detailed land use planning management. Implementing local-level planning measures is crucial for mitigating the ecological and economic damages caused by landslides.

CONCLUSIONS

Land use and land cover patterns play a fundamental role in determining both the occurrence of landslides and the morphology of resulting landslide surfaces. Our research reveals that changes in surface geomorphological features, particularly landslide boundaries, are intimately connected to temporal variations in land use and cover patterns. Analysis of two-dimensional landslide boundaries in the study area revealed distinct elliptical and non-elliptical configurations. The ellipticity index in the study area ranged from -0.74 to 0.96.

Small, active landslides, often elliptical in shape, occur alongside larger dormant ones, reflecting ongoing but diminishing geomorphic processes that result in steeper slopes and distinct geometric characteristics, with elliptical slides showing smaller area, perimeter, and width than non-elliptical types. Landslide geometry, as indicated by ellipticity and an average slope angle of 18.4°, is not solely dependent on shape but is influenced by local factors such as slope height, soil material, and slip surface, with steeper slopes generally associated with smaller landslide volumes, highlighting the inverse relationship between slope angle and landslide size under consistent material strength.

Landslide direction in the study area predominantly follows south to southwest orientations, reflecting the influence of regional geological structures and local lineaments, as further evidenced by river flow patterns shaped by structural controls that channel water from north to south and then southwest.

Elliptical landslide boundaries (-0.74 to 0.63) predominantly occur in areas of bare land, mixed gardens, rain-fed rice fields, and intermittent rivers. These elliptical patterns are typically associated with lands characterized by minimal management intensity and unstable slope conditions, underscoring the need for careful land management strategies that prioritize slope stability. In contrast, non-elliptical landslide boundaries (0.63 to 0.96) are primarily observed in areas with significant infrastructure development, including roads, housing developments, irrigated rice fields with canal systems, and managed mixed gardens. These non-elliptical patterns generally correspond to areas with intensive land management and relatively stable slope conditions. However, even in these more stable areas, careful attention must be paid to factors that could trigger slope failure, particularly the saturation of potential slip surfaces that might reactivate dormant landslides. ■

REFERENCES

Bartelletti C., Giannecchini R., D'Amato Avanzi G., Galanti Y., and Mazzali A. (2017). The influence of geological-morphological and land use settings on shallow landslides in the Pogliaschina T. Basin (northern apennines, Italy). Journal of Maps, 13(2). DOI: 10.1080/17445647.2017.1279082.

Bruschi V. M., Bonachea J., Remondo J., Gómez-Arozamena J., Rivas V., Méndez G., Naredo, J. M., and Cendrero A. (2013). Analysis of geomorphic systems' response to natural and human drivers in northern Spain: Implications for global geomorphic change. Geomorphology, 196. DOI: 10.1016/j.geomorph.2012.03.017.

Cevasco A., Pepe G., and Brandolini P. (2014). The influences of geological and land use settings on shallow landslides triggered by an intense rainfall event in a coastal terraced environment. Bulletin of Engineering Geology and the Environment, 73(3). DOI: 10.1007/s10064-013-0544-x.

Chen A., Darbon J. Ô., and Morel, J. M. (2014). Landscape evolution models: A review of their fundamental equations. Geomorphology, 219, 68–86. DOI: 10.1016/J.GEOMORPH.2014.04.037.

Chen L., Guo Z., Yin K., Pikha Shrestha D., and Jin S. (2019). The influence of land use and land cover change on landslide susceptibility: A case study in Zhushan Town, Xuan'en County (Hubei, China). Natural Hazards and Earth System Sciences, 19(10). DOI: 10.5194/nhess-19-2207-2019.

Davies T. (2015). Landslide Hazards, Risks, and Disasters: Introduction. In Landslide Hazards, Risks, and Disasters. DOI: 10.1016/B978-0-12-396452-6.00001-X.

Dhingra R. D., Barnes J. W., Hedman M. M., and Radebaugh J. (2019). Using Elliptical Fourier Descriptor Analysis (EFDA) to Quantify Titan Lake Morphology. The Astronomical Journal, 158(6). DOI: 10.3847/1538-3881/ab4907.

Donnarumma A., Revellino P., Grelle G., and Guadagno F. M. (2013). Slope angle as indicator parameter of landslide susceptibility in a geologically complex area. Landslide Science and Practice: Landslide Inventory and Susceptibility and Hazard Zoning, 1. DOI: 10.1007/978-3-642-31325-7_56.

Fan X. Y. (2010). Comparative study of movement behaviors of seismic and non-seismic induced landslides. Yantu Lixue/Rock and Soil Mechanics, 31(SUPPL. 2).

Fang Y., Xu Q., Chen J., Li J., and Zhang T. (2024). Investigating landslide-induced tsunamis using a low-dissipation Riemann weakly compressible smoothed particle hydrodynamics model. Elsevier. DOI: 10.1016/J.OCEANENG.2024.118512.

Fattet M., Fu Y., Ghestem M., Ma W., Foulonneau M., Nespoulous J., Le Bissonnais Y., and Stokes A. (2011). Effects of vegetation type on soil resistance to erosion: Relationship between aggregate stability and shear strength. Catena, 87(1). DOI: 10.1016/j.catena.2011.05.006.

Fell R., Corominas J., Bonnard C., Cascini L., Leroi, E., and Savage W. Z. (2008). Guidelines for landslide susceptibility, hazard, and risk zoning for land use planning. Engineering Geology, 102(3–4). DOI: 10.1016/j.enggeo.2008.03.022.

Garcia-Chevesich P., Wei X., Ticona J., Martínez G., Zea J., García V., Alejo F., Zhang Y., Flamme H., Graber A., Santi P., McCray J., Gonzáles E., and Krahenbuhl R. (2021). The impact of agricultural irrigation on landslide triggering: A review from chinese, english, and Spanish literature. In Water (Switzerland) (Vol. 13, Number 1). DOI: 10.3390/w13010010.

Garrity D. (2012). Agroforestry and the Future of Global Land Use. DOI: 10.1007/978-94-007-4676-3_6.

Guo W. Z., Chen Z. X., Wang W. L., Gao W. W., Guo M. M., Kang H. L., Li P. F., Wang W. X., and Zhao M. (2020). Telling a different story: The promote role of vegetation in the initiation of shallow landslides during rainfall on the Chinese Loess Plateau. Geomorphology, DOI: 10.1016/j.geomorph.2019.106879.

Hooke J. M. (2020). Changing landscapes: Five decades of applied geomorphology. Geomorphology, 366. DOI: 10.1016/j.geomorph.2019.06.007.

Jaafari A., Najafi A., Rezaeian J., Sattarian A., and Ghajar I. (2015). Planning road networks in landslide-prone areas: A case study from the northern forests of Iran. Land Use Policy, 47. DOI: 10.1016/j.landusepol.2015.04.010.

Jaboyedoff M., Michoud C., Derron M. H., Voumard J., Leibundgut G., Sudmeier-Rieux K., Michoud C., Nadim F., and Leroi, E. (2016). Human-Induced Landslides: Toward the analysis of anthropogenic changes of the slope environment. Landslides and Engineered Slopes. Experience, Theory and Practice, 1. DOI: 10.1201/b21520-20.

Jin Z., Hilton R. G., West A. J., Li G. K., Zhang F., Wang J., Li G., Fan X., and Hsieh M. L. (2021). The role of earthquake-induced landslides in erosion and weathering from active mountain ranges: Progress and perspectives. In Science China Earth Sciences (Vol. 64, Number 12). DOI: 10.1007/s11430-021-9832-3.

Kimura T. (2024). Effects of Land Cover Changes and Rainfall Variation on the Landslide Size–Frequency Distribution in a Mountainous Region of Western Japan. Geosciences (Switzerland), 14(3). DOI: 10.3390/geosciences14030059.

Li H., He Y., Xu Q., Deng J., Li W., and Wei Y. (2022). Detection and segmentation of loess landslides via satellite images: a two-phase framework. Landslides, 19(3). DOI: 10.1007/s10346-021-01789-0.

Li X. Z., and Kong J. M. (2010). Runout distance estimation of landslides triggered by "5·12" Wenchuan earthquake. Sichuan Daxue Xuebao (Gongcheng Kexue Ban)/Journal of Sichuan University (Engineering Science Edition), 42(5).

Liu X., Peng Y., Lu Z., Li W., Yu J., Ge D., and Xiang W. (2023). Feature-Fusion Segmentation Network for Landslide Detection Using High-Resolution Remote Sensing Images and Digital Elevation Model Data. IEEE Transactions on Geoscience and Remote Sensing, 61. DOI: 10.1109/TGRS.2022.3233637.

Lombardo U. (2014). Quantitative morphometric analysis of lakes using GIS: Rectangularity R, ellipticity E, orientation O, and the rectangularity vs. ellipticity index, REi. Cartography and Geographic Information Science, 41(4). DOI: 10.1080/15230406.2014.919540.

Martel S. J. (2004). Mechanics of landslide initiation as a shear fracture phenomenon. Marine Geology, 203(3–4). DOI: 10.1016/S0025-3227(03)00313-X.

Massey C. I., Petley D. N., and McSaveney M. J. (2013). Patterns of movement in reactivated landslides. Engineering Geology, 159. DOI: 10.1016/j.enggeo.2013.03.011.

Meneses B. M., Pereira S., and Reis E. (2019). Effects of different land use and land cover data on the landslide susceptibility zonation of road networks. Natural Hazards and Earth System Sciences, 19(3). DOI: 10.5194/nhess-19-471-2019.

Noviyanto A., Sartohadi J., and Purwanto B. H. (2020). The distribution of soil morphological characteristics for landslide-impacted Sumbing Volcano, Central Java - Indonesia. Geoenvironmental Disasters, 7(1). DOI: 10.1186/S40677-020-00158-8.

Ogita S., Sagara W., and Higaki D. (2017). Shapes and Mechanisms of Large-Scale Landslides in Japan: Forecasting Analysis from an Inventory (WCoE 2014–2017). In Advancing Culture of Living with Landslides. DOI: 10.1007/978-3-319-59469-9_26.

Otgonbayar M., Tsedevdorj S.-O., Bumtsend B., Enkhjargal O., and Tovuudorj R. (2023). Spatial Relationships between Topographic Variables and Their Interactions with Natural and Climatic Zonalization in Mongolia. European Modern Studies Journal, 7(5). DOI: 10.59573/emsj.7(5).2023.10.

Patanduk A. (2020). Integrasi Metode Foto udara, GNSS dan Seismik Refraksi Untuk Identifikasi Karakteristik Retakan Permukaan Pada Area Longsor Kalisari, Sub DAS Bompon, Magelang, Jawa Tengah. Tesis. Universitas Gadjah Mada. https://etd.repository.ugm.ac.id/penelitian/detail/187118.

Pirajno F. (2009). Hydrothermal processes and mineral systems. Springer.

Pratiwi E. S., Sartohadi J., and Wahyudi. (2019). Geoelectrical Prediction for Sliding Plane Layers of Rotational Landslide at the Volcanic Transitional Landscapes in Indonesia. IOP Conference Series: Earth and Environmental Science, 286(1). DOI: 10.1088/1755-1315/286/1/012028.

Pudasaini S. P., and Krautblatter M. (2021). The mechanics of landslide mobility with erosion. Nature Communications, 12(1). DOI: 10.1038/s41467-021-26959-5.

Pulungan N. A., and Sartohadi J. (2018). New Approach to Soil Formation in the Transitional Landscape Zone: Weathering and Alteration of Parent Rocks. Journal of Environments, 5(1). DOI: 10.20448/journal.505.2018.51.1.7.

Purwaningsih R., Hartanti H., Hatta D., Ariyanti S., Putri H., Paradisa D., Dewi I., Priyawati E., Stakhis E., Salsabila G., Mashum H., Pulungan N. A., Handayani S., Sartohadi J., and Gomez C. (2025). Subsoil Mineralogy as Environmental Factors in Controlling Topsoil Material in the Menoreh–Sumbing Volcanic Structural Transitional Landscape, Central Java. Journal of Water and Land Development, 67 (X–XII): 73–85(67). DOI: 10.24425/jwld.2025.156041.

Purwaningsih R., Sartohadi J., and Anggri M. (2020). Trees and crops arrangement in the agroforestry system based on slope units to control landslide reactivation on volcanic foot slopes in Java, Indonesia. Land, 9(9). DOI: 10.3390/LAND9090327.

Ran Q., Hong Y., Li W., and Gao J. (2018). A modelling study of rainfall-induced shallow landslide mechanisms under different rainfall characteristics. Journal of Hydrology, 563. DOI: 10.1016/j.jhydrol.2018.06.040.

Samodra G., Ramadhan M. F., Sartohadi J., Setiawan M. A., Christanto N., and Sukmawijaya A. (2020). Characterization of displacement and internal structure of landslides from multitemporal UAV and ERT imaging. Landslides, 17(10). DOI: 10.1007/s10346-020-01428-0.

Sartohadi J., Rahma A. D., and Nugraha S. S. (2024). Productive conservation at the landslide prone area under the threat of rapid land cover changes. Open Geosciences, 16(1). DOI: 10.1515/GEO-2022-0700/MACHINEREADABLECITATION/RIS.

Schmittbuhl M., Allenbach B., Le Minor J. M., and Schaaf A. (2003). Elliptical descriptors: Some simplified morphometric parameters for the quantification of complex outlines. Mathematical Geology, 35(7). DOI: 10.1023/B: MATG.0000007783.72366.0c.

Segoni S., Ajin R. S., Nocentini N., and Fanti R. (2024). Insights gained from the review of landslide susceptibility assessment studies in Italy. REMOTE SENSING, 16(23), 1–31. DOI: 10.3390/RS16234491.

Shen W., Berti M., Li T., Benini A., and Qiao Z. (2022). The influence of slope gradient and gully channel on the run-out behavior of rockslide-debris flow: an analysis on the Verghereto landslide in Italy. Landslides, 19(4). DOI: 10.1007/s10346-022-01848-0.

Tarolli P., and Sofia G. (2016). Human topographic signatures and derived geomorphic processes across landscapes. Geomorphology, 255, 140–161. DOI: 10.1016/J.GEOMORPH.2015.12.007.

Taylor F. E., Malamud B. D., Witt A., & Guzzetti F. (2018). Landslide shape, ellipticity and length-to-width ratios. Earth Surface Processes and Landforms, 43(15). DOI: 10.1002/esp.4479.

Thorn C. E., and Welford M. R. (1994). The Equilibrium Concept in Geomorphology. Annals of the Association of American Geographers, 84(4), 666–696. DOI: 10.1111/J.1467-8306. 1994.TB01882. X.

Tort A. (2003). Elliptical Fourier functions as a morphological descriptor of the genus Stenosarina (Brachiopoda, Terebratulida, New Caledonia). Mathematical Geology, 35(7). DOI: 10.1023/B: MATG.0000007784.18452.73.

Trantham G. (2023). Improving Detail in Shaded Relief. Cartographic Perspectives, 2023(101). DOI: 10.14714/CP101.1789.

Vuillez C., Tonini M., Sudmeier-Rieux K., Devkota S., Derron M. H., and Jaboyedoff, M. (2018). Land use changes, landslides and roads in the Phewa Watershed, Western Nepal from 1979 to 2016. Applied Geography, 94. DOI: 10.1016/j.apgeog.2018.03.003.

Wida W. A., Maas, A., and Sartohadi J. (2019). Pedogenesis of Mt. Sumbing Volcanic Ash above The Alteration Clay Layer in The Formation of Landslide Susceptible Soils in Bompon Sub-Watershed. Ilmu Pertanian (Agricultural Science), 4(1). DOI: 10.22146/ipas.41893.

Wu Y. Bin, Duan Z., Peng J. B., and Zhang Q. (2024). Implications of sand grains' mobility and inundating area to landslides at different slope angles. Granular Matter, 26(1). DOI: 10.1007/s10035-023-01387-y.

Yin Z., Qin X., and Zhao W. (2014). Characteristics of Landslides from Sigou Gorge to Lagan Gorge in the Upper Reaches of Yellow River. In Landslide Science for a Safer Geoenvironment (Vol. 1). DOI: 10.1007/978-3-319-04999-1_56.

RESEARCH PAPER

GEOSTATISTICAL REGULARITIES OF SOIL ACIDITY DIFFERENTIATION ON FOREST, ARABLE, AND MEADOW LANDS IN THE BEREZINA RIVER VALLEY (BELARUS)

Arkadzy L. Kindeev*
Belarusian State University, Nezavisimosti av., 4, Minsk, 220030, Belarus
*Corresponding author: AKindeev@tut.by

Received: April 2nd 2025 / Accepted: January 27th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3978

ABSTRACT. This article investigates the regularities of the spatial distribution of soil acidity on a detailed scale. A description of the soil cover and the rationale for selecting key areas within forest, arable, and meadow lands in the Berezina River valley (Belarus) are provided. A universal scheme of geostatistical diagnostics of soil cover properties is proposed, describing theoretical aspects of geostatistics, which can be applied to solve soil-geographical problems in various specialized contexts. Variogram analysis was employed to ascertain the distances at which similar soil-geochemical processes occur within the studied landscapes. In forested areas, a reduction in dispersion is observed at distances of 140–180 m, which aligns with the slope length and the spacing between ravines with temporary watercourses. Acidity on arable lands is characterized by high dispersion values at small distances (70–80 m) and decreases at large distances (more than 250 m). Meadow lands show a sharp jump in dispersion at distances of 130–170 m, which corresponds to the width of floodplain ridges. For quantitative assessment of variation, we propose a new indicator "variation per meter", which allows us to move from comparisons of absolute values to relative ones, thus removing the influence of site size The values obtained for the new indicator elucidate classical concepts regarding the distribution of soil acidity and the transformation of natural landscapes due to anthropogenic impact. The "variation per meter" is approximately 2% for forest lands (minimum anthropogenic transformation), 0.1–0.2% for arable lands (maximum transformation), and about 1% for meadow lands (intermediate transformation).

KEYWORDS: soil cover, spatial structure, variogram, soil-geochemical processes

CITATION: Kindeev A. L. (2026). Geostatistical Regularities Of Soil Acidity Differentiation On Forest, Arable, And Meadow Lands In The Berezina River Valley (Belarus). Geography, Environment, Sustainability, 1 (19), 86-96
https://doi.org/10.24057/2071-9388-2026-3978

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

The study of spatial differentiation of soil properties is undoubtedly one of the most important directions of digital soil cartography and the basis for optimizing agricultural activities, taking into account natural factors. To this end, the scientific and methodological framework of data analysis based on probability theory, known as geostatistics, was employed. This apparatus enables the expression of random phenomena in space and/or time through mathematical and statistical indicators and the classification of the regularities of their distribution.

The extensive range of geostatistical methods, when applied in a dynamic and innovative manner, has enabled the advancement of soil geography, utilizing state-of-the-art technologies and contemporary software solutions. This has facilitated the development of methods for precise, direct, or "mathematical" mapping of soil cover and its characteristics [Webster 2007; Krasilnikov & Targulyan 2019]. The development of such methods has enabled the establishment of a new research direction in soil science and soil geography, known as pedometrics [Savin et al. 2019; McBratney et al. 2018].

It is essential to note that direct (mathematical) mapping and geostatistics are not in opposition to the classical approach to land cover mapping. However, as demonstrated by chaos theory, the presence of order and regularity at one hierarchical level does not preclude their absence at another [Esmaeilizad et al. 2024]. The concept of a transition to a "new soil geography" was introduced in 2019, representing a synthesis of two divergent approaches [Krasilnikov & Targulyan 2019]. Presently, considerable attention is being directed towards the utilization of machine learning algorithms for the prediction of soil properties. However, methodologies predominantly based on regression matrices overlook the spatial dimension inherent in the data being studied [Heuvelink & Webster 2022; Vaysse & Lagacherie 2017; Wadoux A.M.J.C., et al. 2021]. In certain instances, the integration of machine learning with residual kriging has been shown to yield optimal accuracy [Molchanov et al. 2015]. Concurrently, the analysis of spatial data employing variogram analysis facilitates not only the consideration of data distribution in space but also the identification of patterns and consequences of their differentiation.

Over recent decades, the number of studies in geostatistics and pedometrics aimed at investigating these properties has increased significantly (a summary of these works is given in [Oliver 2010] and [Biswas 2024]). Most of them, however, are limited to the study of the accuracy of the instruments [Mohamed et al. 2023; Soropa et al. 2021; Vaysse & Lagacherie 2017] or to the interpretation of the maps obtained [Suleymanov et al. 2023; Xiao et al. 2023], although several studies dedicated to a comprehensive analysis of the causal relationships underlying the distribution of acidity [Helfenstein et al. 2022] and organic carbon [Szatma´ri & P´asztor 2019] should be noted. The studies face a number of difficulties, the most significant of which are, firstly, the labor-intensive process of obtaining data sets, and secondly, the financial expense associated with soil sample analysis. Concurrently, the selection of soil properties for analysis ought to be informed by the available options for their examination. While moisture and granulometric composition appear to be obvious choices for analysis, their study is very labor-intensive. Conversely, analysis of acidity remains relatively straightforward and does not necessitate substantial financial expenditure. It is also important to note that acidity is one of the most significant soil properties.

The degree of soil acidity exerts a significant influence on the eventual soil fertility, with the process of soil acidification representing a form of chemical degradation [Molchanov et al. 2015]. In the context of leaching water management, soil acidification emerges as a pivotal factor contributing to hazardous agricultural practices in the Republic of Belarus [Klebanovich & Vasilyuk 2003].

The consequences of soil acidification include a decline in the productivity of cultivated plants, a disruption in the growth of natural vegetation, and a deterioration in the viability of aquatic ecosystems. The main causes are deficiencies in Ca, Mg, P, and Mo [Guo et al 2015]. These deficiencies may be exacerbated by a flushing water regime, which depletes ions in the soil.

Although studies that investigate the causes and consequences of acidity's spatial distribution in soil are important, a significant gap remains in understanding how soil properties vary within the classification units themselves.

The objective of this study is to identify the patterns of spatial differentiation of soil acidity on a detailed scale under different types of land in the Berezina River valley (Belarus), and to assess the degree of transformation of the spatial pattern of acidity distribution under the influence of human agricultural activities.

MATERIALS AND METHODS

The following section outlines the general characteristics of the territory. The geographical area under scrutiny is delineated by the geographical station (GS) "Western Berezina", situated between the Oshmianskaya and Minsk uplands in the valley of the Berezina River. This territory is situated within the Volozhinsky administrative district of the Republic of Belarus (Fig. 1). The rationale behind the selection of this territory is twofold. Firstly, the high dissection of the relief leads to a high heterogeneity of the soil cover, which in turn contributes to the active course of erosion-denudation processes. Second, the area has undergone significant agricultural development.

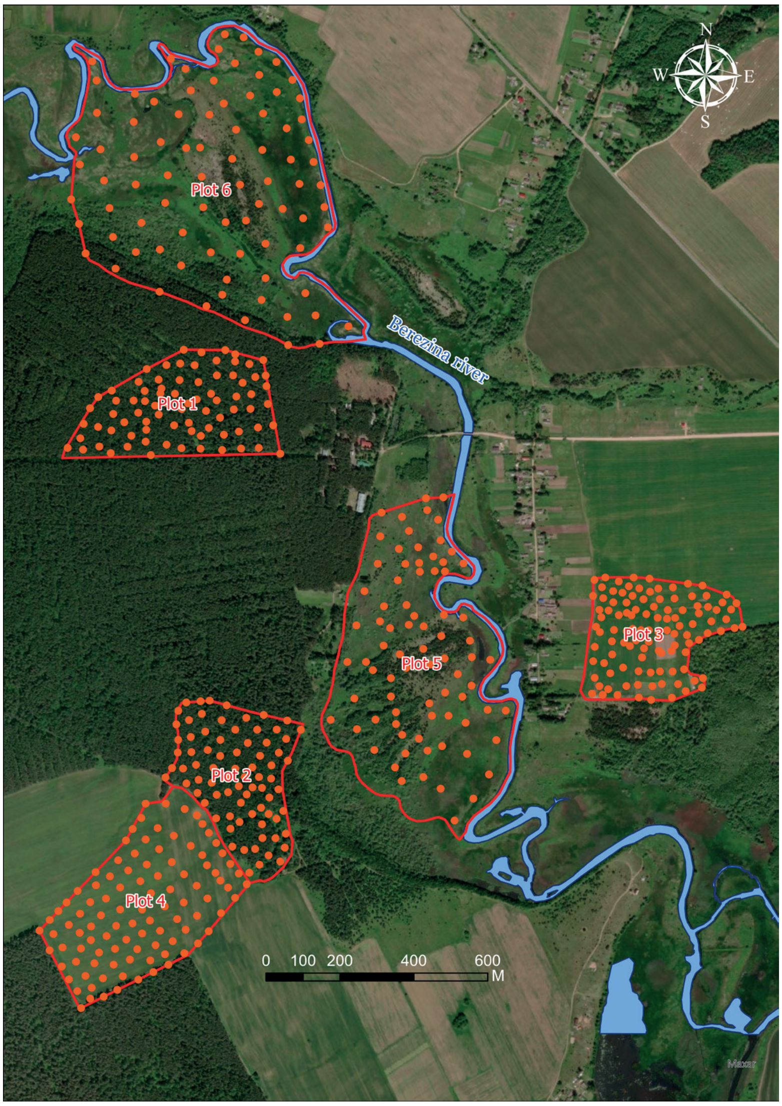


Fig. 1. Overview map of the study area location with sampling points

The objects of the study are six plots of land, categorized as follows: forest (plots 1 and 2), arable (3 and 4), and meadow (5 and 6). The plots were selected to include all major soil types in the study area (Fig. 2), ensuring thorough representation of hydromorphism and granulometric composition. This approach ensures the representativeness of the selected plots for the study's objectives.

The sampling process was conducted synchronously at all selected sites throughout May 2023, in accordance with the prescribed methodology (GOST 17.4.3.01-83 "General requirements for sampling"). The sampling depth was maintained between 0–20 cm. The humus-accumulative horizon capacity in forest plot soils was 10–15 cm thick. Soils on arable land had a 30 cm humus-accumulative horizon. Meadow plots in the flooded part of the floodplain had a 25–30 cm horizon, while those on aeolian [Kurlovich et al. 2024] ridges had a 13–15 cm horizon.

Forest plots of 12.6 and 11.3 ha were selected for analysis, from which 80 and 78 samples were collected. The soil cover of the forest plots consists of sod-podzolic soils (Retisols) and sod-podzolic waterlogged soils (Gleyic Retisols), which more often develop on cohesive and loose loams and light loams. In the northern part of both sites, the vegetation is represented by pine orchard-grass on sod-podzolic automorphic soils. In the southern part, it is represented by mossy spruce, which is confined to sod-podzolic, weakly gleyey soils.

In contrast, the arable plots, measuring 10.6 and 17.6 hectares, exhibit less heterogeneity in soil composition. From these plots, 110 and 102 samples were taken for further analysis. The soils are classified as sodpodzolic, developed on cohesive and loose loams and on light loams. At the time of fieldwork, wheat was sown on the plots. Notably, no crop rotation had been implemented on plot 3 in the previous year. The precursor crops on plot 4 were corn in 2022 and fallow in 2021.

The meadow plots, situated within the maned floodplain, exhibited the most extensive area (30.7 and 44.2 ha, respectively) from which 79 and 85 samples were collected. The soil cover is represented by alluvial soddy-gleyey soils (Gleyic Fluvisols), soddy-gleyey soils (Gleyic Fluvisols) and silty-peaty-peaty-gleyey soils (Hemic/Fibric Histosols).

The soil types present at these locations encompass 42 varieties, indicative of automorphic, semi-hydromorphic, and hydromorphic soils. This distinguishes these sites as unique "catenas", illustrating the sequential transition among diverse soil types. This facilitates geostatistical analysis and interpretation of the horizontal distribution of acidity across the predominant soil categories.

Research Methods. Research methods. This study is based on the use of keys and on the principles of continuity of soil cover properties, probability theory, chaos theory, and the entropy approach [4], which underpin geostatistical interpolators.

The construction of the sampling grid was informed by two criteria: the geomorphological features of the area and the uniformity of plot coverage. This approach was adopted to circumvent the occurrence of both clustering of data and "empty" areas devoid of sampling.

The analysis of soil acidity (pHKCl) was conducted in the laboratory using potentiometric titration in accordance with the standards outlined in GOST 26484-85. "Soils. Method of determination of exchangeable acidity".

The geostatistical analysis of the obtained results was conducted using ArcGIS ArcMap software, following a complex flowchart that was developed based on the analysis of numerous studies [Lark 2012; Samsonova & Meshalkina 2020; Biswa 2024] and empirical research [Klebanovich et al. 2021; Klebanovich et al. 2018] (Fig. 3).

The distinguishing characteristic of the proposed geostatistical analysis algorithm is its integration of theoretical geostatistics with a decision-making algorithm

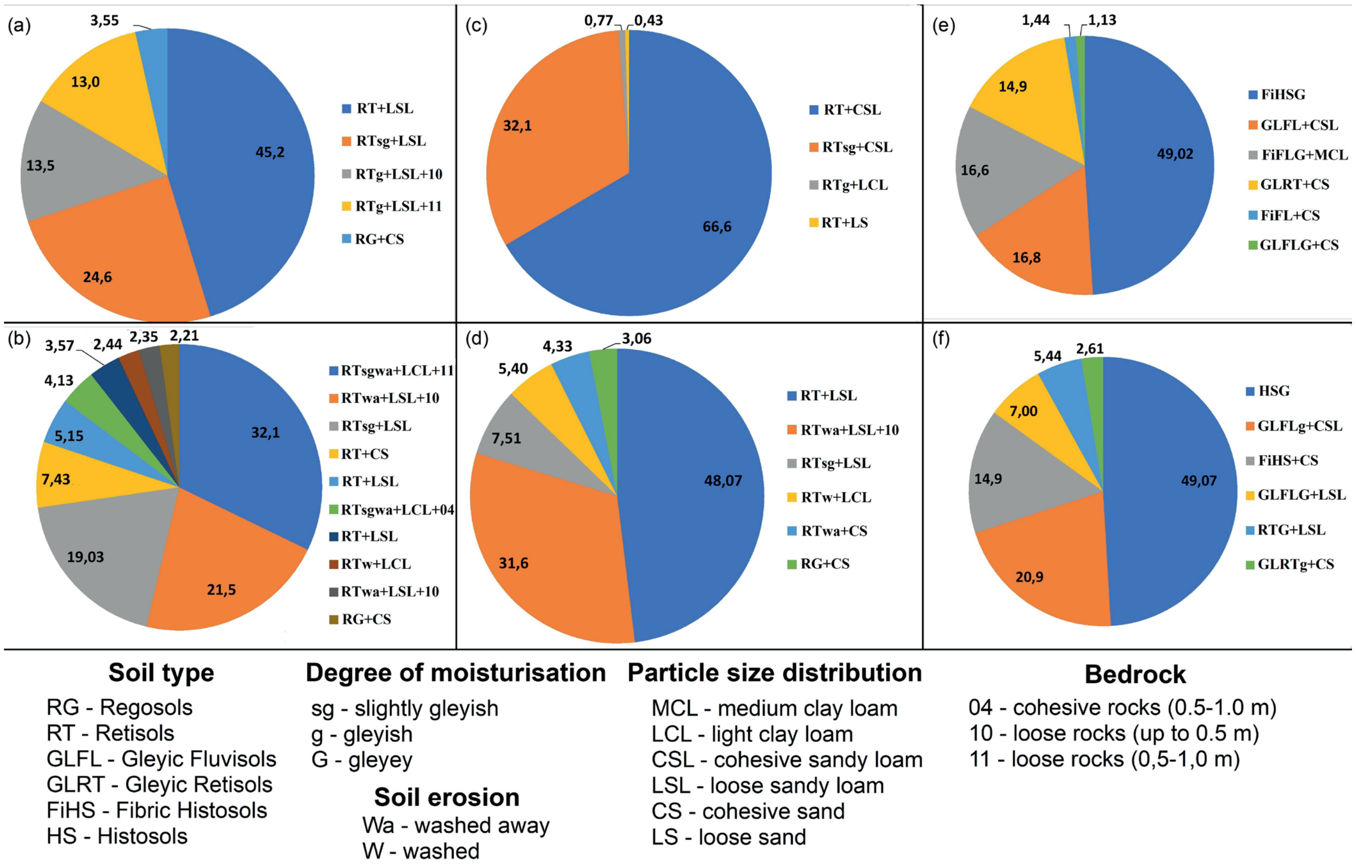


Fig. 2. Land cover composition of the study plots: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

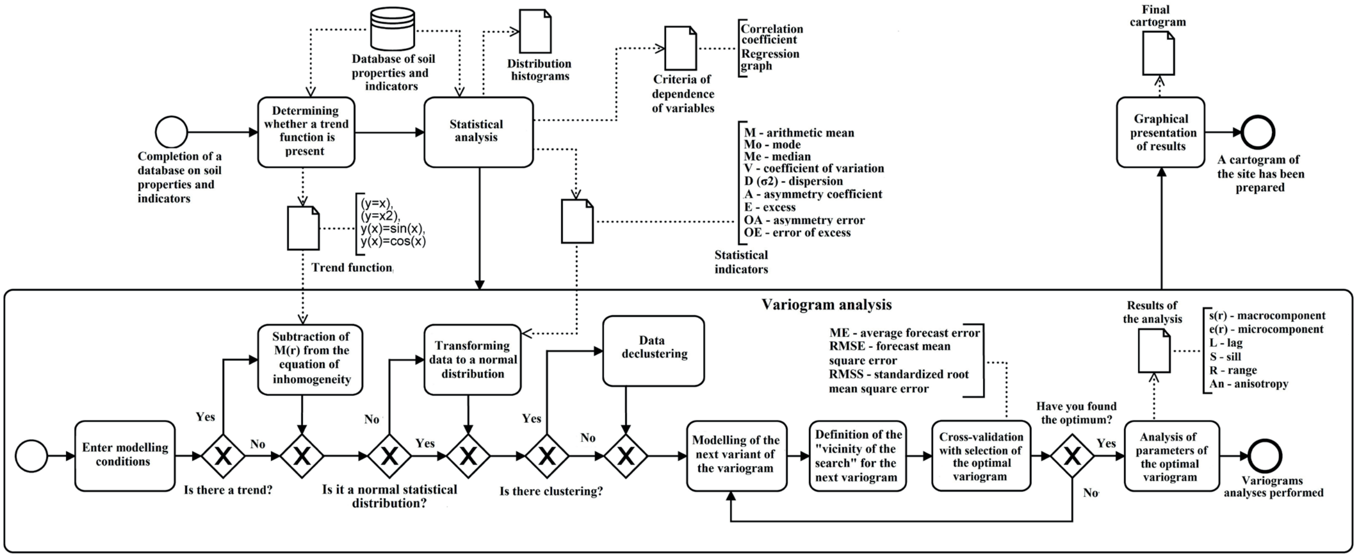


Fig. 3. Block diagram of the model of geostatistical diagnostics of soil properties variability

for modeling the studied phenomenon in solving applied geographical problems. Additionally, it offers the possibility of using it to create separate applications for geostatistical analysis.

The initial block delineates the primary phases of spatial heterogeneity analysis. Spatial heterogeneity *(f(r))* is defined as the sum of three components: macrocomponents *(m(r))*, mesocomponents *(s(r))* and microcomponents *(e(r))* [Yakushev et al. 2010]. The radius vector *(r)* defines the position of an arbitrary point in the chosen coordinate system *(x; y), i.e., r = r (x, y)*. The statistical distribution of the data is determined by the center of distribution (mode *(Mo)* and median (Me)), variant dispersion (variation *(V)* and dispersion *(D, σ2)*), and distribution shape (asymmetry *(A)* and excess *(E)*). The deterministic trend (macrocomponent) is also identified, in addition to variogram analysis and graphical presentation of results.

The primary processing and analysis of the obtained statistical results were carried out in Microsoft Excel software products, where the data were checked for normality of distribution according to the method of E.A. Pustylnik [Pustylnik 1963] (Eqs. 1 and 2). E. A. Pustylnik's method was chosen because Microsoft Excel uses its formulas to calculate the skewness and kurtosis coefficients. To maintain uniformity, errors were also calculated using his method.

$$O_A=3\times\sqrt{\frac{6\times(n-1)}{(n+1)\times(n+3)}} \qquad (1)$$

$$O_E=5\times\sqrt{\frac{24\times n\times(n-2)\times(n-3)}{(n+1)^2\times(n+3)\times(n+5)}} \qquad (2)$$

where *OA* – asymmetry coefficient error; O_E – excess error; *n* – sample size.

The second block is devoted to the most important component of geostatistical analysis: variography, which is the selection of the mathematical function of the variogram to the variogram constructed from real data (empirical variogram). The calculation of the main parameters of the variogram is then undertaken. The nugget-effect *(e(r))*, representing the "noise" component, is defined as the unaccounted variability. The partial sill *(s(r))*, which corresponds to the variance included in the model, is another key element. The sill *(s(r)+e(r))*, representing the maximum level of variance at the variogram rank level, is also a crucial factor. The variogram range (R, m) is defined as the distance at which points cease to correlate with each other. Lag size (L, m) is the average distance between sampling points; residual variance (nugget to sill ratio) (D, %) is the fraction of noise and interpolation accuracy.

It encompasses supplementary operations implemented solely when required: trend removal, data transformation, and declustering. The term also includes anisotropy (An, °), nearest neighbor point, and cross-validation (cross-validation) determinations.

The accuracy of variogram fitting was checked by cross-validation, where each point in turn is removed from the calculation, and the remaining points are used to calculate the predicted value at the location of the removed point (the process is repeated for all input points) and the calculation of interpolation errors (Eq. 3–5):

$$ME=\frac{\sum_{i=1}^{n}\left(\hat{Z}(s_i)-z(s_i)\right)}{n} \qquad (3)$$

where ME – average forecast error;

$\hat{Z}(s_i)-z(s_i)$ the average difference between the measurement and the predicted value;

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}\left(\hat{Z}(s_i)-z(s_i)\right)^2}{n}} \qquad (4)$$

where *RMSE* – forecast mean square error.

$$RMSS=\sqrt{\frac{\sum_{i=1}^{n}\left[\left(\hat{Z}(s_i)-z(s_i)/\sigma(s_i)\right)\right]^2}{n}} \qquad (5)$$

where *RMSS* – standardized root mean square error.

RESULTS

The following table (Table 1) presents the indicators of descriptive statistics based on the results of laboratory analyses of soil acidity in the selected plots.

The primary distinction among forest, arable, and meadow (floodplain) plots can be identified using the coefficient of variation. Forest plots exhibit variation levels of 11.9 and 16.5%, while arable plots demonstrate the lowest values of 9.52 and 8.74%. The meadow (floodplain) plots, on the other hand, exhibit a coefficient of variation of 14.8%.

Site 1 is distinctly different, where asymmetry and kurtosis reach very high values of 2.40 and 8.16, which are an order of magnitude higher than their errors (0.41 and 0.76, respectively). A slight right-sided asymmetry is observed in

Table 1. Statistical indicators of pH_{KCl} variation of study sites

Parameters	The forest plots		The arable plots		The meadow plots (floodplain)	
	1	2	3	4	5	6
n	80	78	110	102	79	85
Minimum	4.19	4.02	4.98	4.82	3.89	3.67
Maximum	7.75	8.17	7.41	7.63	7.05	6.96
Mean (M)	4.98	5.80	5.91	6.10	5.44	5.26
Median (Me)	4.88	5.83	5.80	6.06	5.32	5.35
Mode (Mo)	4.41	6.79	5.80	5.91	6.11	5.04
Mean square deviation, √D	0.59	0.95	0.57	0.53	0.81	0.78
Coefficient of variation (V, %)	11.9	16.5	9.52	8.74	14.8	14.8
Asymmetry (A)	2.40	0.27	0.70	0.52	0.32	-0.03
Asymmetry error (OA)	0.41	0.42	0.37	0.38	0.41	0.41
Excess (E)	8.16	-0.64	-0.01	0.99	-0.69	-0.79
Error of excess (OE)	0.76	0.77	0.72	0.69	0.77	0.76

plots on arable land, thus showing a greater proportion of high pH values, which is due to the application of chemical inputs. Regarding kurtosis, except for the first plot, slight leptokurtosis is evident in plot 4, while plot 6 shows slight platykurtosis.

During macrocomponent determination, no trends were observed in plots 1 and 6, while plots 2 and 4 exhibited a second-order trend and plot 5 displayed a first-order trend; these were taken into account when selecting variograms (Fig. 4). The selection of variograms was carried out on the basis of forecast errors and residual variance. The primary variogram types analyzed are outlined in Table 2.

There is a high degree of similarity in variogram error rates; however, a substantial discrepancy is observed

Table 2. Prediction error rates of variograms of the study sites

Plot	Variograms	ME	RMSE	RMSS	Residual Dispersion, %
1	Gauss	-0.01	0.57	1.04	50.9
	Exponential	-0.01	0.57	1.07	13.2
	Spherical	-0.01	0.57	1.05	40.4
2	Pentaspherical	-0.02	0.95	1.01	0
	Exponential	-0.01	0.95	1.01	36.3
	Spherical	-0.02	0.95	1.01	43.9
3	Gaussian	0.01	0.41	1.01	36.9
	Exponential	0.01	0.41	1.01	10.0
	Spherical	0.01	0.41	0.99	24.3
4	Gaussian	-0.001	0.50	1.08	67.2
	Exponential	-0.001	0.51	1.07	45.5
	Pentaspherical	-0.003	0.49	1.05	0.00
5	Gaussian	-0.002	0.56	1.02	37.0
	Exponential	0.004	0.59	1.06	2.37
	Spherical	0.001	0.57	1.01	22.3
6	Gaussian	-0.01	0.63	1.05	48.1
	Exponential	-0.01	0.64	1.06	7.12
	Spherical	-0.01	0.63	1.05	37.5

in residual variance, which was used as the primary criterion for model selection. An exponential variogram was selected for the first forest area, and a pentaspherical variogram was selected for site 2. The spatial distribution of acidity in plots 4 and 5 is described by exponential and pentaspherical variograms, respectively. The soil acidity of the meadow (floodplain) plots is primarily characterized by the exponential variogram (Fig. 4).

The analysis of soil acidity in the forest plots reveals a decrease in dispersion at distances of 140–180 m (more clearly expressed in the variogram of the second plot), which corresponds to the length of the slope in plot 1 and the distance between ravines with temporary watercourses in plot 2. The increase in variance at distances of 220–280 m and the subsequent decline (more clearly evident in the variogram of the first plot) indicates the presence of general deterministic dependencies emerging at distances greater than 400 m.

Variograms of soil acidity on arable plots exhibit a single distinction, namely the presence of anisotropy at plot 4. Otherwise, variograms demonstrate a smooth increase in dispersion at small distances, oscillations in the central part, and a decrease at the end.

It is also noteworthy that the variograms of sites 2 and 4 are almost identical, a phenomenon attributable to their close proximity to each other and, consequently, to geochemical processes.

It is pointed out at site 5 that anisotropy is observed, along with a sharp jump in variance at distances of 130–170 m, which corresponds to the width of the floodplain mane in the central part of the site; this emphasizes significant differences between acidity in the flooded part of the floodplain and on the elevation.

The experimental variogram at these sites shows similar behavior, evidenced by a slight decrease in variance at medium distances (370–500 m). The decrease is attributable to similar natural conditions at sampling points at these distances (points near the terrace and the mane, or flooded areas separated by the mane).

The main parameters obtained from the variograms are presented in Table 3.

The obtained results allow for the determination of the main geostatistical differences among the sites. First, note that the average distance between points (L) for sites within the same group differs insignificantly, allowing comparison of the results.

The sill values for forest soil acidity *(s(r)+e(r))* are observed in the plots, ranging from 0.9 to 1.3, for arable soils from thousandths and hundredths of a fraction (0.006–0.01), and for meadow (floodplain) soils from tenths of a fraction (0.53–0.54).

To confirm these differences, a measure of "variation per meter" was derived, since the threshold reflects the variance and the range the distance, it is possible to

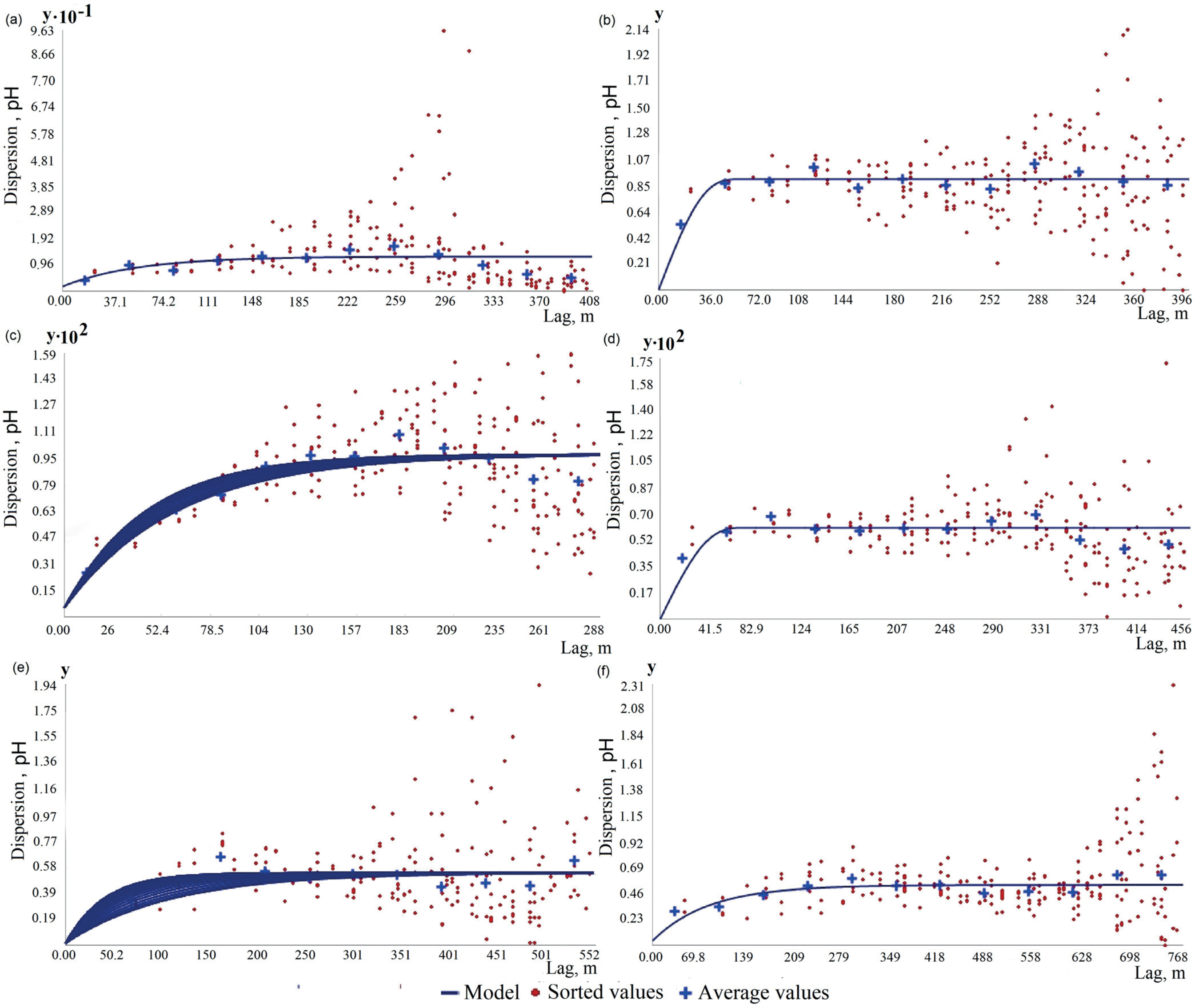


Fig. 4. Soil acidity variograms of the study sites: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

Table 3. Parameters of variograms of the investigated sites in the territories adjacent to HS "Western Berezina"

Parameters	The forest plots		The arable plots		The meadow plots (floodplain)	
	1	2	3	4	5	6
Nugget (e(r))	0.17	0.002	0.001	0.00	0.01	0.04
Partial sill (s(r))	1.11	0.91	0.01	0.01	0.53	0.50
Sill (s(r)+e(r))	1.28	0.91	0.01	0.01	0.55	0.53
Lag (L, m)	34.0	33.0	24	38	46	64
Range (R, m)	172	58.9	182	71.3	160	241
Residual Variance (D, %)	13.2	0.20	10.0	0.00	2.37	7.12
Anisotropy (An, °)	0.00	0.00	21.4	0.00	52.5	0.00
Mean	4.98	5.8	5.91	6.1	5.44	5.26
Dispersion/meter (s(r)+e(r)/R)	0.007	0.015	0.00005	0.00008	0.003	0.002
Mean square deviation/meter	0.086	0.124	0.007	0.009	0.06	0.05
Variation/meter (V/R, %)	1.73	2.14	0.13	0.15	1.08	0.89
Variation coefficient, %	11.9	16.5	9.52	8.74	14.8	14.8

calculate the variance attributable to one unit of distance. The calculation was performed as follows: 1) The ratio of threshold (maximum variance) to range (distance (m) of autocorrelation) was calculated – "variance/meter"; 2) From "variance/meter", the root (standard deviation dissection) was extracted – "mean square deviation/meter"; and 3)"mean square deviation/meter" was divided by the arithmetic mean of plot pH – "variation per meter". This indicator allows moving from comparisons of absolute values to relative values, thus removing the influence of site size. The values obtained for this indicator confirm the previously described differences: variation per meter is approximately 2% for forest soils, 0.1–0.2% for arable soils, and about 1% for meadow soils.

The observed low values of residual dispersion indicate the high accuracy of the obtained cartograms (Fig. 5).

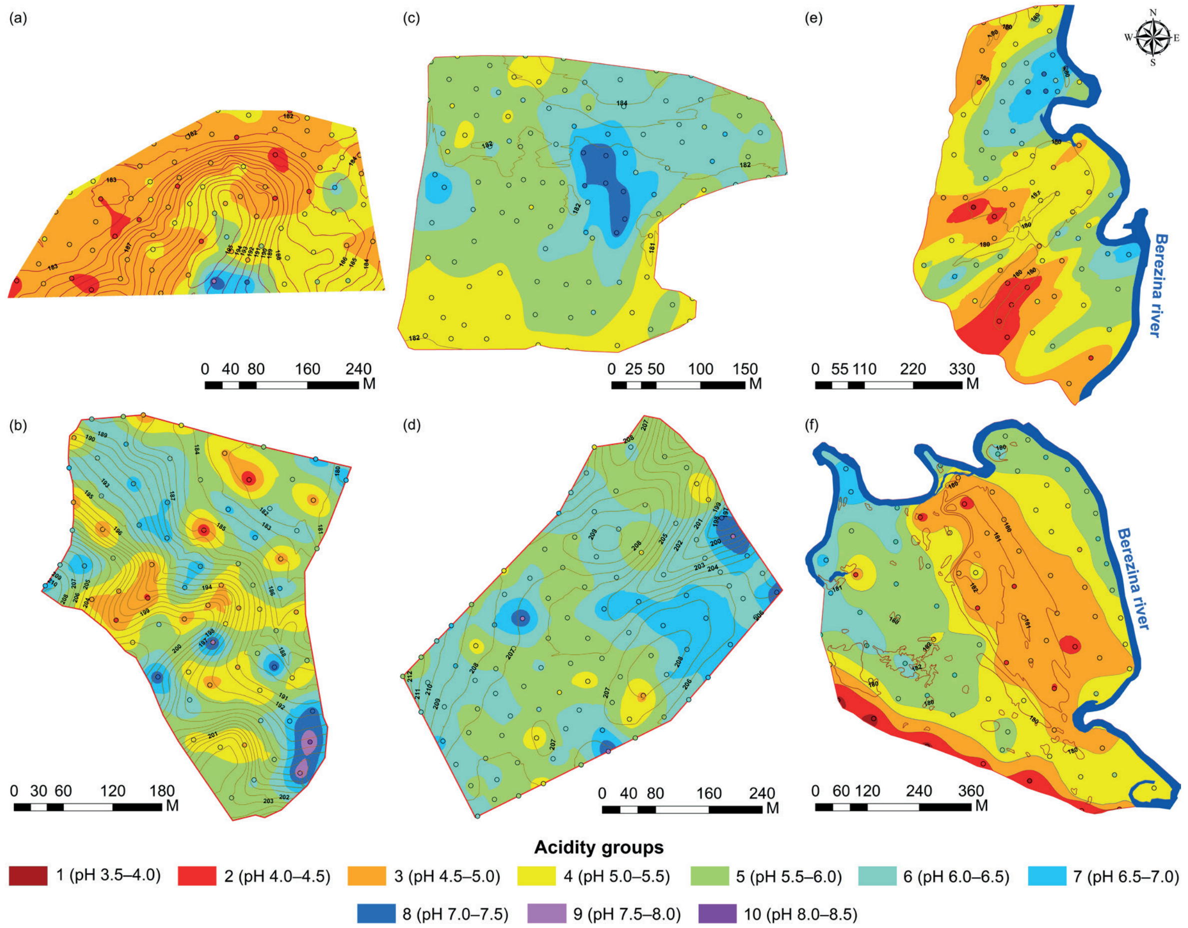


Fig. 5. Soil acidity cartograms of the study sites: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

Almost all cartograms demonstrate a high level of heterogeneity in acidity. In forest plots, values ranging from 4.0 to 8.5 are evident. The first plot is characterized by a smaller spread of pHKCl values (4.0–7.0) and a standard error of interpolation ranging from 0.20–0.37 (2.79 ha (16.8 % of the total plot area) to 3.36 ha (20.2 %)). The remaining part of the plot (10.5 ha) has another area of 0.37–0.50 (63.0 %). The second forest plot has a variety value range of 0.15–0.95. Only 5.7% (0.64 ha) of this area has an error of less than 0.56, while the remaining areas with values from 0.92 to 0.95 account for 26.7% (3.02 ha).

There is significantly lower differentiation of soil acidity, both within and between fields, in arable land (from 5.0 to 7.5). Due to the smaller scatter of values, prediction errors at site 3 do not exceed 0.43 (1.03 ha / 9.11 %). Most of the area has an index between 0.27 and 0.42 (7.41 ha / 65.5 %), while the remainder belongs to values below 0.27 (0.18 ha / 25.4 %). A similar situation is observed at site 4, where only 2.5 % (0.44 ha) of the entire study area has a prediction error greater than 0.45, with a maximum error of 0.63. The minimum – 0.04 units, with an average range of 0.22 to 0.44 (16.75 % to 95.2 %).

The meadow plots (5 and 6), which are located on the maned floodplain, are characterized by high soil acidity heterogeneity. With an average pHKCl value of 5.0–6.0 across both sites, two areas of approximately equal size can be distinguished: 1) areas with acidic and moderately acidic reactions (pHKCl 4.5–5.5), and 2) areas with slightly acidic and close-to-neutral reactions (pHKCl 5.5–6.5). The RMS error of prediction is similar at both sites. The respective areas are 38.1 % (11.9 ha) and 44.5% (20.4 ha) between 0.47 and 0.55. The minimum (0.20–0.34) and maximum (0.55–0.70) values at both sites account for less than 1.5 ha.

DISCUSSION

The primary statistical processing indicates a substantial differentiation of acidity at all sites, despite relatively modest values of the coefficient of variation (8.74 to 14.8%), which can be attributed to the logarithmic values of H^+ ion content. However, when assessing the values of the coefficient of variation qualitatively, it is important to recognize the difficulty in determining which coefficients to consider as large and which as small [3]. Given that pH values are already logarithmic scale, it can be posited that these values may indeed signify a medium to high degree of heterogeneity in acidity.

To address the anomalously high kurtosis index for acidity values at site 1, a rather uncommon method of data transformation was employed, namely the Box-Cox method of the second degree (used for exponential distribution). Conversely, no such transformation was necessary for plot 2. Plots 3 and 4 show right-skewed distributions, necessitating logarithmic transformation of the data to approximate normality. The floodplain plots (5 and 6) demonstrate approximately normal distributions and thus do not require transformation.

It is quite typical to observe trends in arable plots (3 and 4), which are often a consequence of human impact on the soil, such as fertilizer application and plowing. The key plot 3 polynomials indicate an increase in values in the central part, with an increase followed by a decline in the north-south and west-east directions (more pronounced in the north-south direction), which corresponds to a parabolic dependence. A similar polynomial describes the spatial distribution of acidity at site 4; however, in the west-east direction, there is a decrease in values in the central part. The acidity of meadow plot 5 is characterized by a linear dependence, reflecting an increase in pHKCl values from west to east (from terrace to channel) and from south to north (from peaty soils to alluvial soddy and soddy waterlogged soils).

The selection of the variogram is central to any geostatistical analysis, and it is this that has a significant impact on the final interpolation result. As demonstrated in this study, an exclusive focus on standard criteria can yield skewed and inconsistent results. While differences in errors as small as hundredths or thousandths may appear negligible, disparities in residual variance prove to be of paramount importance to the ultimate outcomes. For instance, at site 1, the RMS normalized error between variograms differs by a mere 0.02 (all other indices being equal), while the differences in residual variance reach 30–40%. If the Gaussian variogram is chosen, the cartogram is found to be 50.9% different from the actual one.

Therefore, utilizing a carefully curated selection of models, variogram analysis enables the clear determination of the distances at which analogous natural processes and idiosyncrasies in geochemical conditions manifest. These conditions pertain to the redistribution of chemical elements within the landscape. Forest plots demonstrate a decline in dispersion at distances of 140–180 m, which corresponds to the length of the slope (plot 1) and the distance between ravines with temporary watercourses at plot 2. The arable plots, on the other hand, demonstrate an increase in variance at short distances (70–80 m) and a decrease at longer distances (250 m at plot 3 and 370 m at plot 4). This indicates strong differences between adjacent points and similarity between distant ones, reflecting the disturbances in spatial structure caused by anthropogenic influence on acidity indices. Consequently, these experimental variograms delineate the anthropogenic influence on acidity heterogeneity, with the "natural component" of heterogeneity incorporated only at the conclusion. The experimental variograms for meadow (floodplain) sites (5 and 6) demonstrate a pronounced increase in variance at distances of 130–170 m, which corresponds to the width of the floodplain mane in the central region of the site. This variation highlights the substantial disparities between acidity levels directly in the flooded portion of the floodplain and those in areas of higher topography. The slight decrease in variance at medium distances (370–500 m) is attributable to the similarity of natural conditions at sampling locations at these distances (points near the terrace and the mane or the flooded areas separated by the mane).

The distribution of pH_{KCl} values in forest plots exhibits striking variations. Plot 1 is distinguished by the prevalence of medium acid soils (pH_{KCl} 4.5–5.0) with isolated areas of strongly acid soils (pH_{KCl} < 4.5) in the northwestern part of the study area, restricted to sod-podzolic weakly gleyey and gleyey soils. Conversely, the southeastern and eastern regions of the territory are predominantly characterized by acidic soils (pH_{KCl} 5.0–5.5), though there are areas exhibiting soils with reactions approaching neutral (pH_{KCl} 6.0–6.5) and slightly alkaline (pH_{KCl} 7.0–7.5), confined to a hill with soddy-carbonate soils (fig. 1.A.).

Site 2 is characterized by a wide variety of acidity groups, from strongly acidic (pH_{KCl} 4.00–-4.50) soils to medium alkaline (pH_{KCl} 7.50–8.00), which is explained by the cumulative heterogeneity of vegetation. In the northern part of the site, vegetation consists of an eagle-grass pine forest on sod-podzolic automorphic soils, while in the southern part, there is a mossy spruce forest. The presence of all acidity groups is observed at this site, indicative of cumulative heterogeneity in vegetation. The northern region is characterized by the presence of eagle-grass pine forest on sod-podzolic automorphic soils, while the southern region is dominated by mossy spruce forest. The latter is confined to sod-podzolic, weakly gleyey soils.

The presence of ravines with temporary watercourses contributes to the transfer of particles from adjacent arable land. The bedrock is hypothesized to be composed of carbonate moraine, which is susceptible to erosion, resulting in the removal of the upper soil horizons. The outcrop of carbonate rocks has led to the formation of moderately alkaline soils (pH > 7.50, up to 8.17) at Plot 2, which is atypical for sod-podzolic soils. Furthermore, the leaching of fertilizers from the adjacent arable land (Plot 4) likely contributed to localized alkalinization, further enhancing the acidity heterogeneity at this site. On plots of arable land, soil acidity has demonstrated a lesser degree of differentiation, both within individual fields and between fields. The majority of the area under both plots is characterized by soils with slightly acidic pH levels (pH_{KCl} 5.5–6.0) and those approaching neutrality (pH_{KCl} 6.0–6.5). Site 3 also contains an array of acidic soils in the northern part of the site, adjacent to the woodland and aeolian till. A contour of elevated pHKCI values (7.0–7.5) was identified in the center, as noted in the trend analysis. In general, the arable plots exhibit acidity levels consistently above 5.0, with the presence of micro-contours, alkaline soils (0.98 ha, constituting 9.17% of the plot area) and a contour of near-neutral soils (29.3 ha, accounting for 27.6%) that do not necessitate lime application. However, these micro-contours and near-neutral soils remain undetected during the determination of lime requirement, potentially leading to additional material costs and a decline in crop yield on microplots with pHKCI >6.

The topography of meadow lands is characterized by the presence of plots situated on the floodplain, in conjunction with forest areas. These regions are distinguished by a pronounced degree of differentiation in soil acidity. The topography of these regions is predominantly level, with the presence of areas characterized by acidic and medium-acidic soils, which are confined to floodplain ridges. In these areas, the processes of acidification are attributable to the predominance of moisture movement in a downward direction, resulting in weakened hydrological links with the river. Additionally, there are areas of poor aeolian sands. Other notable features include areas of weakly alkaline and neutral soils with close groundwater tables, as well as areas remote from the channel, where peat formation processes occur and medium- and strongly acidic soils are formed.

CONCLUSIONS

The complex spatial distribution of acidity in the forested areas near the "Western Berezina" GS is driven by vegetation heterogeneity (eagle-grass pine forest in the north and mossy spruce forest in the south) and topography. Ravines with temporary watercourses facilitate the transfer of particles from adjacent arable land, and hypothetical outcrops of carbonate moraine occur where the upper soil horizons have been eroded. In the meadow areas, acidic and moderately acidic soils are confined to floodplain ridges and areas distant from the river channel where peat formation occurs. In contrast, weakly alkaline and neutral soils are found in areas with a high groundwater table. The low spatial variation of acidity in arable land results from a homogenizing anthropogenic impact.

The spatial differentiation of soil acidity is more accurately assessed through geostatistical indicators that clarify the classical ideas about the formation of acid-alkaline geochemical barriers and transformation of natural landscapes under human influence. The lowest degree of soil acidity homogeneity is observed on forest lands with insignificant anthropogenic transformation and is characterized by a dispersion range from 0.90 to 1.30 (± 5%) and relative variation per meter of 1.73–2.14%; the increase of anthropogenic impact on meadow soils is characterized by the increase of anthropogenic impact on the soils. he human impact on the environment includes 40-50 years of liming, as well as plowing and applying basic nutrients at these sites and across the country. This leads to the flattening of natural geochemical processes and the prevention of soil acidity's natural heterogeneity.

In conjunction with geostatistical indicators, the presented cartograms demonstrate that conventional notions concerning the specific acidity group or range of pHKCI values attributed to particular soil types necessitate re-evaluation and elucidation through the utilization of contemporary geoinformation and geostatistical analysis methodologies. ■

REFERENCES

Biswas A. (2024). Geostatistics in Soil Science: A Comprehensive Review on Past, Present and Future Perspectives. Journal of the Indian Society of Soil Science, 72(1), 1–22. DOI: https://doi.org/10.5958/0972-0228.2024.00017.1

Demyanov V.V., Savelyeva-Trofimova A. (2010). Geostatistics: theory and practice. (in Russian)

Dmitriev E. A. (2009). Mathematical statistics in soil science. (in Russian)

Esmaeilizad A., Shokri R., Davatgar N. and Dolatabad H.K. (2024). Exploring the driving forces and digital mapping of soil biological properties in semi-arid regions Computers and Electronics in Agriculture, 220, 108831, DOI: https://doi.org/10.1016/j.compag.2024.108831

Kurlovich D. M. et al. (2024). Geographical station «Western Berezina»: manual: in 2 parts. Part 1. (in Russian)

Guo P.T. et al. (2015). Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach. Geoderma. 237–238, 49–59, DOI: https://doi.org/10.1016/j.geoderma.2014.08.009

Helfenstein A., Mulder V.L., Heuvelink G.B., and Okx J.P. (2022). Tier 4 maps of soil pH at 25 m resolution for the Netherlands. Geoderma, 410, 115659, DOI: https://doi.org/10.1016/j.geoderma.2021.115659

Heuvelink G.B., Webster R. (2022). Spatial statistics and soil mapping A blossoming partnership under pressure. Spatial statistics, 50, 12, DOI: https://doi.org/10.1016/j.spasta.2022.100639

Ibáñez J.J., Saldaña A. (2007) Dilemma of continuum in pedometrics. Geostatistics and geography of soils : monograph, 109-120. (in Russian)

Kanevskiy M.F., Demyanov V.V. (1999). Introduction to the methods of analyzing environmental data. Problems of environment and natural resources, Vol. 11. 2-12. (in Russian)

Klebanovich N.V., Vasilyuk G.V. (2003). Liming of Belarus soils. (in Russian)

Klebanovich N.V., Kindeev A.L., Sazonov A.A. (2021). Geostatistical analysis in mapping the spatial heterogeneity of soil moisture and acidity. Geospheric studies, 3, 80-91, DOI: https://doi.org/10.17223/25421379/20/6 (in Russian)

Klebanovich N.V., Prokopovich S.N., Kindeev A.L. (2018) Geospatial characterization of heterogeneity of soil properties. Land of Belarus, Vol. 2. 40-48. (in Russian)

Krasilnikov P.V., Targulyan V.O. (2019). On the way to the "new geography of soils": challenges and solutions (review). Soil Science, Vol. 2. 131-139. DOI: https://doi.org/10.1134/S0032180X19020096 (in Russian)

Lark R.M. (2012). Towards soil geostatistics. Spatial Statistics, 1, 92–99, DOI: https://doi.org/10.1016/j.spasta.2012.02.001

Loginov V.F., Kolyada V.V. (2015). Problems of increasing the adaptive capacity of the Republic of Belarus to climate change. Nature Management, Vol. 28. 5-13. (in Russian)

McBratney A.B., Minasny B., Stockmann U. (2018). Pedometrics.

Mohamed S.A., Metwaly M.M., Metwalli M.R., AbdelRahman M.A., and Badreldin N. (2023). Integrating active and passive remote sensing data for mapping soil salinity using machine learning and feature selection approaches in arid regions. Remote Sensing, 15, 7, 1751, DOI: https://doi.org/10.3390/rs1507175

Molchanov E.N., Savin I. Yu., Yakovlev A.S., Bulgakov D.S., and Makarov O.A. (2015). Domestic approaches to assessing the degree of soil and land degradation. Soil Science, Vol. 11. 1394-1394. (in Russian) DOI: https://doi.org/10.7868/S0032180X15110118

Oliver M.A. (2010). Geostatistical Applications for Precision Agriculture. DOI: https://doi.org/10.1007/978-90-481-9133-8

Pustylnik E.A. (1968). Statistical methods of analysis and processing of observations. (in Russian)

Samsonova V.P., Meshalkina Y.L. (2020). Frequently occurring inaccuracies and errors in the application of statistical methods in soil science. Bul. of the Soil Institute named after V. V. Dokuchaev, Vol. 102. 164-182, (in Russian) DOI: https://doi.org/10.19047/0136-1694-2020-102-164-182

Savin I.Yu., Zhogolev A.V., Prudnikova E.Yu. (2019). Modern trends and problems of soil cartography. Soil Science, Vol. 5. 517-528. (in Russian) DOI: https://doi.org/10.1134/S0032180X19050101

Soropa G. et al. (2021). Spatial variability and mapping of soil fertility status in a high-potential smallholder farming area under sub-humid conditions in Zimbabwe. SN Applied Sciences, 3, 19, DOI: https://doi.org/10.1007/s42452-021-04367-0

Suleymanov A.R. et al. (2024). Drivers of Soil Organic Carbon Spatial Distribution in the Southern Ural Mountains: A Machine Learning Approach // Eurasian Soil Science, 57, 11, 1942–1949, DOI: https://doi.org/10.1134/S1064229324602014

Suleymanov A.R et al. (2023). Spatial prediction of soil properties using random forest, k-nearest neighbors and cubist approaches in the foothills of the Ural Mountains, Russia. Modeling Earth Systems and Environment, 9, 3, 3461–3471, DOI: https://doi.org/10.1007/s40808-023-01723-4

Szatma´ri G., P´asztor L. (2019). Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms. Geoderma, 337, 1329–1340, DOI: https://doi.org/10.1016/j.geoderma.2018.09.008

Vaysse K., Lagacherie P. (2017). Using quantile regression forest to estimate un- 690 certainty of digital soil mapping products. Geoderma, 291, 55–64, DOI: https://doi.org/10.1016/j.geoderma.2016.12.017

Wadoux A.M.J.C., et al. (2021). Ten challenges for the future of pedometrics. Geoderma, 401, 11, DOI: https://doi.org/10.1016/j.geoderma.2021.115155

Webster R. (2007). Soil Science and Geostatistics. Geostatistics and Geography of Soils : Monograph, 8-18.

Xiao S., Ou M., Geng Y., and Zhou T. (2023). Mapping soil pH levels across Europe: An analysis of LUCAS topsoil data using random forest kriging (RFK). Soil Use and Management, 39, 2, 900–916, DOI: https://doi.org/10.1111/sum.12874

Yakushev V.P., Zhukovsky E.E., Kabanets A.L., Petrushin A.F., and Yakushev V.V. (2010). Variogram analysis of spatial heterogeneity of agricultural fields for the purposes of precision farming : method. manual, 47 p.

Appendix

Fig. A.1. Main soil types of the study sites: (a) – sod-podzolic (Retisols) (forest plots 1 and 2); (b) – soddy-carbonate (Rendzic calcaric) (forest plot 1); (c) – sod-podzolic arable (Anthrosols) (cropland plots 3 and 4); (d) – sod-podzolic on aeolian sands (aeolian ridges on the floodplain); (e) and (f) – soddy-gleyey (Gleyic Fluvisols) soils and alluvial soddy-gleyey (meadow plots 5 and 6)

Check for updates

cc BY

RESEARCH PAPER

COMPREHENSIVE ASSESSMENT OF LAKE GUSINOYE, REPUBLIC OF BURYATIA, BASED ON WATER, SUSPENDED PARTICULATE MATTER AND BOTTOM SEDIMENTS GEOCHEMISTRY

Lyudmila E. Efimova[1], **Natalia E. Kosheleva**[2], **Anna N. Lukyanova**[1], **Daria G. Sycheva**[2], **Vasilii A. Efimov**[1*]
[1]Department of Hydrology, Faculty of Geography, Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119991, Russian Federation
[2]Department of Landscape Geochemistry and Soil Geography, Faculty of Geography, Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119991, Russian Federation
*Corresponding author: Roxifixat@yandex.ru

Received: April 4th 2025 / Accepted: December 30th 2025 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-3979

ABSTRACT. Lake Gusinoye, the second largest lake in the Republic of Buryatia, is a key source of domestic and industrial water supply for the city of Gusinoozersk and nearby areas. It also serves as a recreational site and as a cooling reservoir for the coal-fired Gusinoozerskaya Thermal Power Plant. The lake is exposed to prolonged anthropogenic impact from adjacent pollution sources, including the power plant, overburden dumps of the Kholboldzhinsky coal mine, and a municipal solid waste landfill. This study aims to assess the current hydrochemical state of Lake Gusinoye based on a field survey conducted in July 2019. Sampling of water, suspended matter, and bottom sediments was carried out along three transverse and one longitudinal profiles covering the entire lake area, as well as in inflowing rivers and the plant's discharge channel. Parameters analyzed included physico-chemical characteristics, major ions, nutrients, heavy metals, and metalloids (HMMs). Aeration of the water column observed during the study period was caused by wind-driven mixing and upwelling, while oxygen depletion was noted near the wastewater discharge zone. Nutrient levels indicated mesotrophic conditions, approaching eutrophic. Wind forcing facilitated a more uniform distribution of water mineralization and major ions across the greater part of the lake's surface area. Increased water mineralization and sulfate, chloride, and sodium concentrations were confined to the northern part of the lake, where the main anthropogenic sources of readily soluble salts are located. Among HMMs, dissolved Zn, Sr, Mo, and Cu predominated, with maximum concentrations exceeding sanitary norms in the northeastern part of the lake, adjacent to pollution sources. Suspended matter was depleted in HMMs relative to regional lithospheric clarks, whereas bottom sediments showed higher concentrations, indicating continuous pollutant accumulation, especially in the eastern part of the lake.

KEYWORDS: cooling reservoir, major ions, heavy metals and metalloids, nutrients, Gusinoozerskaya TPP, wastewater

CITATION: Efimova L. E., Kosheleva N. E., Lukyanova A. N., Sycheva D. G., Efimov V. A. (2026). Comprehensive assessment of Lake Gusinoye, Republic of Buryatia, based on water, suspended particulate matter and bottom sediments geochemistry. Geography, Environment, Sustainability, 1 (19), 97-114
https://doi.org/10.24057/2071-9388-2026-3979

ACKNOWLEDGEMENTS: The study was conducted as part of the state assignment of the Department of Geochemistry of Landscapes and Soil Geography and the Department of Land Hydrology of the Geographical Faculty of Lomonosov Moscow State University. The authors are grateful to the staff of the Scientific Station "Lake Gusinoye" of the Baikal Institute of Nature Management of the SB RAS and its director, academician of RAS E. Zh. Garmaev, for their support of the field work.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Lake Gusinoye is the second largest water body in the Republic of Buryatia. It serves as an important source of domestic and industrial water supply, a recreational site, and a cooling reservoir for the coal-fired Gusinoozerskaya Thermal Power Plant (TPP), which operates on lignite extracted from local deposits. The lake's catchment area belongs to one of the most densely populated parts of the Selenga River basin (Ulzetueva et al. 2015). The Gusinoozersk industrial complex, one of the largest in Buryatia, has developed in the vicinity of Gusinoozersk and includes energy, processing, and transport enterprises. Persistent pollution sources for the lake include drainage water from unreclaimed overburden dumps and abandoned mines of the Kholboldzhinsky coal deposit, atmospheric emissions and wastewater discharges from the TPP, and municipal sewage from the city of Gusinoozersk and the village of Gusinoye Ozero. The Gusinoozerskaya TPP alone accounts for 86.65% of total surface water consumption in the Republic of Buryatia. In addition to significant thermal impact on the northern part of the lake, the TPP operation results in chemical pollution through stormwater runoff and discharges from ash disposal sites. In recent years, increasing recreational and tourist pressure on the lake and its watershed has also been reported (Babikov et al., 2018).

The diverse and intensive use of Lake Gusinoye has prompted studies of its chemical composition. Initial research dates back to the 1930s, coinciding with the development of local lignite deposits on the lake's northeastern shore. Scientific interest in the lake increased significantly following the launch of the Gusinoozerskaya TPP. Hydrological and hydrochemical data are available in works by Samarina & Khudyakova, 1969; Bogdanov, 1977; Obozhin et al. 1984; Adushinov et al., 1994; Domysheva et al., 1995. Over the past decade, environmental studies have intensified in response to growing anthropogenic pressure, as evidenced by numerous publications on various aspects of the lake's hydrology and hydrochemistry (Chebunina et al., 2016; Tsydypov et al., 2017; Khazheeva & Plyusnin, 2018; Zhigzhitzhapova et al., 2019; Dagurova et al., Lukyanova et al., 2020; Tsybekmitova et al., 2020; Bazarova & Kuklin 2021; Kosheleva et al., 2022). However, many of these studies have focused on limited areas of the lake and examined only a narrow range of hydrochemical indicators.

A full-lake hydrochemical survey, including bottom sediment sampling, was conducted in 2020 (Radnaeva et al., 2022; Bazarzhapov et al., 2023). This research allowed for the assessment of changes in water levels and physico-chemical characteristics of Lake Gusinoye from 1951 to 2021. Elevated concentrations of total dissolved solids, sulfates, sodium, fluoride, and oxidant-resistant organic fractions were recorded near wastewater discharge points. Seasonal peaks of iron and manganese concentrations were also observed. Nevertheless, a more comprehensive understanding of the lake's current condition requires integrated studies of all components of the aquatic landscape, including not only dissolved substances but also their presence in suspended matter and bottom sediments.

To address this need, the present study sets out the following objectives: to assess the physico-chemical parameters of lake waters during the summer period and the hydrological factors affecting their spatial distribution; to determine the concentrations and spatial variability of major ions; to evaluate contamination levels by heavy metals and metalloids (HMMs) in water, suspended matter, and bottom sediments; and to assess nutrient levels and the trophic status of the lake during the summer.

A field survey was conducted in summer 2019, during which physico-chemical parameters, major ion concentrations, nutrients, and HMMs (in dissolved and suspended forms, as well as in bottom sediments) were determined, allowing for a comprehensive assessment of the current environmental status of Lake Gusinoye.

STUDY OBJECT

Lake Gusinoye is located in the center of the Gusinoozersk intermountain basin. The reservoir stretches from northeast to southwest. The length of the lake is 24.8 km, the average width is 8 km, the average depth is 15 m, with a maximum depth of 26 m. The catchment area of the lake is 924 km^2, the area of the water surface is 164 km^2, and the ratio of the catchment area to water surface (5.7) characterizes the lake as a reservoir with a small specific watershed. The location of the lake in an intermountain basin determines the predominance of winds of the northeastern or southwestern directions, coinciding with the longitudinal axis of the lake.

The central pool of the reservoir, 22–24 m deep, has a rounded shape with a sharp drop in depth. At a distance of 100 m from the shore, the depth is 15–19 m. The maximum depth of the southern pool reaches 21 m, but the southern coast is relatively flat and has a beach. From the south, the Tsagan-Gol River flows into the lake. It is the largest tributary in terms of water discharge. The village of Gusinoye Ozero, with a population of about 2.5 thousand people, whose municipal wastewater is discharged into the Tsagan-Gol River, is located along the southwestern coast.

The northern pool of the lake is smaller in area and its maximum depths do not exceed 10 m. The northeastern shore is occupied by the Gusinoozerskaya TPP with a capacity of 1190 MWh and Gusinoozersk, a town with 25 thousand inhabitants. Domestic wastewater from the biological treatment plant in Gusinoozersk is discharged into the Zagustai River, which is the longest tributary (44 km) flowing into the lake near the TPP. To the west of the power plant, the Tobkhor River, which flows near the power plant ash dump cards, enters the lake.

The rivers that belong to the catchment of Lake Gusinoye flow down from the slopes of the Khambinsky ridge. The Tsagan-Gol River flows into the lake from the southwest, and the Zagustai and the Tobkhor rivers, as well as smaller rivers and numerous streams, flow into the lake from the northeast. The only river that flows out of the lake is the Bayan-Gol River, which flows into the Selenga, the largest tributary of Lake Baikal. The groundwater system of the study area belongs to the Gusinoozersk artesian basin of the Transbaikal type. The groundwater is formed mainly due to atmospheric precipitation and, partly, due to the infiltration of river water. The upper strata, composed of river alluvium, contain low-mineralized (0.1–0.3 mg/l) water of the Ca-HCO_3 type. The lower strata of the Lower Cretaceous age contain artesian water of the Na-SO_4 type with mineralization up to 1.2–3 g/l.

According to the ratio of water inflow to water volume (Kw < 0.5), Lake Gusinoye belongs to the reservoirs of the slow water exchange type. The water balance and morphometric characteristics of the lake contribute to the accumulation of the entering chemical substances and also define the importance of intra-reservoir processes in the transformation of these chemicals.

Lakes located near sources of anthropogenic pollution typically experience negative impacts on both their water bodies and catchment areas (Basha et al., 2010; Kara et al., 2014; Al Naggar et al., 2018; Ćipranić et al., 2019; Kosheleva et al., 2022; Zhao et al., 2022; Vithanage et al., 2022). In 2020, emissions from energy sector enterprises in Gusinoozersk amounted to approximately 45–50 thousand tons, accounting for 84% of the city's total atmospheric emissions (Bityukova et al., 2021). The main pollutants include ash particulates containing a wide range of HMMs, carbon and nitrogen oxides, sulfur dioxide, volatile organic compounds, and benzo[a]pyrene. A comparison of the actual and calculated ash composition (based on the coal's properties and ash content) showed that over half of the studied HMMs can condense onto aerosols and volatilize with the flue gases emitted by the thermal power plant (Sycheva & Kosheleva, 2023).

The intensity of airborne deposition of contaminated fine particulate matter over the lake depends on meteorological conditions – wind direction and speed, as well as the amount and duration of precipitation (Potemkin et al., 2011; Khazheeva & Plyusnin, 2016; Bortin et al., 2023). Pollution of the lake's water surface and its catchment occurs both during dry periods, due to wind erosion, and during snowmelt or rainfall events.

Atmospheric precipitation filtering through ash dumps contaminates surface and groundwater with suspended solids, petroleum products, and HMMs.

Lake Gusinoye is the source of water supply for the city of Gusinoozersk and the village of Gusinoye Ozero, as well as smaller settlements located on the coast of the reservoir. Wastewater discharge into the lake and diffuse runoff from the catchment area into the lake are a constant source of pollution. The treatment plants use physical, biological treatment and chlorination, but due to equipment deterioration and overload, the quality of treatment is low.

MATERIALS & METHODS

Physico-chemical parameters of water (temperature, electrical conductivity, dissolved oxygen content) were determined by probing the water column every 1 m with a YSI Pro30 thermoconductometer and a Pro YSI Pro ODO oximeter. The probing was carried out at 19 stations along 3 cross-sections and one longitudinal profile, in the discharge channel of the TPP, in the Bayan-Gol River flowing from the lake, and in the major lake tributaries (the Zagustai, Tobkhor, and Tsagan-Gol, in their upper and lower reaches) (Fig. 1).

Major ions were determined using the Kapel capillary electrophoresis system (Komarova & Kamentsev, 2006). Alkalinity was determined by the acidimetric method. Water samples for the analysis of the total and dissolved forms of nutrients (phosphorus, nitrogen, and silicon) were taken from three horizons (from surface horizon, below the thermocline, and from bottom horizon). The contents of nutrients were analyzed using the spectrophotometric method following standard procedures (Sapozhnikov et al., 2003). The calculation of the trophic index based on total phosphorus **TSI(TP)** it was carried out according to Eq. (1):

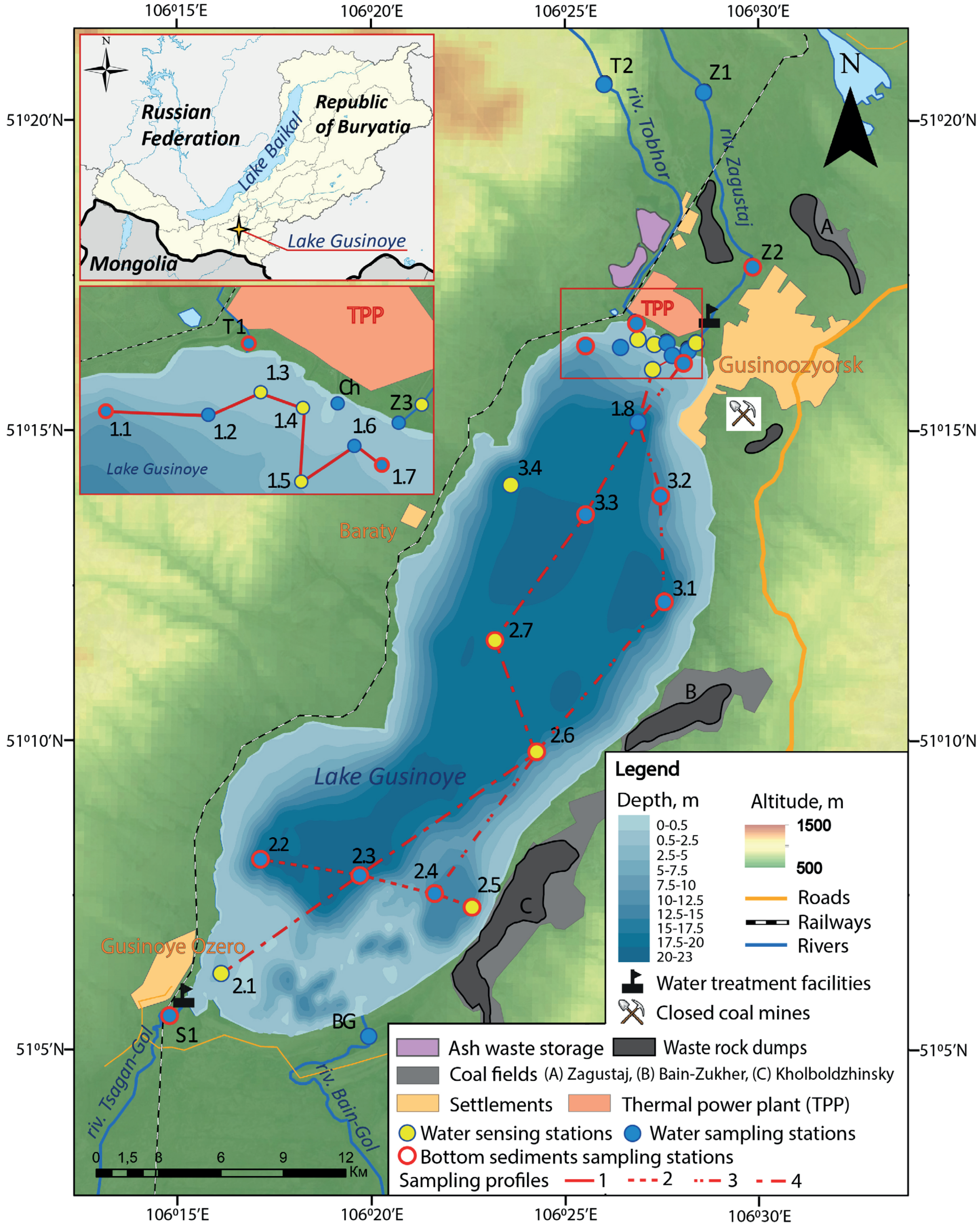


Fig. 1. Sampling stations on survey profiles in Lake Gusinoye, July 2019

$$TSI(TP) = 14.427In(TP) + 4.1504 \tag{1}$$

where *TP* is the total phosphorus content in the water, mg/m^3 (Carlson, 1977).

The total nitrogen trophic index *TSI(TN)* was calculated using Eq. (2):

$$TSI(TN) = 54.45 + 14.43 \times In(TN) \tag{2}$$

where *TN* is the total nitrogen content in the water, mg/m^3 (Kratzer & Brezonik, 1981).

Water samples for the determination of HMMs concentrations were taken at 10 stations, and the bottom sediments were sampled at 13 stations along the same profiles. Water samples were divided into liquid and solid phases: suspended particles were separated from the true solution using the Millipore vacuum filtration system with a membrane filter of 0.45 µm pore diameter. The filters were then dried to determine the content of suspension and the concentrations of suspended HMMs. Bulk concentrations of As, Cd, Pb, Zn, Co, Mo, Cu, Sb, Cr, V, Mn, Sr, and Ag in water, suspension, and bottom sediments were determined in the certified laboratory of the Institute of Microelectronics Technology Problems and Particularly Clean Materials of the Russian Academy of Sciences. The inductively coupled plasma mass spectrometry (ICP-MS) and inductively coupled plasma – atomic emission spectrometry (ICP-AES) methods were applied according to the certified protocols (NSAM 499-AES/MS, 2015) using the X-7c mass spectrometer (Thermo Elemental, the USA) and the iCAP-6500 atomic emission spectrometer (Thermo Scientific, the USA).

For the ecological and geochemical assessment of the trace element composition of bottom sediments, clarkes of concentration:

$$CC = Ci / K \tag{3}$$

and clarkes of dispersion:

$$DC = K / Ci \tag{4}$$

were calculated, where *Ci* is the concentration of a pollutant in the lake bottom sediments, mg/kg, and *K* is the clarke of the lithosphere (Rudnick & Gao, 2003). The coefficients of environmentally hazardous contamination of water, suspension, and bottom sediments by particular HMMs:

$$Kh = Ci / MPCi \tag{5}$$

were calculated, showing the excess of the concentration *Ci* of an element *i* over the corresponding standard. The standards for soils (HN 2.1.3684-21, 2021) were used, since the MPCs of pollutants are not yet developed for bottom sediments.

RESULTS & DISCUSSION

Distribution of physico-chemical parameters in the lake

Water temperature and hydrological structure of the lake. The thermal regime of a body of water largely depends on weather conditions. In July 2019, the weather in Buryatia was hot and dry, with an average daily temperature of 25–34°C. Rainfall began only in the third decade of July, and the temperature dropped to 17–20°C. The analysis of the water temperature distribution in Lake Gusinoye during the expedition showed that the temperature reached 22.9°C at the outlet of the thermal power plant discharge channel, decreased to 20°C at a distance of 300 m from the outlet (station 1.7), and was 18–19°C on the surface layer of the northern basin (stations 1.1–1.5). At the same time, the water temperature at the surface was 21–22°C in the central (stations 2.5, 2.6, 3.1–3.3) and southern basins (2.1–2.4). The warming effect of water discharge from the TPP is clearly visible in the surface layer up to 300–500 m in the coastal shallow zone (Fig. 2).

During the expedition, the water mass of most of the lake was mixed to depths of 10–12 m. At the same time, near the northeastern shores, the depth of the thermocline zone was only 1–3 m (stations 1.2–1.5). In the central pool (stations 2.6–3.4), the thermocline was located at the depth of 11–12 m; in the southern pool (stations 2.2–2.5) it was located at the depth of 13–14 m (Fig. 2), i.e., there was a significant decrease in the thermocline depth in the northern pool and an increase of its depth in the south.

The northeast wind caused a flow of warm water masses to the southern pool and upwelling of colder masses of the hypolimnion into the epilimnion. A similar phenomenon is typical for large lakes elongated in the direction of the prevailing winds (Moiseenko et al., 2002). The temperature of the epilimnion of Lake Gusinoye was 20–23°C. Thus, water of the same temperature occupied most of the lake's volume. The even water temperature down to fairly deep layers and the absence of a density gradient in this volume contribute to the mixing of the entire epilimnion with a steady wind.

A deepening of the thermocline up to 20 m in cooling reservoirs is observed in summer (Suzdaleva & Goryunova, 2014), since warm waters entering the lake from the canal have a lower density and spread in the surface layer of water, increasing the volume of the epilimnion. Deep water intakes of the Gusinoozerskaya TPP, located in the northeastern part of the lake, lead to a decrease in the volume of the hypolimnion, and the discharge of warm water to the surface leads to an increase in the thickness of the epilimnion. The even water temperature and the absence of a density gradient to a depth of 10–12 m contribute to mixing and uniform distribution of the components of the chemical composition of water.

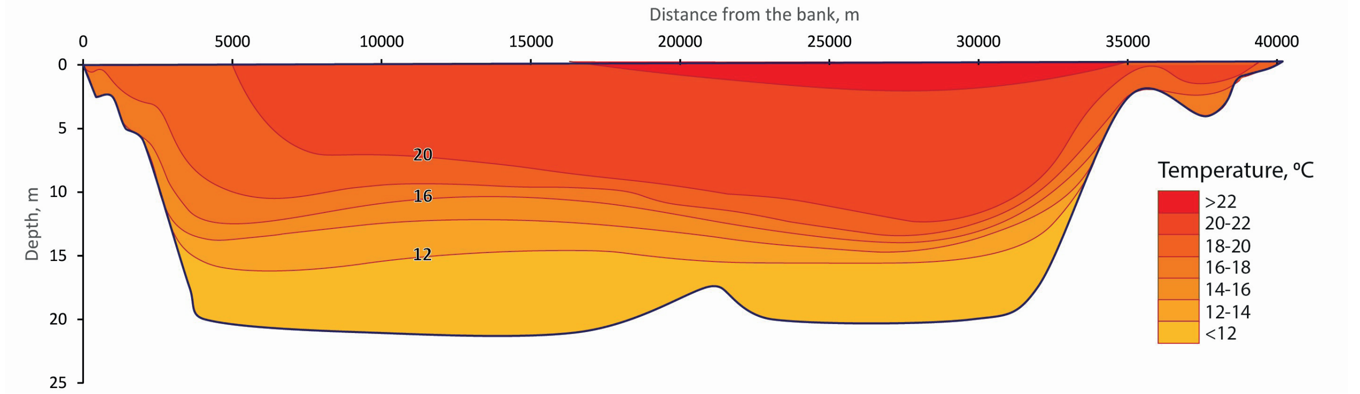


Fig. 2. Distribution of water temperature in Lake Gusinoye along profile 3 (stations 1.6, 1.5, 1.8, 3.3, 2.7, 2.6, 2.3, 2.1), July 25, 2019

In addition to the northeast wind, there are also winds from the south and southwest. Their long-term impact is manifested in the surge of water into the northern part of the lake. There is an effect of "locking" warm waste waters in the northern pool and overheating of the water column above the standards, i.e. above 28°C (Tsydypov et al., 2017).

According to field measurements conducted by scientists from the Baikal Institute of Nature Management of the SB RAS, the difference in water temperature between the discharge canal and the lake water in summer was 11–14 °C, and in winter 14–16°C. The vertical distribution of temperature in the discharge channel itself depends on the season of the year. The difference in water temperatures in the discharge canal near the power plant and at its mouth is minimal in summer (1.8–2.5°C), and maximal in winter (11.2–12.8°C). In winter, in the area where warm waters are discharged, a "polynya" (large air hole in ice cover) with an area of 0.04 to 0.9 km^2 remains unfrozen until February with a water temperature of 14°C; towards the edges of the "polynya" the temperature gradually decreases, but remains positive (Chebunina et al., 2016; Tsydypov et al., 2017). Reverse temperature stratification, typical of the winter period, does not form; in the "polynya" convection encompasses the entire volume of water to the bottom. Such a distribution was referred to as "winter homothermy" (Lapin et al., 2014).

Oxygen and pH. In the upper layer, mixed to a depth of 10–12 m, aeration reached 89–93% (7.9–8.4 mg/l); below the depth of 12 m, the saturation of water with oxygen decreased to 65–70%; in the bottom layer of the hypolimnion, it was about 60%. The lowest oxygen content (5.6 mg/l) was recorded at stations 2.4 and 2.5 in the near-bottom layers at depths of 15 m and 12 m.

The long period of high temperatures that preceded the observations contributed to the development of photosynthesis in the surface layer, an increase in pH and oxygen content, and the maintenance of oxidizing conditions in the epilimnion. Destruction processes in the near-bottom layer caused a decrease in oxygen content and a decrease in pH.

Near the mouth of the Tsagan-Gol River, the oxygen saturation level dropped below 62%, which was caused by the flow of wastewater from the treatment facilities of the village of Gusinoye Ozero. Pollution of the coastal zone in the southwestern part of the lake is enhanced by insufficiently treated municipal wastewater and diffuse runoff from landfills and private residential buildings, especially during the period of snow melting and rainfall. Insufficient treatment of wastewater containing large amounts of easily oxidized organic substances led to a decrease in the amount of oxygen that is consumed in oxidation processes. Note that as oxygen reserves in the bottom layers of the lake are depleted, at depths of more than 20 m, oxidizing conditions can change to reducing conditions, which, as shown in numerous studies (Zhao et al., 2017; Efimova et al., 2019; Grechushnikova et al., 2020), might cause a release of inorganic phosphorus, iron, and manganese from bottom sediments into the water.

The concentrations of dissolved oxygen in the zone of thermal impact from the Gusinoozerskaya TPP in all hydrological seasons were higher than the background values. In the warm period of the year, this may be due to the increased rate of photosynthesis, and in winter, during the period of freeze-up, constant aeration is provided by the "polynya" in the area of warm water discharge (Tsydypov et al., 2017).

The water of cooling ponds located in a zone of moisture deficit is sometimes subjected to alkalinization. One of the most important reasons for this process is the technogenic input of nutrients and easily oxidized organic matter (Shavrak et al., 2012). In the water of Lake Gusinoye, despite the slow water exchange, alkalinization was not observed owing to the strong wind and wave effects on the surface layers of the reservoir, preventing the concentration of organic compounds. The pH values in the lake water during the research period ranged from 7.6 to 8.45, decreasing from the surface to the bottom. The patterns of the vertical pH distribution are similar in different parts of the reservoir, which is due to the similarity of biochemical reactions and water exchange processes in the lake.

Mineralization and content of the main ions of the salt composition

According to the classification of (Alekin, 1970), the waters of Lake Gusinoye and its tributaries are characterized by average mineralization; during the survey period, water mineralization was 310–355 mg/l with an electrolytic conductivity of 433–451 μS/cm. The parameters varied across the water area and in the water column of the lake. In the surface layer of water (up to 1 m), the highest mineralization was observed in the northern pool (Fig. S1), which coincided with the temperature distribution. Near the northeastern shore of the lake, insignificant depth of the thermocline and wind-induced negative surge were observed, which might contribute to the entry of more mineralized cold waters of the hypolimnion into the epilimnion. The mineralization of water in the bottom layers of the lake was higher than in the surface layers, increasing with depth. The highest mineralization of water was recorded not at the deepest stations, but in the bottom layers of the reservoir near the settlements – Gusinoozersk (station 1.8) and Gusinoye Ozero (station 2.1).

Hydrocarbonates dominated among the anions; the cationic composition was mixed, with a slight predominance of sodium. The content of major ions in the water column of the lake is distributed fairly evenly with depth. Despite the small variations, the spatial distribution of dissolved substances reflects the hydrological processes occurring in the lake (Table S1). The higher concentrations of dissolved substances were found in the northern part of the lake (stations 1.1–1.8), where technogenic sources of easily soluble salts are concentrated. There, increased mineralization and higher levels of sulfates and chlorides were registered, which may be related to wastewater discharges from Gusinoozersk treatment facilities and groundwater discharge from closed pits and mines.

A specific feature of the chemical composition of the water of cooling reservoirs located in a zone of moisture deficit is a gradual increase in the mineralization of water, derived not only from evaporation processes but also from the changes in the ratios between the major ions caused by the inflow of highly mineralized waters in the bottom layers. The chemical composition of the water of Lake Gusinoye is formed in the process of interaction in the "water–rock" system, therefore it reflects the geochemical specialization of the host rocks of underground aquifers. Due to the arid climate, the contribution of surface runoff and precipitation is relatively small compared to underground runoff into the lake, which is supplied by the discharge of confined aquifers associated with coal seams and groundwater under the overburden rocks of the Kholboldzhinsky coal mine (Zhambalova et al., 2020).

The increased content of sulfates in precipitation falling on the lake surface in winter is usually associated not only with emissions from the Gusinoozerskaya Power Plant but also with the burning of coal and oil products in residential areas located along the shores of the reservoir. After snowmelt, sulfates enter the lake, increasing the mineralization of its surface layers. Despite the remediation, Kholboldzhinsky mine dumps undergo weathering and generate surface runoff. The influx of meltwater from the mine causes a rise in mineralization in the surface layers of the lake along the coastline. For example, during the flood period in 2015, the maximum water mineralization (365 mg/l) was observed in the zone influenced by meltwater runoff derived from coal mine dumps; at other stations, the mineralization varied in the range of 217–278 mg/l (Khazheeva & Plyusnin, 2018).

A local increase in sulfate content was found in the area affected by the TPP. During the expedition, an increase in diffuse runoff into the lake was observed due to precipitation, which led to a decrease in sulfate content in the lake’s surface layer near station 3.3. A local maximum in sulfate and sodium (Fig. 3) content was observed in the southwestern part of the lake, where a railway line, industrial facilities, fuel and lubricant stores, and various waste dumps are located close to the lake, from which these ions enter the lake through diffuse runoff. Near station 3.2, increased concentrations of the major cations Na, Ca, and K were observed in both the surface and bottom layers; the increase in ion concentrations was likely due to the influx of mineralized water from deep underground horizons along faults and fractures. The tributaries flowing into Lake Gusinoye form a mixing zone of river and lake waters, where, as a result of dynamic and convective mixing, the content of the major ions equalizes to the values characteristic of lake waters. In the southern part of the lake, the influence of the Tsagan-Gol River inflow is detectable (Table 1, Fig. 3).

Water in the Tsagan-Gol River is much less (5 times) mineralized than the lake water; therefore, when it flows

Table 1. Concentrations of the major ions (mg/l) in waters of the tributaries of Lake Gusinoye

Major ions	Tsagan-Gol mouth	Zagustay upper reaches	Zagustay mouth	Tobkhor upper	Tobkhor mouth	Bayan-Gol, source from the lake	Discharge from the TPP
hydrocarbonates	45.9	47.6	202	reaches	140.3	170.2	180.6
sulfates	15.9	13	66.7	28.1	54.1	47.7	52.2
chlorides	3.4	0.6	7.9	11.6	7.1	7.8	8.6
calcium	12.4	13.6	32.6	0.3	31.8	30.3	31.3
magnesium	2.8	2.3	15.8	8.2	12.3	14.5	14.6
sodium	3.1	3.7	44.5	2.3	37.5	45.8	44.1
potassium	1.2	0.3	2.4	4.6	3	3.9	3.7
mineralization	84.6	81.1	371	0.6	286	320	335
discharge, m^3/sec	2.87		1.09	55.7		0.113	

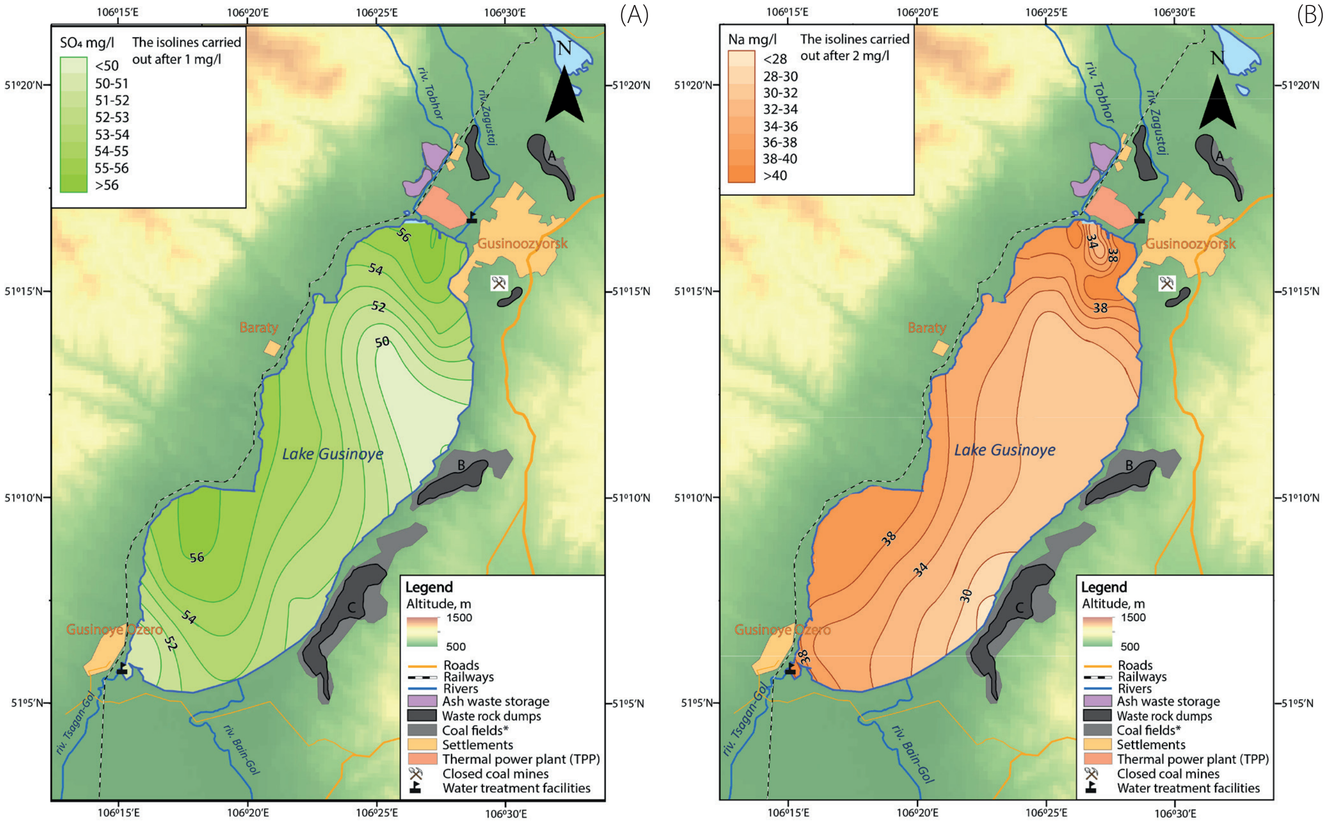


Fig. 3. Distribution of sulfates (A) and sodium (B) (mg/l) in the surface layer of Lake Gusinoye, July 2019

into a reservoir, it spreads for some time in the surface layer. Being colder and denser, the water of the Tsagan-Gol River descends below the epilimnion and spreads into the bottom layer, where it mixes with lake water.

The Zagustay and Tobkhor rivers flow into Lake Gusinoye in its northern part. The water in the Tobkhor River is less mineralized than the water of the Zagustay River (Table 1), which may be caused by the shallower depth of the channel incision and the volume of groundwater entering its bed. In the downstream direction, the mineralization and the content of major ions in the water of the Tobkhor River increase almost 5 times. Water mineralization of the Zagustay River also increases 4.5 times. The two rivers are characterized by a change in the ratio between calcium and sodium ions: the proportion of the latter increases during the low-water period (Fig. 4).

In addition, in the middle reaches of the Zagustay River, the channel is partially blocked by overburden dumps from the Zagustay coal mine, from which sulfates, chlorides, and sodium migrate into the river, increasing the concentration of these ions towards the river mouth. The Tobkhor River in its lower reaches goes around the cards of the ash and slag dumps of the Gusinoozerskaya TPP, from which drainage waters of sludge pulp containing high concentrations of easily soluble salts enter the river through a collector channel.

The flow of water from the main tributaries into Lake Gusinoye creates local zones of relatively low (the Tsagan-Gol River) or relatively high (the Zagustay River) water mineralization and major ion concentrations. The size of the mixing zone of river and lake waters depends on seasonal fluctuations in water flow.

In the 1980s and 1990s, changes in water mineralization and the ratio of major ions were caused by the development and subsequent closure of the Kholboldzhinsky open-pit mine and coal mine, as well as the opening of the Gusinoozerskaya TPP and the impact of its wastewater on the lake (Bogdanov, 1977; Obozhin et al.,1984; Adushinov et al., 1994). From the 1990s to the present, no noticeable increase in water mineralization has been registered (Radnaeva et al., 2022). However, an increase in the content of sulfates and chlorides was noted, which led to a change in the ratio between the major anions (Khazheeva & Plyusnin, 2018; Lukyanova et al., 2020). The relative content of sulfates increased from 6.5 to almost 13%-eq, and the relative content of chlorides rose from 1.7 to 3%-eq (Fig. S2). This finding is confirmed by the studies of Radnaeva et al. (2022), who noted a gradual increase in sulfate and sodium concentrations in the lake water, which may be caused by rising groundwater levels. Thus, the impact of multiple anthropogenic factors on the ionic composition and content of highly soluble salts has so far manifested only in certain parts of the lake and has not affected the entire water body.

Heavy metals and metalloids

Dissolved forms of heavy metals and metalloids. Among the dissolved forms of Mn, Cu, Zn, Sr, Mo, Pb, and As analyzed in the water of Lake Gusinoye, the highest concentrations were shown by Sr: the concentrations varied from 958 to 991 µg/l, increasing with depth. The concentrations of Mo in the surface horizon varied from 14.9 to 16.0 µg/l and in most of the lake's water area they increased with depth, reaching a maximum value of 31.1 µg/l in the eastern part of Lake Gusinoye, at station 3.1 (Fig. 5). The exception was station 1.2 located in the impact zone of the Tobkhor River, where the highest Mo content (16.0 µg/l) was registered in the surface horizon. The concentrations of dissolved Zn were in the range of 0.73–36.1 µg/l, they varied quite strongly throughout the lake (Fig. 5).

Over the major part of the reservoir, the concentration of Zn decreased with depth. The maximum value in the surface horizon (3.4 µg/l) was registered near the Gusinoozerskaya TPP, at station 1.1, and in the bottom horizon (36.1 µg/l), at station 2.4, in the southeastern part of the lake due to discharge of groundwaters draining overburden dumps of the Kholboldzhinsky coal mine (Table S2). The Mn concentration varied in the range of 0.24–1.7 µg/l, the maximum value of 1.7 µg/l was found in the northern part of the water area, at the confluence of the Zagustay River. In the northern part of the lake, the Mn concentration decreased with increasing depth; in the southern part of the lake, on the contrary, the Mn concentration, as a rule, increased with depth: at stations 2.4, 3.1, the concentrations of this element at a depth of 21 m were 3 times as much as in the surface horizon.

The content of dissolved Cu varied between 0.87–1.5 µg/l with a maximum value of 1.5 µg/L in the surface layer at station 1.6 in the northern part of the lake. The lowest concentrations among the studied HMMs were observed for As and Pb. The concentrations of As displayed very low variability within the water area and little change with respect to depth; As concentrations were in the range of 1.0–1.2 µg/l. The Pb concentrations varied from 0.03 µg/l in the surface layer at station 3.3 to a maximum of 0.18 µg/l

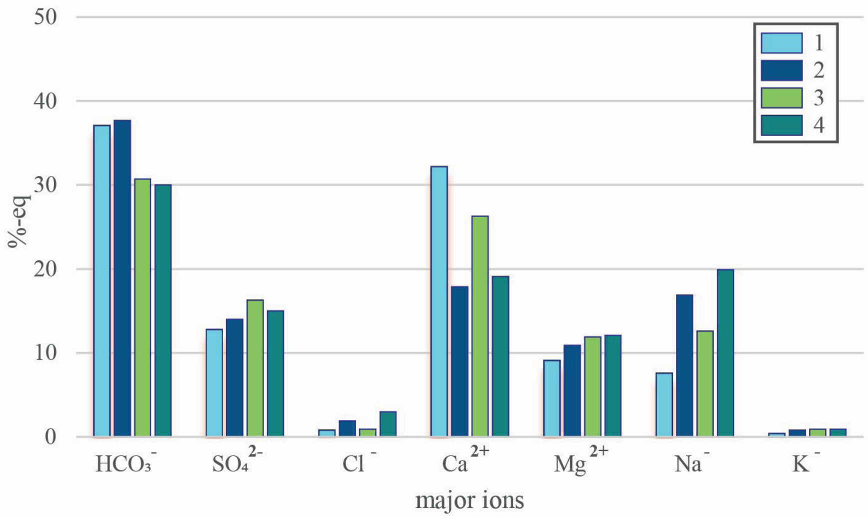


Fig. 4. Relative content of the major ions (%-eq) in the water of the Zagustay River in its upper reaches (1) and at the mouth (2), and in the water of the Tobkhor River in its upper reaches (3) and at the mouth (4)

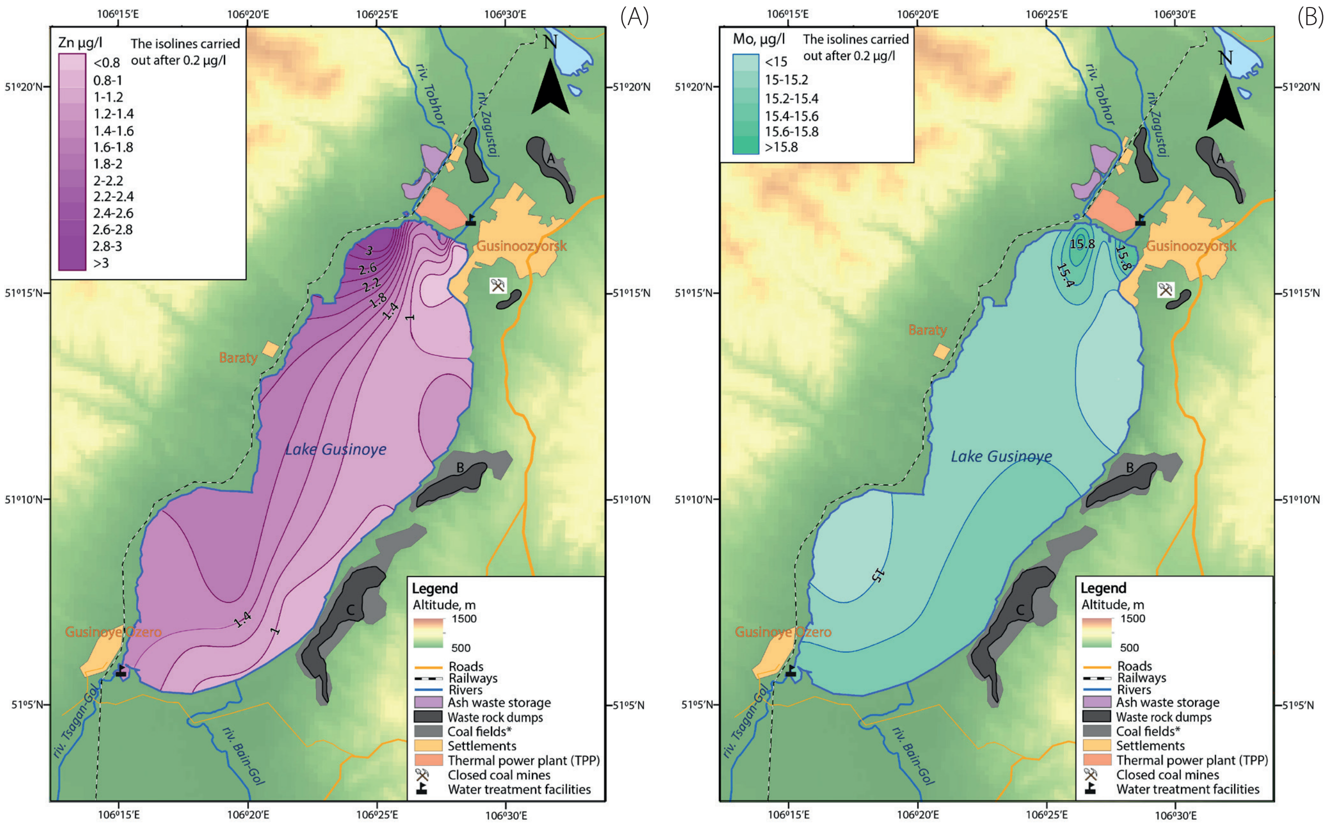


Fig. 5. Distribution of Mo (A) and Zn (B) (μg/l) in the surface layer of Lake Gusinoye, July 2019

at station 1.2 in the northern part of the lake at a depth of 5 m, which are related to the influence of the tributaries and discharges from the TPP.

The comparison of the results of the present study with the findings of the research conducted in 2021 (Bazarzhapov et al., 2023) which focused on seasonal changes of chemical composition of water and bottom sediments of Lake Gusinoye, showed that in the summer of 2021, the average concentrations of dissolved forms of Pb and Mn were higher, and the concentrations of Zn and Cu, on the contrary, were lower. The Mn concentrations in the lake in 2021 were 2.5 times as much as the concentrations observed two years earlier. Since 2019 they increased from 0.81 to 1.9 μg/l. The increase in Pb concentrations (from 0.08 to 0.11 μg/l) was less distinct. The concentration of Cu in 2021 was much lower compared to the previous period (the difference with 2019 observations was almost 2 times); they dropped from 1.09 to 0.58 μg/l, and Zn decreased from 3.07 to 2.5 μg/l (the difference was 1.2 times). At the same time, the results of the research of (Bazarzhapov et al., 2023) showed that the summer period is characterized by rather low concentrations of dissolved HMMs compared to other seasons of the year (autumn, spring, and winter).

Suspended forms of heavy metals and metalloids. In the suspended sediments, all the studied elements – Mn, Cu, Zn, Sr, Mo, Pb – were below the clarke values estimated for the upper part of the Earth's crust. The highest concentrations were displayed by Mn; its concentration varied in a wide range – from 24.4 to 395 μg/g. In general, the northern part of Lake Gusinoye is characterized by higher levels of Mn: in the area where the Tobkhor River flows into the lake at station 1.6, its content reached 212 μg/g. At this station, Cu and Sr also showed the maximum concentrations. The Mn content in suspended sediments increased with the depth of the reservoir and degree of dispersion of bottom sediments, with the exception of station 3.2, where the concentration of Mn in the surface layer was 395 μg/g, and at a depth of 21 m it decreased to 43.4 μg/g. A similar vertical distribution for station 3.2 was found for Cu (surface-bottom concentrations, 4,0–1.1 μg/g), Zn (8.0–2.5 μg/g), Sr (15.3–7.7 μg/g) and Mo (0.63–0.19 μg/g). These patterns in distribution may be associated with the influence of surface runoff generated from the waste landfill and the abandoned coal mine, whose locations are the closest to station 3.2. The increase in the influx of polluted water into the lake was driven by heavy rainfall in June-July 2019, when 117 mm of rain was recorded over 22 days.

The concentrations of Sr in the suspended matter varied from 7.4 to 17.9 μg/g; the highest concentration was detected in the area influenced by the discharge zone of the power plant, at station 1.6. The concentration of Zn varied between 3.2–11.7 μg/g and, as a rule, increased with depth, with the exception of stations 1.2 and 3.2. The highest concentrations of Zn 10.5–11.7 μg/g were found in the bottom layer in the northern part of the lake. The content of Pb in the suspended matter was 1.1–11.4 μg/g, with the highest concentrations observed along profile 1, where the Tobkhor and Zagustay rivers reach the lake, and at the discharge zone of the Gusinoozerskaya Power Plant. Over most of the water area, the Pb concentrations showed slight variations in the range of 1.1–3.2 μg/g. The Cu concentrations varied in the range of 1.0–7.4 μg/g. The patterns of Cu distribution have not been identified either with respect to the depth or to the water area. The lowest concentrations in suspended matter were recorded for Mo (0.17–0.75 μg/g); the concentrations of this element increased with depth throughout the entire water area, except for station 3.2.

Heavy metals and metalloids in bottom sediments. The bottom sediments of Lake Gusinoye were analyzed to determine Ag, Mo, As, Cu, Sr, Zn, Sb, Pb, V, Co, and Cr concentrations. Compared to the clarke values for the upper part of the earth's crust, bottom sediments are enriched in $Mo_{9.2}Ag_{3.8}Sb_{2.5}Sr_{1.5}Zn_{1.5}Pb_{1.4}As_{1.3}Cu_{1.1}$ (lower index is CC value), and depleted in $Cr^{3.1}Co^{1.7}V^{1.3}$ (upper index is *DC* value). The maximum relative enrichment was observed for Ag at station 2.6 (*CC* = 26) in the eastern part of Lake Gusinoye, which might be due to the influence of overburden lignite dumps located on the shore.

In general, the northern part of the lake showed the lowest concentrations of almost all studied HMMs, which can be explained by the high speed of water delivery through the discharge channel of the power plant and the deposition of elements far from the discharge site. The exceptions are As and Sb, for which the highest concentrations, reaching 12 mg/kg and 1.3 mg/kg, respectively, were found at station 1.1, in the northwestern part of the water area.

The highest concentrations in bottom sediments of Lake Gusinoye were displayed by Sr (222–772 mg/kg), Zn (39–117 mg/kg), Cr (4.6–44 mg/kg), Cu (3.5–46 mg/kg), V (21–96 mg/kg) and Pb (15–29 mg/kg). Mo and As concentrations were in the range of 2.0–14 mg/kg, and the lowest concentrations were found for Ag (0.07–1.4 mg/kg) and Sb (0.41–1.3 mg/kg). The maximum Sr concentration, equal to 772 mg/kg, was detected in the eastern part of the lake, where the Kholboldzhinsky coal mine is located, and decreased towards the western part of the reservoir to 370 mg/kg. The highest concentrations of Ag, Mo, Pb, and Cr were found at station 2.6, in the eastern part of the lake. The bottom sediments in the central part of the lake, at stations 2.3 and 2.7, displayed the maximum concentrations of Zn, V, and Co. The highest concentrations of Cu were determined in the northeastern part of Lake Gusinoye, at station 3.2.

The analysis of the spatial distribution of the HMMs in bottom sediments revealed the accumulation of pollutants in numerous depressions of the bottom, while in shallow water the accumulation of HMMs was less obvious. Bottom sediments in the deep-water areas of the lake have a greater capacity for the accumulation of HMMs, since they are represented by black silts with a high proportion (up to 51%) of fine particles (less than 10 μm in diameter). These sediments contain a relatively high amount of organic matter (5.1–6.9%) in contrast to sandy or sandy loam sediments of shallow areas with a lower content of organic matter (0.9–4.3%).

The comparison of the results obtained for bottom sediments by (Bazarzhapov et al., 2023) revealed that the average content of Pb, Mn, Cu, and Zn in sediments in July 2021 was significantly lower than in 2019: the concentration of Pb decreased by 3.4 times, Cu by 1.8 times, and Zn and Cr by 1.4–1.5 times. No significant seasonal changes in HMMs concentrations in bottom sediments were reported.

The study of the HMMs concentrations in bottom sediments of Lake Gusinoye and its tributaries conducted by (Dampilova et al., 2022) showed significant discrepancies with our results, which might indicate high geochemical heterogeneity of the lake and river sediments. In the eastern part of the lake, at the location where wastewater enters the lake from the overburden dumps of the Kholboldzhinsky coal mine, the concentrations of Cu, Zn, and Cr in the bottom sediments were reported to be higher (2.9, 2.8, and 2.3 times, respectively) than the concentrations established in the present research. At the same time, the concentrations of Pb turned out to be twice lower. A comparison of the trace element composition of the bottom sediments in the central part of Lake Gusinoye showed that, according to the results of our study, the concentration of Pb was 7.2 times, Cu 1.4 times, and Zn 1.6 times as high as the concentrations reported by (Dampilova et al., 2022). The only element that showed comparable concentrations in the bottom sediments was Cr.

Environmental hazard of HMMs pollution in Lake Gusinoye. In the water of the lake, an excess of the dissolved forms of Cu, Zn, Sr, and Mo over the sanitary standards (maximum permissible concentrations) was revealed (MPC, 2016). The greatest excess over the standards was observed for Mo and Sr; their concentrations were higher than the MPC values in 100% of the analyzed water samples. For Mo, the environmental hazard coefficient *Kh* varied from 15 to 31, for Sr, it varied from 2.4 to 2.5. For these elements, the maximum excesses over the MPC (31 and 2.5 times, respectively) were found in the eastern part of the lake, at station 3.1, in its bottom layer. For Cu, the excess over the MPC was detected in 52% of the examined water samples; the maximum value of *Kh* = 1.5 was recorded in the northern part of the lake, in the zone influenced by the Zagustay River. The concentration of dissolved Zn exceeds the MPC only in the bottom layer (*Kh* = 3.6) of the lake, at station 2.4 located in its eastern part, which is affected by the Kholboldzhinsky coal deposit.

Nutrients

During the field research, silicon concentration in the epilimnion of Lake Gusinoye varied between 1.2–1.7 mg/l, and in the water of near-bottom layers, it was nearly twice as high. As a rule, a decrease in silicon concentrations in the surface layers of dimictic water bodies occurs from May to August as a result of the great silicon requirements of diatoms, which dominate the species composition of phytoplankton at the beginning of the growing season (Xiao et al., 2019; Sharapova et al., 2020). The waters of the lake's tributaries contained silicon in concentrations of 4–7 mg/L due to the predominance of their underground feeding. The silicon content increased from the upper reaches to the mouths of the lake's tributaries. The influence of the tributaries is clearly evident in the surface layers of the lake, where they form areas of elevated silicon concentrations (Fig. 6A).

In summer, almost all the nitrogen and phosphorus were present in organic form, and the content of inorganic phosphorus and nitrate nitrogen was close to analytical zero (Table S3). Despite the upwelling caused by the influence of constantly blowing winds, a summer thermocline formed in the lake, which prevented complete mixing of the water column and hampered the supply of nutrients from the bottom layers. The predominance of organic forms of nitrogen and phosphorus was recorded not only in the surface but also in the bottom horizons of the lake.

They probably created competition for phytoplankton for nutrients (inorganic forms of phosphorus and nitrogen) in the substrate. The content of ammonium was noticeably higher in the area where warm waters were discharged from the TPP. The increased water temperature contributed to more active mineralization of macrophytes and dead planktonic organisms and easily oxidized organic substances that entered the lake with the water of the Zagustay River and with wastewater from the treatment facilities of Gusinoozersk (Fig. 6A).

The distributions of phosphates and nitrates along profile 3 were rather similar (S3). The enrichment with phosphates was observed in the near-bottom layers of the lake, at depths of 15 m or more. The replenishment of the epilimnion with nutrients in the summer could occur due to the influx from the tributaries or with diffuse surface runoff from the catchment area (Fig. 6B, S4).

The patterns of the distributions showed that in the area of the power plant impact there is a zone of local enrichment in nutrients, but a much larger amount of these elements comes with the Tsagan-Gol River discharge in the southern pool (Figs. 6, 7, S4).

The distribution of organic nitrogen corresponded to the distribution of temperature: the highest contents of nitrogen were observed in places with warm water in the

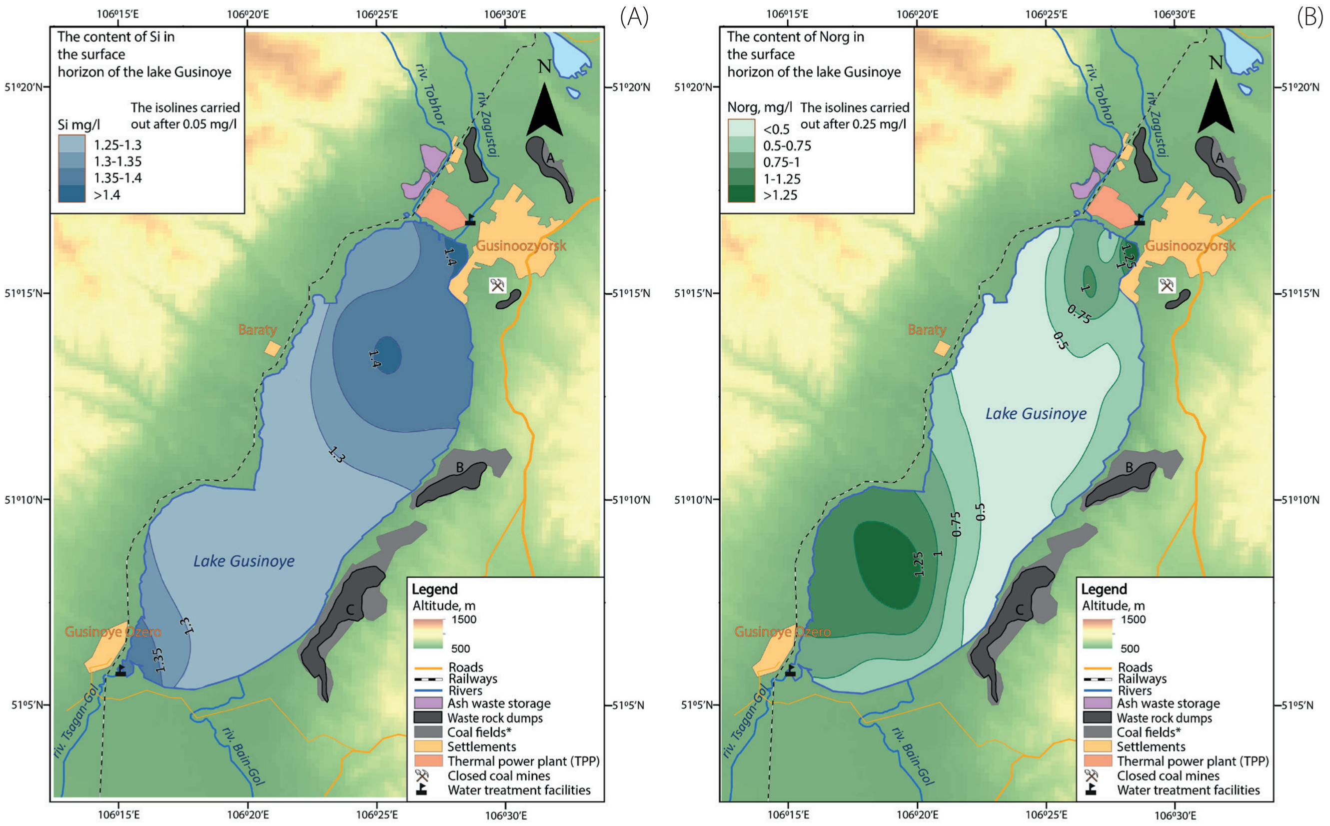


Fig. 6. Distribution of Si (A) and Norg (B) in the surface layer of Lake Gusinoye, July 2019July 2019

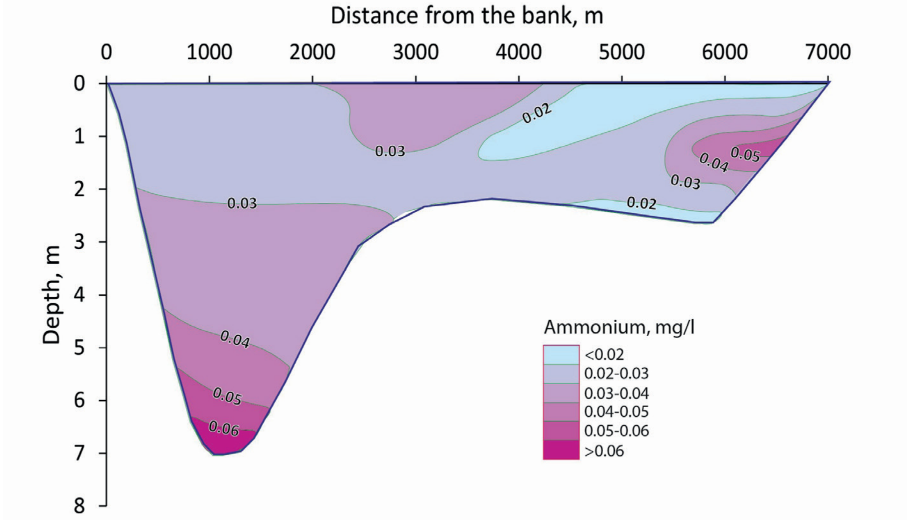


Fig. 7. Distribution of ammonium (mg/l) in Lake Gusinoye in August 2019 along profile 1 (stations 1.1–1.7)

center of the southern pool and in the northern pool in the area of discharge of warm water from the power plant, as well as in the impact zone of the Tsagan-Gol River (Fig. 7). The utilization of mineral nitrogen and phosphorus compounds was also facilitated by the presence of extensive macrophyte thickets in the lake's littoral zone, as confirmed by the distribution of organic nitrogen along the eastern shore of the reservoir. Furthermore, the most important sources of organic compounds of nitrogen and phosphorus are wastewater from the Gusinoozerskaya TPP, municipal wastewater from Gusinoozersk and the village of Gusinoye Ozero, and diffuse runoff from the southwestern shore of the lake. The similar distribution of these elements on the reservoir's surface (Fig. 3, 6 B) indicates similar sources of nitrogen, phosphorus, sulfates, and sodium.

The contribution of phosphorus and nitrogen to the production capacity of aquatic ecosystems has been assessed in many studies (Soranno et al., 1997; Wilhelm & Adrian, 2008; Adamovich et al., 2016; Erina et al., 2020; Walumona et al., 2021; Sapelko et al., 2025). For this purpose, the ratio of total nitrogen to total phosphorus (TN:TP) is often used. At a value <10, the development of algae is limited by nitrogen, and at a value >17, the growth of algae is limited by phosphorus (Smith, 1982). The TN:TP ratio observed in the epilimnion of Lake Gusinoye varied between 24–64, which indicates that phosphorus was the limiting element. The highest values of this ratio were typical of the epilimnion of the reservoir at stations 1.7 and 1.8. These stations are under the impact of Gusinoozersk wastewater and showed increased levels of organic nitrogen, accounting for 86–99% of its total concentrations (Fig. 7).

The assessment of the nutrient limitation using the ratio of inorganic substances that are most accessible to microorganisms – nitrates and phosphates – indicates a more pronounced limitation by inorganic nitrogen. The NO_3:PO_4 ratio in the northern pool ranged from 0.3 to 1, which may imply a deficit of inorganic nitrogen, which is necessary for optimal phytoplankton productivity. The highest ratio of nitrates to phosphates (3–7) was observed in the surface layers of the reservoir at stations 3.1–3.3, where the influence of leachates from the coal dumps located on the eastern shore is possible.

There are a lot of lakes where the organic matter is created not as a result of the photosynthetic activity of phytoplankton but mainly as a result of the activity of submerged macrophytes (Pokrovskaya et al., 1983; Sapelko et al., 2025). In the coastal areas of the shallow northern part of Lake Gusinoye, we observed thickets of underwater macrophytes. The increase in the overgrowth of Lake Gusinoye indicates that the ecosystem is developing according to the macrophyte type and, despite the anthropogenic load, is still capable of maintaining the overall level of productivity (Bazarova & Kuklin, 2021; Bazarova, 2024).

The trophic status of a lake is the most important hydro-ecological characteristic determined by the physico-chemical and biological processes occurring in the aquatic system (Wetzel, 2001). The trophic status of Lake Gusinoye was characterized by Trophic State Indices (TSI), calculated using the contents of total phosphorus TSI(TP) (Carlson, 1977) and total nitrogen TSI(TN) (Kratzer & Brezonik, 1981). The range of values that limit bioproductivity in mesotrophic water bodies is 40–50 (Carlson, 1977). The calculations performed indicate that in the summer of 2019, Lake Gusinoye can be classified as a mesotrophic reservoir because index values were TSI(TP) = 49.4 and TSI(TN) = 50. Such values indicate that the lake lies on the boundary between mesotrophic and eutrophic water bodies and also imply its probable transition to a eutrophic state. Close values of the TSI(TP) and TSI(TN) indices, as well as the ratio of nitrogen and phosphorus in the epilimnion of the lake, indicate that co-limitation of nitrogen and phosphorus is observed, which is typical of water bodies of the temperate zone.

CONCLUSIONS

During the open-water period, the distribution of physico-chemical characteristics in the surface layers of the lake is shaped by wind influence, which generates drift currents and causes wind-driven downwelling, leading to the formation of upwelling zones. Wind activity promotes a more uniform distribution of water mineralization and major ions across most of the lake's area. Elevated levels of water mineralization and major ion concentrations are confined to the northern part of the lake, where the main anthropogenic sources of highly soluble salts are located. Sulfates, chlorides, and sodium enter the lake with wastewater discharges from Gusinoozersk and the village of Gusinoye Ozero, with groundwater inflow from closed coal pits and mines, with diffuse runoff from ash dumps located in the channel of the Zagustay River, and from the territory of the Gusinoozerskaya TPP.

During the study period, pollution of the lake with HMMs was identified. Among the dissolved forms of HMMs in the lake water, Sr, Mo, Cu, and Zn predominated. The highest concentrations of these elements, exceeding MPC values, were found in the northern part of the lake and in the area where drainage water from the overburden dumps of the Kholboldzhinsky coal mine discharges into the water body. Suspended matter was depleted in all studied HMMs compared to the Clarke values for the upper continental crust; the highest contents in suspended sediments were found for Mn and Sr, which is attributed to surface runoff from the municipal solid waste landfill and an abandoned coal mine. At the same time, bottom sediments were enriched in Mo, Ag, Sb, Sr, Zn, Pb, and As, compared to average concentrations in the upper continental crust. Concentrations of Cu, Zn, Sr, Mo, and Pb in bottom sediments were higher than in suspended matter. HMMs entering the lake from various sources are deposited and accumulated in bottom sediments, increasing their pollution level and forming anthropogenic HMM anomalies.

The greater part of the lake's water mass is well aerated. However, in coastal zones, oxygen content decreases due to the decomposition of organic matter entering the water body with insufficiently treated wastewater from the city of Gusinoozersk and the village of Gusinoye Ozero, as well as with diffuse runoff from landfills and private households within the watershed. The decomposition of remnants of coastal aquatic vegetation also contributes to oxygen depletion. During the study period, almost all nitrogen and phosphorus in Lake Gusinoye were present in organic form. Despite wind-induced mixing and upwelling, the replenishment of the euphotic zone with inorganic nutrients was limited due to their low content in the hypolimnion. The utilization of mineral forms of nitrogen and phosphorus was further facilitated by the presence of extensive macrophyte beds in the littoral zone, which is corroborated by the distribution of organic nitrogen along the eastern coast. In the summer of 2019, Lake Gusinoye was classified as a mesotrophic water body.

A comprehensive study of the lake's water mass together with suspended sediments showed that the current anthropogenic load has not yet led to significant changes in the content of highly soluble salts and HMMs; exceedances of MPCs for individual metals were identified only in the northern part of the lake. However, the gradual accumulation of HMMs in bottom sediments indicates that in the near future they may become a source of secondary pollution for the lake's ecosystem. An unfavorable trend is also observed in the dynamics of nutrient content. Under the influence of increasing municipal load associated with wastewater discharge and diffuse runoff, as well as unregulated tourism, the waters of the northern reach of the lake may transition to a eutrophic state. This could negatively affect the ecosystem of Lake Gusinoye and the quality of water used for domestic and drinking water supply. ■

REFERENCES

Adamovich B.V.; Zhukova T.V.; Mikheeva T.M.; Kovalevskaya R. Z.; Lukyanova E. V. (2016). Long-term variations of the trophic state index in the Narochanskie Lakes and its relation with the major hydroecological parameters. Water Resources., 43, 809–817, doi:10.1134/S009780 81605002X.

Adushinov A.; Borisenko I.; Zonkhoev E.; Rezanov I.; Parfenov I.; Shaibonov B.; Abasheeva N; Ayushina T.; Boikov T.; Litvinov A. (1994). Ecology of Lake Gusinoye. Buryat Academic Press: Ulan-Ude, Russia.

Al Naggar Y.; Khalil M.S.; Ghorab M.A. (2018). Environmental Pollution by Heavy Metals in the Aquatic Ecosystems of Egypt. Open Access Journal of Toxicology, 3, 1–9. doi:10.19080/OAJT.2018.03.555603.

Alekin O.A. (1970). Fundamentals of hydrochemistry. Gidrometeoizdat: Leningrad, Russia.

Babikov V.A.; Sanzheev E.D.; Tsyrendorzhieva T.B.; Shirapova S.D. (2018). Environmental risks of tourist and recreational activity on the northeast coast of Lake Gusinoye. Scientific notes of the V.I. Vernadsky Crimean Federal University. Geography. Geology, 4, 3–13.

Basha Sh.; Jhala J.; Thorat R.; Goel S.; Trivedi R.; Shah R.; Menon G.; Gaur P.; Mody K.H.; Jha B. (2010). Assessment of heavy metal content in suspended particulate matter of coastal industrial town, Mithapur, Gujarat, India. Atmospheric Research, 97, 257–265.

Bazarzhapov T.Z.; Shiretorova V.G.; Radnaeva L.D.; Nikitina E.P.; Bazarsadueva S.V.; Shirapova G.S.; Dong S.; Li Z.; Liu S.; Wang P. (2023). Distribution of Heavy Metals in Water and Bottom Sediments in the Basin of Lake Gusinoe (Russia): Ecological Risk Assessment. Water, 15, 3385, doi:0.3390/w15193385.

Bazarova B.B.; Kuklin A.P. (2021). On the current state and long-term dynamics of the flora and vegetation of Lake Gusinoe (Republic of Buryatia). Ekosistemy. 25, 72–81.

Bazarova B.B. (2024). Assessment of Lake Gusinoe Ecosystem Based on Macrophyte Analysis (Republic of Buryatia) Proceedings of the Russian Geographical Society. 156, 3, 181-193.

Bityukova V.R., Dekhnich V.S., Petukhova N.V. (2021). The impact of thermal power plants on air pollution in Russian cities. Vestnik of Moscow University. Series 5: Geography. 4, 38–51.

Bogdanov V.T. (1977). Hydrochemical state of the waters of Lake Gusinoe // Hydrochemistry of rivers and lakes in conditions of a sharply continental climate. Vladivostok: FVNTs AN USSR. 113-123.

Bortin N.N.; Vasilevsky D.N.; Vasilevskaya L.N.; Lisina I.A. (2023). Evaluation of the dynamics of pollutant fallout with atmospheric precipitation on the watershed areas and water areas of Peter the Great Bay. Water sector of Russia: problems, technologies, management. Water management of Russia, 3, 17–35, doi:10.35567/19994508_2023_3_2.

Carlson R.E. (1977). A trophic state index for lakes. Limnol.Oceanog, 11, 361–369.

Chebunina N.S.; Pakhakhinova Z.Z.; Beshentsev A.N.; Batoev V.B. (2016). Assessment of the impact of the discharge of warm waters from the Gusinoozerskaya TPP on the dynamics of the ice regime of Lake Gusinoe (Western Transbaikalia). Sci. Journ. KubSAU, 116, 301–308.

Ćipranić I.; Marković R.; Đorđievski S.; Stevanović Z.; Stevanović M. (2019).. Journal of the Serbian Chemical Society, 84, 527–530, doi:10.2298/JSC190129012C.

Dagurova O.P.; Zaitseva S.V.; Tsydenova B.V.; Dambaev V.B.; Matafonov D.V. (2019). Monitoring of physical and chemical parameters the water of Lake Gusinoe. Vestnik Buryatskogo Gos. Universiteta, 4, 3–8.

Dampilova B.V.; Khazheeva Z.I.; Plyusnin A.M. (2022). Heavy Metal Species in the Bottom Sediments of the Aquatic System of Lake Gusinoe (Buryatia). Geochem. Int., 60, 279–285, doi:10.1134/S0016702922010037.

Domysheva V.M.; Sinyukovich V.N.; Hodger T.V. (1995). Water regime and hydrochemistry of Lake Gusinoe in the modern period. Geogr. Nat. Resour., 2, 73–80.

Erina O.N.; Tereshina M.A.; Efimova L.E.; Sokolov D.I. (2020). Actual dynamics of nutrients in the Mozhaisk reservoir. Vestnik Moskovskogo universiteta. Seriya 5, Geografiya, 3, 81–90.

Efimova L.E.; Lomov V.A.; Grechushnikova M.G.; Lomova D.V.; Kremenetskaya E.R. (2019). Influence of hydrological structure on exchange processes between water and bottom sediments in Mozhaysk reservoir. IOP Conference Series: Earth and Environmental Science, 321, 1-5, doi 10.1088/1755-1315/321/1/012015

Grechushnikova M.G.; Sokolov D.I.; Erina O.N.; Tereshina M.A.; Lomov V.A.; Efimova L.E. (2020). Seasonal and synoptic variability of hydroecological characteristics of a low-flow valley reservoir. Meteorology and hydrology, 8, 92–101.

Kara M.; Dumanoglu Y.; Altiok H.; Elbir T.; Odabasi M.; Bayram A. (2014). Seasonal and spatial variations of atmospheric trace elemental deposition in the Aliaga industrial region, Turkey. Atmospheric Research, 149, 204–216, doi:10.1016/j.atmosres.2014.06.009.

Khazheeva Z.I.; Plyusnin A.M. (2016). Variations in climatic and hydrological parameters in the Selenga River basin in the Russian Federation. Russian Meteorology and Hydrology, 41, 640–647, doi.org/10.3103/S1068373916090077.

Khazheeva Z.I.; Plyusnin A.M. (2018). The current state of water resources of Lake Gusinoe (Western Transbaikalia). Water Resources, 45, 68–74, doi:10.1134/S0097807818010116.

Komarova N.V.; Kamentsev Ya.S. (2006). Practical guide to using the systems of capillary electrophoresis «Kapel». Veda LLC: St. Petersburg, Russia.

Kosheleva N.; Efimova L.; Efimov V.; Sycheva D. (2022). Potentially toxic elements in the Gusinoye lake (Republic of Buryatia, Russia). Environmental Science and Pollution Research, 29, 1–16, doi:10.1007/s11356-022-18848-1.

Kratzer C.R.; Brezonik P.L. (1981). A Carlson-type trophic state index for nitrogen in Florida Lakes. Journal of the American Water Resources Association, 17, 713–715.

Lapin S.A.; Gangnus I.A.; Zozulya N.M. (2014). Features of seasonal variability in hydrological and hydrochemical characteristics of the Desnogorsk Reservoir. Fisheries issues, 15, 277–284.

Lukyanova A.N.; Lukyanova O.N.; Efimova L.E.; Efimov A.V. (2020). Influence of natural and anthropogenic factors on water quality in cooling ponds (by the example of Lake Gusinoe). Water industry of Russia: problems, technologies, management, 2, 146–161. doi:10.35567/1999-4508-2020-2-10.

Moiseenko T.I.; Dauvalter V.A.; Lukin A.A.; Kudryavtseva L.P.; Ilyashchuk L.I.; SandimirovS.S.; Kagan L.Ya.; Vandysh O.I.; Sharov A.N.; Sharova Yu.N.; Koroleva I.N. (2002). Anthropogenic modifications of the ecosystem of Lake Imandra. Nauka: Moscow, Russia.

Obozhin V.N., Bogdanov, V. T., Klikunova, O.F. (1984). Hydrochemistry of Rivers and Lakes in Buryatia. Nauka: Novosibirsk, Russia, 151.

Pokrovskaya T.N.; Mironova N.Ya.; Shilkrot G.S. (1983). Macrophytic lakes and their eutrophication. Nauka: Moscow, Russia.

Potemkin V.L.; Potemkina T.G.; Guseva E.A. (2011). Regional transfer of contaminanats as a geoecological problem of the Baikal region. Bulletin of Irkutsk State Technical University, 6, 103–107.

Radnaeva L.D.; Bazarzhapov T.Zh,; Shiretorova V.G.; Zhigzhitzhapova S.V.; Nikitina E.P.; Dylenova E.P.; Shirapova G.S.; Budaeva O.D.; Beshentsev A.N.; Garmaev E. Zh.; Wang P.; Dong S.; Li Z.; Tulokhonov A.K. (2022). Ecological State of Lake Gusinoe—A Cooling Pond of the Gusinoozersk GRES. Water, 14(1), 4, https://doi.org/10.3390/w14010004

Rudnick R.L., Gao S. (2003). Composition of the continental crust. Treatise on geochemistry. Elsevier Science, 3, 1–64.

Samarina, A.V., Khudyakova, R.V. (1969). Chemical Characteristics of Lake Gusinoe. In Proceedings of the 2nd Meeting on the Matter and Energy Cycle in Lakes. Novosibirsk.

Sapelko T.V., Lapenkov A.E., Guzivaty V.V., Karetnikov S.G., Gazizova T.V., Ignatyeva N.V., Kuznetsov D.D., Rusanov A.G. (2025). Features Of Unique Lakes Development On The Kurgalsky Peninsula In The Southeast Baltic Sea Based On The Results Of Interdisciplinary Research. Geography, Environment, Sustainability, 65-74. https://doi.org/10.24057/2071-9388-2025-3512

Sapozhnikov V.V.; Agatova A.I.; Arzhanova N.V.; Mordasova N.V.; Lapina N.M.; Zubarevich V.L.; Lukyanova O.N.; Torgunova N.I. (2003). Guidelines for the chemical analysis of marine and fresh water in environmental monitoring of fishery water bodies and areas of the World Ocean that are promising for fishing. VNIRO press: Moscow Russia.

Sharapova E.; Efimova L.; Denisova I.; Ermekova A.; Lychagin M.; Chalov S. (2020). Spatial variability of biogenic elements and organic carbon content in the tributaries of lake Baikal. E3S Web of Conferences St. Petersburg, Russia, March 20-27. doi:10.1051/e3sconf/202016305012.

Shavrak E.I.; Gorskaya O.I.; Generalenko I.A.; Shavrak D.S. (2012). Features of the formation of the salt composition of the cooling reservoir of the Rostov NPP. Problems of nuclear radiation and environmental safety, 1, 32–37.

Smith V.H. (1982). The nitrogen and phosphorus dependence of algal biomass in lakes: an empirical and theoretical analysis. Limnol. Ocearnogr, 23, 1248–1255.

Soranno P.A.; Carpenter S.R.; Lathrop R.C. (1997). Internal phosphorus loading in Lake Mendota: response to external loads and weather. Canadian Journal of Fisheries and Aquatic Sciences, 54, 1883–1893.

Suzdaleva A.L.; Goryunova S.V. (2014). Technogenesis and degradation of surface water bodies; Publishing House Energy: Moscow Russia.

Sycheva D.G.; Kosheleva N.E. (2023). Ecological and geochemical state of the soil cover of Gusinoozersk in the impact zone of coal thermal power plant. Eurasian Soil Science, 56, 1114–1129. doi:10.1134/S1064229323600896.

Tsybekmitova G.Ts.; Matveeva M.O. (2020). Biogenic elements (nitrogen and phosphorus) in the Kharanor reservoir. Ekosistemy, 24, 142–151.

Tsydypov B.Z.; Andreev S.G. Ayurzhanaev A.A.; Sodnomov B.V.; Gurzhapov B.O.; Batotsyrenov E.A.; Pavlov I.A.; Shiretorova V.G.; Gabeeva D.A.; Radnaeva L.D.; Garmaev E.Zh. (2017). Influence of discharges from the Gusinoozerskaya GRES on the thermal and hydrochemical regime of Lake Gusinoe. Izv. Irkutsk. Gos. Univ. Ser. Earth Sciences, 22, 135–150.

Ulzetueva I.D.; Gomboev B.O.; Zhamyanov D.Ts.-D.; Molotov V.S. (2015). Assessment of anthropogenic impact on water objects of the Selenga river transboundary basin (the Russian part). Bulletin of Buryat State University, 4, 68–75.

Vithanage M.; Bandara P.Ch.; Novo L.A.B.; Kumar A.; Ambade B.; Naveendrakumar G.; Ranagalage M.; Magana-Arachchi D.N. (2022). Deposition of trace metals associated with atmospheric particulate matter: Environmental fate and health risk assessment. Chemosphere, 303, 135051. doi:10.1016/j.chemosphere.2022.135051.

Walumona J.R.; Odoli C.O.; Raburu P.; Amisi F.M.; Murakaru M.J.; Kondowe B.N.; Kaunda-Arara B. (2021). Spatio-temporal variations in selected water quality parameters and trophic status of Lake Baringo Kenya. Lakes and Reservoirs: Research and Management 16 e12367. doi:10.1111/lre.12367.

Wetzel R.G. (2001) Limnology. Lake and River Ecosystems. Gulf professional publishing: Oxford United Kingdom.

Wilhelm S.; Adrian R. (2008). Impact of summer warming on the thermal characteristics of a polymictic lake and consequences for oxygen nutrients and phytoplankton. Freshwater Biology, 53, 226–237.

Xiao W.; Huang Yu.; Mi W.; Wu H.; Bi Y. (2019). Variation of Diatoms and Silicon in a Tributary of the Three Gorges Reservoir: Evidence of Interaction. Water, 11, 1369. doi:10.3390/w11071369.

Zhambalova D.I.; Plyusnin A.M.; Chernyavskii M.K.; Peryazeva E.G.; Ukraintsev A.V. (2020). Changing the hydrochemical regime of lake Gusinoe under the influence of coal mining Water-Rock Interaction. Geological Evolution, 224-226. doi:10.31554/978-5-7925-0584-1-2020-224-226.

Zhigzhitzhapova, S.V.; Pavlov, I.A.; Shiretorova, V.G.; Dylenova, E.P.; Radnaeva, L.D.; Tulokhonov A.K. (2019). Metal content in aquatic plants of Lake Gusinoye. Water Chem. Ecol., 1, 34–40.

Zhao Sh.; Shi X.; Li Ch.; Zhang Sh.; Sun B.; Wu Y.; Zhao Sh. (2017). Diffusion flux of phosphorus nutrients at the sediment–water interface of the Ulansuhai Lake in northern China. Water Science and Water Science & Technology, 75, 1455–1465.

Zhao X.; Xu Zh.; Li P.; Dong Zh.; Fu P.; Liu Chandra C.-Q.; Pavuluri M. (2022). Characteristics and seasonality of trace elements in fine aerosols from Tianjin North China during 2018-2019. Environmental Advances, 9, 100263. doi:10.1016/j.envadv.2022.100263.

Appendix

Supplementary materials to the article by Lyudmila Efimova, Natalia Kosheleva, Anna Lukyanova, Daria Sycheva and Vasiliy Efimov: "Comprehensive assessment of Lake Gusinoye, Republic of Buryatia, based on water, suspended particulate matter and bottom sediments geochemistry"

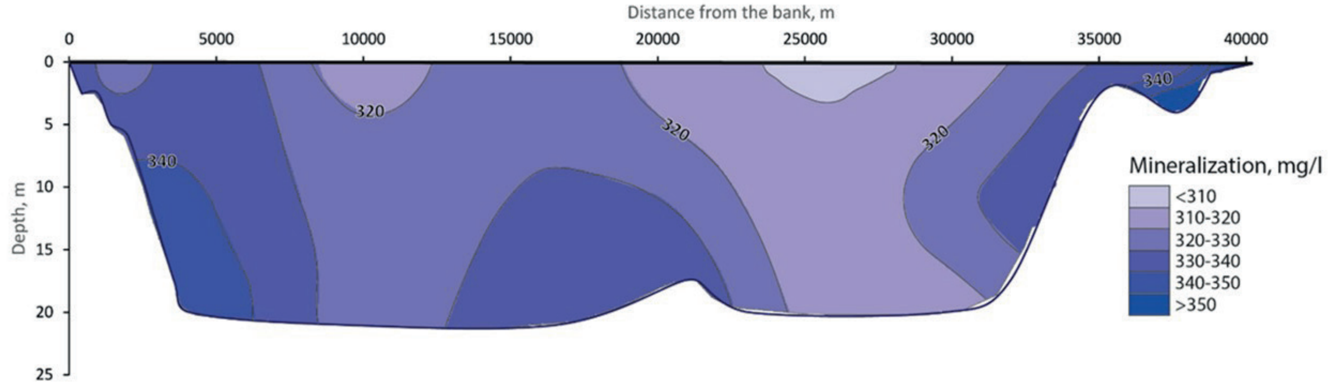


Fig. S1. Distribution of water mineralization (mg/l) along profile 3 (stations 1.6, 1.5, 1.8, 3.3, 2.7, 2.6, 2.3, 2.1), July 25, 2019

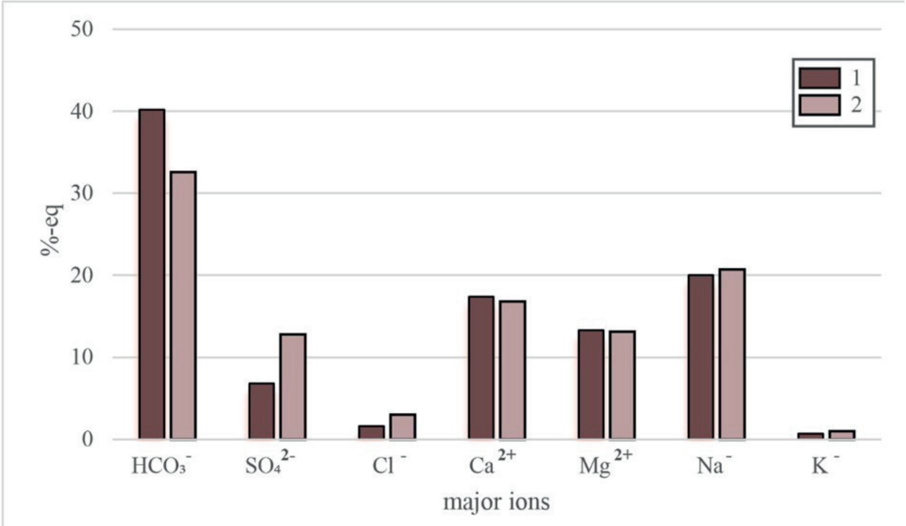


Fig. S2. Relative content of major ions in the water of Lake Gusinoye, %-eq:
a – August 1969 (Adushinov et al., 1994); b – July 2019 (data from MSU expedition, 2019)

Table S1. The content (surface-bottom) of major ions (mg/dm^3) in the water of Lake Gusinoye at water sampling stations

Station No	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	2.1
The depth, m	6	5	2.5	2.0*	4.5	1.8*	2.5*	11	4.5
pH	8.33-8.28	8.32-8.34	8.23-8.23	8.28	8.34-8.28	8.3	8.32	8.37-8.26	8.29-7.60
HCO_3^-	179.6-179.6	168.7-180.2	187.1-190.1	182	180.7-184.9	181.8	187	182.5-192.5	169-192.2
SO_4^{2-}	55.8-55.9	56.2-56.8	56.4-56.6	57.1	57.0-55.7	56.0	55.3	55.9-55.6	50.1-22.6
Cl^-	9.7-9.2	9.4-9.1	11.2-9.1	9.4	9.1-8.9	8.8	8.8	8.9-9.5	8.1-2.8
Ca^{2+}	28.5-28.0	31.3-27.8	24.7-26.4	26.4	23.9-30.5	28.7	29.0	28.5-28.4	28.6-27.4
Mg_2^{2+}	16.5-16.2	17.8-16.2	14.0-15.5	14.9	13.8-16.6	17.0	16.7	16.8-17.1	14.9-7.1
Na^+	39.0-38.7	41.8-38.3	32.5-35.8	37.2	32.2-39.8	40.6	40.5	41.6-40.8	36.2-14.1
K^+	2.6-2.7	2.8-2.6	2.1-2.5	2.7	2.1-2.8	3.1	2.7	2.8-2.9	2.8-1.8
M	336-339	329-334	332-335	318	319-324	326	331	337-347	330-355
The station number	2.2	2.3	2.4	2.5	2.6	2.7	3.1	3.2	3.3
The depth, m	20	22	15.6	12	19	21	19.6	16	21
pH	8.44-7.90	8.44-8.27	8.44-8.0	8.42-8.36	8.44-7.94	8.42-7.91	8.34-7.91	8.42-8.17	8.45-7.91
HCO_3^-	177.8-189.5	175.4-178.6	180.3-190.4	186.5-189.8	182.2-206	191.5-202.7	183.7-183.9	183.1-184.2	190.9-199.3
SO_4^{2-}	55.5-55.7	55.1-51.1	52.9-52.9	51.9-52.4	52.4-51.7	53.8-52.5	49.7-53.1	54.0-54.1	49.7-53.3
Cl^-	8.7-9.0	9.3-9.3	8.8-8.9	8.6-8.8	8.5-8.7	9.1-8.8	8.2-8.4	8.2-8.1	7.3-7.7
Ca^{2+}	27.6-28.1	22.0-23.7	22.1-22.1	20.4-22.4	22.2-23.5	23.1-23.5	21.3-22.2	21.7-22.4	21.6-21.2
Mg_2^{2+}	16.3-16.4	13.8-13.6	13.1-12.9	12.6-13.6	13.7-14.0	14.2-14.1	13.0-14.4	14.5-14.9	14.0-13.5
Na^+	39.0-39.0	34.1-32.1	31.2-29.9	27.9-31.6	31.9-32.4	32.7-32.6	30.8-32.4	32.6-32.2	30.0-28.8
K^+	2.7-2.7	2.7-2.8	2.6-2.6	2.1-2.5	2.8-2.7	2.6-2.7	2.6-2.6	3.2-3.0	2.9-2.7
M	328-340	312-327	311-320	310-321	314-329	327-337	310-318	318-320	317-327

Note: *The content of major ions at stations 1.4, 1.6, 1.7 was measured only at the surface.

Table S2. The concentrations of dissolved forms of HMMs in water (surface – bottom) of Lake Gusinoye and in the suspended and bottom sediments at water sampling stations. The MPCs for fishery and clarke values for the upper part of the earth's crust (Rudnick, Gao., 2003)

Station No	1.1	1.2	1.6	1.7	1.8	2.2	2.3	2.4	3.1	3.2	3.3	
Depth, m	6	2.5	2.0*	2.5*	20	21	21	17	21	21	20	
Concentrations of dissolved forms of HMMs, µg/l`												MPC for fishery
Mn	0.66-0.29	1.1-1.0	1.7	0.8	0.52-1.2	0.63-0.85	0.64-1.3	0.24-0.73	0.43-1.2	1.2-0.74	0.55-0.27	10
Cu	1.4-1.0	1.3-0.93	1.5	0.92	1.1-1.4	0.95-1.0	1.1-1.2	0.98-1.1	1.1-1.2	1.0-0.94	1.1-0.87	1
Zn	3.4-1.2	1.5-0.8	1.7	0.75	0.75-1.5	1.4-0.73	1.7-1.0	1.0-36.1	1.3-2.3	1.0-0.79	1.1-1.0	10
Sr	962-990	961-969	990	971	968-983	969-979	968-985	984-960	963-991	958-970	979-982	400
Mo	15.2-15.8	16.0-15.0	15.4	15.8	15.2-16.2	14.9-15.6	15.1-15.9	15.3-15.4	14.9-31.1	14.9-15.8	15.1-15.4	1
Pb	0.11-0.09	0.13-0.18	0.11	0.06	0.04-0.13	0.08-0.04	0.13-0.04	0.06-0.12	0.08-0.1	0.04-0.05	0.03-0.08	6
As	1.07-1.08	1.1-1.2	1.1	1.1	1.0-1.1	1.0-1.1	1.1-1.2	1.0-1.2	1.1-1.2	1.0-1.1	1.0-1.1	50
Concentrations of HMMs in suspended sediments, µg/g												Clarkes
Mn	31-154	121-162	212	199	41-158	47.6-133	24.4-110	39.7-87.7	-	395-43.4	46.3-74.6	900
Cu	2.4-2.3	3.3-1.1	7.4	2.4	2.5-3.0	1.2-2.3	1.0-1.8	1.3-1.2	-	4.0-1.1	1.1-3.0	28
Zn	9.1-10.5	7.4-4.9	9.2	7.1	4.6-5.5	4.8-5.0	3.2-3.9	4.2-4.4	-	8.0-2.5	7.3-11.7	67
Sr	7.8-13.5	7.8-10.8	17.9	14.6	9.8-13.7	11.6-12.9	8.1-11.4	12.0-15.3	-	15.3-7.7	7.4-9.4	330
Mo	0.17-0.20	0.27-0.62	0.29	0.37	0.28-0.59	0.31-0.56	0.12-0.27	0.17-0.75	-	0.63-0.19	0.17-0.47	1.1
Pb	4.9-11.4	3.9-6.6	10.1	4.6	3.3-4.1	2.8-3.1	2.0-2.2	2.9-3.2	-	1.1-1.2	1.3-2.1	17
Concentrations of HMMs in bottom sediments, µg/g												Clarkes
Ag	0.07	0.01	0.09	0.1	0.06	0.09	1.4	0.28	0.12	0.33	0.1	0.053
Mo	2.3	2.3	10.6	13.4	12.2	9.6	13.9	10.4	10.2	12.7	13.6	1.1
As	11.5	2	5.8	7.4	5.3	5.9	6.5	5.7	4.2	6.9	6.9	4.8
Cu	17.6	3.5	36.5	39.9	28.2	27.1	42.4	34	35.4	46	39.7	28
Sr	222	492	370	462	760	772	467	424	351	705	390	330
Zn	90.8	39	109	118	70.9	84.9	117	121	104	106	112	67
Sb	1.3	0.41	1.1	1.2	0.8	0.92	1.2	1.2	1.0	1.13	1.21	0.4
Pb	22.7	15.7	24.6	25.5	20.2	21.9	29	28.7	21.7	26.1	25.9	17
V	76.2	21.2	83.6	92.4	56.6	68	88.9	95.9	64.4	73.1	86	97
Co	9.5	4.5	11.1	13.1	8.2	10.9	11.8	13.2	8.3	10.3	11.2	17.3
Cr	22.6	4.6	32.2	39	22.2	27.1	44.3	41.1	24.9	30.2	35.3	92

Note: *The concentrations of HMMs were measured only at the surface; "-"no data available.

Table S3. The concentrations (surface-bottom) of nutrients (mg/dm^3) in the water of Gusinoye Lake at water sampling stations. The MPCs for fishery

Station No	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	2.1
The depth, m	6	5	2.5	2.0*	4.5	1.8*	2.5*	11	4.5
Si	1.26-1.34	1.36-1.40	1.38-1.36	1.40	1.26-1.44	1.42	1.42	1.28-1.44	1.38-2.78
PO_4^{3-}	0.021-0.012	0.017-0.018	0.014-0.014	0.014	0.014-0.017	0.029	0.017	0.015-0.015	0.019-0.045
$P\text{-}PO_4^{3-}$	0.0068-0.004	0.0035-0.0033	0.0047-0.0047	0.0047	0.0045-0.0054	0.0094	0.0054	0.0049-0.0049	0.0063-0.0148
P total	0.021-0.024	0.026-0.025	0.025-0.022	0.022	0.021-0.022	0.029	0.027	0.021-0.025	0.022-0.052
$N\text{-}NO_3^-$	0.004-0.002	0.004-0.003	0.002-0.003	0.002	0.001-0.004	0.004	0.002	0.002-0.008	-
$N\text{-}NH_4^-$	0.022-0.063	0.022-0.037	0.048-0.031	0.016	0.010-0.009	0.047	0.042	0.024-0.063	0.015-0.020
$N\text{-}NO_2^-$	0.003-0.003	0.004-0.003	0.003-0.004	0.004	0.003-0.003	0.004	0.004	0.003-0.004	0.001-0.006
N total	0.5-0.6	1.0-0.6	0.6-0.5	0.6	0.6-0.5	0.6	1.3	1.1-0.7	1.0-0.8
N org	0.47-0.53	0.90-0.56	0.55-0.46	0.58	0.58-0.48	0.55	1.25	1.07-0.63	0.98-0.77

Station No	2.2	2.3	2.4	2.5	2.6	2.7	3.1	3.2	3.3
The depth, m	20	22	15.6	12	19	21	19.6	16	21
Si	1.26-2.23	1.26-2.29	1.28-2.09	1.32-1.54	1.26-2.09	1.26-2.13	1.36-2.21	1.38-1.76	1.48-2.31
PO_4^{3-}	0.012-0.062	0.013-0.064	0.015-0.019	0.014-0.018	0.011-0.044	0.013-0.054	0.011-0.049	0.011-0.016	0.012-0.052
$P\text{-}PO_4^{3-}$	0.004-0.0202	0.0042-0.0209	0.0049-0.0063	0.0047-0.0059	0.0035-0.0143	0.0042-0.0176	0.0035-0.0159	0.0035-0.0052	0.0037-0.017
P total	0.027-0.041	0.021-0.039	0.021-0.023	-	-	-	0.016-0.039	0.019-0.022	0.017-0.032
$N\text{-}NO_3^-$	-	0.003-0.008	0.004-0.007	-	-	-	0.043-0.009	0.067-0.026	0.020-0.134
$N\text{-}NH_4^-$	0.016-0.026	0.017-0.017	0.015-0.023	0.033-0.017	0.011-0.018	0.015-0.022	0.029-0.021	0.019-0.019	0.028-0.034
$N\text{-}NO_2^-$	0.000-0.001	0.000-0.000	0.001-0.001	0.001-0.001	0.000-0.001	0.001-0.001	0.000-0.000	0.001-0.001	0.000-0.001
N total	1.2-0.5	1.4-0.9	0.6-0.6	-	-	-	0.6-0.8	0.6-0.5	0.5-0.7
N org.	1.18-0.47	1.38-0.87	0.58-0.57	-	-	-	0.53-0.77	0.51-0.45	0.45-0.53

Note: *The concentrations of the nutrients at stations 1.4, 1.6, 1.7 was measured only at the surface; "-"no data available.

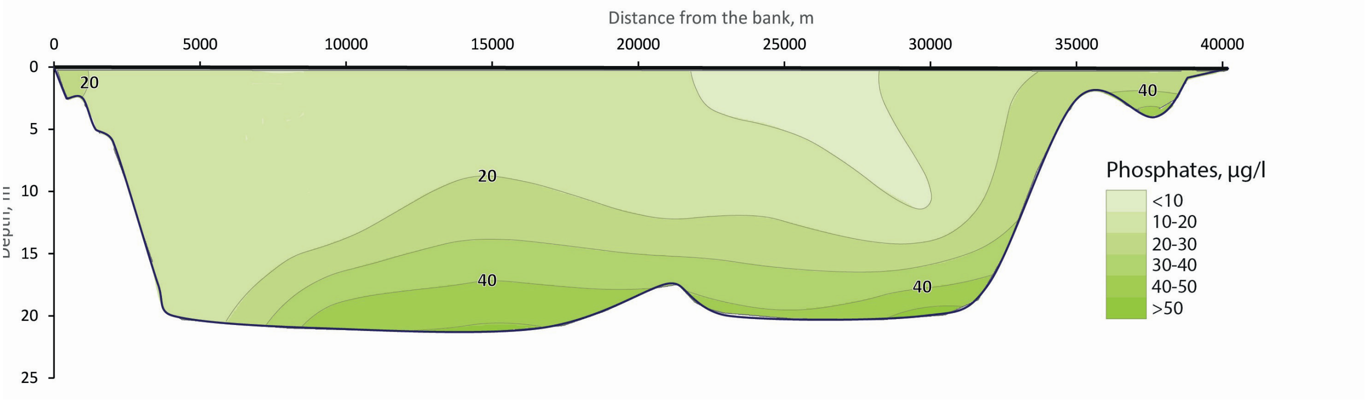


Fig. S3. Distribution of phosphates (PO_4, µg/l) in water of Lake Gusinoye in July 2019 along profile 3 (stations 1.6, 1.5, 1.8, 3.3, 2.7, 2.6, 2.3, 2.1)

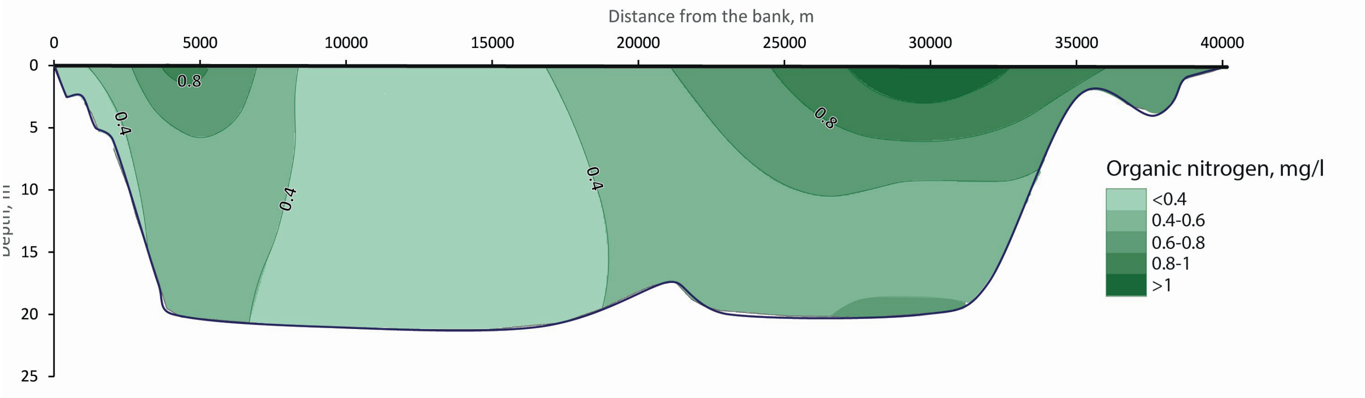


Fig. S4. Distribution of organic nitrogen (mg/l) in water of Lake Gusinoye in July 2019 along profile 3 (stations 1.6, 1.5, 1.8, 3.3, 2.7, 2.6, 2.3, 2.1)

RESEARCH PAPER

VARIATION OF AEROSOL PROPERTIES BASED ON SATELLITE DATA OVER THE STUDY AREA OF LHAASO

Andleeb Masood[1*], Fengrong. Zhu[1], Zulfaqar Sa'adi[2], Irfan Ullah[3],
Md. Abdullah Al Mamun Hridoy[4], Huan-Yu Jia[1]
[1]School of Physical Science and Technology, Southwest Jiaotong University, Chengdu, 611756, China
[2]Centre for Environmental Sustainability and Water Security, Research Institute for Sustainable Environment, Universiti Teknologi Malaysia, UTM Skudai, Johor Bahru, 81310, Malaysia
[3]College of Hydrology and Water Resources, Hohai University, Nanjing, Jiangsu, 210098.China
[4]Department of Fish Biology and Genetics, Faculty of Fisheries, Sylhet Agricultural University, Sylhet-3100, Bangladesh
*Corresponding author: andleebmasood238@gmail.com

Received: April 2ndt 2025 / Accepted: February 26th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4063

ABSTRACT. Atmospheric aerosols play a crucial role in modulating Earth's radiation balance and climate systems, yet they remain a major source of uncertainty in climate modeling. This study aims to characterize aerosol dynamics over the Large High Altitude Air Shower Observatory (LHAASO; 29.36° N, 100.14° E; ~4,410 m above sea level) region by analyzing satellite-based measurements of Aerosol Optical Depth (AOD), Aerosol Index (AI), Single Scattering Albedo (SSA), and Extinction Coefficient (EC) from Moderate Resolution Imaging Spectroradiometer (MODIS), Sentinel-5P, CALIPSO, OMI, and MERRA-2 Reanalysis for the period 2017–2023. The mean AOD peaked at 0.24 ± 0.02, primarily due to long-range dust transport from the Taklamakan and Gobi deserts. The AI varied between 0.68 and 1.26 (standard deviation SD = 0.31 ± 0.26), indicating moderate to significant aerosol activity, particularly during dust events. Aerosol mass concentration (MC) ranged from 4.14 to 31.07 μg/m^3 with a SD of 0.67 ± 8.14, reflecting influences from dust transport, meteorological conditions, and local emissions. The EC generally decreased with altitude, consistent with reduced aerosol concentrations and particle sizes. These findings advance our understanding of aerosol behavior in complex terrain and provide essential perspectives for improving climate projections and radiative forcing assessments.

KEYWORDS: Aerosol index, AOD, Extinction coefficient, Satellite remote sensing

CITATION: Masood A., Zhu F., Sa'adi Z., Ullah I., Md. Abdullah Al Hridoy M., Huan-Yu Jia (2026). Variation Of Aerosol Properties Based On Satellite Data Over The Study Area Of LHAASO. Geography, Environment, Sustainability, 1 (19), 115-129
https://doi.org/10.24057/2071-9388-2026-4063

ACKNOWLEDGEMENTS: This work was supported by the National Key R&D Program of China (2021YFA0718403) and the National Natural Science Foundation of China for International Scientists under Grant No. 42350410438. We acknowledge the NASA Langley Research Center Atmospheric Sciences Data Center.

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

The prediction of global climate change is hindered by significant uncertainties arising from the highly variable spatio-temporal distribution of atmospheric aerosols and our limited understanding of their optical properties (IPCC 2023). In the atmosphere, tiny suspended particles known as aerosols range in size from approximately 0.001 to 100 μm (Mahowald et al. 2014). These particles exist in various types, including intermediate, fine, and coarse modes, depending on their dimensions. Aerosols influence solar radiation by scattering and absorbing it, with their impact varying based on their chemical composition and size (Delene & Ogren 2002; Gautam et al. 2021; Jacobson 2002). By interacting with radiation, aerosols affect the Earth's radiation budget both spatially and temporally, a process known as radiative forcing. This alters the Earth's radiative balance and impacts the global climate (Ramachandran et al. 2006; Ranjan et al. 2007; Satheesh & Ramanathan, 2000). Due to their short lifespans and uneven geographic distribution, which is influenced by emissions (Bhowmik et al. 2025), atmospheric chemical reactions, and weather patterns, aerosol properties exhibit significant spatial and temporal variability (Andrews et al. 2011; Boucher et al. 2013; Kumar et al. 2022b; Park et al. 2019). Submicron-sized aerosols may also play a key role in the formation of mist and fog in certain locations (Lillis et al. 1999). Aerosols interact with solar radiation directly through scattering and absorption, while their indirect effects involve influencing cloud processes. Studying aerosols not only deepens researchers' understanding of their environmental effects but also offers essential theoretical support for decision-makers in developing effective environmental protection policies (Edenhofer & Seyboth, 2013; Gong et al. 2015; Zhang et al. 2014a).

Aerosol Optical Depth (AOD) is the fundamental optical property of atmospheric aerosols, serving as a key indicator of atmospheric clarity. It plays a critical role in characterizing atmospheric turbidity and is essential for assessing the effects of aerosols on climate (Zhang et al. 2014b). AOD is frequently used in studies that examine how aerosols influence regional climate's spatial and temporal variation in atmospheric characteristics (Jing et al. 2018; Wang et al. 2019). Moderate Resolution Imaging Spectroradiometer (MODIS) AOD products produce high retrieval accuracy, extensive time series, and excellent spatial coverage. These attributes have been validated through numerous studies conducted by international researchers, highlighting their significant value in aerosol research (Jie et al. 2017; Perumpully, Gautam, & M, 2024; Tian et al. 2018). In recent years, the escalating environmental challenges in China have prompted researchers to intensively investigate the optical, physical, and distribution characteristics of aerosols across key developed regions, including the Yangtze River Delta, Beijing-Tianjin-Hebei, and the Pearl River Delta (Bilal et al. 2019; Zhao et al. 2023). On the other hand, the Beijing urban area helps assess the efficacy of clean air action and provide crucial optical parameters for assessing radiative forcing within two size ranges (Hu et al. 2023; Ren et al. 2022; Zhao et al. 2019). The accurate monitoring of aerosol optical and physical characteristics and their spatiotemporal distributions over the Large High Altitude Air Shower Observatory (LHAASO) holds significant practical implications for atmospheric environmental protection.

Satellite remote sensing (SRS) is essential, as it enables the global measurement of highly variable aerosol fields over extended periods (Blaga & Gautam, 2024; Gautam et al. 2024; Ginoux et al. 2004; Perumpully & Gautam 2025). However, despite advancements in satellite observation techniques for aerosols facilitated by the utilization of increasingly sophisticated instruments, discrepancies persist among various satellite products, indicating ongoing inaccuracies in SRS techniques (Hridoy & Paul 2024). Satellite sensors often have limited spatial resolution, which can hinder the ability to detect fine-scale aerosol features and variability. Hence, a comprehensive analysis of the optical properties and other attributes of aerosols cannot be achieved solely using existing satellite data. The LHAASO study area, because of its distinctive high-altitude environment, is fundamental for understanding the spatial and temporal variations of aerosol properties. The findings of this study are expected to contribute to a deeper understanding of aerosol dynamics in high-altitude regions and provide valuable insights for improving climate models, air quality forecasts, and environmental management strategies. The main objective of this study is to examine aerosol properties using satellite data collected over LHAASO for an extended period (~7 years), considering variations across all four seasons. In this study, we used all available satellite data from 2017 to 2023 to analyze AOD, aerosol index (AI), aerosol extinction coefficient (EC), aerosol concentration, single scattering albedo (SSA), dust surface density, and dust mass concentration.

MATERIALS AND METHODS

The datasets were integrated through a pixel-to-station matching procedure, with explanatory weights determined statistically rather than assigned arbitrarily. Errors in satellite data were minimized through atmospheric correction, cloud masking, and validation against in situ measurements using root-mean-square error (RMSE) and bias statistics. Preprocessing included filling small gaps in the measurement series using linear interpolation between adjacent time points, filtering unreliable satellite pixels flagged by quality assurance bands, and removing spurious ground measurements outside instrument calibration ranges. Outliers were detected via the interquartile range method and excluded only when confirmed as artifacts (e.g., sensor malfunction or accidental contamination), to ensure the dataset reflected genuine environmental variability.

Site Description

The study area encompasses LHAASO (29.35°N, 100.13°E, 4,410.0 m a.s.l.) situated within the Hengduan Mountains. These mountains serve as the junction between the southeastern edge of the Tibetan Plateau, the Yunnan-Guizhou Plateau, and the Sichuan Basin. With an average elevation of 4,410 m a.s.l., the Hengduan Mountains represent a unique and sensitive ecological zone. The location map of LHAASO is presented in Fig. 1. The lack of readily available data collection has hindered comprehensive analysis of AOD dynamics in the area. Despite sporadic ground-based measurements historically recorded at sites like Yunling Baimang Snow Mountain and the Shangri-La atmosphere background station, there exists a gap in continuous monitoring. Recent efforts, including ground measurements conducted at the Litang station, approximately 70 km from LHAASO (Masood et al. 2025), have started shedding light on AOD variations, particularly during the summer months. The elevation map of the study area reveals a diverse topography characterized by significant variations in altitude, as shown in Fig. 2. Spanning from a towering 4,883 m at its highest point to a lower elevation of 4,036 m at its lowest, the map showcases the rugged terrain of the region. These elevation fluctuations play a crucial role in shaping local climate patterns, hydrological processes, and ecosystem dynamics. The depiction of such elevation gradients underscores the importance of understanding the spatial context in which AOD variations occur. As aerosols interact with the atmosphere, their impact can be influenced by the complex interplay of altitude, terrain features, and atmospheric conditions. Therefore, integrating elevation data into the analysis provides valuable information for the spatial distribution and potential drivers of AOD across the study area. Digital Elevation Models (DEMs) with a 30 m resolution were used to establish a comprehensive understanding of the terrain. These DEMs, sourced from the Shuttle Radar Topography Mission (SRTM) through the USGS Earth Explorer platform, allowed for precise calculation and visualization of the elevation profile surrounding the LHAASO observatory.

Data Acquired

The study relied on a multi-satellite and reanalysis approach, where each dataset contributed complementary aerosol-related parameters. Rather than applying fixed weights, the datasets were integrated according to their specific strengths: MODIS for spatially extensive AOD, Sentinel-5P for high-resolution AI, CALIPSO for vertical extinction profiles, OMI for single scattering albedo, and MERRA-2 for dust concentration estimates. Inter-comparison between overlapping parameters (e.g., MODIS AOD with Sentinel-5P AI) was used to ensure consistency. Satellite retrieval errors, particularly those influenced by surface reflectance and cloud contamination, were minimized through preprocessing, quality flag filtering, and validation against MERRA-2 reanalysis.

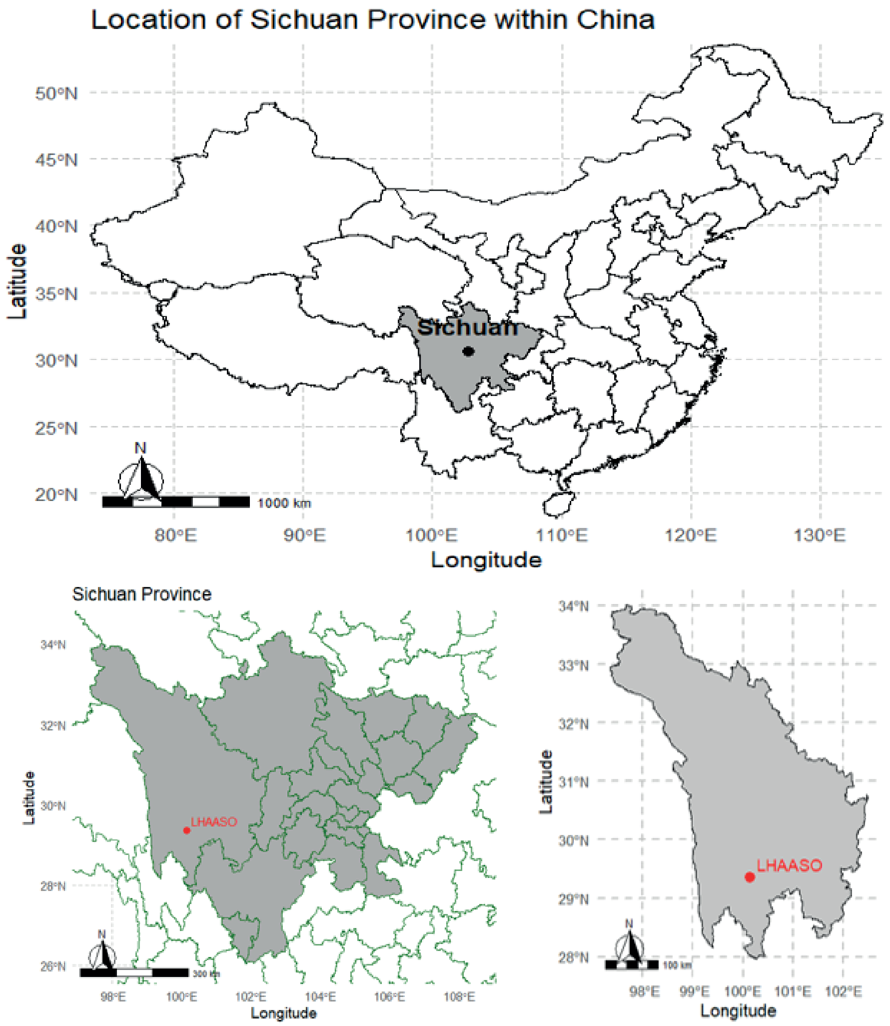


Fig. 1. Location of the LHAASO site within Sichuan Province, China. Sichuan Province highlighted within the map of China (top). Administrative boundaries of Sichuan Province with the LHAASO site marked (bottom left). Enlarged view showing the precise location of LHAASO within Sichuan (bottom right)

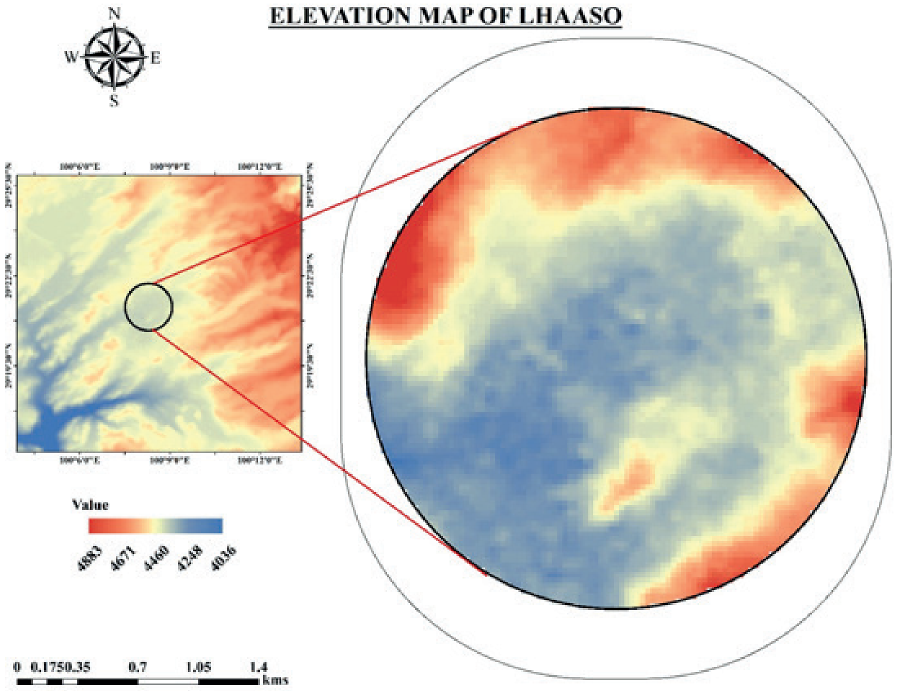


Fig. 2. Digital Elevation Map of LHAASO. The map was created with ArcGIS (version 10.8[1])

---

[1] https://www.arcgis.com/

MODIS

National Aeronautics and Space Administration (NASA)'s Earth Observing System (EOS) launched the Terra and Aqua satellites to monitor the Earth's properties, including the atmosphere and oceans. The Moderate Resolution Imaging Spectroradiometer (MODIS), a key instrument on the Terra satellite, has provided highly accurate observations of aerosols and clouds since its operation began in February 2000. MODIS is renowned for its exceptional spatio-temporal resolution, as it measures radiation reaching the top of the atmosphere across 36 spectral channels, ranging from 0.41 µm to 14.2 µm. The instrument has a wide swath of 2,330 km, with spatial resolutions of 250 m, 500 m, and 1,000 m, and offers a temporal resolution of 1 to 2 days. Alam et al. (2014) found that MODIS's capability to track various fluctuations and compile data has enhanced scientists' understanding of atmospheric, terrestrial, and oceanic phenomena. This accumulated data is invaluable for investigating aerosol characteristics and understanding aerosol–cloud interactions. Using the Deep Blue algorithms (Hridoy et al. 2025), the MODIS instrument aboard the Terra satellites provides comprehensive global coverage of aerosol optical properties over land. The Level 3 MODIS Terra data (MYD08_D3v6) of Aerosol Optical Depth (AOD) used during the period from 2017 to 2023 is retrieved from the Giovanni website[2]. Fig. 3 shows the range of AOD and Aerosol Index (AI) over the LHAASO region from 2017 to 2023.

Sentinel-5P

The Sentinel-5 Precursor (Sentinel-5P) satellite, part of the European Space Agency's Copernicus program, is indeed essential for monitoring atmospheric composition, including aerosol properties through the AI. The AI serves as a quantitative measure that assesses the presence and distribution of aerosols in the atmosphere, providing crucial insights into air quality and climate change. Sentinel-5P Near Real-Time ultraviolet (UV) AI data was incorporated into the analysis. It offers detailed examination of aerosol concentrations and distributions within the study region.

CALIPSO

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) mission is a collaborative effort between NASA's Langley Research Center (LaRC) and the National Space Research Center of France, launched in 2006. This satellite plays a crucial role in providing a three-dimensional view of clouds and aerosols on a global scale, with observations made every 16 days. One of its main tools, the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), uses a 1,064 nm wavelength channel and two polarization channels at 532 nm. It can effectively capture backscattering information about clouds and aerosols across latitudes ranging from 82° north to south. CALIOP enhances the accuracy of vertical distributions of clouds and aerosols by analyzing backscatter data collected at various atmospheric levels. Notably, as an active remote sensing instrument, CALIOP functions effectively both day and night, independent of the Earth's surface conditions and without the reliance on short-wave solar radiation that characterizes passive remote sensing techniques. As a result, CALIPSO provides a robust and comprehensive dataset for studying the vertical structure and transmission of aerosols in the atmosphere. The collected data included the CALIPSO Level 2 aerosol extinction profiles generated from LIDAR measurements, including aerosol extinction coefficients at various vertical levels. These profiles offer valued insights into the spatial distribution and vertical structure of aerosols within the study area. By analyzing the vertical aerosol extinction profiles obtained from CALIPSO, we are able to characterize the vertical distribution of aerosols and investigate their variability over time.

OMI

The Ozone Monitoring Instrument (OMI) is an important satellite-based sensor that is part of NASA's Aura satellite, launched in 2004. OMI is designed to monitor and collect data on atmospheric composition, particularly ozone and other trace gases. For the data collection of the aerosol Single Scattering Albedo (SSA), we obtained time series data from the OMI (Ozone Monitoring Instrument) satellite sensor. The dataset titled "Time Series, Area-Averaged of SSA 500 nm daily 1 deg." provides daily measurements of SSA at a wavelength of 500 nm, averaged over a spatial resolution of 1 degree. The OMI instrument captures information about aerosol properties in the Earth's atmosphere, including their ability to scatter light, which is quantified by the single scattering albedo. By utilizing this dataset, we aimed to analyze the temporal variations and spatial distribution of SSA over the study area.

Merra-2 Reanalysis

By utilizing the dataset, we aimed to examine the temporal variability and spatial distribution of dust surface mass concentration in the study area. For the dust column mass concentration/density, we utilized data from the MERRA-2 Reanalysis dataset. The dataset titled "Time Series, Area-Averaged of Dust Column Mass Density monthly 0.5 × 0.625 deg." provides monthly measurements of dust column mass density at a spatial resolution of 0.5 × 0.625 degrees. This dataset integrates information about the vertical distribution of dust within the atmosphere, offering an understanding of the total mass of dust particles present in the atmospheric column.

Procedure

The methodology adopted for this study involved several key steps, beginning with the acquisition, preprocessing, and analysis of data related to DEMs and AOD dynamics. The DEM and AOD data were preprocessed to improve quality and eliminate potential inconsistencies, ensuring accurate data for further analysis. First, the DEM data were acquired to characterize the complex topography of the high-altitude region (~4,410 m a.s.l.). The DEM was projected to a common geographic coordinate system to match the spatial reference of the satellite-derived AOD data. Resampling techniques were applied where necessary to harmonize spatial resolution between datasets. Any voids or missing elevation values were filled using interpolation algorithms embedded within the GIS processing software. Slope and elevation layers were derived from the DEM to assess topographic influences on aerosol distribution and atmospheric transport. This step is particularly important in mountainous regions such as Daocheng County, Sichuan, China, where terrain-induced variability can significantly influence aerosol dispersion. Similarly, preprocessing was applied to the aerosol index data obtained from the Sentinel-5P Near Real-Time UV

---

[2] https://giovanni.gsfc.nasa.gov/giovanni/

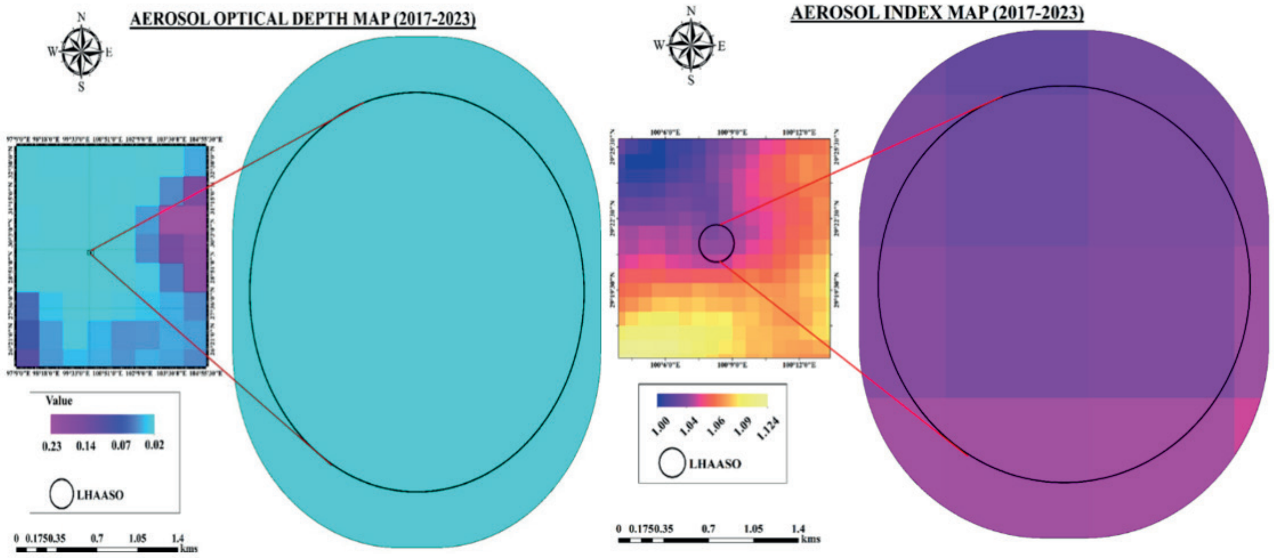


Fig. 3. Mapping of the range of AOD and AI over the LHAASO region from 2017 to 2023

AI dataset. For both the AOD and AI data, pre-processing included: (i) spatial interpolation of all datasets to a common 0.5° × 0.5° grid using bilinear interpolation, which balances computational efficiency with the preservation of spatial gradients; (ii) filtering based on the quality assurance (QA) flags provided by each sensor (e.g., MODIS cloud/surface reflectance flags, Sentinel-5P retrieval confidence flags, and OMI row anomaly filters), thereby excluding low-quality retrievals; and (iii) statistical outlier removal, where values exceeding ±3 standard deviations from the monthly mean were discarded to reduce the influence of spurious satellite retrievals. These steps ensured that only consistent, high-quality data were retained. Importantly, no actual emission sources were removed; instead, the filtering process was designed solely to eliminate data contaminated by cloud interference, snow cover, or algorithm retrieval artifacts.

By following these methodological steps, a robust dataset was compiled, integrating aerosol-related parameters. This comprehensive pre-processing improved the reliability and integrity of the datasets, allowing for an in-depth analysis of the relationship between topography and aerosol dynamics in the study area. CALIPSO vertical profiles were used as an independent constraint to verify the vertical distribution of aerosols, while OMI single scattering albedo was employed to complement MODIS AOD in assessing aerosol radiative properties. No explicit statistical weights were assigned, but each dataset was used in its optimal domain (e.g., MODIS for wide coverage, CALIPSO for vertical profiles), thereby creating a complementary framework.

In addition to descriptive statistics, Pearson correlation was used to evaluate linear relationships, while Spearman rank correlation was employed to capture potential non-linear associations between the study parameters. These statistical tests provide quantitative support for the observed spatial and temporal variability. Statistical tests, including Shapiro–Wilk, Kruskal–Wallis, and Levene's tests, were applied to assess the normality, month-wise variability, and homogeneity of variance of the studied variables. Aerosol vertical profiles were analyzed using linear regression and altitude-based binning to quantify terrain effects on aerosol distribution from 2017 to 2023. Data visualization techniques were employed to create graphs and visual representations, facilitating the interpretation of spatial and temporal trends in AOD and AI values. By following these methodological steps, a robust dataset was compiled, integrating both elevation data and aerosol-related parameters. This integration of multiple data sources, together with systematic pre-processing and cross-validation, ensures a more reliable representation of aerosol dynamics in the study area. The seasonal distribution of AOD is primarily driven by local emissions, boundary-layer dynamics, humidity effects, and wet scavenging during rainfall. Snow cover at higher altitudes may affect retrieval accuracy but does not directly increase AOD. Long-range transport of dust and regional pollution, particularly in the pre-monsoon (Hridoy et al. 2025), also plays a role, while secondary aerosol formation (e.g., sulfate, nitrate) contributes during warmer months. Similar seasonal patterns and drivers have been documented in other high-altitude regions, including the Himalaya, Tibetan Plateau, and observatories such as LHAASO, supporting our interpretation.

RESULTS AND DISCUSSION

Descriptive statistics of monthly variable

The descriptive statistics of the studied monthly variables, namely AOD, AI, and MC are shown in Table 1. The analysis of AOD values across the months reveals significant variations in minimum, maximum, mean, range, standard deviation (SD), coefficient of variation (CV), and skewness. Results demonstrate that January, April, October, and December have the lowest minimum AOD values (0.03, 0.03, 0.00, and 0.03), indicating clearer air, while July records the highest minimum (0.22), reflecting peak aerosol levels. April shows the highest maximum AOD (0.18), with July (0.27) and August (0.21) also reflecting increased aerosols, while November and December have the lowest maximum AOD values (0.04 and 0.05), indicating cleaner air. The mean AOD follows a seasonal trend, with the highest in July (0.24), followed by June (0.17) and August (0.14), and the lowest in November, December, and January (0.03, 0.03, and 0.04), reflecting lower aerosol levels. April shows the highest AOD range (0.15), indicating significant variability, while July shows the smallest range (0.04), indicating more stable conditions. The SD shows the greatest AOD variability in April (0.05), with minimal variation in July (0.02) and December (0.01). The CV highlights extreme variability in October (0.98), while July has the lowest CV (0.07), indicating more stability. Skewness values indicate distribution patterns; April has a strong negative skew (-1.31), suggesting lower AOD values are more frequent. September has strong positive skewness (1.23), indicating higher AOD values are more common. Moderate positive skewness is observed in July (0.76) and August (0.86), while March (-0.47) and May (-0.34) show negative skewness, indicating lower AOD values. Our results support previous work, which found that using MODIS-based observations recorded a strong association between cloudiness and AOD across many regions of the globe, particularly during the winter and summer seasons (Goto et al. 2019; Singh et al. 2020; Yoon et al. 2012). Overall, AOD levels demonstrate a seasonal trend, with higher concentrations and greater stability in mid-year (June, July, August), particularly in July (mean AOD 0.24), and lower levels with more variability at the start and end of the year (January, November, December), where mean AOD values drop to 0.03-0.04. The AI data demonstrates distinct seasonal variations throughout the year. January shows a minimum AI value of 0.32, while the highest minimum value occurs in August at 1.01, indicating elevated aerosol presence during the late summer. Maximum AI values peak in May at 1.86, reflecting a significant increase in aerosol activity, while the lowest maximum is recorded in January (1.23), suggesting relatively clearer conditions at the year's start. The mean AI values show a steady rise from January (0.68) to May (1.14), with a peak in June (1.19), followed by slightly lower values in July (1.07) and August (1.16). The lowest mean values occur in November (0.81) and December (0.72), reflecting reduced aerosol concentrations at the year's end. The SD is highest in November (0.41), indicating more variability in aerosol levels, while July (0.17) has the lowest SD, reflecting stable conditions. The range of AI values fluctuates, with the largest range in March (1.06) and the smallest in July (0.40), indicating more stable aerosol levels during mid-year. Such varied fluctuation has also been documented in different regions of the continents of Asia and Africa (Wang et al. 2014), and similar findings concur with a 10-Year Record of Aerosol Optical Properties and Radiative Forcing Over Three Environmentally Distinct AERONET Sites in Kenya, East Africa. Specifically, our observed summer AOD peak of 0.27 at LHAASO is comparable to pre-monsoon AOD values reported over the Tibetan Plateau. For instance, Xia et al. (2011) and Liu et al. (2015) reported pre-monsoon AOD values ranging between 0.20 and 0.30 at high-altitude stations on the Plateau, which is consistent with our observations. Similarly, studies over the Andes region (e.g., Segura et al. 2016) have documented seasonal AOD values in the range of 0.18–0.28, highlighting the influence of both local sources and long-range transport in mountainous terrains.

The contribution of different aerosol types to direct radiative forcing over distinct environments of Pakistan inferred from the AERONET data. The CV highlights that November exhibits the highest relative variability (0.51), while July (0.16) shows the lowest, indicating consistency in aerosol presence during mid-year. Skewness values generally indicate a positive skew, with May (1.04) showing the most substantial skewness, suggesting a higher frequency of elevated AI values, while June (0.24) and July (0.23) display the lowest skewness, indicating more balanced distributions of AI levels. Overall, these results show that aerosol activity increases during mid-year with higher stability, while the start and end of the year are characterized by more variability and lower concentrations. The mean MC of particulate matter shows significant seasonal variation across the year. Starting in January, the mean MC is relatively low at 5.83 × 10^{-6}kg m^{-2}, but it increases gradually through February (12.77 × 10^{-6} kg m^{-2}) and March (24.12 × 10^{-6} kg m^{-2}). A peak is observed in April (31.07 × 10^{-6} kg m^{-2}) and May (26.08 × 10^{-6} kg m^{-2}), suggesting higher moisture and particulate levels during the mid-year. Following this, a decline in MC is evident, particularly in June (13.51 × 10^{-6} kg m^{-2}) and July (7.06 × 10^{-6} kg m^{-2}), reflecting drier conditions. August (6.65 × 10^{-6} kg m^{-2}) and September (4.14 × 10^{-6} kg m^{-2}) experience further decreases, reaching their lowest values in September. However, the MC slightly rises again in October (4.80 × 10^{-6} kg m^{-2}) and November (6.57 × 10^{-6} kg m^{-2}) before falling to 4.88 × 10^{-6} kg m^{-2} in December. The range of values is widest in March (22.89 × 10^{-6} kg m^{-2}) and April (24.50 × 10^{-6} kg m^{-2}), indicating considerable fluctuation during these months, while the smallest range occurs in December (1.71), suggesting more consistent conditions. The SD also shows greater variability in March, April, and May, with values peaking at 8.14 μg m^{-3} in April. In contrast, the CV reflects a high degree of variability in February (0.37) and November (0.22), with the lowest in December (0.11), suggesting a more stable pattern towards the year's end. The skewness values indicate a generally positive skew, especially in February (0.83) and March (0.43), suggesting a tendency for higher MC values in these months, while a negative skew in December (-0.53) indicates a shift toward lower concentrations. This data reflects a clear seasonal fluctuation in MC, with the highest levels observed during the mid-year, followed by a decrease toward the year-end.

Aerosol Optical Depth

AOD is a measure of how much sunlight is blocked or scattered by particles such as dust, smoke, and pollution in the atmosphere (Akinyoola et al. 2024; Ruiz-Arias et al. 2016). Basically, AOD quantifies the extinction of solar radiation due to aerosol particles, indicating the extent to which direct sunlight is prevented from reaching the Earth's surface (Goto et al. 2019; Singh et al. 2020). The monthly AOD data from 2017 to 2023 were analyzed to assess the temporal variation of AOD in the study area, as shown in Fig. 4. The data reveals noticeable fluctuations in AOD levels throughout the year, across different years. In 2017, for example, AOD values start relatively low in January, gradually increasing to peak levels around March or April before declining toward the end of the year. However, the highest recorded AOD values occurred in June and July, with a subsequent decrease after October. This seasonal pattern is consistent across the following years, although the specific magnitude of AOD values varies slightly from year to year. In recent studies, it was found that AOD performs well in characterizing seasonal differences across the desert and urban/built environments across the world, where the AOD retrievals reveal underestimation in the desert area of Southeast Asia. Besides, the frequent reoccurrences of the dry/hot boundary layer driven by large-scale circulation patterns tend to lead to the AOD underestimation (Stirnberg et al. 2018).

In addition, monthly AOD values demonstrate a degree of consistency across the years, particularly in certain months. For instance, July consistently exhibits higher AOD values compared to other months, suggesting a seasonal trend or environmental factors that influence aerosol concentrations during this period. By analyzing the individual monthly AOD graphs, a detailed examination of AOD dynamics over time can be conducted, revealing significant variations and anomalies. Comparing AOD trends across different years reveals important patterns in long-term trends and seasonal variability of aerosol concentrations within the study area. The average annual AOD values range from approximately 0.097 to 0.120, indicating relatively consistent aerosol concentrations in the atmosphere. However, noticeable fluctuations in the maximum and minimum AOD values reflect the dynamic nature of aerosol distribution and atmospheric conditions. Notably, 2023 shows the highest average AOD value of

Table 1. Descriptive statistics of the studied monthly variables, namely, AOD, AI, and MC (10^{-6} kg m^{-2})

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
AOD	Min	0.03	0.05	0.11	0.03	0.05	0.13	0.22	0.11	0.11	0.00	0.03	0.03
	Max	0.06	0.10	0.17	0.18	0.13	0.19	0.27	0.21	0.15	0.07	0.04	0.05
	Range	0.03	0.05	0.06	0.15	0.08	0.07	0.04	0.10	0.04	0.06	0.01	0.02
	Mean	0.04	0.07	0.14	0.14	0.10	0.17	0.24	0.14	0.12	0.03	0.03	0.03
	SD	0.01	0.02	0.02	0.05	0.03	0.03	0.02	0.04	0.02	0.03	0.01	0.01
	CV	0.24	0.24	0.14	0.35	0.27	0.16	0.07	0.28	0.14	0.98	0.17	0.20
	Skewness	0.56	0.41	-0.47	-1.31	-0.34	-0.31	0.76	0.86	1.23	0.26	0.25	0.65
AI	Min	0.32	0.44	0.49	0.63	0.83	0.80	0.89	1.01	0.96	0.71	0.31	0.37
	Max	1.23	1.39	1.56	1.45	1.86	1.58	1.29	1.54	1.71	1.51	1.43	1.25
	Range	0.91	0.96	1.06	0.82	1.03	0.78	0.40	0.53	0.75	0.79	1.12	0.89
	Mean	0.68	0.80	0.89	0.95	1.14	1.19	1.07	1.16	1.26	1.07	0.81	0.72
	SD	0.31	0.35	0.37	0.31	0.35	0.29	0.17	0.20	0.26	0.30	0.41	0.33
	CV	0.45	0.44	0.41	0.33	0.31	0.25	0.16	0.17	0.21	0.28	0.51	0.45
	Skewness	0.61	0.56	0.67	0.55	1.04	0.24	0.23	0.96	0.45	0.38	0.36	0.57
MC	Min	4.49	7.95	14.43	18.46	18.43	8.02	3.92	4.34	3.28	3.94	5.22	3.88
	Max	8.32	22.04	37.32	42.97	34.46	19.13	10.68	10.49	5.21	7.19	9.21	5.60
	Range	3.83	14.09	22.89	24.50	16.04	11.11	6.76	6.15	1.93	3.24	3.99	1.71
	Mean	5.83	12.77	24.12	31.07	26.08	13.51	7.06	6.65	4.14	4.80	6.57	4.88
	SD	1.29	4.73	7.45	8.14	5.91	4.01	2.39	2.25	0.67	1.08	1.48	0.55
	CV	0.22	0.37	0.31	0.26	0.23	0.30	0.34	0.34	0.16	0.22	0.22	0.11
	Skewness	0.83	0.83	0.43	0.01	-0.02	0.07	0.21	0.51	0.26	1.47	0.67	-0.53

Note: Descriptive statistics of the studied monthly variables, namely AOD, AI, and MC, derived from combined satellite observations (MODIS, Sentinel-5P) and reanalysis datasets (MERRA-2). The table summarizes the mean, SD, minimum (Min), maximum (Max), and range, coefficient of variation (CV), and skewness for each variable.

0.12, along with a peak AOD value of 0.27. This suggests a potential increase in aerosol loading during that year, highlighting the need for further investigation into the factors contributing to this rise. Our results support the findings of previous studies conducted across many parts of the region. More than 30 critical fields concentrated on and directed across North America, Europe, and China have researched the science of tropospheric ozone, meteorological examples, forerunner emanations, and displaying endeavors (Solomon et al. 2000). Moreover, research on black carbon has zeroed in on BC and PM10 levels, close by air contamination files, in the locales of Beijing and Lhasa between January and December 2006 (Gao et al. 2007).

The year 2018 stands out with the lowest minimum AOD value of 0.003, indicating relatively clear atmospheric conditions during that period. Such variations in AOD extremes highlight the significance of closely monitoring aerosol dynamics to assess air quality and their broader environmental impacts. The annual summary of AOD trends provides essential insights into the temporal variability of aerosol concentrations in the study area, setting the stage for further analysis and interpretation in the subsequent sections. Distinct patterns in AOD levels emerge across the months, with periods like July showing higher average AOD values, suggesting elevated aerosol concentrations during this season. Conversely, months like October demonstrate lower average AOD values, indicating cleaner atmospheric conditions. Recognizing these seasonal fluctuations in AOD is critical for understanding air quality environmental impacts and informing effective mitigation strategies.

Aerosol Index

The AI is a qualitative measure indicating the presence of elevated aerosol layers with significant absorption, primarily from desert dust, biomass burning, and volcanic ash plumes. Analyzing monthly AI data from 2017 to 2023 presented in Fig. 5 reveals substantial variations in aerosol concentrations throughout the year, with peak values typically observed in the summer months of May, June, and July and lower values in winter, particularly January and February. Inter-annual variability is also notable, with the year 2020 exhibiting elevated AI values across several months. From 2017 to 2023, the annual average mean AI values range from 0.730 to 1.330, reflecting the average aerosol loading in the atmosphere each year. Seasonal changes, regional air pollution, and meteorological conditions can influence variations in these values. Notably, 2020 recorded the highest annual average AI

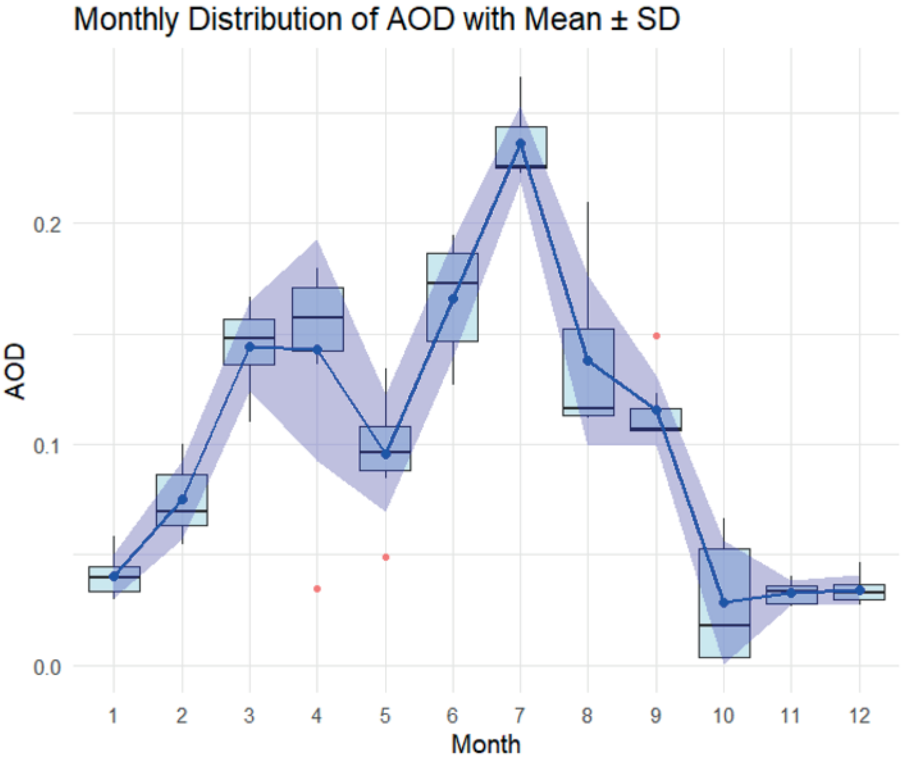


Fig. 4. Mean monthly distribution of AOD. The light blue boxplots represent the full monthly range, including outliers (red points), while the blue line and points indicate the monthly mean. The shaded blue area around the mean represents SD, highlighting the variability within each month

value of 1.330, suggesting elevated aerosol concentrations, possibly linked to increased anthropogenic activities or environmental events. Maximum AI values, which indicate the peak aerosol concentrations within a year, ranged from 1.028 to 1.855. The year 2019 had the highest maximum AI value of 1.474, marking a significant aerosol event. Minimum AI values ranged from 0.315 to 0.940, with 2022 showing the lowest minimum AI of 0.315, indicating a period of relatively cleaner air. The aerosol distributions during the interannual variability and seasonal changes could be attributed to winter large fractions of the Asian dust and Indian Ocean with a strong correlation with the physical atmospheric mechanism (Das Mahapatra et al. 2022; Gandham et al. 2022; Jin et al. 2021). However, such associated controlling phenomena can be attributed to dust emission from the dryland of the region, which is in agreement with previous studies (Duniway et al. 2019; Huang et al. 2017).

The calculated mean AI values across the months from January to December reveal distinct seasonal patterns. The winter months of January and February show lower mean AI values, ranging from 0.679 to 0.799, likely due to reduced biological and anthropogenic aerosol emissions and stable atmospheric conditions. In the spring, mean AI values gradually increase to a range of 0.888 to 1.138, driven by factors such as increased vegetation activity, dust storms, and biomass-burning events. Summer months, from June to August, see higher mean AI values, ranging from 1.138 to 1.187, which can be attributed to increased atmospheric instability, higher temperatures, and intensified human activity. During autumn, mean AI values remain elevated, ranging from 1.067 to 1.257, influenced by agricultural practices, regional wildfires, and atmospheric circulation patterns. In December, mean AI values slightly decrease to a range of 0.724 to 0.805, reflecting lower biomass burning and reduced anthropogenic emissions as temperatures drop and seasonal activities change.

Aerosol Extinction Coefficient

The results of aerosol extinction coefficients at various altitudes from 2017 to 2023 indicate fluctuations in aerosol concentration and distribution within the atmosphere over the study period. At lower altitudes, ranging from 1 km to 3 km, the aerosol extinction coefficients exhibit relatively low values, typically below 0.01 km-1. However, there are sporadic increases observed in certain years, particularly in 2020 and 2023, where slightly elevated values are noted, possibly attributed to localized aerosol events or meteorological conditions influencing aerosol transport and dispersion.

Moving to higher altitudes, between 4 and 7 km, the aerosol extinction coefficients show more pronounced variability across the years. Multiple time series extinction coefficients at various altitudes from 2017 to 2023 are seen in Fig. 6. In some instances, such as at 5 km altitude, significant increases in aerosol extinction coefficients are observed, notably in 2017 and 2019, suggesting the presence of elevated aerosol layers or enhanced atmospheric aerosol loading during those periods. Conversely, there are instances of decreased aerosol extinction coefficients in certain years, indicating temporal variability in aerosol distribution and concentration profiles within the mid-tropospheric region. At altitudes exceeding 7 km, aerosol extinction coefficients generally remain relatively low, with occasional minor fluctuations observed across the years. However, it's worth noting the substantial increase in aerosol extinction coefficients observed at 6 km altitude in 2018, primarily attributed to localized aerosol events or transient atmospheric phenomena impacting aerosol concentrations at that altitude level. The results highlight the dynamic aerosol distribution and concentration profiles driven by atmospheric dynamics, meteorological conditions, and aerosol emission sources.

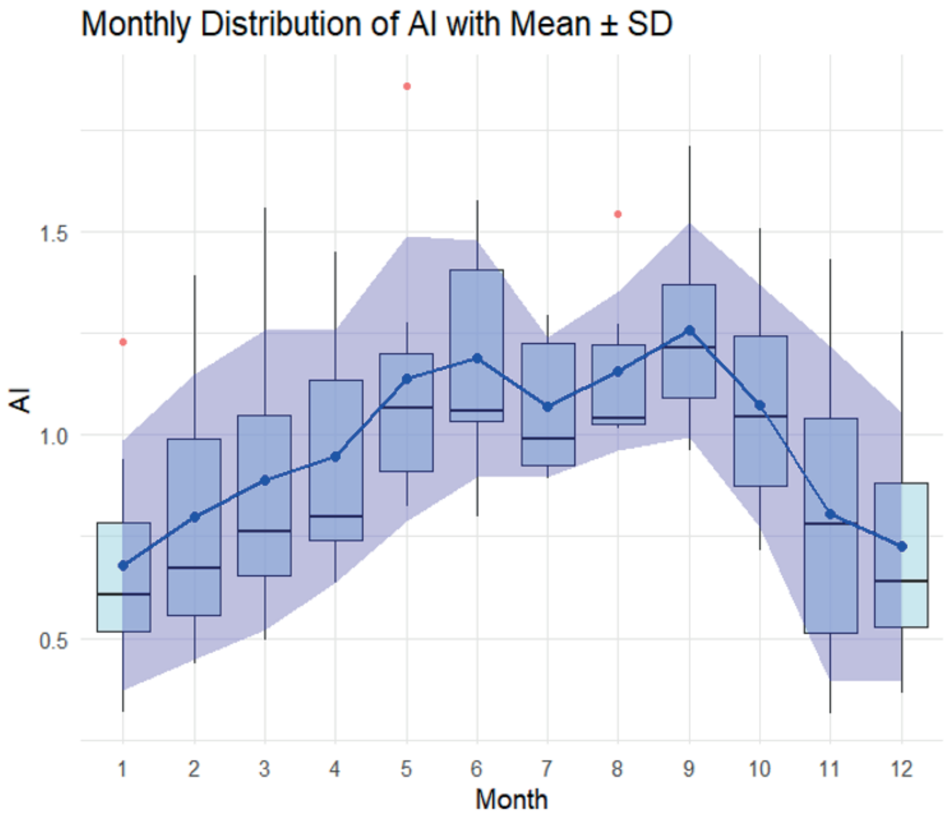


Fig. 5. Mean monthly distribution of AI. The light blue boxplots represent the full monthly range, including outliers (red points), while the blue line and points indicate the monthly mean. The shaded blue area around the mean represents SD, highlighting the variability within each month

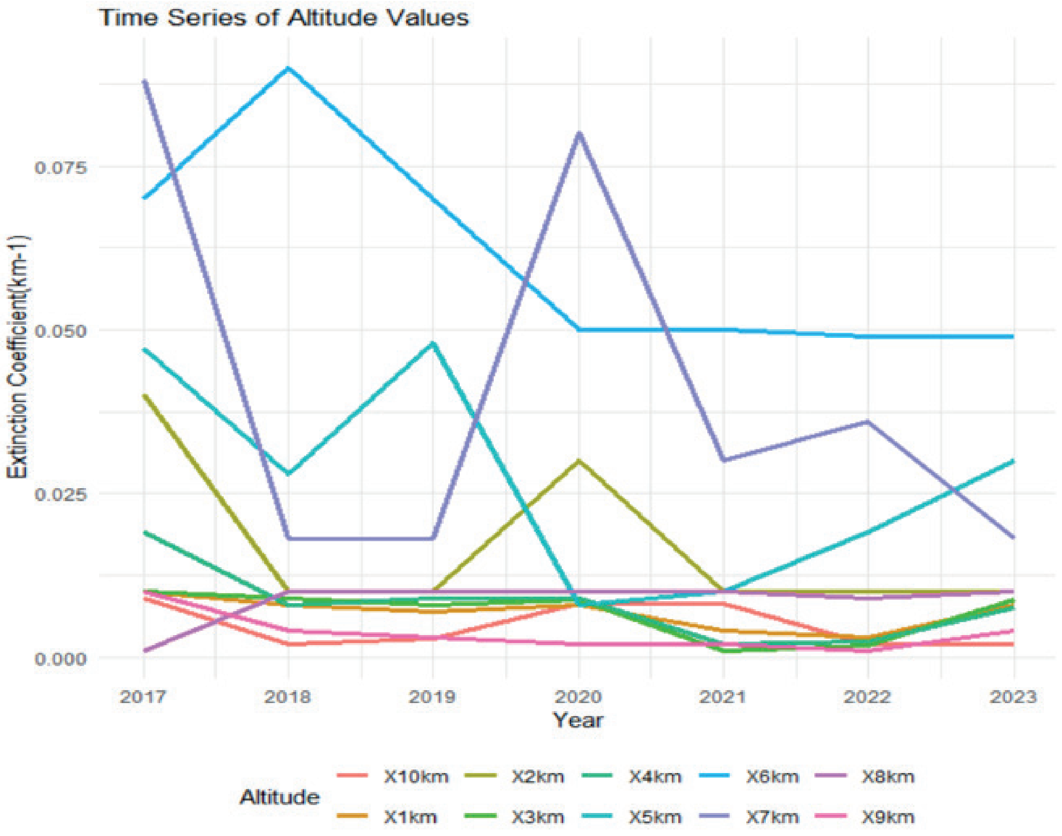


Fig. 6. Multiple time series extinction coefficients at various altitudes from 2017 to 2023

Mass concentration & Single scattering albedo

For the study period, the aerosol mass concentration exhibits notable fluctuations, indicative of changes in aerosol emissions, atmospheric transport, and deposition processes. In 2018, a significant increase in aerosol concentration was observed, with a value of 15.15 kg m^{-2} 10^{-6}, representing a peak in aerosol loading during the analyzed years. The increase in aerosol concentration indicates either stronger emissions or greater atmospheric transport of particles, likely driven by regional sources, meteorological factors, or human activities.

Conversely, a decline in aerosol concentration was noted in 2020, with a value of 9.49 kg m^{-2} 10^{-6}, representing a notable reduction compared to preceding years. On the other hand, the monthly averaged plot did not show a significant increase between 1980 and 2018. Along the North African coast, the values range from 0.24×10^{-6} kg m^{-2} month^{-1} (eastern Egypt) to 0.63×10^{-6} kg m^{-2} month^{-1} (southern Tunisia), while in marine areas, they vary from 0.26×10^{-6} kg m^{-2} month^{-1} (Balearic Islands) to 0.48×10^{-6} kg m^{-2} month^{-1} (Malta) (Bibi et al. 2020). This decrease in aerosol concentration may be attributed to changes in emissions patterns, atmospheric circulation patterns, or meteorological conditions leading to reduced aerosol transport and deposition within the study area. The monthly patterns of mass concentration and single scattering albedo over the LHAASO from 2017- 2023 are presented in Fig. 7. The subsequent years, 2021 to 2023, witnessed a moderate increase in aerosol concentration compared to 2020 but remained lower than the peak observed in 2018. These fluctuations in aerosol concentration highlight the complex interplay of various factors influencing aerosol dynamics, including emission sources, atmospheric circulation patterns, precipitation events, and atmospheric stability. Overall, the temporal variability in aerosol concentration underscores the dynamic nature of aerosol processes within the study area and emphasizes the importance of continuous monitoring and analysis to better understand the drivers and implications of aerosol variability on atmospheric composition, air quality, and climate dynamics (Li et al. 2024; Perumpully & Gautam 2024).

The analysis of SSA revealed consistent patterns across the study period from 2017 to 2023. The SSA values ranged from approximately 0.935 to 0.942, indicating the fraction of light scattered by aerosol particles relative to their total extinction. Over the years, there was a slight downward trend in the SSA values, with a gradual decrease observed from 2017 to 2023. This decline suggests potential changes in the composition or optical properties of aerosol particles in the atmosphere. Factors such as variations in aerosol sources, transport mechanisms, and atmospheric conditions could contribute to these temporal trends in SSA. The relatively stable SSA values across the study period indicate consistent aerosol scattering characteristics within the region. However, the observed decline over time may reflect shifts in aerosol composition or size distribution, which can influence their scattering properties (Kaufman et al. 1994). The results obtained in the present studies are similar to those of Pakistan using clustering techniques between different major cities of Pakistan, identifying distinct classifications of aerosol over the region.

Dust surface and Column density

The dust surface density, representing the mass of dust per unit area at the Earth's surface, exhibited a relatively stable pattern across the years, with values ranging from approximately 1.1×10^{-11} to 1.5×10^{-11} kg m^{-2}. Similarly, the dust column density, which indicates the total mass of dust per unit area within a vertical column of the atmosphere, displayed consistent values ranging from approximately 2.8×10^{-5} to 4.5×10^{-5} kg m^{-2}. The monthly patterns of dust column density are presented in Fig. 8. The temporal stability observed in both dust surface and column mass concentrations suggests a relatively constant presence of airborne dust particles in the atmosphere over the study period. Besides, high groupings of residue particles can lead to respiratory issues and increment non-inadvertent mortality (Pu & Ginoux 2018). However, minor fluctuations in dust density values from year to year may reflect variations in dust emission sources, atmospheric circulation patterns, and meteorological conditions. Factors such as wind speed, land surface properties, and regional climate dynamics can influence the generation, transport, and deposition of dust particles, contributing to the observed variations in dust mass concentrations. On the other hand, it has been observed that the long-term dust concentration measured over a specific region is attributed to weak correlations between the North Atlantic Oscillation and surface concentration during the winter season (Ginoux et al. 2004). Thus, such extended imitation highlights key information to explore the North Atlantic Oscillation effects on dust distribution rather than limited and/or affected regions at the edge of that phenomenon (Tao et al. 2017; Wang et al. 2007; Xie et al. 2011; Yang et al. 2022).

Correlation analysis and statistical evaluation of the studied variables

The Pearson and Spearman correlation analyses revealed distinct relationships between aerosol and studied variables from 2017 to 2023 (Fig. 9). For Pearson correlations, AOD exhibited moderate positive associations with SSA ($r = 0.571$), AI ($r = 0.292$), dust column density ($r = 0.274$), and

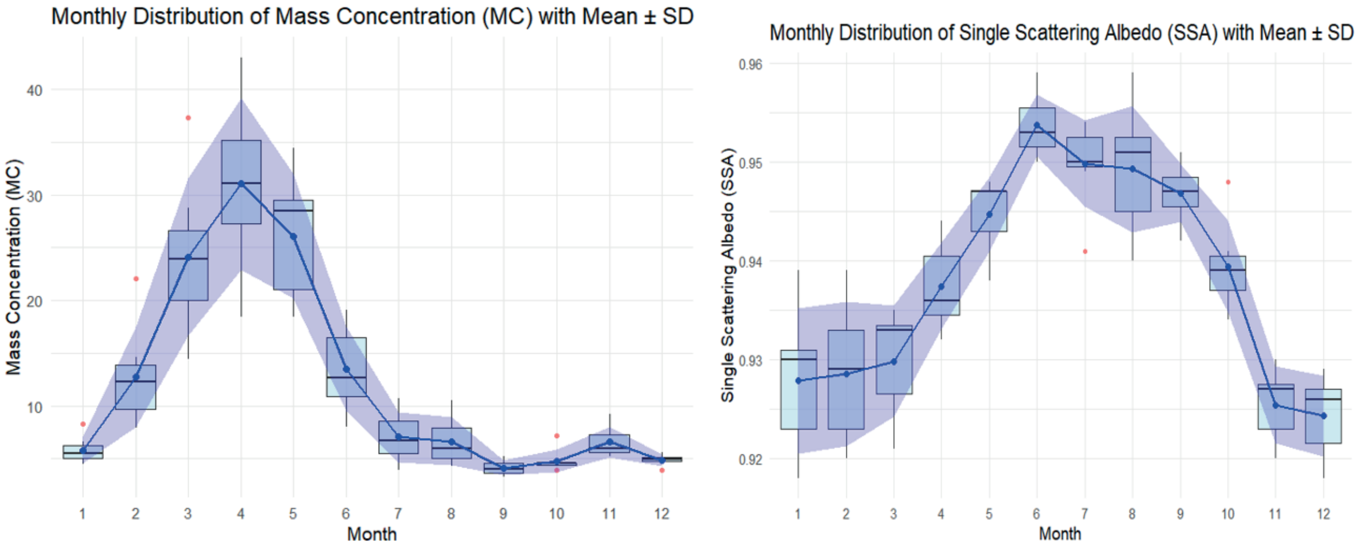


Fig. 7. Mean monthly distribution of MC (kg m^{-2} 10^{-6}) and SSA. The light blue boxplots represent the full monthly range, including outliers (red points), while the blue line and points indicate the monthly mean. The shaded blue area around the mean represents SD, highlighting the variability within each month

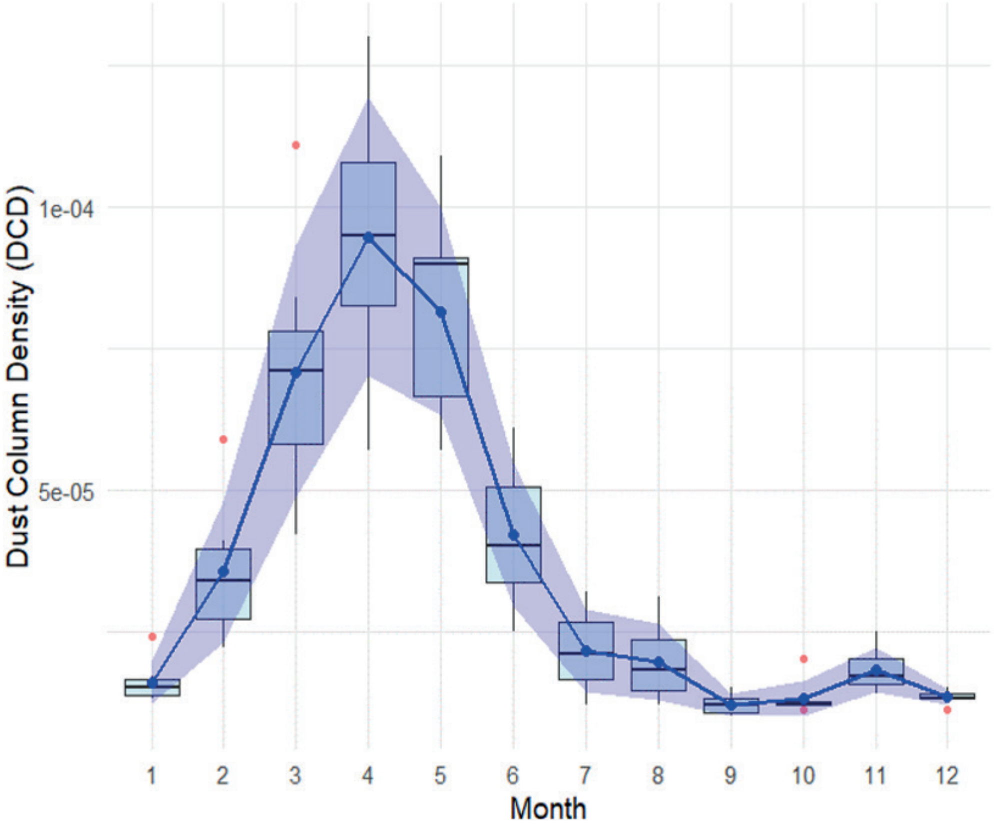


Fig. 8. Mean monthly distribution of dust column density (kg m^{-2}). The light blue boxplots represent the full monthly range, including outliers (red points), while the blue line and points indicate the monthly mean. The shaded blue area around the mean represents SD, highlighting the variability within each month

MC (r = 0.258), suggesting that higher AOD coincides with higher aerosol scattering efficiency and slightly increased mass and columnar dust. AI showed a moderate positive correlation with SSA (r = 0.399) but negligible correlations with MC (r = –0.028) and DCD (r = –0.011), indicating limited linear association between atmospheric aerosol presence and dust quantities. As expected, MC and dust column density were strongly correlated (r = 0.998), reflecting the nearly identical patterns of dust mass and column density in the study area. SSA displayed very weak correlations with MC (r = 0.001) and DCD (r = 0.034), indicating that aerosol scattering properties are largely independent of absolute dust concentrations.

Spearman rank correlations, which capture monotonic relationships, were largely consistent with the Pearson results but slightly stronger for some variables. AOD correlated positively with SSA (ρ = 0.593), MC (ρ = 0.354), dust column density (ρ = 0.394), and AI (ρ = 0.313), reflecting consistent rank-order associations. AI again showed a moderate positive correlation with SSA (ρ = 0.409) but weak negative correlations with MC (ρ = –0.090) and DCD (ρ = –0.069). The MC– dust column density relationship remained extremely strong (ρ = 0.995), confirming their near-identical monotonic behavior. SSA correlations with MC (ρ = 0.006) and DCD (ρ = 0.054) remained negligible, reinforcing that aerosol scattering efficiency is not closely associated with dust mass or column density. Overall, these results suggest that AOD and SSA are moderately related, whereas dust quantity variables (MC and dust column density) are strongly interrelated but largely independent of aerosol scattering properties.

Statistical tests were conducted to assess the variability and distribution characteristics of aerosol and dust-related variables from 2017 to 2023 as shown in Table 2. The Shapiro–Wilk test indicated that AI was normally distributed (p = 0.508), whereas AOD (p = 0.0017), dust column density (p < 0.001), MC (p < 0.001), and SSA (p = 0.0121) significantly deviated from normality, suggesting non-parametric methods were appropriate for further analysis. The Kruskal–Wallis test, examining month-wise variability, revealed that all variables except AI exhibited significant differences across months, with AOD, dust column density, MC, and SSA all showing p-values < 0.001, while AI demonstrated moderate variability (p = 0.0107). This confirms that aerosol optical depth, dust mass, dust column density, and scattering properties exhibit strong temporal variability on a monthly scale. Levene's test for homogeneity of variance showed that AI (p = 0.969), AOD (p = 0.330), and SSA (p = 0.685) had relatively stable variances across months, whereas dust column density (p = 0.0003) and MC (p = 0.0006) displayed significant heterogeneity, reflecting greater fluctuations in dust quantities compared to optical properties. Results indicate that AOD, dust column density, MC, and SSA significantly deviate from normality and exhibit substantial monthly variability, whereas AI shows moderate variability and homogenous variance. Collectively, these results provide robust statistical evidence that aerosol and dust characteristics vary significantly over time, with dust mass and column density exhibiting the most pronounced variability, while aerosol optical properties such as AI and SSA are comparatively more stable.

Relationship between aerosol extinction coefficient (EC) and terrain

A linear regression analysis was performed to assess the relationship between EC values and altitude from 1 to 10 km over the period 2017–2023 (Fig. 10) The regression model yielded an intercept of 0.0175 (±0.0057, p = 0.00298), indicating the expected EC value at sea level. The slope for altitude was 0.000185 ± 0.000917 (p = 0.841), suggesting a very weak and statistically non-significant increase in the

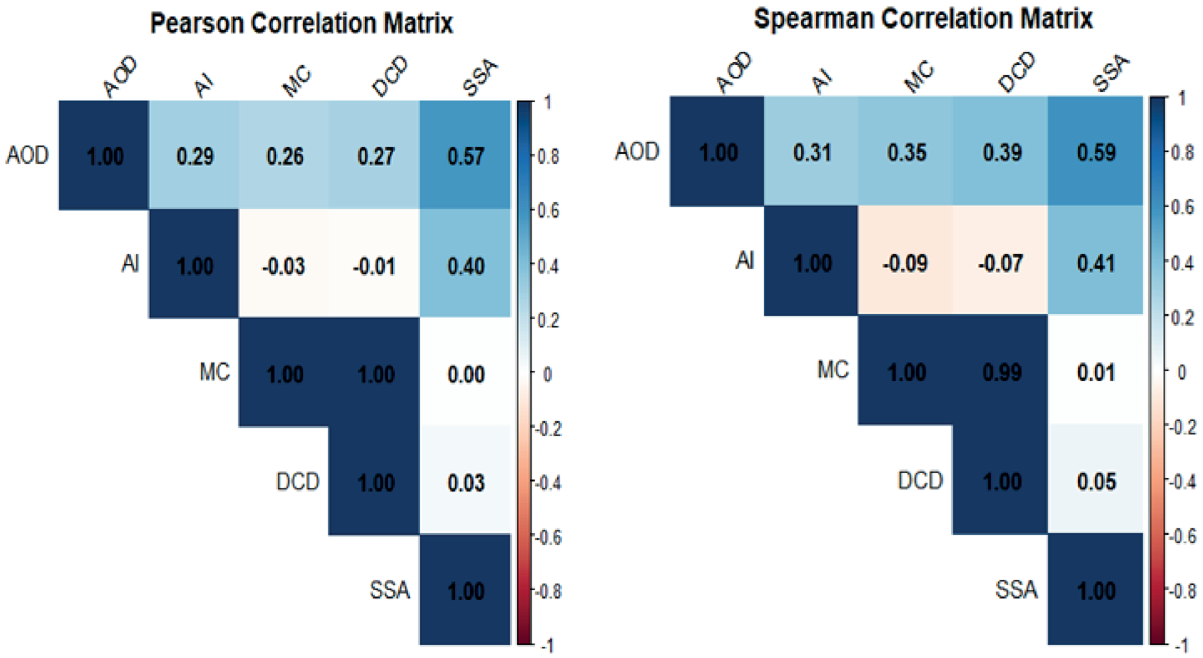


Fig. 9. Matrices showing pairwise correlations among aerosol and studied variables from 2017 to 2023. (a) Pearson correlation matrix, illustrating linear relationships between AOD, AI, MC, dust column density (DCD), and SSA. (b) Spearman rank correlation matrix, capturing monotonic associations between the same variables

Table 2. Summary of statistical tests for aerosol and dust-related variables from 2017 to 2023. Shapiro–Wilk test assesses the normality of each variable, with p < 0.05 indicating deviation from normality. Kruskal–Wallis test evaluates significant month-wise differences in variable values, with p < 0.05 indicating significant variability. Levene's test examines homogeneity of variance across months, with p < 0.05 indicating heterogeneity

Variable	Shapiro-Wilk test	Kruskal-Wallis test	Levene's test
AI	0.5078	0.01070	0.96910
AOD	0.0017	0.00000	0.32980
Dust column density	0.0000	0.00000	0.00030
MC	0.0000	0.00000	0.00060
SSA	0.0121	0.00000	0.68500

values with increasing altitude. The residuals ranged from –0.0182 to 0.0714, with a median of –0.0090, reflecting minor deviations from the fitted line. The model explained only a negligible fraction of the variance in EC values, with a multiple R-squared of 0.0006 and an adjusted R-squared of –0.0141. The F-statistic was 0.041 ($p = 0.841$), confirming that altitude does not significantly predict aerosol concentrations in this dataset. These results indicate that, over the study region, EC values are largely independent of altitude, and any terrain-related effects on vertical aerosol distribution are minimal within the observed altitude range.

The aerosol extinction coefficient values were further analyzed by stratifying the data into 2 km altitude bins to examine potential terrain effects (Fig. 11). The mean EC value in the lowest altitude bin (0–2 km) was 0.012 with a standard deviation of 0.0102 ($n = 14$), indicating relatively low and moderately variable concentrations. In the 2–4 km bin, the mean decreased to 0.00743 with a standard deviation of 0.00464 ($n = 14$), reflecting the lowest aerosol concentrations observed across all bins. Extinction coefficient values increased substantially in the 4–6 km bin, with a mean of 0.0441 and a standard deviation of 0.0234 ($n = 14$), representing the highest concentrations and greater variability, suggesting localized accumulation at mid-altitudes. The 6–8 km bin exhibited a mean EC value of 0.0249 with a higher standard deviation of 0.0267 ($n = 14$), indicating considerable temporal variability at these altitudes. Finally, the 8–10 km bin had the lowest mean EC value of 0.0043 with a standard deviation of 0.00312 ($n = 14$), highlighting minimal aerosol presence at higher altitudes. Overall, the stratified analysis demonstrates that aerosol concentrations peak at mid-altitudes (4–6 km) and decrease toward both lower and higher altitudes, with variability generally increasing at intermediate altitudes, suggesting that terrain and atmospheric processes may influence vertical aerosol distribution.

CONCLUSION

Seven years of satellite-based observations over LHAASO from 2017 to 2023 were used to analyze aerosol optical features AOD, AI, SSA, EC, MC, and dust surface and column density. This study concludes that peak AOD levels typically occur in the spring and summer when dust transport and biomass burning are most prevalent. AI values showed significant seasonal variation. SSA values suggest that aerosols were predominantly scattering in nature, with some absorption components. EC reflects variations in aerosol loading and composition over time. Mass concentration values also exhibited seasonal shifts, particularly during the dust season when higher concentrations were recorded, compared to lower concentrations during non-dust periods. Column densities

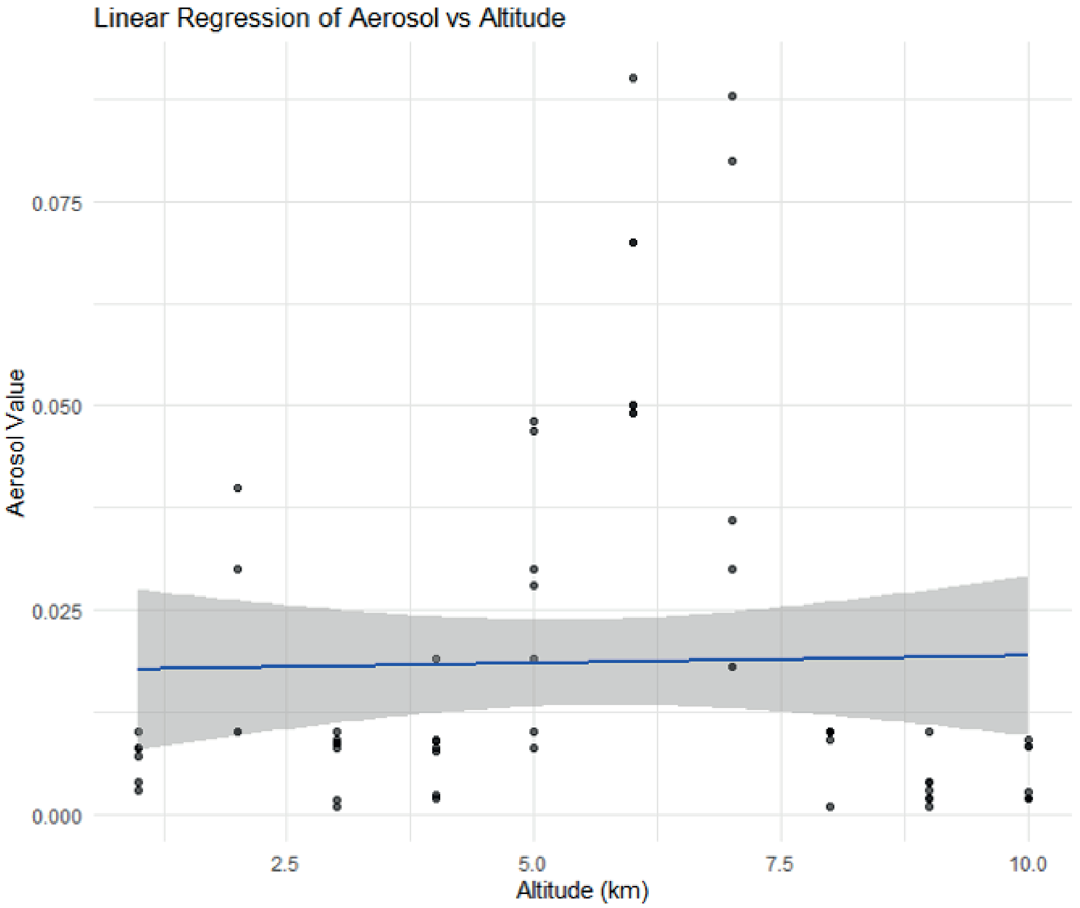


Fig. 10. Linear regression of aerosol extinction coefficient values against altitude (1–10 km) from 2017 to 2023. The regression line (blue) with a 95% confidence interval shows a non-significant relationship (slope = 0.000185, p = 0.841), indicating that aerosol concentrations are largely independent of altitude in the study region

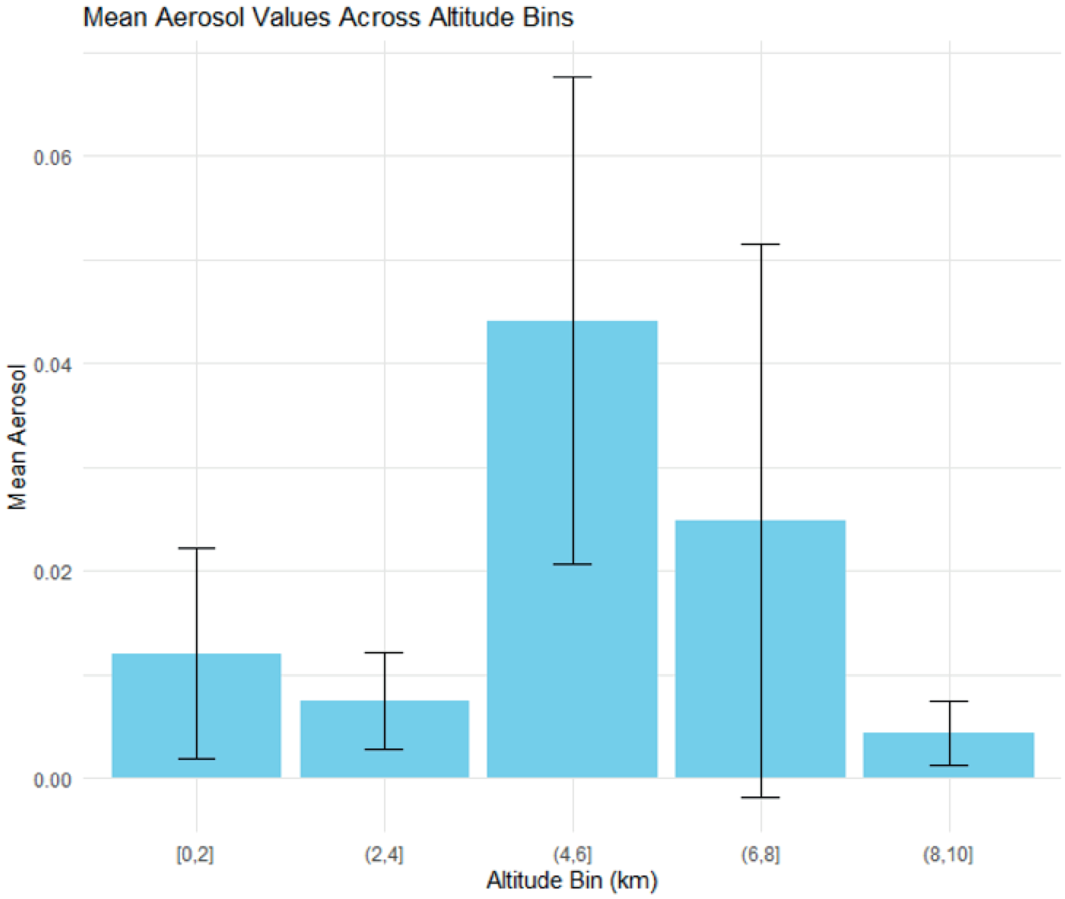


Fig. 11. Mean extinction coefficient values and standard deviations across 2 km altitude bins from 0 to 10 km over the period 2017–2023. Aerosol concentrations peak at mid-altitudes (4–6 km) and are lowest at both lower and higher altitudes, highlighting potential influences of terrain and atmospheric processes on vertical aerosol distribution

showed an increase of up to 30% during dust transport events, illustrating the influence of desert sources on local aerosol composition. These results enhance our understanding of local aerosol characteristics over LHAASO and their climatic implications, and they provide a basis for future improvements in regional climate modeling and air-quality assessments. ■

REFERENCES

Akinyoola, J. A., Oluleye, A., & Gbode, I. E. (2024). A Review of Atmospheric Aerosol Impacts on Regional Extreme Weather and Climate Events. Aerosol Science and Engineering, 8(3), 249-274.

Andrews E, Ogren JA, Bonasoni P, Marinoni A, Cuevas E, Rodríguez S et al (2011) Climatology of aerosol radiative properties in the free troposphere. Atmos Res 102(4): 365–393.

Bhowmik, R. C., Islam, S., Roy, S. K., Hridoy, M. A. A. M., Yeamin, M. B., Hoque, M. A., ... & Roshid, M. M. (2025). The Role of Hydropower, Clean Energy, Renewable Energy, FDI, and Capital Formation in Affecting CO2 Emissions and Environmental Sustainability in South Asia. Environment, Innovation and Management, 1, 2550021.

Bibi, H., Alam, K., & Bibi, S. (2016). In-depth discrimination of aerosol types using multiple clustering techniques over four locations in Indo-Gangetic plains. Atmospheric Research, 181, 106-114.

Bibi, M., Saad, M., Masmoudi, M., Laurent, B., & Alfaro, S. C. (2020). Long-term (1980–2018) spatial and temporal variability of the atmospheric dust load and deposition fluxes along the North-African coast of the Mediterranean Sea. Atmospheric Research, 234, 104689.

Bilal M et al (2019) Evaluation of Terra-MODIS C6 and C6.1 aerosol products against Beijing, XiangHe, and Xinglong AERONET Sites in China during 2004–2014. Remote Sens 11:486

Blaga, R., & Gautam, S. (2024). Improving PM10 sensor accuracy in urban areas through calibration in Timișoara. NPJ Climate and Atmospheric Science, 7(1), 268.

Boucher O, Randall D, Artaxo P, Bretherton C, Feingold G, Forster P et al (2013) Clouds and aerosols. In: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, p 571–657.

Das Mahapatra, P., Abhilash, S., & Ruchith, R. D. (2022). Aerosol distribution during spring and summer and its relationship with Inter-Annual variability of summer monsoon rainfall over Indian region. Geocarto International, 37(26), 11441-11455.

Delene DJ, Ogren JA (2002) Variability of aerosol optical properties at four North American surface monitoring sites. J Atmos Sci 59(6):1135–1150.

Duniway, M. C., Pfennigwerth, A. A., Fick, S. E., Nauman, T. W., Belnap, J., & Barger, N. N. (2019). Wind erosion and dust from US drylands: a review of causes, consequences, and solutions in a changing world. Ecosphere, 10(3), e02650.

Edenhofer O, Seyboth K (2013) Intergovernmental panel on climate change (IPCC).

Gautam S, Gautam AS, Singh K, James EJ, Brema J (2021) Investigations on the relationship among lightning, aerosol concentration, and meteorological parameters with specific reference to the wet and hot humid tropical zone of the southern parts of India. Environ Technol Innov. https:// doi. org/ 10. 1016/j. eti. 2021. 101414.

Gautam, S., Blessy, A., Abhilash, P., Yadav, A., & Justin, A. (2024). Exploring radiative forcing sensitivity to aerosol optical properties across varied geographical regions in India. Air Quality, Atmosphere & Health, 17(8), 1689-1700

Gandham, H., Dasari, H. P., Karumuri, A., Ravuri, P. M. K., & Hoteit, I. (2022). Three-dimensional structure and transport pathways of dust aerosols over West Asia. npj Climate and Atmospheric Science, 5(1), 45.

Gao, R., Niu, S., Zhang, H., Guo, J., Meng, D., Ma, J., ... & Zhang, Y. (2007, October). A comparative study on black carbon aerosol observations in regions of Beijing and Lhasa in 2006. In Remote Sensing and Modeling of Ecosystems for Sustainability IV (Vol. 6679, pp. 533-540).

Ginoux, P., Prospero, J. M., Torres, O., & Chin, M. (2004). Long-term simulation of global dust distribution with the GOCART model: correlation with North Atlantic Oscillation. Environmental Modelling & Software, 19(2), 113-128.

Gong W, Zhang M, Han G, Ma X, Zhu Z (2015) An investigation of aerosol scattering and absorption properties in Wuhan, Central China. Atmosphere 6:503–520.

Goto, D., Kikuchi, M., Suzuki, K., Hayasaki, M., Yoshida, M., Nagao, T. M., ... & Nakajima, T. (2019). Aerosol model evaluation using two geostationary satellites over East Asia in May 2016. Atmospheric Research, 217, 93-113.

Hridoy, M. A. A. M., & Paul, P. B. (2024). Assessing Total Dissolved Oxygen and Electrical Conductivity Using Sentinel-2 Remote Sensing: A Multivariable Approach to Flood Detection in Flood-Prone Urban area Sylhet.

Hridoy, M. A. A. M., Akter, P., Bordin, C., Acharjee, M. R., Baki, A. O., Neogi, S., ... & David, G. S. (2025). Integrated Assessment of Heavy Metal Contamination and Human Health Risks in Granitic Soils of South India: A Multi-Index Approach to Pollution and Ecological Impacts. Results in Surfaces and Interfaces, 100628.

Hridoy, M. A. A. M., Bordin, C., Masood, A., & Masood, K. (2025). Predictive modelling of aquaculture water quality using IoT and advanced machine learning algorithms. Results in Chemistry, 16, 102456.

Huang, J., Li, Y., Fu, C., Chen, F., Fu, Q., Dai, A., ... & Wang, G. (2017). Dryland climate change: Recent progress and challenges. Reviews of Geophysics, 55(3), 719-778.

Hu, X., Sun, J., Xia, C., Shen, X., Zhang, Y., Liu, Q., ... & Zhang, X. (2023). Measurement report: Rapid decline of aerosol absorption coefficient and aerosol optical property effects on radiative forcing in an urban area of Beijing from 2018 to 2021. Atmospheric Chemistry and Physics, 23(9), 5517-5531.

IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647.

Jacobson MZ (2002) Control of fossil-fuel particulate black carbon and organic matter, possibly the most effective method of slowing global warming. J Geophys Res Atmos 107(D19): ACH-16.

Jie Z et al (2017) Validation of MODIS C6 AOD products retrieved by the dark target method in the Beijing–Tianjin–Hebei urban agglomeration, China. Adv Atmos Sci 34:993–1002.

Jin, Q., Wei, J., Lau, W. K., Pu, B., & Wang, C. (2021). Interactions of Asian mineral dust with Indian summer monsoon: Recent advances and challenges. Earth-Science Reviews, 215, 103562.

Jing W, Lin S, Bo H, Bilal M, Wang L (2018) Verification, improvement and application of aerosol optical depths in China. Part 1: inter-comparison of NPP-VIIRS and Aqua-MODIS. Atmos Environ 175:221–233.

Kaufman, Y. J., Gitelson, A., Karnieli, A., Ganor, E., Fraser, R. S., Nakajima, T., ... & Holben, B. N. (1994). Size distribution and scattering phase function of aerosol particles retrieved from sky brightness measurements. Journal of Geophysical Research: Atmospheres, 99(D5), 10341-10356.

Kumar RP, Perumpully SJ, Samuel C, Gautam S (2022b) Exposure and health: a progress update by evaluation and scientometric analysis. Stoch Environ Res Risk Assess. https:// doi. org/ 10. 1007/ s00477-022-02313-z.

Lillis D, Cruz CN, Collet J Jr, Richards LW, Pandas SN (1999) Production and removal of aerosol in a polluted fog layer: model evaluation and fog effect on PM. Atmos Environ 33(29):4797–4816.

Li, Y., Wang, T., Wang, Q. G., Li, M., Qu, Y., Wu, H., ... & Xie, M. (2024). Deciphering the seasonal dynamics of multifaceted aerosol-ozone interplay: Implications for air quality management in Eastern China. Science of The Total Environment, 946, 174327.

Mahowald N, Albani S, Kok JF, Engelstaeder S, Scanza R, Ward DS, Flanner MG (2014) The size distribution of desert dust aerosols and its impact on the Earth system. Aeol Res 15:53–71.

Masood, A., Zhu, F. R., Sa'adi, Z., Mamun Hridoy, M. A. A., & Ullah, I. (2025). Study of meteorological parameters and classification of aerosols using remote sensing over LHAASO. Theoretical and Applied Climatology, 156(7), 375.

Park T, Liu MY, Wang TC, Zhu JY (2019) Semantic image synthesis with spatially-adaptive normalization. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, p 2337–2346.

Perumpully, S. J., & Gautam, S. (2024). Impact of aerosols on atmospheric processes and climate variability: A synthesis of recent research findings. Geosystems and Geoenvironment, 100317.

Pu, B., & Ginoux, P. (2018). Climatic factors contributing to long-term variations in surface fine dust concentration in the United States. Atmospheric Chemistry and Physics, 18(6), 4201-4215.

Perumpully, S. J., Gautam, S., & M, S. (2024). Evaluating the impact of personal exposure to emissions from sustainable commercial heating and cooking fuels on women in Rural Southern India and their alignment with sustainable development goals. Water, Air, & Soil Pollution, 235(1), 54.

Perumpully, S. J., & Gautam, S. (2025). Impact of aerosols on atmospheric processes and climate variability: A synthesis of recent research findings. Geosystems and Geoenvironment, 4(1), 100317.

Ramachandran S, Rengarajan R, Jayaraman A, Sarin MM, Das SK (2006) Aerosol radiative forcing during clear, hazy, and foggy conditions over a continental polluted location in north India. J Geophys Res Atmos. https:// doi. org/ 10. 1029/ 2006J D0071 42.

Ranjan RR, Joshi HP, Iyer KN (2007) Spectral variation of total column aerosol optical depth over Rajkot: a tropical semi-arid Indian station. Aerosol Air Qual Res 7(1):33–45.

Ren, X., Wu, J., Gong, C., Gao, W., Zhao, D., Ma, Y., & Xin, J. (2022). The relationship between PM2. 5 pollution and aerosol radiative forcing in a heavy industrial city, Taiyuan, in China. Atmospheric Research, 267, 105935.

Ruiz-Arias, J. A., Gueymard, C. A., Quesada-Ruiz, S., Santos-Alamillos, F. J., & Pozo-Vázquez, D. (2016). Bias induced by the AOD representation time scale in long-term solar radiation calculations. Part 1: Sensitivity of the AOD distribution to the representation time scale. Solar Energy, 137, 608-620.

Satheesh SK, Ramanathan V (2000) Large differences in tropical aerosol forcing at the top of the atmosphere and Earth's surface. Nature 405(6782):60–63.

Singh, P., Vaishya, A., Rastogi, S., & Babu, S. S. (2020). Seasonal heterogeneity in aerosol optical properties over the subtropical humid region of northern India. Journal of Atmospheric and Solar-Terrestrial Physics, 201, 105246.

Solomon, P., Cowling, E., Hidy, G., & Furiness, C. (2000). Comparison of scientific findings from major ozone field studies in North America and Europe. Atmospheric Environment, 34(12-14), 1885-1920.

Stirnberg, R., Cermak, J., & Andersen, H. (2018). An analysis of factors influencing the relationship between satellite-derived AOD and ground-level PM10. Remote Sensing, 10(9), 1353.

Tao, M., Chen, L., Wang, Z., Wang, J., Che, H., Xu, X., ... & Hou, C. (2017). Evaluation of MODIS Deep Blue aerosol algorithm in desert region of East Asia: Ground validation and intercomparison. Journal of Geophysical Research: Atmospheres, 122(19), 10-357.

Tian X, Liu Q, Li X, Wei J (2018) Validation and comparison of MODIS C6.1 and C6 aerosol products over Beijing. China. Remote Sens 10:2021.

Wang, L., Xin, J., Wang, Y., Li, Z., Liu, G., & Li, J. (2007). Evaluation of the MODIS aerosol optical depth retrieval over different ecosystems in China during EAST-AIRE. Atmospheric Environment, 41(33), 7138-7149.

Wang, P. X., Wang, B., Cheng, H., Fasullo, J., Guo, Z. T., Kiefer, T., & Liu, Z. Y. (2014). The global monsoon across timescales: coherent variability of regional monsoons. Climate of the Past, 10(6), 2007-2052.

Wang Y, Yuan Q, Li T, Shen H, Zheng L, Zhang L (2019) Evaluation and comparison of MODIS Collection 6.1 aerosol optical depth against AERONET over regions in China with multifarious underlying surfaces. Atmos Environ 200:280–301.

Xie, Y., Zhang, Y., Xiong, X., Qu, J. J., & Che, H. (2011). Validation of MODIS aerosol optical depth product over China using CARSNET measurements. Atmospheric Environment, 45(33), 5970-5978.

Yang, L., Tian, X., Liu, C., Ji, W., Zheng, Y., Liu, H., ... & Che, H. (2022). Evaluation and comparison of MODIS C6 and C6. 1 deep blue aerosol products in arid and semi-arid areas of northwestern China. Remote Sensing, 14(8), 1935.

Yoon, J., von Hoyningen-Huene, W., Kokhanovsky, A. A., Vountas, M., & Burrows, J. P. (2012). Trend analysis of aerosol optical thickness and Ångström exponent derived from the global AERONET spectral observations. Atmospheric Measurement Techniques, 5(6), 1271-1299.

Zhang MM, Liu ZB, Yun-Jian GE, Basin EY (2014a) Spatio-temporal distribution of atmospheric aerosol optical depth in Jiangsu Province.

Zhang M, Gong W, Zhu Z (2014b) Aerosol optical properties of a haze episode in Wuhan based on ground-based and satellite observations. Atmosphere 5:699–719.

Zhao, S., Hu, B., Du, C., Tang, L., Ma, Y., Liu, H., ... & Wang, Y. (2019). Aerosol optical characteristics and radiative forcing in urban Beijing. Atmospheric Environment, 212, 41-53.

Zhao, S., Liu, M., Tao, M., Zhou, W., Lu, X., Xiong, Y., ... & Wang, Q. (2023). The role of satellite remote sensing in mitigating and adapting to global climate change. Science of the Total Environment, 166820.

Check for updates

CC BY

RESEARCH PAPER

GEOSPATIAL MODELING OF WALKING ROUTES COMBINING ON-ROAD AND OFF-ROAD MOVEMENTS

Elena S. Zaslavskaya, Andrey M. Karpachevskiy*
Lomonosov Moscow State University, Geography faculty, Department of cartography and geoinformatics, Leninskie Gory 1, Moscow, 119991, Russia
***Corresponding author:** karpach-am@yandex.ru

Received: May 30th 2025 / Accepted: February 19th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4068

ABSTRACT. This paper presents a geographic information system (GIS)-based methodology for the automated planning of walking routes in areas with sparse road networks. This approach overcomes the limitations of traditional routing methods that rely solely on either vector road graphs or raster-based cost surfaces by combining both types of representations into a single flexible model. This combined model supports routing on both roads and off-road terrain, taking into account land cover, slope, and movement direction. A hierarchy of network edges prioritizes more passable segments, such as paved roads, while still allowing transitions to natural surfaces where necessary. The methodology has been implemented in ArcGIS Pro and tested using field data from the Satino educational station in the Kaluga region. Three types of regular grid-based network models (square with queen and knight adjacency, and hexagonal) were compared, and routing was simulated under different weather conditions using variable edge weights. The results demonstrate that the square grid with queen adjacency provides the best balance between accuracy and computational efficiency. Compared to the traditional least-cost path (LCP) method, the proposed approach reduces raster-induced distortions and allows for greater adaptability to user preferences. This methodology has practical applications in tourism, environmental research, and emergency route planning, particularly in regions with limited formal transportation infrastructure.

KEYWORDS: least cost path, irregular network model, land cover passability, hiking routes modeling, Satino

CITATION: Zaslavskaya E. S., Karpachevskiy A. M. (2026). Geospatial Modeling of Walking Routes Combing On-road and Off-road Movements. Geography, Environment, Sustainability, 1 (19), 130-138
https://doi.org/10.24057/2071-9388-2026-4068

ACKNOWLEDGEMENTS: The study was supported by the state task of the Department of Cartography and Geoinformatics of the Lomonosov Moscow State University “Digital technologies and methods of cartography, geoinformatics and remote sensing in geographical research”.

Conflict of interests: The authors report no potential conflict of interests.

INTRODUCTION

The traditional approach to solving routing problems is to model the road network as a weighted graph and apply algorithms to determine the shortest path (Miller, Shaw 2001). While effective for finding optimal routes within road networks, this approach is not suitable for calculating the shortest paths across continuous surfaces, such as geographic terrain. In geographic information systems (GIS), an alternative approach uses a raster to represent the cost surface in the form of a regular grid. Each cell in the raster is assigned a cost value, reflecting the difficulty of traversing that cell. These costs can be expressed in terms of money, time, or other units, taking into account factors such as travel time, distance, and safety. See also comparative algorithmic reviews in [Ashish 2021; Susanto 2021]

Finding a route with the minimum accumulated cost from the start cell to the target cell is known as the least-cost path (LCP) method. The cost raster is treated as a planar graph, with each cell center representing a vertex, and each node is connected to neighboring cells according to a predefined neighborhood pattern (Kourtz and O’Regan 1971; Tomlin 1990; ESRI 1991). The most common neighborhood patterns are the “knight”, “rook”, and “queen” patterns, with the “queen” and “rook” patterns being the most frequently used. Some studies suggest that the knight pattern is optimal for pathfinding (Choi 2013; Yu et al. 2010).

The LCP method is often affected by distortions caused by the raster structure. Paths may appear zigzagged due to turn-angle restrictions, particularly close to the starting point, which can introduce significant errors. These distortions are independent of raster resolution because the path in a regular model always proceeds from one vertex to an adjacent one in discrete increments, rather than traversing continuous space (Tomlin 1990). This type of deviation and elongation error was first formalized by Goodchild (1977). Post-processing can reduce these distortions; for example, the zigzag segments of the path can be simplified using the Douglas–Peucker algorithm (Choi et al. 2009).

Several studies have attempted to reduce the geometric distortions inherent in raster-based least-cost path modeling. Antikainen (2013) demonstrated that placement of nodes either at cell centers (CC) or along cell boundaries

(BC), as well as the choice of neighborhood structure, significantly influences path accuracy and computational performance. Subsequently, Seegmiller et al. (2020) built upon this approach by formulating an extended raster model in which nodes are positioned on cell boundaries. This improves topological consistency and mitigates the impact of raster-induced artifacts. Nevertheless, these approaches still operate strictly within the raster domain and do not fully resolve the limitations associated with integrating vector networks into continuous cost surfaces.

To address the issue of zigzagging, one option is to increase neighborhood connectivity, which minimizes the possible turning angles. However, this approach violates the theoretical principle of the shortest path, since a straight line passing through cells with varying costs is not necessarily the shortest. The optimal path should therefore follow the boundaries between cells with different values, in a manner similar to the laws of refraction in optics (Warntz 1957) and to later refinements of cost-surface refraction modeling (Tomlin 2010). The raster representation only approximates real-world landscapes, whose boundaries are difficult to define precisely. Consequently, cell boundaries may be disregarded.

Large neighborhoods create additional issues. Minimal turning angles may be blocked if an edge intersects a high-cost cell or an obstacle. Paths with a shared endpoint may also intersect incorrectly. Errors associated with traversing cells with varying costs and with intersecting paths can be addressed using a method proposed by Bemmelen et al. (1993). This method involves constructing a graph in which nodes are located along cell boundaries (boundary-connected, BC), rather than at cell centers (center-connected, CC). This extended raster configuration enables connections between nodes within the same homogeneous cell, thereby facilitating calculations and improving flexibility in route planning. However, this approach does not entirely eliminate distortions associated with raster structures.

The least-cost path method, which is based on conventional raster representation, is widely used and effective. The search for an optimal path is crucial not only for off-road route planning, as shown by Balstrøm (2002) and Rees (2004), but also in the design and construction of infrastructure, such as highways (Yu et al. 2010), railways, pipelines (Rylsky 2009), power lines (Novakovsky et al. 2017; Bagli et al. 2011), channels (Collischonn & Pilar 2000), and other transportation systems.

However, a separate issue arises when linear features, such as roads and waterways, are included in a raster model. While these features can be converted to a raster format and incorporated into a cost surface using overlay operations, this approach has limitations. Notably, integrating vector objects into a regular model can make it challenging to distinguish between different types of intersection (e.g., bridges, overpasses, and road crossings). Consequently, topological relationships cannot be preserved correctly (Choi et al. 2013). Accurately representing vector networks and their properties in a raster-based model is important because the optimal path often depends on the location of existing linear network features. For instance, waterways can restrict movement, whereas road networks usually provide low-cost travel routes. Anisotropic costs must also be considered, as these account for changes in edge costs based on movement direction, which is often influenced by slope orientation and steepness.

The scientific literature offers few approaches to solving the problem of combining irregular and regular network models. While Choi et al. (2013) propose a method that integrates these two classical models, their algorithm only permits access to or exit from a vector network at graph vertices and not at any point along the edges. Overcoming this limitation could make the integrated methodology more widely applicable.

Many geospatial studies have successfully applied algorithms to determine optimal paths based on both regular and irregular network models. However, the challenge of developing an efficient method for integrating these models while preserving their respective advantages remains unresolved. Existing routing approaches demonstrate that modeling walking routes in sparsely connected areas requires a method that integrates the two models and allows flexible parameterization based on user needs. This study aims to develop and test a methodology that integrates raster-based cost surfaces with vector road network data in order to simulate realistic walking routes in sparsely connected areas while accounting for varying environmental conditions and user-defined priorities.

MATERIALS AND METHODS

The initial data for creating the network model include the road network, field track-data, and remote sensing imagery used for surface segmentation. The continuous terrain surface was first segmented into land cover units with different passability conditions based on image interpretation. These units were subsequently assigned walking speeds derived from field observations, forming a passability map representing terrain suitability for transportation or walking (Pokonieczny 2020). The passability map based on satellite imagery can be created in two ways: (1) through automated classification or (2) through expert interpretation. To identify different land types in our case, we used a mosaic of ultra-high-resolution satellite scanner images available through the ESRI map viewer online service for expert interpretation.

Table 1 was used as the basis for identifying potential land cover types. The classification follows a functional approach to passability mapping, where surface categories are distinguished according to vegetation structure, obstacle presence, and expected pedestrian mobility conditions. The selected classes reflect both established approaches to terrain suitability assessment and local field knowledge of movement conditions within the study area, which, together, justify their use for routing impedance modeling. Initially, the entire area was divided into large and homogeneous groups based on the surface or the prevailing type of vegetation: open ground areas characterized by the presence of a predominantly grassy layer; forests; sparsely populated land; bodies of water; and settlement areas.

The next stage involved dividing the main groups into more specific categories, as some areas had significant variations in the dominant type of flora and its density. For instance, logging has been carried out in the study area for several years, resulting in a significant number of old and relatively recent cuttings with various structures and at different stages of restoration. Therefore, additional land cover categories were required. The mapped objects vary widely in terms of vegetation type, density, and succession stage, all of which affect passability. This is determined by the type of vegetation cover, the presence and nature of secondary vegetation tiers, and the degree of density of the stand.

An average travel speed is assigned to each identified land cover type and road. This determines the weight for travel along edges associated with that area or network

Table 1. Classification of land types by passability conditions for Satino territory

Land type		Pedestrian Passability	Description
Open areas	Meadow	From moderate to high	Open areas with low herbaceous vegetation. Movement is facilitated by the absence of dense tree and shrub cover. Mobility may be somewhat hindered by surface irregularities resulting from plowing. Passability becomes more difficult during and after rain.
	Arable land		
	Fallow land		
	Wasteland		
Forest areas	Deciduous forest	Variable (from moderate to very low)	Passability conditions vary depending on the presence and density of understory vegetation as well as the dominant tree species. The understory of deciduous forests is often denser than that of spruce forests, which can make movement more difficult. However, spruce forests have lower branches that create more problematic obstacles. The passability of mixed forests depends on the dominant vegetation type and its density. Birch groves and pine forests are easier to pass through due to their less dense undergrowth and thinner trunks. Movement becomes more difficult during and after rain, and strong winds may increase the risk of falling branches and trees, especially in spruce forests, which further complicates pedestrian movement.
	Spruce forest		
	Mixed forest		
	Pine forest		
	Birch grove		
Areas with other woodland communities	Clearcuts	Variable (from moderate to very low)	Although fresh clearcuts may offer moderate passability due to the absence of a tree canopy, movement can be hindered by logging debris, such as uprooted stumps and roots. Over time, secondary succession occurs, leading to the emergence of new dense and difficult-to-pass young growth. Passability at this stage depends on the stage of succession. Floodplain vegetation often includes tall and dense herbaceous cover and woody plants, resulting in moderate passability for pedestrians. As with other types of terrain, rainfall reduces passability due to increased soil moisture.
	Shrub thickets and young growth		
	Floodplain vegetation		
Water bodies	Rivers, streams, lakes, ponds	From very low to impossible	While movement across bogs is extremely difficult, it is possible. In general, bodies of water are obstacles to pedestrian movement. However, in winter, these waterways and bogs become accessible.
	Bogs		
Modified areas	Settlements and individual building complexes	Impossible (inaccessible) and adaptive	Movement is possible along the existing road-path network.
	Industrial zones		
	Quarries and extraction sites		

segment. These speeds can be calculated analytically, based on territorial knowledge, or empirically, using field data obtained from satellite navigation receivers. In our case, we used the second approach – the resulting table of speeds can be found in the Results section. The resulting network model combines a regular grid-based surface model with a road network model. The spatial resolution of this model is determined by the regular point grid step.

According to the chosen neighborhood pattern (e.g. queen or knight), each vertex is connected to neighboring cells. These edges form a non-planar graph, with geometrically identical duplicates excluded. The impedance of each edge is calculated based on its length and the time difference in seconds between the connected vertices. Additionally, a hierarchy of edges is established to reduce computation time and limit the search for paths by prioritizing higher-level edges. This approach makes it easier to determine more convenient routes, including transitions between roads and unpaved surfaces when necessary. The hierarchy is algorithmically implemented by assigning each edge a priority class and multiplying its base impedance by a hierarchy coefficient. During pathfinding, higher-priority edges have lower effective costs, forcing the search algorithm to explore them first unless there is a substantially shorter alternative outside the hierarchy.

Route planning also accounts for terrain irregularities by incorporating elevation and movement direction using a digital elevation model (DEM). We used a DEM from the Satino local educational geoinformation system with 5-meter resolution. We used the hiking exponential Tobler function (Tobler 1993), which models speed based on slope angle, to calculate a terrain-based coefficient.

$$v(\alpha) = 6e^{-3.5 \times \left| tg(\alpha) + 0.05 \right|} \quad (1)$$

According to this formula, the maximum speed is achieved at an angle of inclination of $\alpha=-2.86^0$, while an increase or decrease in this angle results in an exponential decrease in speed.

Obstacles such as water bodies, fenced areas, and buildings are marked as inaccessible, and edges crossing them are removed from the graph. Weather conditions are factored in by adjusting edge weights with additional coefficients.

This methodology was tested at the Satino educational field station in the Borovsky district of the Kaluga region. A set of field data was collected using a satellite navigation receiver. Three network models were initialized based on the passability surface to represent different regular lattice configurations and connectivity schemes commonly used in raster-based path modeling. The square grid with queen adjacency (8 neighbors) was used as a conventional baseline configuration; the knight adjacency (16 neighbors, including queen adjacency and L-shaped neighbors) (Yu et al. 2010) was included to increase directional flexibility and angular resolution of movement, while the hexagonal grid represents an alternative lattice geometry often considered to reduce directional bias. Together, these models enable comparison of routing behavior under varying structural properties of the regular network (Fig. 1).

For each network model, three experimental routes were created using different parameters, based on the modified Dijkstra's algorithm, including edge hierarchy (ESRI 2005). Similar start and end points were also used for routing, which was performed using the least-cost path (LCP) method on a raster cost surface. Weather influence modeling was tested by applying additional coefficients to edge weights. The start and end points were located on opposite sides of a river, enabling the analysis of various crossing options under different conditions.

RESULTS

Three routes were created based on the assumption of comfortable weather conditions: cloudy weather without precipitation and a temperature of +17°C. To compare the results, similar routes were generated using the traditional least-cost path (LCP) method. In this method, the cost surface is a rasterized representation of different terrain types, including a rasterized road network. The cost of each raster cell represents the time it takes to traverse that area. The resolution of the raster surface is 10 meters, similar to the modeled one (Fig. 2).

An analysis of the generated routes revealed significant differences in the modeling results, depending on the initial conditions. One key factor influencing these differences is the parameter related to the hierarchy of network elements. Routes generated based on the hierarchical model prioritize the road network when searching for paths. The algorithm selects the shortest route from the starting point to the nearest road and follows it for as long as possible. This results in a more convenient path, as people generally prefer to travel on roads and avoid natural obstacles. In contrast, routes generated using a non-hierarchical model

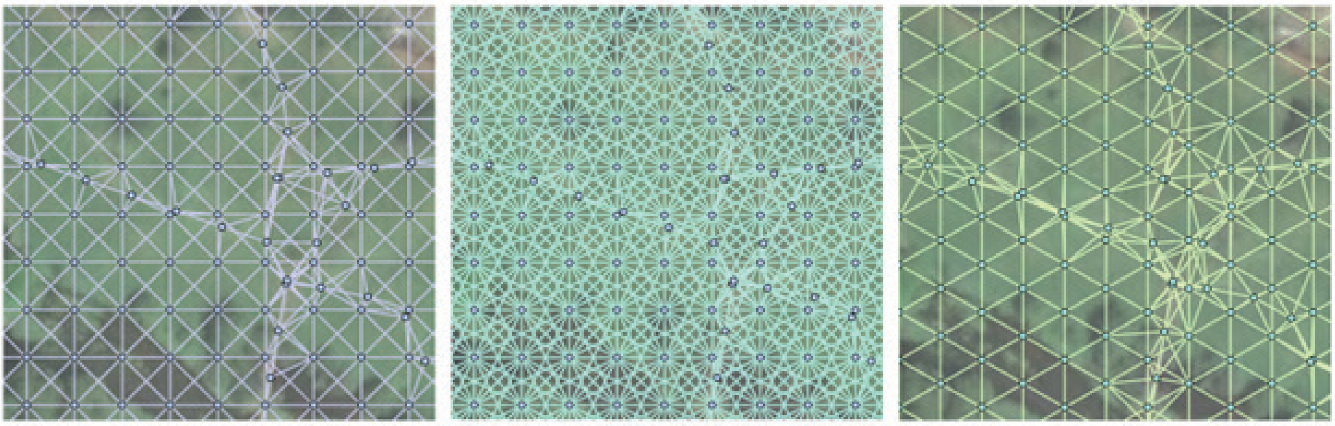


Fig. 1. Types of regular network patterns (queen, knight, hexagon)

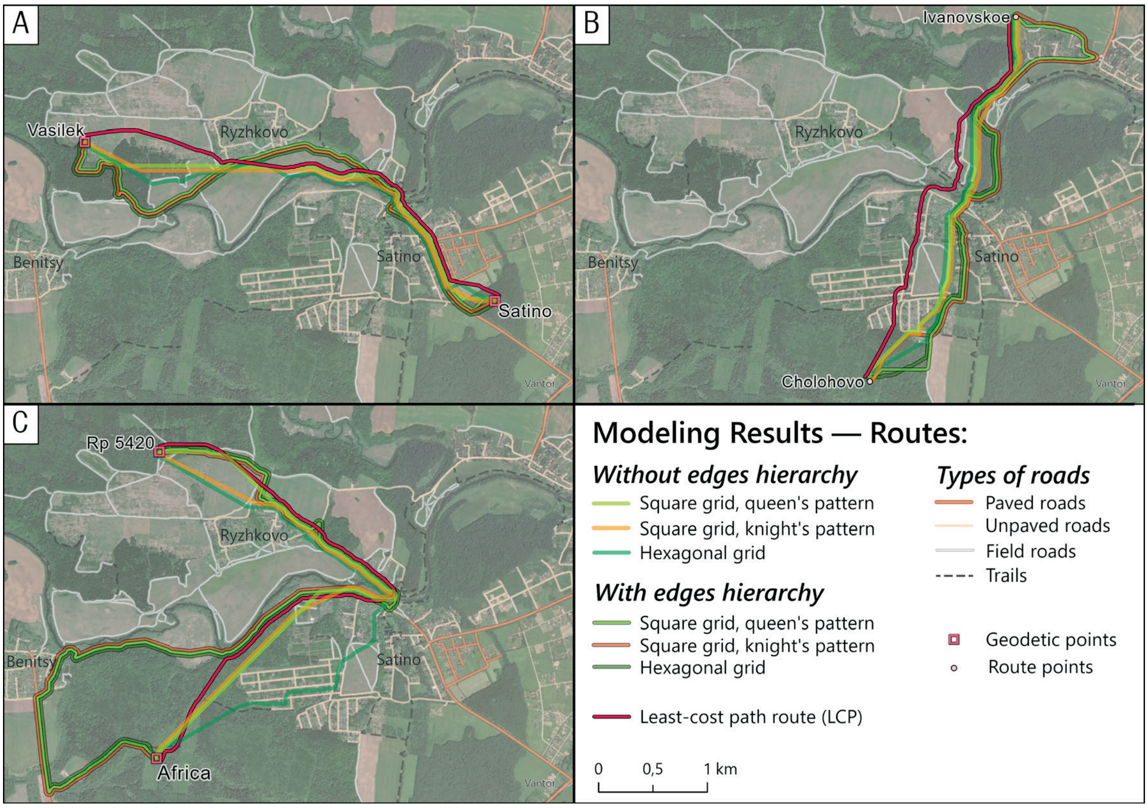


Fig. 2. Modeling results with different parameters of the network for three different routes: A) Satino – Vasilek, B) Cholohovo – Ivanovskoe, C) Africa – RP 5420

focus more on minimizing total path length or optimizing movement time, which may not always prioritize the road network.

Routes generated by the LCP method are conceptually similar to those produced by the non-hierarchical model. However, LCP results depend strongly on raster resolution, as seen in our experiments. For example, the LCP-generated path suggests crossing a river directly instead of using a nearby bridge, because the river is too narrow to be represented in the raster at its given scale, leading to a misinterpretation by the model. Furthermore, the LCP method lacks the ability to assign hierarchical importance to individual raster pixels, which highlights the advantage of the proposed non-hierarchical approach over LCP.

Various weather conditions were modeled using hierarchical network models. Edge weights were adjusted using coefficients to simulate different scenarios (see Table 2). Two scenarios were tested: rainy weather at +7°C and extremely hot weather at +32°C. In the extreme heat scenario, the use of a ford at a designated location was permitted. The edge weight coefficients were defined analytically for each scenario, based on the assumption that movement speed generally decreases on most surface types and roads, as the energy required to move increases under these conditions (Fig. 3).

Route modeling under these scenarios shows that, in rainy conditions, both the hexagonal and square grids with queen adjacency patterns produce shorter routes through forested areas and the road network. Meanwhile, the square grid model with knight adjacency produced a longer route around forested areas initially, prioritizing use of the road network wherever possible.

Rainy weather produced the most significant variation in route options. In contrast, the routes were almost identical in hot weather because users tended to choose the ford, which is closer than the bridge. To evaluate the practicality of the proposed routes, we conducted a survey of individuals familiar with the area. They were asked to design their preferred routes based on the same start and end points used in the modeling. Analyzing the results, we found that the preferred routes were similar to the modeled ones: in hot weather, people tended to take a shorter route by using the ford, which is closer to the starting point than the bridge (Fig. 4).

When we compared different network models based on a regular grid structure, we found that the square configuration had certain advantages over the hexagonal one. This is because significant geometric distortions are absent from the resulting routes, which are often seen in the hexagonal model. These distortions arise from the grid's connectivity type, resulting in a zigzag effect, similar to that seen in the least-cost path (LCP) method, where raster structure causes artificial elongation and increased path costs. In terms of adjacency patterns, the knight pattern offers 16 possible movement directions compared to 8 for the queen pattern. In most weather scenarios, except rainy conditions, the knight pattern produced routes that were slightly more advantageous in terms of travel time. This advantage is due to its connectivity structure, with longer edges, as the knight pattern extends into a 5×5 neighborhood, whereas the queen pattern extends into a 3×3 neighborhood. This allows for more diverse path options.

To complement the qualitative assessment, we performed a quantitative comparison of route similarity across all modeling configurations. Fig. 5 shows pairwise percentages of spatial identity between routes generated using different grid structures, adjacency patterns, and hierarchical settings. Models with edge hierarchy demonstrate very high internal consistency (97–100%) across all three test routes, while agreement between hierarchical and non-hierarchical models drops to 13–36%, reflecting their different movement priorities. The LCP paths exhibit only 11–35% spatial overlap with hierarchical models, confirming that raster-based routing deviates substantially from more realistic on-road/off-road transitions.

However, these advantages are offset by the data and processing time required for the knight pattern. It requires

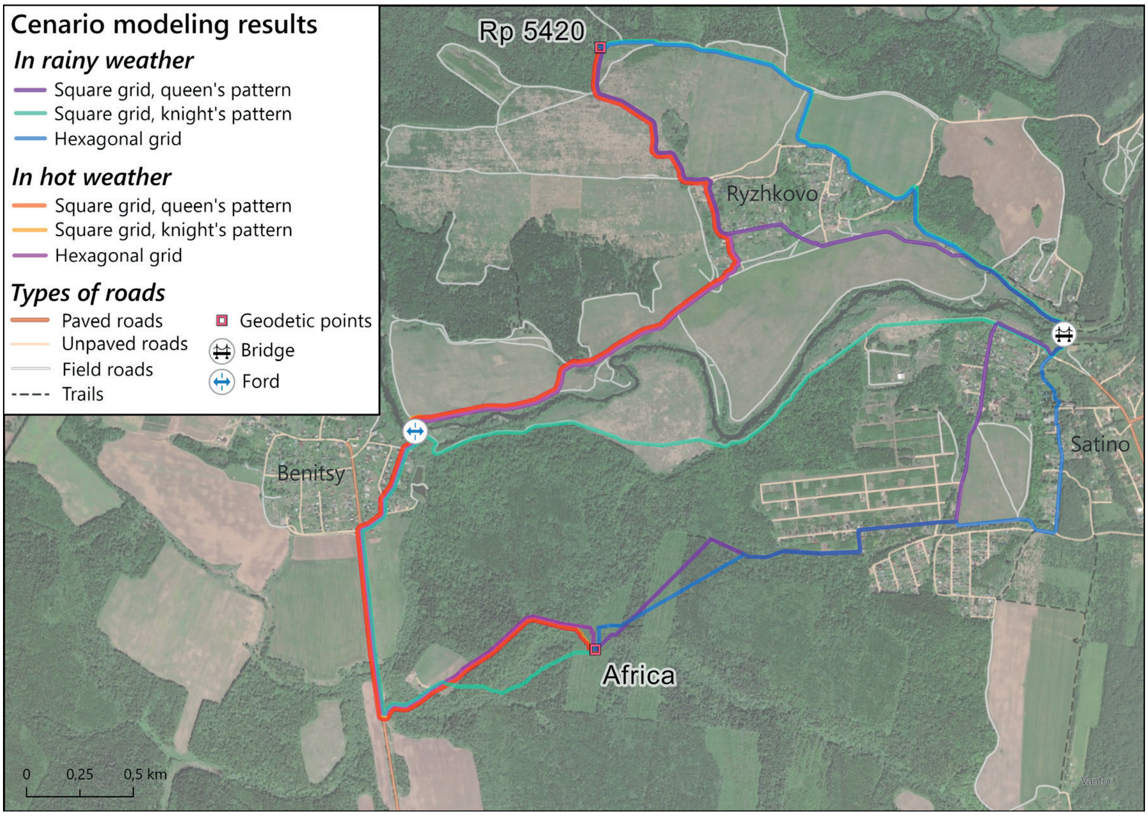


Fig. 3. Resulting routes in different scenario models and grid patterns

Table 2. Average walking speeds and weather coefficients for different surface types based on the field experimental data

Surface Type	Average Speed, km/h	Rainy Weather Coefficient, t = 7°C	Extremely Hot Weather Coefficient, t = 32°C
Scenario-based modeling on road network			
Asphalt roads	4,21	0,9	0,85
Dirt roads	4,13	0,85	0,85
Field roads	4,10	0,7	0,8
Trails	3,82	0,6	0,8
Scenario-based modeling on terrain surface			
Meadow	3,53	0,6	0,7
Ploughed field	3,64	0,5	0,7
Fallow land	3,56	0,5	0,7
Mixed forest	2,78	0,7	0,95
Deciduous forest	2,49	0,65	0,9
Spruce forest	2,96	0,9	0,95
Dense mixed forest	2,55	0,55	0,8
Dense deciduous forest	2,26	0,5	0,7
Clear-cut area 1	2,98	0,7	0,8
Clear-cut area 2	2,15	0,6	0,7
Clear-cut area 3 (mixed)	2,08	0,4	0,65
Clear-cut area 3 (deciduous)	2,01	0,4	0,6
Clear-cut area 3 with dense deciduous undergrowth	1,97	0,4	0,5
Undergrowth 1	3,49	0,7	0,7
Undergrowth 2	3,03	0,5	0,7
Undergrowth 3	2,01	0,4	0,5
Floodplain vegetation	2,78	0,4	0,7
Water body	Movement prohibited	Movement prohibited	Ford crossing possible
Swamp	2,06	0,3	0,8
Gully bottom	1,99	0,4	0,9
Garden plots (Dacha zone)	Off-road movement prohibited	Off-road movement prohibited	Off-road movement prohibited
Quarry	1,84	0,4	0,7

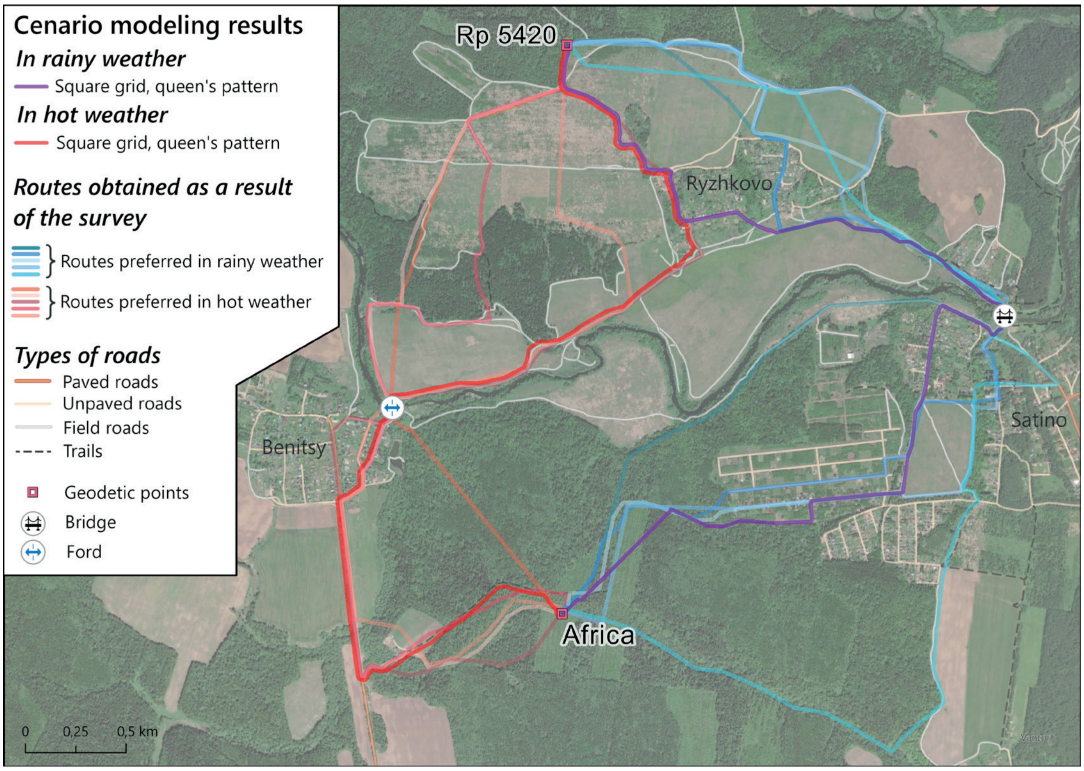
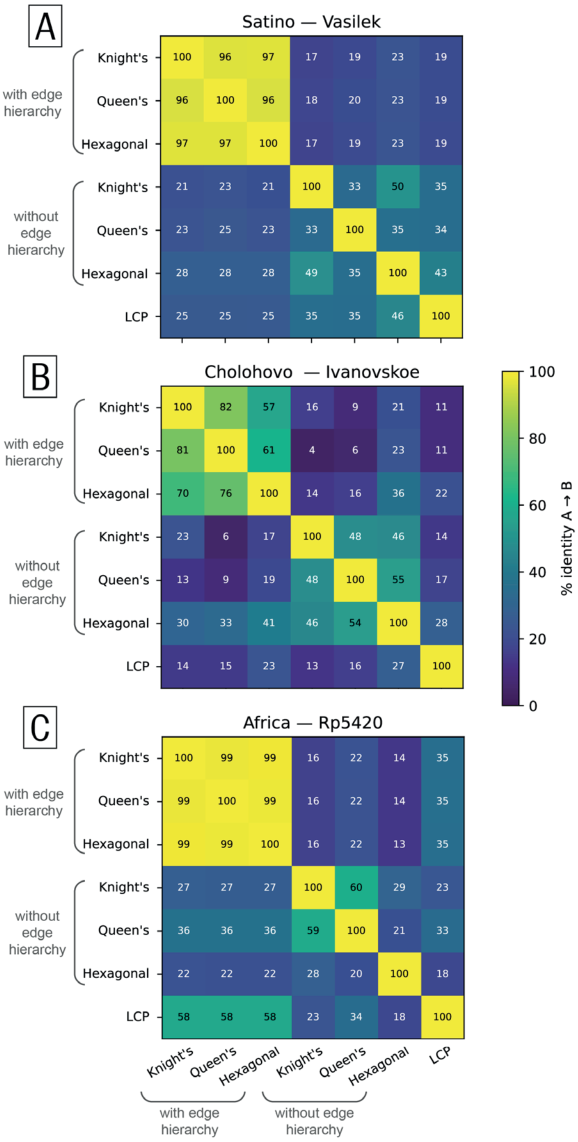


Fig. 4. Resulting routes in different scenario models and social survey resulting routes



A Satino — Vasilek

		Knight's (with edge hierarchy)	Queen's (with edge hierarchy)	Hexagonal (with edge hierarchy)	Knight's (without edge hierarchy)	Queen's (without edge hierarchy)	Hexagonal (without edge hierarchy)	LCP
with edge hierarchy	Knight's	100	96	97	17	19	23	19
	Queen's	96	100	96	18	20	23	19
	Hexagonal	97	97	100	17	19	23	19
without edge hierarchy	Knight's	21	23	21	100	33	50	35
	Queen's	23	25	23	33	100	35	34
	Hexagonal	28	28	28	49	35	100	43
	LCP	25	25	25	35	35	46	100

B Cholohovo — Ivanovskoe

		Knight's (with edge hierarchy)	Queen's (with edge hierarchy)	Hexagonal (with edge hierarchy)	Knight's (without edge hierarchy)	Queen's (without edge hierarchy)	Hexagonal (without edge hierarchy)	LCP
with edge hierarchy	Knight's	100	82	57	16	9	21	11
	Queen's	81	100	61	4	6	23	11
	Hexagonal	70	76	100	14	16	36	22
without edge hierarchy	Knight's	23	6	17	100	48	46	14
	Queen's	13	9	19	48	100	55	17
	Hexagonal	30	33	41	46	54	100	28
	LCP	14	15	23	13	16	27	100

C Africa — Rp5420

		Knight's (with edge hierarchy)	Queen's (with edge hierarchy)	Hexagonal (with edge hierarchy)	Knight's (without edge hierarchy)	Queen's (without edge hierarchy)	Hexagonal (without edge hierarchy)	LCP
with edge hierarchy	Knight's	100	99	99	16	22	14	35
	Queen's	99	100	99	16	22	14	35
	Hexagonal	99	99	100	16	22	13	35
without edge hierarchy	Knight's	27	27	27	100	60	29	23
	Queen's	36	36	36	59	100	21	33
	Hexagonal	22	22	22	28	20	100	18
	LCP	58	58	58	23	34	18	100

Fig. 5. Qualitative assessment of resulting routes overlap

more than twice the data and computational resources of the queen pattern, rendering it less efficient. Therefore, we conclude that a network model based on a square grid with queen connectivity offers greater efficiency in terms of processing time and routing.

DISCUSSION

Compared to previous attempts at integrating vector networks with raster cost surfaces, such as the extended raster approach (Antikainen 2013; Seegmiller et al. 2020) and the method proposed by Choi et al. (2013), the methodology presented in this study introduces several significant improvements. While the extended raster and boundary-connected (BC) models do overcome some of the distortions inherent in regular raster structures, they still rely on cell-based geometry and do not allow for flexible transitions between road networks and continuous terrain.

While Choi et al. (2013) take an essential step towards integrating regular and irregular models, their method restricts entry to and exit from the vector network to predefined graph vertices. Consequently, transitions can only occur at road intersections or endpoints, which restricts the realism of simulated routes, particularly in sparsely connected areas where pedestrians often join a road at various positions along its length.

The approach developed in this paper builds on this research by enabling transitions between roads and off-road terrain at any point along an edge, rather than just at vertices. This significantly increases the flexibility of the integrated model and better reflects real pedestrian movement. Additionally, the incorporation of edge hierarchy – with prioritization of high-quality or more passable segments – is a methodological advancement lacking in previous raster- or vector-based approaches. This hierarchy enables the model to approximate realistic walking preferences rather than merely identifying the shortest possible route.

Undoubtedly, the framework combines these improvements with anisotropic movement costs, weather-dependent coefficients, and a consistent treatment of slopes using Tobler's function. Together, these features offer a more comprehensive and adaptable routing methodology. These contributions distinguish the proposed approach from previous work, providing a more robust tool for simulating walking behavior in heterogeneous environments.

Despite the demonstrated advantages of the proposed methodology, several limitations must be acknowledged. Firstly, the model is inherently dependent on the spatial resolution of the regular grid used to generate the network. While finer grids improve the representation of terrain heterogeneity and reduce geometric distortions, they significantly increase data volume and computational requirements. Conversely, coarser grids simplify calculations but may overlook important landscape features that influence walking behavior.

Secondly, assigning walking speeds to land cover types and calibrating weather-related coefficients involves a degree of subjectivity. Although these coefficients were based on analytical estimates and expert knowledge, they may not fully capture the variability of real-world movement conditions. Their accuracy can vary between regions and user groups, and additional empirical field measurements would help to refine these parameters.

Thirdly, the computational cost of the method increases substantially when it is applied to large regions or when complex adjacency schemes are used. The number of edges increases rapidly with grid density and connectivity, which affects processing time and potentially limits the approach's applicability in extensive territories without additional optimization.

Finally, the model assumes simplified walking behavior, focusing primarily on minimizing travel time or distance and prioritizing roads through hierarchical edge weighting. However, human route choices are influenced by many additional factors, such as comfort, safety, visibility, fatigue, and terrain preferences – that are not fully represented in the current implementation. This may result in simulated routes lacking realism in certain contexts.

CONCLUSIONS

Given the need for routing methods capable of integrating road networks with continuous terrain representation while overcoming the limitations of conventional raster-based approaches, the experiments we conducted have allowed us to determine the advantages of the developed technique compared to the current pathfinding method using the LCP algorithm.

In particular, the LCP approach is similar to the resulting non-hierarchical network model. However, the LCP output is accompanied by a significant number of distortions due to the raster structure, and its accuracy depends heavily on its resolution. Furthermore, the LCP technique does not permit the formation of a hierarchy among individual pixels, preventing the prioritization of movement along the road network.

Following analysis, we identified the square regular grid network model with a "queen" adjacency pattern as the optimal option in terms of both computational complexity and route efficiency. This model has the shortest implementation time and is virtually free from geometric errors that could distort routes. Overall, the proposed methodology represents terrain more accurately than a raster-based approach. It is also more flexible and can be adapted to different movement conditions.

Depending on user priorities, the model can either emphasize more convenient routes through hierarchy or focus solely on the shortest (or fastest) path by removing hierarchy. The decision on which approach to use depends on the goals of the proposed model.

The developed methodology can be applied in several real-world contexts. In tourism route planning, for example, the model can generate comfortable walking routes that take into account weather conditions, terrain complexity, and user preferences regarding road versus off-road movement. In environmental and field research, it supports assessment of site accessibility, the planning of field campaigns, and the evaluation of how land cover changes affect mobility. In search and rescue operations, the hierarchical network structure enables the quick identification of feasible routes, optimal access points, and alternative paths under adverse conditions, providing a practical advantage over raster-based LCP approaches. Together, these use cases demonstrate the broader applicability and flexibility of the proposed routing framework. ■

REFERENCES

Antikainen H. (2013). Comparison of different strategies for determining raster-based least-cost paths with a minimum amount of distortion. Transactions in GIS, 17(1), 96-108, DOI: 10.1111/j.1467-9671.2012.01355.x.

Ashish D.D.V.S., Munjal S., Mani M., Srivastava S. (2021). Path finding algorithms. In: Emerging Technologies in Data Mining and Information Security. Proceedings of IEMIS 2020, 1, 331-338. Advances in Intelligent Systems and Computing, DOI: 10.1007/978-981-15-9927-9_33.

Bagli S., Geneletti D., Orsi F. (2011). Routing of power lines through least-cost path analysis and multicriteria evaluation to minimize environmental impacts. Environmental Impact Assessment Review, 31, 234-239, DOI: 10.1016/j.eiar.2010.10.003.

Balstrøm T. (2002). On identifying the most time-saving walking route in a trackless mountainous terrain. Geografisk Tidsskrift, 102, 51-58, DOI: 10.1080/00167223.2002.10649465.

Bemmelen J., Quak W., Hekken M., Oosterom P. (1993). Vector vs. raster-based algorithms for cross country movement planning. Proceedings of Auto-Carto, 11, 50-59.

Choi Y., Park H.-D., Sunwoo C., Clarke K. (2009). Multi-criteria evaluation and least-cost path analysis for optimal haulage routing of dump trucks in large scale open-pit mines. International Journal of Geographical Information Science, 23, 1541-1567, DOI: 10.1080/13658810802385245.

Choi Y., Um J.G., Park M.H. (2013). Finding least-cost paths across a continuous raster surface with discrete vector networks. Cartography and Geographic Information Science, 41(1), 75-85, DOI: 10.1080/15230406.2013.850837.

Collischonn W., Pilar J.V. (2000). A direction dependent least-cost-path algorithm for roads and canals. International Journal of Geographical Information Science, 14, 397-406, DOI: 10.1080/13658810050024304.

ESRI (1991). Performing analysis with GRID: Cell-based modeling with GRID. ARC/INFO User's Guide 6.0. Redlands, CA: Environmental Systems Research Institute.

ESRI (2005). Hierarchical routes in ArcGIS Network Analyst. ESRI White Paper. Redlands, CA: ESRI.

Goodchild M.F. (1977). An evaluation of lattice solutions to the problem of corridor location. Environment and Planning A, 9, 727-738, DOI: 10.1068/a090727.

Kourtz P.H., O'Regan W.G. (1971). A model for a small forest fire to simulate burned and burning areas for use in a detection model. Forest Science, 17(2), 163–169, DOI: 10.1093/forestscience/17.2.163.

Miller H., Shaw S.-L. (2001). Geographic Information Systems for Transportation. Oxford: Oxford University Press.

Novakovsky B.A., Kargashin P.E., Karpachevskiy A.M. (2017). Geoinformation analysis of territory for automated selection of power line routing. Geoinformatics, (2), 30-38 (in Russian with English summary).

Pokonieczny K. (2020). The methodology of creating variable resolution maps based on the example of passability maps. ISPRS International Journal of Geo-Information, 9(12), 738, DOI: 10.3390/ijgi9120738.

Rees W.G. (2004). Least-cost paths in mountainous terrain. Computers & Geosciences, 30, 203-209, DOI: 10.1016/j.cageo.2003.11.001.

Rylsky I.A. (2009). The influence of geographical environment factors on automated pipeline routing. Moscow: MAX Press, 192 p. (in Russian).

Seegmiller L., Shirabe T., Tomlin C. (2020). A method for finding least-cost corridors with reduced distortion in raster space. International Journal of Geographical Information Science, 35, 1-22, DOI: 10.1080/13658816.2020.1850734.

Susanto A. (2021). Performance comparison of Dijkstra, A*, and Floyd–Warshall algorithms in shortest path problems. Journal of Applied Mathematics and Computation, 5(2), 89-98.

Tobler W. (1993). Three presentations on geographical analysis and modeling. Santa Barbara, CA: National Center for Geographic Information and Analysis.

Tomlin C.D. (1990). Geographic Information Systems and Cartographic Modeling. Englewood Cliffs, NJ: Prentice-Hall.

Warntz W. (1957). Transportation, social physics, and the law of refraction. The Professional Geographer, 9(4), 2-7, DOI: 10.1111/j.0033-0124.1957.094_2.x.

Yu C., Lee J., Munro-Stasiuk M.J. (2010). Extensions to least-cost path algorithms for roadway planning. International Journal of Geographical Information Science, 17, 361-376, DOI: 10.1080/1365881031000072645.

Check for updates

CC BY

RESEARCH PAPER

INTEGRATION OF AI IN GIS FOR IDENTIFYING AND LOCATING ILLEGAL WASTE DEPOSITS IN ALGERIAN MUNICIPALITIES

Soumia Chergui[1*], **Abdallah Farhi**[2], **Bouthina Saib**[3]
[1*]Earth Sciences, Geography, and Land Planning Faculty , University of Science and Technology Houari Boumediene (USTHB), BP 32 Bab Ezzouar, Algiers 16111, Algeria.
[2]Architecture and Urban Planning Department, University of Mohamed Khider Biskra, BP 145, Biskra 07000, Algeria.
[3]Computer Science Department, Mathematics and Computer Science Faculty, Mustapha Ben Boulaid University, Batna 05000, Algeria.
***Corresponding author:** cherguibenarieb@gmail.com

Received: July 7th 2025 / Accepted: February 19th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4152

ABSTRACT. Like many municipalities around the world, Algerian municipalities are faced with the challenge of managing the collection of illegal waste deposits located across their territories. These deposits may occur both at authorized household waste collection points and at unauthorized locations. The management of these deposits presents a challenge, as their handling is incompatible with refuse trucks typically used for household waste collection or because they must be covered by collection services other than those provided by the municipality. It is therefore essential to identify and locate these deposits to ensure appropriate handling.

This article aims to address this issue through an innovative solution that integrates artificial intelligence (AI) into geographic information systems (GIS). The method is based on transfer learning combined with MobileNetV2 to generate a classification model for images of illegal waste deposits at authorized and unauthorized points. This model is integrated into a plugin created with QGIS software to perform image classification, enabling the location and identification of these deposits. The model achieved an accuracy of 98% during training, and its application to images from Biskra municipality illustrates its potential effectiveness. Beyond this case study, the approach offers a scalable and adaptable solution for improving illegal waste deposit management practices in diverse municipal contexts.

KEYWORDS: Illegal waste deposits., AI, GIS, Transfer Learning, MobileNetV2, Image classification Model

CITATION: Chergui S., Farhi A., Saib B. (2026). integration Of AI In GIS For Identifying And Locating Illegal Waste Deposits In Algerian Municipalities. Geography, Environment, Sustainability, 1 (19), 139-151
https://doi.org/10.24057/2071-9388-2026-4152

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

Municipalities in Algeria are facing an increasing problem of illegal waste deposits within their territory. These deposits, abandoned by residents, can be found in both unauthorized locations and at authorized waste collection points, where they disrupt regular collection processes. As in many municipalities around the world, particularly in developing countries, managing these deposits represents a major challenge for local authorities (Adipah and Kwame 2019; AND 2021; United Nations 2023; Del Carmen-Niño et al. 2023; Marange et al. 2023). The difficulties encountered may lie in the organizational shortcomings of local waste management systems, as noted by Dregulo and Khodachek (2022). In Algerian municipalities, this organizational deficiency may be attributed to the services in charge of collecting these deposits or the inadequacy of collection equipment, such as refuse trucks, when illegal waste deposits are dumped at household waste collection points. This leads to slowness in collecting illegal deposits, which in turn causes environmental problems and public health risks (Jakeni et al. 2024; Pathak et al. 2024). Moreover, identification, and geolocation of illegal waste deposits remain critical issues in municipal waste management (Devesa and Brust 2021), which requires effective detection methods to remedy the problem as noted by Yu et al. (2024).

Several studies have emphasized the usefulness of geographic information systems (GIS) in managing illegal waste deposits at the municipal level (Paunović and Krstić 2014; Fatunmibi and Gbopa 2021; Jimoh et al. 2019; Karimi and Ng 2022; Syafrudin et al. 2023; Bošković et al. 2024). GIS facilitate spatial mapping and analysis of such data (Abdulai et al. 2015; Zainun et al. 2016; Krsmanović et al. 2022; Jakeni et al. 2024), helping local authorities identify areas vulnerable to illegal dumping (Thompson et al. 2013; Bodea et al. 2014). Moreover, GIS supports municipal decision-makers in planning and managing waste collection operations (Caputo and Pelagagge 2000; Iacoboaea and Aldea 2011; Awasare and Sutar 2015; Brus et al. 2016; Hua et al. 2016; Rao et al. 2020; Asefa et al. 2022; Sakshi et al. 2023).

However, traditional GIS approaches face limitations in processing the growing volume and complexity of spatial data (Ahmad 2023; Choi 2023; Yang et al. 2024). This is also true for data related to illegal waste deposits, which present similar challenges due to their exponentially increasing volume. Consequently, there is interest in integrating other technologies with GIS technology, such as artificial intelligence (AI). This integration involves the application of AI techniques to enhance the capabilities of GIS software (Ahmed 2024). With AI's computational power, this combination improves the ability of GIS to process complex spatial data. (Cai and Ge 2023) AI, which aims to replicate human cognitive functions (Vozenilek 2009; Panesar 2020; Ghosh and Thirugnanam 2021; Raj 2024), enables the emulation of human intelligence and problem-solving abilities by computers and machines (Saxena et al. 2023; Shrivastava et al 2024). Its effectiveness in managing illegal waste deposits has been demonstrated in numerous studies (Torres and Fraternali 2021; Shahab and Anjum 2022; Kim and Cho 2022; Inamdar et al. 2023; Ulloa-Torrealba et al. 2023), particularly through the application of Transfer Learning (Padubidri et al. 2022; Yu et al. 2024). The latter is based on Deep Learning, which uses neural networks to process large and complex datasets (Shafik 2024; Narayanan and Arjun 2024), and it requires significant computing and memory resources, as noted by Talaei Khoei et al. (2023). DL models are generally based on multilayer neural architectures such as convolutional neural networks (CNNs) (Nandi 2023; Mohanta et al. 2024).

TL uses pre-trained deep learning models (Talaei Khoei et al. 2023; Narayanan and Arjun 2024), allowing for reduced computational demands in terms of CPU and RAM (Puigcerver 2020; Khan et al. 2024; Zorić et al. 2024). MobileNetV2, a lightweight convolutional neural network (CNN) architecture trained on the ImageNet dataset, exemplifies this efficiency by significantly reducing computational load while maintaining high performance in image classification tasks (Howard et al. 2018; Sandler et al. 2018; Gondhalekar et al. 2024). These features make it particularly suitable for use in municipalities of developing countries where technical infrastructure is limited. In this context, combining TL with MobileNetV2 has proven effective in detecting illegal waste deposits (Shahab and Anjum 2022; Inamdar et al. 2023).

This article aims to address the issue of identifying and locating illegal waste deposits at the municipal level through an innovative solution that integrates artificial intelligence (AI) into geographic information systems (GIS). The proposed method relies on transfer learning combined with the MobileNetV2 architecture to develop an image classification model that distinguishes between two categories of illegal waste deposit locations: authorized and unauthorized points. The model is then integrated into a plugin developed within the QGIS software environment, enabling the automatic classification and localization of these points. This integration is intended to assist municipal authorities in managing such deposits by helping them deploy adequate resources and transmit the relevant information to the responsible collection services. It is important to note that the generated model has been tested on real data from the municipality of Biskra in Algeria, allowing us to assess its operational robustness in a concrete territorial context.

MATERIALS AND METHODS

Study area

The municipality of Biskra, chosen as the study area, is the capital of Biskra Province, located in the southeastern region of Algeria. It covers an area of 127.7 km^2 (Monographie de Biskra, 2024). According to the most recent official census (2008), the population of Biskra was approximately 205,608 inhabitants. Geographically, the municipality is situated between longitudes 5°39´22´´ and 5°46´44´´ East and latitudes 34°54´58´´ and 34°47´05´´ North as shown in Fig. 1.

Data

To enhance the diversity of the dataset and improve the model's generalization capabilities, images were collected from three sources: Algerian municipalities (58.6%), Turkish municipalities (27.3%), and publicly accessible online platforms (14.1%). The Turkish images were gathered during field visits, and selected for their relevance and visual similarity to illegal waste deposits commonly observed in Algeria. Likewise, online images were selected based on these same criteria to ensure visual and contextual consistency throughout the dataset.

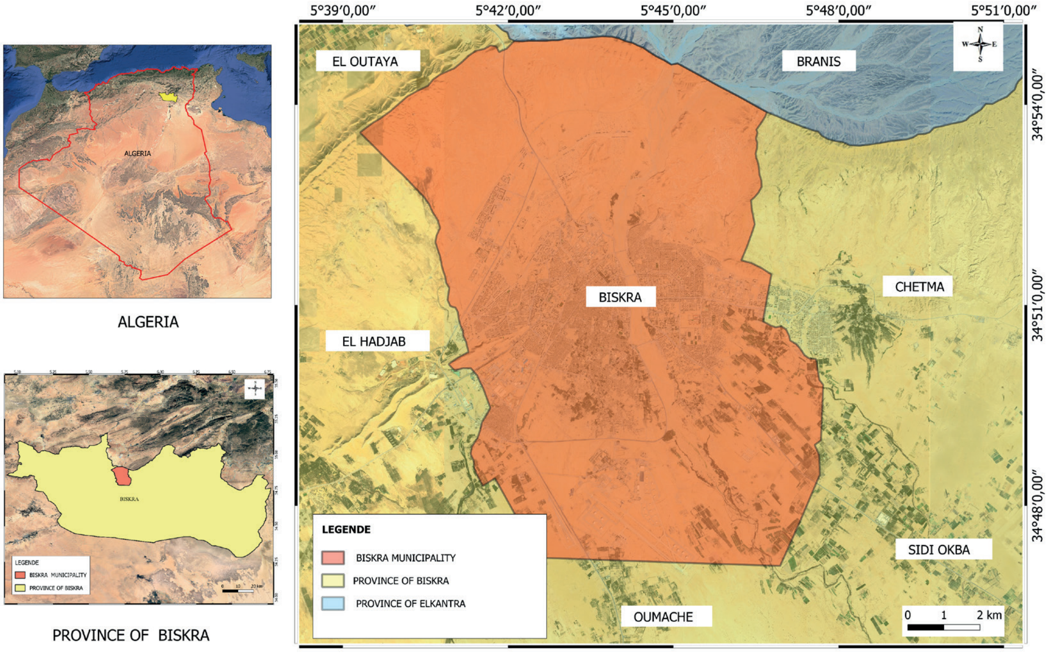


Fig. 1. Location of Biskra municipality

This multi-source approach enabled the construction of a robust image database depicting various instances of illegal waste deposits, whether occurring at authorized or unauthorized locations. Authorized points refer to formally designated areas for household waste collection, recognized in this study by the visible presence of dedicated waste bins, whereas unauthorized points correspond to locations where dumping is strictly prohibited, such as sidewalks, roads, and public spaces. There are two different types of images in the database: authorized and unauthorized points. Each class contains 800 images, for a total dataset of 1,600 images from various sources.

For the first class (authorized points): the waste stored in these collection points is not limited to household waste; however, it has been observed that illegal waste is frequently dumped there. These deposits may come from demolition, construction or renovation work. According to Algerian regulations, this category of waste is classified as inert waste. As stated in Article 37 of Law No. 01-19 of 2001 on the management, control, and elimination of waste, the deposit, dumping, or abandonment of such waste on unauthorized sites is considered an illegal practice and is strictly prohibited. Consequently, failure to comply with this law is punishable by a fine, in accordance with Article 57 of the aforementioned law. The imposition of strict measures to combat these illegal practices is, therefore, aimed at eradicating this phenomenon.

Other types of illegal waste deposits are also abandoned at authorized points, namely bulky waste and special waste. Article 3 of Law No. 01-19 (2001) stipulates that bulky waste is classified as a type of household waste. However, due to its substantial volume, it cannot be collected in the same conditions as household waste. It must be collected, transported and treated separately from household waste. Special waste includes waste generated by industrial, agricultural and service activities. Due to their composition and distinctive nature, these types of waste necessitate specific collection, transportation and treatment methods, which differ from those employed for household, bulky or inert waste. Their management, therefore, requires specific measures to guarantee environmental and health safety.

The management of special waste is governed by a strict regulatory framework that imposes responsibility on the generators and holders of such waste, with the aim of preventing any illegal practices likely to harm the environment and public health. Article 16 of the law specifies that “the generators and/or holders of special waste are required to ensure, or have ensured, at their expense, the management of their waste”. When a holder of special waste fails to comply with the established standards, in particular by abandoning or depositing such waste outside the regulatory procedures, the competent court has the power to intervene and impose the elimination of the waste at the expense of the offender. In this sense, the law ensures that producers and holders of such waste are held accountable, while simultaneously time guaranteeing rigorous application of the law.

For the second class (unauthorized points), these refer to locations not intended or equipped for waste disposal, where dumping is prohibited by municipal regulations. These areas often lack infrastructure for waste management and include spaces where the presence of waste can cause health, environmental, and safety problems. The anarchic accumulation of such waste deteriorates living spaces, obstructs traffic routes, and compromises public health.

In these areas, a variety of illegally dumped waste was identified, including inert waste such as rubble and construction residues, special waste, and household waste discharged by the population outside regulated collection systems.

To combat such practices, Algerian regulations impose financial penalties as a dissuasive deterrent. For example, persons who abandon household waste or refuse to use the facilities set up by local authorities, as well as industrial, commercial and craft operators who fail to comply with waste management systems, are subject to penalties defined by current legislation, Law No. 01-19 (2001). These legal measures serve to reinforce the efforts to combat illegal dumping by making citizens more conscious of their responsibilities, ensuring better waste management and contributing to environmental protection.

The images presented in Fig. 2 illustrate various categories of illegal waste deposits at authorized and unauthorized points in Algeria.

(a) Illegal waste deposits before household waste collection

(b) Illegal waste deposits after household waste collection

(c) Illegal waste deposits

Fig. 2. Different types of illegal waste deposits in Algeria
(a) and (b) at authorized points (c) at unauthorized points

Development environment and system configuration

The development of the classification model and its integration into the GIS environment relied on a combination of Python-based tools and geospatial software. The programming language used was Python 3.10, selected for its compatibility with both deep learning libraries and QGIS scripting interfaces.

Several libraries were employed throughout the pipeline, each serving a specific purpose:

- TensorFlow 2.13 and Keras were used to build, train, and optimize the MobileNetV2-based classification model.
- NumPy supported array manipulation and numerical operations during preprocessing and model structuring.
- OpenCV (cv2) was used for image resizing, applying bilinear interpolation by default via the cv2.resize() function.
- Pillow (PIL) was used for image augmentation tasks such as rotation, zoom, shear, and horizontal flipping, simulating real-world variability in illegal waste deposit imagery.
- PyQGIS enabled seamless integration of the trained model into the QGIS environment, allowing spatial classification and geolocation
- Qt Designer (v5.11.12) was used to design and customize the plugin's graphical interface for municipal use.

Development was carried out in PyCharm Community 2020.3, while QGIS v3.34.12-Prizren served as the GIS platform for plugin deployment and model execution via its Python console.

The system used for development and testing was a standard laptop equipped with an Intel(R) Core (TM) i3-3217U CPU operating at 1.80GHz, 6 GB of RAM, and a 64-bit operating system based on x64 architecture. This modest configuration reflects the practical constraints commonly encountered within municipal infrastructures and demonstrates the operational viability of the proposed solution in low-resource environments.

Methodology

The originality of this work does not lie in the architecture of the classification model itself, which is based on MobileNetV2 and transfer learning, both of which are well-established in the literature, but rather in its integration into a Geographic Information System (GIS). Specifically, the model is embedded within a QGIS plugin designed to automate the classification and geolocation of illegal waste deposits. In addition to this integration, a key novelty of the approach lies in training and deploying the model directly within the QGIS environment, without relying on external platforms or cloud-based services. This training workflow enhances the operational autonomy of municipal users in a spatial context. Together, these elements constitute a novel contribution to the management of illegal waste deposits in municipalities.

Despite the absence of documentation of this particular integration in prior studies, MobileNetV2 was selected due to its demonstrated efficacy in image classification tasks. Its lightweight architecture makes it well-suited for deployment within GIS platforms where computational resources may be limited.

For research on AI into GIS, PyCharm and QGIS were used complementarily. PyCharm constituted the principal environment for developing Python scripts for image splitting, resizing and augmentation, as well as coding the classification plugin. The execution of these scripts was conducted in PyCharm, except for of the plugin.

The QGIS software was used on two levels: firstly, to create the plugin for running the classification within a GIS environment, and secondly, to train and evaluate the model via its Python console. Qt Designer was used to design and customize the plugin's interface. These stages are illustrated in the diagram shown in Fig. 3.

It is important to note that the decision to train the model directly in the QGIS environment stems from the difficulties encountered during tests carried out in PyCharm. Running the classification plugin faced incompatibilities between the versions of TensorFlow used by PyCharm and QGIS, resulting in blocking errors when integrating the model into QGIS.

Data augmentation

Data augmentation, a technique widely adopted in neural network training, enhances the robustness of models and improves their generalization power. As indicated by Hernandez-Garcia (2020) and Zeng (2024), this methodological approach relies on applying image transformations to artificially expand the datasets,

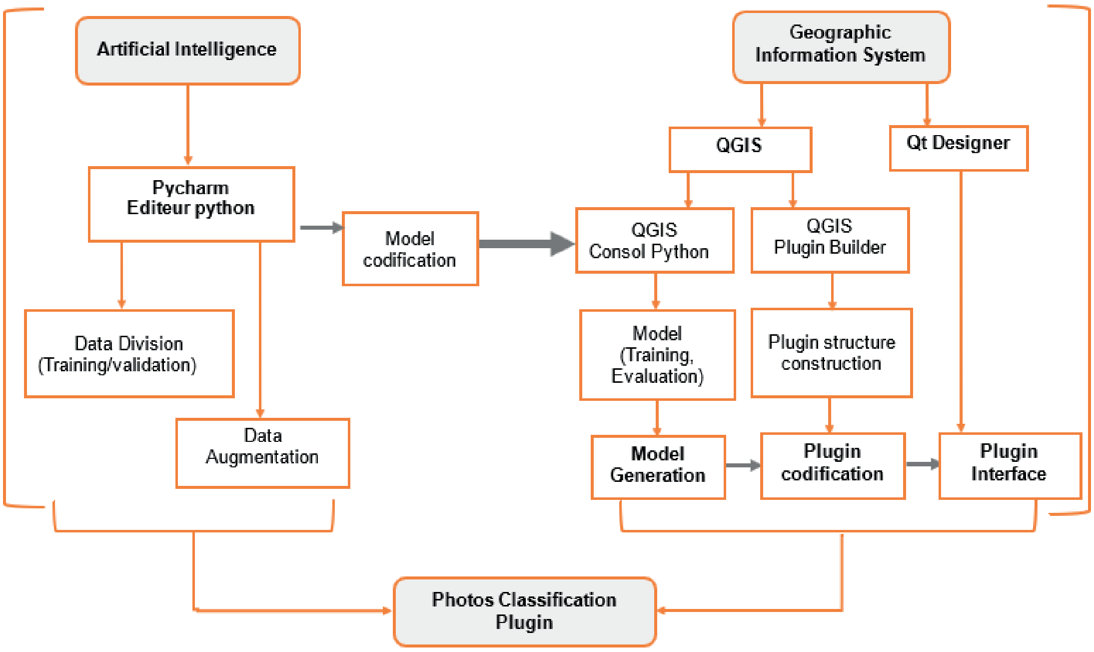


Fig. 3. Diagram illustrating the various stages of the methodology

particularly when datasets are limited. This approach proves effective in improving model performance and reducing the risk of over-fitting.

In the context of this study, prior to any augmentation operation, the initial dataset of 1,600 images was randomly divided into two subsets using an 80/20 ratio: 1,280 images (80%) were allocated to training, while 320 images (20%) were used for validation. This separation was performed programmatically using Python's random.shuffle function to ensure randomized distribution across both predefined classes: one designated for images of authorized points (Authorized_pt), and the other for images of unauthorized points (Unauthorized_pt). The dataset was then organized into a directory structure with two main folders (train and validation), each containing subfolders for the respective classes. This organization was designed to facilitate data access during the model training phase.

The limited size of the training dataset (640 images per class) limits the model's capacity to learn. The latter generally requires a larger volume of data to guarantee satisfactory generalization capability. For this reason, the dataset was augmented, with the objective of enhancing model performance. By augmenting the training data set, each class was expanded to comprise 1,250 images, resulting in a balanced dataset totaling 2,500 images; thus, the final dataset comprises 2,820 images.

Furthermore, a range of data augmentation techniques were employed to augment the size of the training dataset while preserving the essential characteristics of the images. These transformations were selected to simulate the spatial and morphological variability of illegal waste deposits observed across municipalities. As such, each original image was transformed using the following techniques: rotational augmentation, translation augmentation,shear transformation,scale transormation, and horizontal flip.:

As discussed by Muñoz-Aseguinolaza et al. (2023) and Qi et al. (2021), rotational augmentation involves applying angular displacement to the original image around its central axis. This transformation is mathematically represented as Eq. (1):

$$I'(x,y) = I(x\cos\theta - y\sin\theta,\quad x\sin\theta + y\cos\theta) \qquad (1)$$

where *I´* and *I* represent the original and transformed images respectively, and *θ* denotes the rotation angle. The implementation constrains θ within a predefined range (e.g., ±30°) to preserve semantic validity. This transformation induces rotational invariance, a critical property for object recognition systems operating in unconstrained environments.

Translational augmentation implements pixel-wise displacement along horizontal and vertical axes (Kumar et al. 2025; Nanni et al. 2021). This transformation is formally defined as Eq. (2):

$$I'(x,y) = I(x + \Delta x,\quad y + \Delta y) \qquad (2)$$

where *Δx* and *Δy* represent the magnitude of horizontal and vertical translations, typically parameterized as proportions of image dimensions (e.g., $\Delta x \leq 0.2W$, $\Delta y \leq 0.2H$, where *W* and *H* denote image width and height respectively). This augmentation facilitates translational invariance, enabling recognition systems to identify objects regardless of their spatial position within the image frame.

Shear transformation, as described by Awaluddin et al. (2023) and Kumar et al. (2025), applies non-uniform scaling, creating a parallelogram-like distortion characterized by Eqs. (3)-(4):

$$I'(x,y) = I(x + y\tan(\alpha), y) \qquad (3)$$

For horizontal shear with angle *α*

$$I'(x,y) = I(x, y + x\tan(\beta)) \qquad (4)$$

For vertical shear with angle *β*

The shear range parameter (e.g., 0.2 radians) controls the maximum permissible distortion. This transformation simulates perspective variations and viewpoint alterations, enhancing model robustness to affine distortions.

Scale transformation (zoom), according Wang et al. (2025) and Khalil et al. (2023), modifies the apparent size of objects within the image through uniform scaling (Eq.5):

$$I'(x,y) = I(sx, sy) \qquad (5)$$

where *s* represents the scaling factor, typically implemented as a random variable within a predetermined range (e.g., 0.8 ≤ s ≤ 1.2 for zoom range=0.2). This augmentation induces scale invariance, enabling recognition systems to identify objects across variable distances and sizes.

Horizontal flipping is the process of creating a mirror image of the original along the vertical axis (Awaluddin et al. 2023; Kumar et al. 2025) (Eq.6):

$$I'(x,y) = I(w - x, y) \qquad (6)$$

where *W* denotes the image width. This transformation exploits bilateral symmetry, a prevalent feature in numerous natural and artificial objects, effectively doubling the representation of symmetrical features within the training distribution.

Model creation, training and evaluation

This stage of the methodology involves the creation, training, and evaluation of an image classification model using TensorFlow from the QGIS Python console. The model is based on the MobileNetV2 architecture, incorporating pre-trained weights adapted to the specific classification task. As schematically illustrated in Figure 4, the process includes several key steps: data loading and augmentation, image pre-processing, and the definition and training of the model architecture. The figure provides a visual representation of the model pipeline, from the initial input images to the final classification output

Data loading and pre-processing was performed using TensorFlow ImageDataGenerator, which enables data normalization and the automated generation of image batches for training. As described in the data augmentation section, the dataset was augmented and split into training and validation subsets (80/20). No cross-validation was applied; instead, a fixed split was chosen to ensure reproducibility and to accommodate hardware constraints typical of municipal deployments.

All images were resized to 224 × 224 pixels: Validation images were resized prior to training, while training images were automatically resized during the augmentation step, ensuring uniformity across the dataset. Subsequently, all images wererescaled to normalize pixel values between 0 and 1. A batch of 15 images was used during the training process. This configuration enables efficient data processing while ensuring compatibility with MobileNetV2 model requirements.

The adopted model architecture is based on transfer learning from MobileNetV2, a pre-trained convolutional neural network on the ImageNet database. This model was chosen for its computational efficiency and high performance, making it ideal for integration into environments with limited resources.

To adapt MobileNetV2 for binary classification (authorized versus unauthorized points), the following steps were taken to customize its architecture (Figure 5)

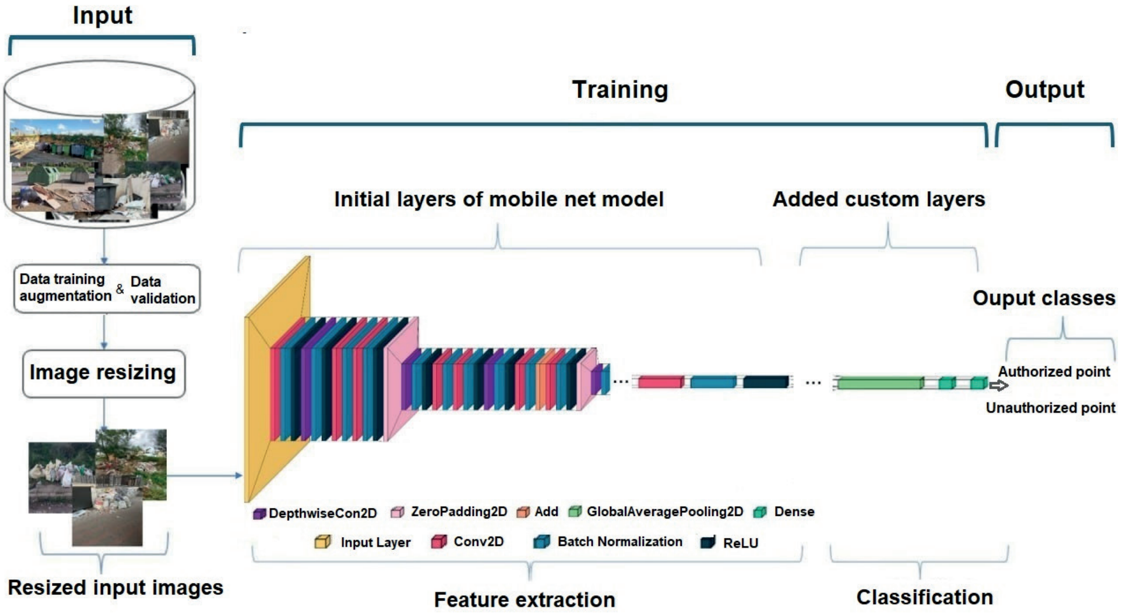


Fig. 4. Image classification pipeline using MobileNetV2 with pre-trained weights: from data input to final classification output

First, MobileNetV2 with pre-trained weights on ImageNet was used as the foundation. Only the last 30 layers were made trainable to enable fine-tuning, and the rest were frozen. This process involves retraining only the final layers of the model so that it can adapt to the specificities of the dataset studied while retaining the general knowledge acquired during the initial training.

Then, feature extraction was achieved using a GlobalAveragePooling2D layer to reduce the spatial dimensions of the feature maps derived from the convolutional network while preserving the essential visual information for classification. These "maps" are internal representations of the image that are automatically generated by the model from patterns detected in previous layers. The pooling layer extracts a compact synthesis, which facilitates learning while limiting the number of parameters to be trained.

Finally, the classification head consits of a dense layer with 128 neurons using ReLU activation, followed by an output layer with 2 neurons using Softmax activation.

The convolutional layers compute features according to Eq. (7):

$$f_l=\sigma\left(w_l\times f_{l-1}+b_l\right) \tag{7}$$

where f_l represents the feature maps at layer l, w_l and b_l are the trainable weights and biases, and σ is the ReLU activation function (Eq.8):

$$p\left(y_i \middle| x\right)=\frac{e^{z_i}}{\sum_j e^{z_j}} \tag{8}$$

The final classification probabilities are computed using the softmax function:

where z_i is the logit corresponding to class i.

the training strategy involved compiling the model with Adam optimizer at a learning rate of 0.0001 and categorical cross-entropy as the loss function, defined by Eq. (9):

$$L=-\sum_i y_i \log\left(\hat{y}_i\right) \tag{9}$$

Where y_i is the true label and $\hat{y}_i$ is the predicted probability.

The model was trained over 20 epochs with the specified hyperparameters. As Toennies (2024) have previously indicated, the number of such epochs is a crucial factor in network training, as it allows network weights to be optimized. Once the training process had been completed, the model was saved in the Keras framework as an h5 - file.

The justification of hyperparameters was guided by both empirical validation and established practices in transfer learning with lightweight convolutional networks. The batch size of 15 was selected to ensure compatibility with limited hardware resources while maintaining stable gradient updates. The learning rate of 0.0001 was chosen to allow gradual fine-tuning of the last 30 layers of MobileNetV2 without disrupting pre-trained weights. The number of epochs (20) reflects a balance between convergence and overfitting risk, as supported by Toennies (2024). The use of categorical cross-entropy as the loss function is standard for multi-class classification tasks, and the Adam optimizer was selected for its adaptive learning capabilities and proven robustness in image classification workflows. The image size of 224 × 224 pixels aligns with MobileNetV2 input requirements, and the 80/20 train-validation split ensures sufficient data for both learning and generalization assessment. These choices collectively support reproducibility, computational efficiency, and model robustness.

Creation of the plugin in QGIS

Prior to elaborating on the technical steps involved in developing the plugin, it is imperative to clarify that its objective extends beyond the simple visualization of spatial data. The plugin developed in this study performs two key functions: it automatically classifies input images as either authorized or unauthorized waste deposit locations, and it determines and displays the geographic location of each classified point directly within the QGIS interface. The integration of the classification model into a GIS environment provides municipal users with an operational tool that combines image analysis and geospatial visualization for the management of illegal waste deposits.

The development process of the plugin is presented in

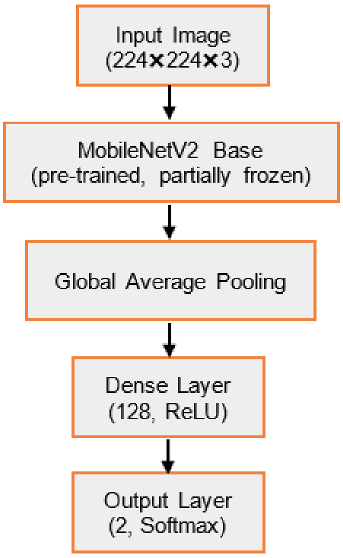


Fig. 5. Architecture of the proposed MobileNetV2-based classification model

three main stages: generating the plugin structure using the Plugin Builder tool, designing the graphical interface with Qt Designer, and integrating the trained classification model into the plugin.

In the first stage, the creation of a QGIS plugin begins with the generation of a standardized and extensible software base, in compliance with the development standards of this open source GIS software. The Plugin Builder tool, developed by the QGIS community automates this process by producing the required file structure for the plugin (Python scripts, metadata, graphic resources, functional directories, etc.). During this phase, various essential metadata elements are entered, including the name, version, author, and description of the project. This ensures effective integration into the QGIS environment and facilitates efficient referencing within the extension manager. This process establishes the technical and documentary foundations for the plugin's maintainability and eventual distribution.

In the second stage, the graphical interface of the image classification plugin was designed using Qt Designer. The interface, defined visually in .ui format, is converted into a Python script using the pyuic5 utility, generating an intermediate file (interface_ui.py) that is compatible with the QGIS PyQt framework. Thereafter, the script is integrated into a class derived from QDialog. The connection between the graphic components and the plugin's internal functions is based on event connectors (self.ui.object.clicked.connect(...)), that link each interactive element to a specific function. In the case of our plugin, these elements activate image classification processes relating to illegal waste deposits points.

In the last stage of the process, the classification model, which had been trained using the TensorFlow library and, exported in .h5 format, was integrated within the QGIS plugin. The aim of this integration is to automatically categorize images into two classes: "authorized_pt" and "unauthorized_pt". The model is then loaded into the plugin upon initialization. The images to be classified are then passed to the prediction model. Thereafter, a binary labeling system is assigned to each image according to the obtained result.

The logic associated with operating the classification model is encapsulated within a modular structure, represented here by the Photoclassification class. This organization guarantees a coherent plugin architecture, facilitating code readability and reusability.

RESULTS AND DISCUSSION

Quantitative Evaluation Results

The proposed approach for classifying illegal waste deposits, distinguishing between authorized and unauthorized categories, has demonstrated a remarkable degree of accuracy, with an overall rate of 98%. This high degree of accuracy highlights the effectiveness of transfer learning, particularly when combined with the MobileNetV2 architecture, in dealing with the challenge of classifying illegal waste deposits. The classification results are summarized in Table 1, and the following key information is highlighted.

Dataset augmentation

The augmentation of the dataset was applied exclusively to the training set, with the objective of enriching the diversity of examples without introducing bias into the validation phase. This strategy proved effective in enhancing the generalizability of the model, particularly in our context where the initial volume of data was limited.

High Precision and Recall for Both Classes:

The model achieves precision values of 0.987 for authorized points and 0.975 for unauthorized points, indicating a low false positive rate. This is essential to ensure that authorized points are not mistakenly marked as unauthorized.

The recall values of 0.975 for authorized points and 0.987 for unauthorized points demonstrate the model's ability to identify most items in both classes effectively.

Balanced F1-Scores

The F1-scores of 0.981 for authorized and unauthorized points reflect a strong balance between precision and recall. This indicates that the model performs consistently well across both classes, without significant bias in favor of one category.

Computational efficiency

The use of MobileNetV2, a lightweight convolutional neural network, guarantees the model's computational efficiency. Consequently, it can be deployed in environments where resources are limited. This deployment can be implemented at the level of deprived municipalities. The efficacy of this measure is especially significant in developing countries, where it facilitates more effective management of illegal waste deposits.

Scalability and practical implications

The high accuracy and efficiency of the MobileNetV2 model, and its seamless integration into a GIS environment, enables automated management of illegal waste deposits. This scalability is of crucial importance to municipal authorities and environmental agencies, as it facilitates large-scale deployment across multiple regions without the need for substantial computational overheads, provided that basic geospatial data are available to ensure the system's operational viability.

Comparative perspective with existing literature

While the integration of AI into GIS has been explored in various contexts (Ahmed, 2024; Cai and Ge, 2023), most existing approaches rely on external platforms or cloud-based services for model training and deployment. Tools such as Mapflow and dzetsaka offer classification capabilities within QGIS, but they do not support in-plugin training workflows tailored to illegal waste detection. Our method addresses this gap by embedding both training and inference directly within QGIS, enabling autonomous and localized decision-making for municipal users. This distinction reinforces the operational relevance of our contribution, particularly in resource-constrained environments.

Environmental and public health impact

By enabling the classification of illegal waste deposits points, this model strengthens improve environmental protection and public health. It offers a proactive instrument for identifying these illegal deposits, linked to environmental degradation and related risks. Once detected, municipal authorities can initiate targeted clean-up operations and implement preventive monitoring protocols. These interventions not only reduce immediate exposure risks for nearby residents but also contribute to long-term environmental recovery. The system supports informed decisionmaking and timely responses to unmanaged illegal waste accumulation, reinforcing its value as a management tool aligning with municipal policy frameworks, including the Waste Management Plan and public health programs on hygiene that emphasize reducing risks for the population.

Confusion matrix an analysis

A confusion matrix analysis reveals the model's high accuracy, thus demonstrating its robustness in classifying illegal waste deposits points. As shown in Figure 6, the results obtained indicate a marginal error rate, resulting in only four (4) cases of confusion where authorized points were incorrectly predicted as unauthorized, and two (2) inverse cases where unauthorized points were wrongly classified as authorized.

The low number of errors observed in this study serves as a reliable indication of the predictive model's efficacy, underscoring its ability to accurately distinguish between the two target categories.

For clarity, Table 2 summarizes the confusion matrix values (TP, FP, FN, TN) corresponding to these results.

Training and validation dynamics

The analysis of the learning dynamics revealed a rapid and stable convergence of the model. As shown in Fig. 7, this is illustrated by the training and validation accuracy curves reaching an asymptote as early as the tenth (10th) epoch. By the end of this phase, the model achieved a training accuracy of 99.97% and a validation accuracy of 98.12%, demonstrating its ability to efficiently generalize to novel data. Additionally, the model's resistance to overfitting is confirmed by the low loss function values, with training and validation losses of 0.0107 and 0.0463, respectively, indicating stability and a balance between performance and generalization. This convergence behavior empirically confirms the appropriateness of the 20 epoch training limit, as previously justified in the methodology section. Extending the training beyond this point would have yielded negligible performance gains while increasing the risk of overfitting and computational cost.

Training and inference time

Leveraging the lightweight architecture of MobileNetV2, the model demonstrated significant computational efficiency with an average inference time of 62ms per image. The complete training process required approximately 38 minutes, which demonstrates the feasibility of training the model in low-resource environments, without reliance on high-performance computing infrastructure, and confirms its suitability for municipal-scale experimentation.

GIS Integration and plugin validation

validation of the developed plugin

The evaluation of the developed plugin was carried out using an image dataset collected in the municipality of Biskra in Algeria. The objective of this experiment was to assess the efficacy of the classification model integrated into the plugin within a concrete territorial context. The results obtained demonstrate the effectiveness of the classification model on these test images. The geolocalized images are displayed in QGIS as vector points (shapefile), each point corresponding to the location of an illegal waste deposit identified in the field.

Table 1. Results of the classification model of the waste disposal points

Class	Precision	Recall	F1-Score	Support
Authorized_point	0.987	0.975	0.981	160
Unauthorized_point	0.975	0.987	0.981	160
Accuracy	0.980			
Macro Avg	0.980	0.980	0.980	320
Weighted Avg	0.980	0.980	0.980	320

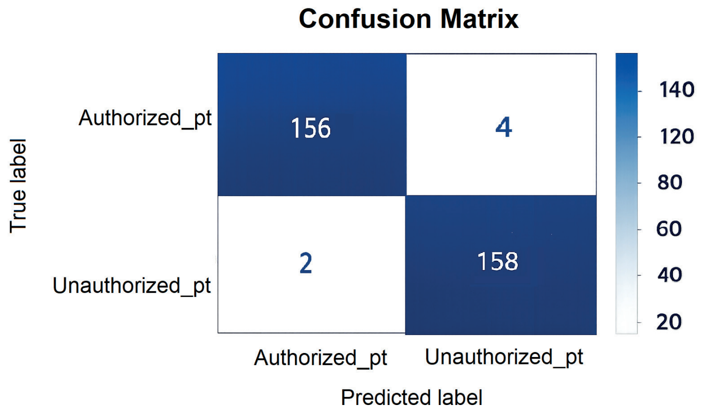


Fig. 6. Confusion matrix of the classification model

Table 2. Confusion matrix summary (TP, FP, FN, TN)

Class	TP	FP	FN	TN
Authorized points	156		4	
Unauthorized points		2		158

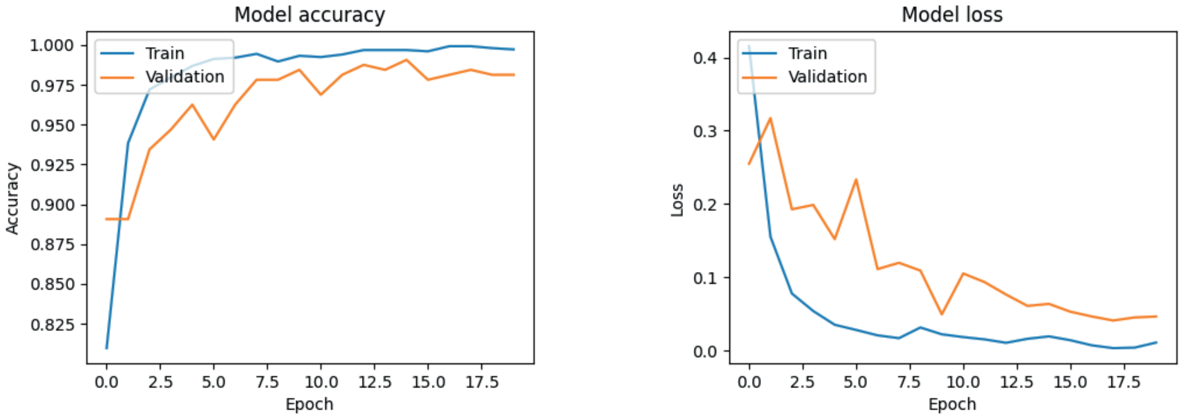


Fig. 7. Training and validation accuracy curves

Potential users are thus able to visualize all the deposits in question as points superimposed on a base map.

As illustrated in Figure 8, when a point is selected, a window opens and displays the corresponding image and the results of the prediction generated by the model. The information displayed includes the predicted class (authorized or unauthorized deposit), the confidence score of the model associated with this prediction, the address of the point, and the date and time the image was taken.It is noteworthy that all geolocated test images, collected independently in the municipality of Biskra and processed through the plugin, were correctly classified into their respective categories, thereby confirming the consistency between the training results and the functional validation in the municipal context.

Operational outputs and GIS visualization

Furthermore, all this information is automatically recorded in the attribute table associated with the shapefile (Figure 9), allowing for immediate exploitation of the data. This structured output supports intervention planning and enhances decision-making processes by providing municipal authorities with precise, contextualized information.

For instance, this data can be directly used to generate a distribution map of illegal waste deposits in Biskra municipality (Figure 10), thus facilitating municipal operations through effective visualization and analysis within a Geographic Information System (GIS) environment.

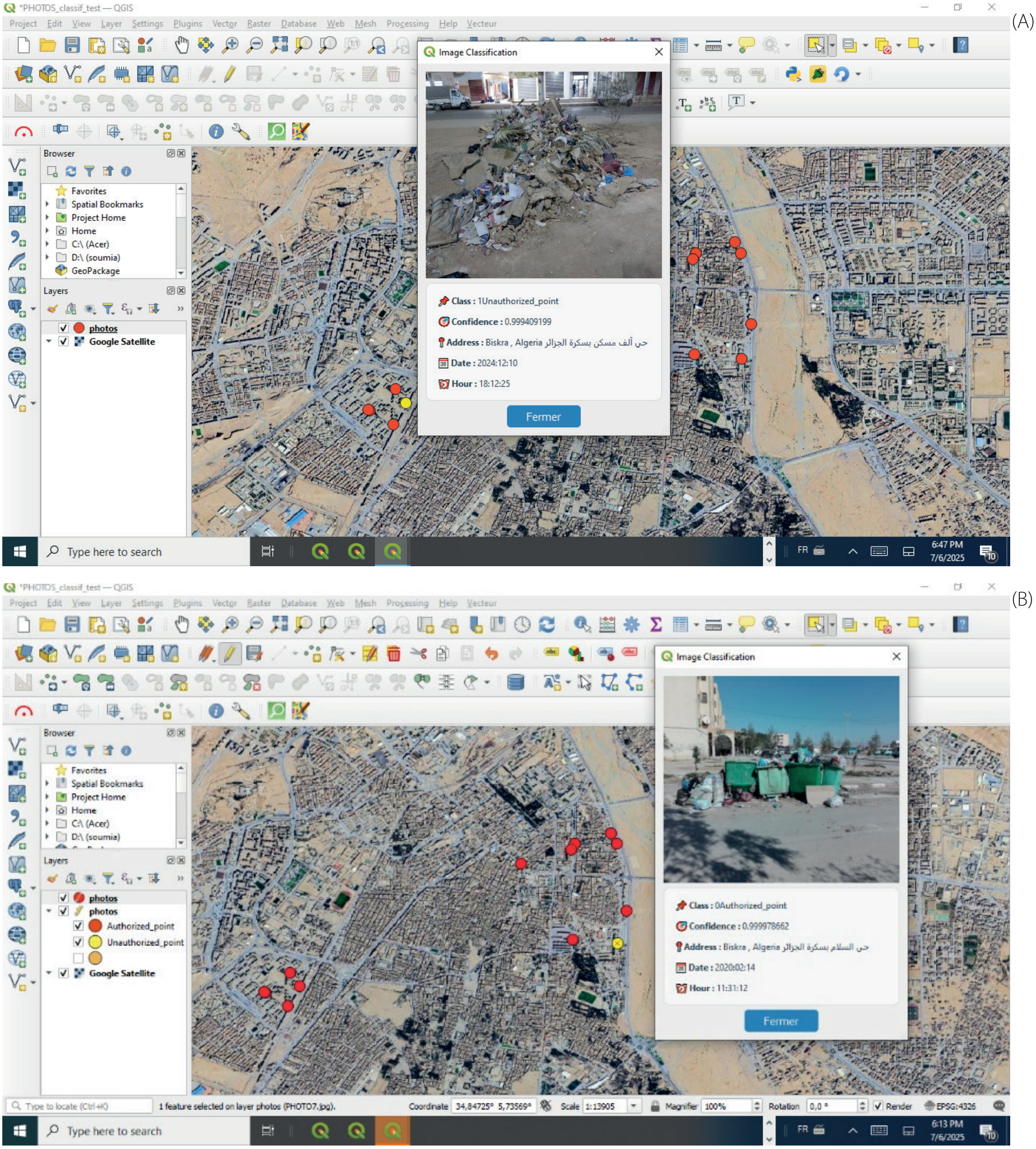
It should be noted that the municipality of Biskra does not possess geospatial data identifying the locations of legally designated waste deposit sites. The outputs generated by the plugin, including classified images, geographic coordinates, and contextual metadata, contribute to the development of such geospatial information for municipal use. This structured dataset can support future planning, monitoring, and management of illegal waste deposits within a GIS framework.

The functional evaluation of the plugin confirms both the technical feasibility and the operational relevance of the proposed approach, which enables automated, geolocated, intelligible, and visually interpretable classification of illegal waste deposits based on geolocated photos collected directly in the field.

CONCLUSION

This study demonstrates the relevance of an innovative approach based on the integration of Artificial Intelligence (AI) into Geographic Information Systems (GIS) for managing illegal waste deposits at the municipal level.

The proposed methodology combines transfer learning with the MobileNetV2 architecture, trained directly within a plugin developed in QGIS, ensuring seamless operation inside the GIS environment. This integration enables automatic identification and precise localization of illegal deposits from georeferenced photos, with visual rendering in an interactive cartographic interface, facilitating operational planning and informed decisionmaking. Tests conducted in Biskra municipality confirm the efficiency of the model in distinguishing between authorized and unauthorized points, highlighting its potential as a scalable tool to support territorial



(A)

Class : 1Unauthorized_point

Confidence : 0.999409199

Address : Biskra , Algeria حي ألف مسكن بسكرة الجزائر

Date : 2024:12:10

Hour : 18:12:25

(B)

Class : 0Authorized_point

Confidence : 0.999978662

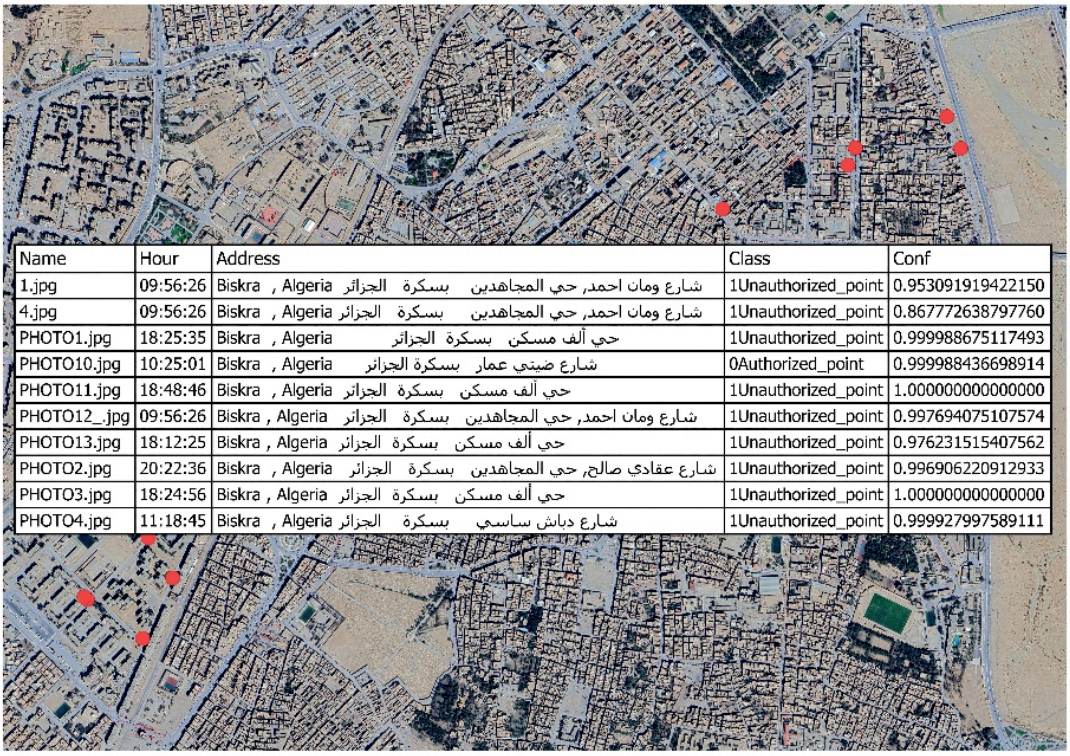
Address : Biskra , Algeria حي السلام بسكرة الجزائر

Date : 2020:02:14

Hour : 11:31:12

1 feature selected on layer photos (PHOTO7.jpg). Coordinate 34,84725° 5,73569° Scale 1:13905 Magnifier 100% Rotation 0,0 ° ☑ Render EPSG:4326

Fig. 8. Plugin image classification results with predictions
(a)illegal waste deposit in authorized point (b) illegal waste deposit in unauthorized point



Name	Hour	Address	Class	Conf
1.jpg	09:56:26	Biskra , Algeria شارع ومان احمد, حي المجاهدين بسكرة الجزائر	1Unauthorized_point	0.953091919422150
4.jpg	09:56:26	Biskra , Algeria شارع ومان احمد, حي المجاهدين بسكرة الجزائر	1Unauthorized_point	0.867772638797760
PHOTO1.jpg	18:25:35	Biskra , Algeria حي آلف مسكن بسكرة الجزائر	1Unauthorized_point	0.999988675117493
PHOTO10.jpg	10:25:01	Biskra , Algeria شارع ضيتي عمار بسكرة الجزائر	0Authorized_point	0.999988436698914
PHOTO11.jpg	18:48:46	Biskra , Algeria حي آلف مسكن بسكرة الجزائر	1Unauthorized_point	1.000000000000000
PHOTO12_.jpg	09:56:26	Biskra , Algeria شارع ومان احمد, حي المجاهدين بسكرة الجزائر	1Unauthorized_point	0.997694075107574
PHOTO13.jpg	18:12:25	Biskra , Algeria حي ألف مسكن بسكرة الجزائر	1Unauthorized_point	0.976231515407562
PHOTO2.jpg	20:22:36	Biskra , Algeria شارع عقادي صالح, حي المجاهدين بسكرة الجزائر	1Unauthorized_point	0.996906220912933
PHOTO3.jpg	18:24:56	Biskra , Algeria حي ألف مسكن بسكرة الجزائر	1Unauthorized_point	1.000000000000000
PHOTO4.jpg	11:18:45	Biskra , Algeria شارع دباش ساسي بسكرة الجزائر	1Unauthorized_point	0.999927997589111

Fig. 9. Plugin output recorded in shapefile attribute table

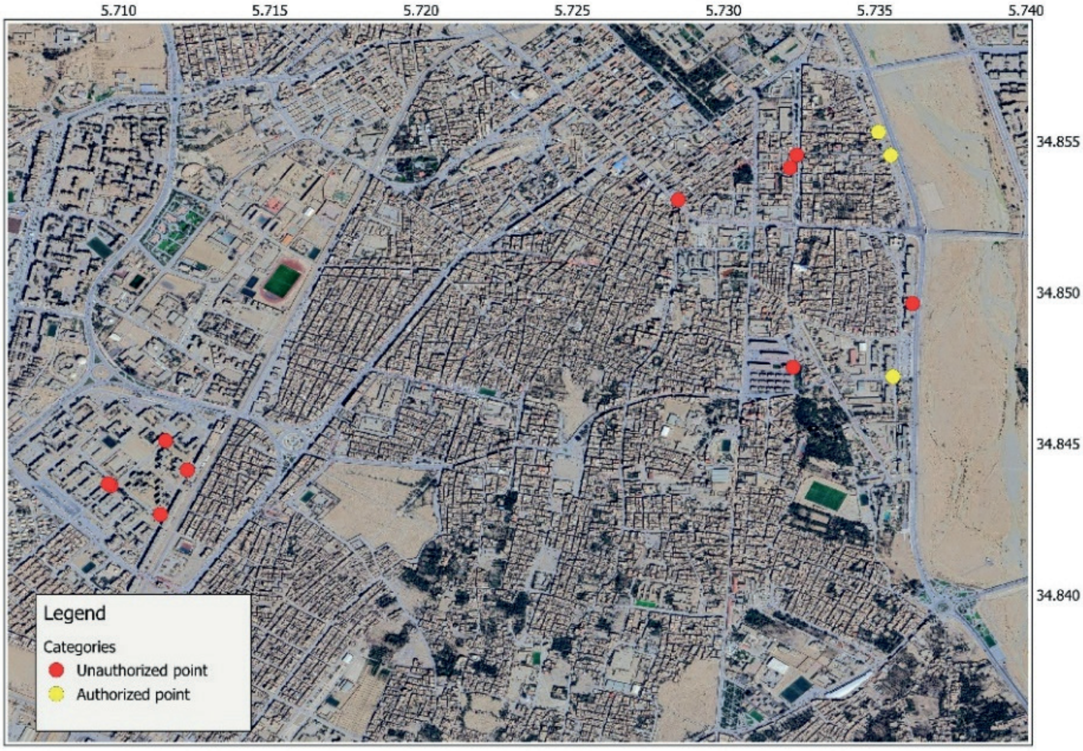


Fig. 10. Distribution of illegal waste deposits by categories in Biskra municipality, based on classified photos

management in municipalities facing illegal dumping challenges.

The main contributions of this work lie in the development of a lightweight and effective classification model that has strong transferability potential and can be implemented in other municipalities worldwide, regardless of their territorial specificities. However, its applicability depends on the availability of georeferenced images, as the model does not support non-geolocated data. This limitation should be considered when planning deployment in municipalities where such data are inconsistently collected. It would be interesting for future research to expand the dataset to further municipalities across the world. ■

REFERENCES

Abdulah S., Alamri F., Nag P., Sun Y., Ltaief H., Keyes D. E., and Genton M. G. (2022). The second competition on spatial statistics for large datasets. arXiv preprint arXiv:2211.03119.

Abdulai H., Hussein R., Bevilacqua E., and Storrings M. (2015). GIS based mapping and analysis of municipal solid waste collection system in Wa, Ghana. Journal of Geographic Information System, 7(2), 85–94. DOI: 10.4236/JGIS.2015.72008.

Adipah S. and Kwame O. N. (2019). A novel introduction of municipal solid waste management. Journal of Environmental Science and Public Health, 3(2), 147–157.

Ahmad M. (2023). AI-Enabled Spatial Intelligence: Revolutionizing Data Management and Decision Making in Geographic Information Systems. In AI and Its Convergence With Communication Technologies, pp. 137–166. IGI Global.

Ahmed Z. Y. (2024). Artificial Intelligence Geographic Information Systems-AI GIS. International Journal of Advanced Engineering and Business Sciences, 5(1).

Agence Nationale des Déchets (2021). Rapport sur l'État de la Gestion des Déchets en Algérie. Agence Nationale des Déchets, 150p.

Asefa E. M., Barasa K. B., and Mengistu D. A. (2022). Application of geographic information system in solid waste management. In Geographic Information Systems and Applications in Coastal Studies. IntechOpen.

Awaluddin M., Hidayat R., and Pratama A. (2023). Enhancing image classification through geometric transformations: A comparative study. Journal of Computer Vision and Applications, 18(2), 45–62.

Awasare S. and Sutar A. (2015). Solid waste management and GIS. International Journal of Research in Environmental Science and Technology, 5(1), 22–28.

Bodea C., Ozunu A., Baciu N., and Măcicăşan V. (2014). Using GIS in waste management – some conceptual considerations. ECOTERRA – Journal of Environmental Research and Protection, 11(1), 61–65.

Bošković G., Cvetanović A. M., Jovičić N., Jovanović A., Jovičić M., and Milojević S. (2024). Digital technologies for advancing future municipal solid waste collection services. In Digital Transformation and Sustainable Development in Cities and Organizations, pp. 167–192. IGI Global. DOI: 10.4018/979-8-3693-3567-3.ch008.

Brus J., Vrkoč J., and Kubásek M. (2016). Design of decision support tools for the quality assessment of illegal dumping notifications based on crowd-sourced data.

Cai T. and Ge J. (2023, July). Design and application of artificial intelligence GIS algorithm based on deep learning technology. In 2023 International Conference on Data Science and Network Security (ICDSNS), pp. 1–5. IEEE. DOI: 10.1109/icdsns58469.2023.10245905.

Caputo A. C. and Pelagagge P. M. (2000). Integrated geographical information system (GIS) for urban solid waste management. WIT Transactions on Ecology and the Environment, 39, 159–169. DOI: 10.2495/URS000181.

Choi Y. (2023). GeoAI: Integration of artificial intelligence, machine learning, and deep learning with GIS. Applied Sciences, 13(6), 3895.

Del Carmen-Niño V., Herrera-Navarrete R., Juárez-López A. L., Sampedro-Rosas M. L., and Reyes-Umaña M. (2023). Municipal solid waste collection: Challenges, strategies and perspectives in the optimization of a municipal route in a southern Mexican town. Sustainability, 15(2), 1083.

Devesa M. R. and Brust A. V. (2021). Mapping illegal waste dumping sites with neural-network classification of satellite imagery. arXiv preprint arXiv:2110.08599.

Direction de la Programmation et du Suivi Budgétaire de Biskra (DPSB). (2024). Monographie de la wilaya de Biskra. Biskra, Algeria.

Dregulo A. M. and Khodachek A. M. (2022). Waste management reform in regions of the Russian Federation: Implementation issues on the way to sustainable development. Geography, Environment, Sustainability, 15(1), 6–13. DOI: 10.24057/2071-9388-2021-078.

Fatunmibi O. and Gbopa A. O. (2021). Dump sites location and its health implications within the Polytechnic, Ibadan using geographical information system approach. International Journal of Research, 8(5), 467–478. DOI: 10.52403/IJRR.20210557.

Ghosh M. and Thirugnanam A. (2021). Introduction to artificial intelligence. In Artificial Intelligence for Information Management: A Healthcare Perspective, pp. 23–44.

Gondhalekar G., Bathala N. K., Merugu N. B., Joshi N., and Kumari P. L. (2024). Enhancing image classification performance through transfer learning and adaptive augmentation: A MobileNetV2 approach.

Hernandez-Garcia A. (2020). Data augmentation and image understanding. arXiv preprint arXiv:2012.14185.

Howard A., Zhmoginov A., Chen L. C., Sandler M., and Zhu M. (2018, June). Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520.

Hua T. M., Nguyen T. K., Van Dinh Thi H., and Thi N. A. N. (2016, December). Towards a decision support system for municipal waste collection by integrating geographical information system map, smart devices and agent-based model. In Proceedings of the 7th Symposium on Information and Communication Technology, pp. 139–146.

Iacoboaea C. and Aldea M. (2011). The assessment of GIS use in waste management. Journal of Applied Engineering Sciences, 1(4), 33–38.

Inamdar S., More P., Chavan M., Pawar A., Patil S., and Raje U. (2023). Smart surveillance system for illegal garbage dumping. International Journal of Advanced Research in Science, Communication and Technology, pp. 293–297. DOI: 10.48175/ijarsct-8599.

Jakeni Y., Maphanga T., Madonsela B. S., and Malakane K. C. (2024). Identification of illegal dumping and community views in informal settlements, Cape Town: South Africa. Sustainability, 16(4), 1429. DOI: 10.3390/su16041429.

Jimoh R., Moradeyo A., Chuma V., Olubukola O., and Yusuf A. (2019). GIS based appraisal of waste disposal for environmental assessment and management in Mainland area of Lagos state, NG. International Journal of Environment and Geoinformatics, 6(1), 76–82.

Karimi N. and Ng K. T. W. (2022). Mapping and prioritizing potential illegal dump sites using geographic information system network analysis and multiple remote sensing indices. Earth, 3(4), 1123–1137.

Khalil R., Benyettou A., and Toumi A. (2023). Zoom-based augmentation for small object detection in urban environments. International Journal of Artificial Intelligence and Smart Systems, 11(1), 77–89.

Khan M. S. A., Husen A., Nisar S., Ahmed H., Muhammad S. S., and Aftab S. (2024). Offloading the computational complexity of transfer learning with generic features. PeerJ Computer Science, 10, e1938.

Kim Y. and Cho J. (2022). AIDM-Strat: augmented illegal dumping monitoring strategy through deep neural network-based spatial separation attention of garbage. Sensors, 22(22), 8819.

Krsmanović M., Šušnjar S., Golijanin J., and Valjarević A. (2022). GIS based vulnerability assessment of illegal waste disposal – case study East Sarajevo. DOI: 10.7251/afts.2022.1427.063k.

Kumar S., Patel R., and Singh A. (2025). Affine transformations for robust image classification: A multi-domain analysis. Pattern Recognition and Machine Intelligence, 39(1), 112–130.

Marange F., Muteveri M., Chipfunde F., and Mapira J. (2023). Challenges confronting local authorities in solid waste management: The case of Dangamvura residential area, Mutare, Zimbabwe. European Journal of Social Sciences Studies, 8(5). DOI: 10.46827/ejsss.v8i5.1477.

Mohanta S. K., Mohapatra A. G., Mohanty A., and Nayak S. (2024). Deep learning is a state-of-the-art approach to artificial intelligence. In Deep Learning Concepts in Operations Research, pp. 27–43. Auerbach Publications.

Muñoz-Aseguinolaza D., García-González A., and Pérez J. (2023). Rotation-invariant convolutional networks for aerial waste detection. Remote Sensing Letters, 14(5), 389–405.

Nandi G. C. (2023). Deep learning. pp. 123–158. DOI: 10.1002/9781394173167.ch4.

Nanni L., Lumini A., and Brahnam S. (2021). Image translation techniques for deep learning: A survey and experimental evaluation. Expert Systems with Applications, 165, 113891.

Narayanan N. and Arjun K. P. (2024). Introduction to deep learning. pp. 32–60. DOI: 10.1201/9781003504900-2.

Office National des Statistiques (ONS). (2008). General Census of Population and Housing. Algiers, Algeria.

Padubidri C., Kamilaris A., and Karatsiolis S. (2022, March). Accurate detection of illegal dumping sites using high resolution aerial photography and deep learning. In 2022 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops), pp. 451–456. IEEE.

Panesar A. (2020). What is artificial intelligence? pp. 1–18. DOI: 10.1007/978-1-4842-6537-6_1.

Paunović S. and Krstić F. (2014). GIS application in the spatial analysis of illegal landfills in big cities: A case study of Belgrade. Bulletin of the Serbian Geographical Society, 94(3), 41–54.

Pathak N., Biswal G., Goushal M., Mistry V., Shah P., Li F., and Gao J. (2024). Smart city community watch — camera-based community watch for traffic and illegal dumping. Smart Cities, 7(4), 2232–2257.

Puigcerver J., Riquelme C., Mustafa B., Renggli C., Pinto A. S., Gelly S., and Houlsby N. (2020). Scalable transfer learning with expert models. arXiv preprint arXiv:2009.13239.

Qi Y., Zhang L., and Chen H. (2021). Rotation-based data augmentation for improved generalization in CNNs. IEEE Transactions on Image Processing, 30, 4567–4579.

Raj A. (2024). Artificial intelligence. International Journal for Science Technology and Engineering, 12(11), 646–655. DOI: 10.22214/ijraset.2024.64695.

Rao K. R., Sreekeshava K. S., Dharek M. S., and Sunagar P. (2020). Issues on planning of solid waste management scheme through evaluation in integrated data information system. In AIP Conference Proceedings, Vol. 2204, No. 1. AIP Publishing. DOI: 10.1063/1.5141548.

Sakshi, Neeti K., and Singh R. (2023). Diverse applications of remote sensing and geographic information systems in implementing integrated solid waste management: A short review. Engineering Proceedings, 56(1), 109. DOI: 10.3390/asec2023-15340.

Sandler M., Howard A., Zhu M., Zhmoginov A., and Chen L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520.

Saxena P., Saxena V., Pandey A., Flato U., and Shukla K. (2023). Multiple aspects of artificial intelligence. Book Saga Publications.

Shafik W. (2024). Deep learning impacts in the field of artificial intelligence. In Deep Learning Concepts in Operations Research, pp. 9–26. Auerbach Publications.

Shahab S. and Anjum M. (2022). Solid waste management scenario in India and illegal dump detection using deep learning: An AI approach towards sustainable waste management. Sustainability, 14(23), 15896.

Shrivastava A., Pandey A., Singh N., Srivastava S., Srivastava M., and Srivastava A. (2024). Artificial intelligence (AI): Evolution, methodologies, and applications. International Journal for Research in Applied Science and Engineering Technology, 12(4), 5501–5505.

Syafrudin S., Ramadan B. S., Budihardjo M. A., Munawir M., Khair H., Rosmalina R. T., and Ardiansyah S. Y. (2023). Analysis of factors influencing illegal waste dumping generation using GIS spatial regression methods. Sustainability, 15(3), 1926.

Talaei Khoei T., Ould Slimane H., and Kaabouch N. (2023). Deep learning: Systematic review, models, challenges, and research directions. Neural Computing and Applications, 35(31), 23103–23124.

Thompson A. F., Afolayan A. H., and Ibidunmoye E. O. (2013, July). Application of geographic information system to solid waste management. In 2013 Pan African International Conference on Information Science, Computing and Telecommunications (PACT), pp. 206–211. IEEE. DOI: 10.1109/SCAT.2013.7055110.

Torres R. N. and Fraternali P. (2021). Learning to identify illegal landfills through scene classification in aerial images. Remote Sensing, 13(22), 4520.

United Nations Development Programme (2023). Harnessing the role of private sector in waste management through South-South and Triangular Cooperation for inclusive urbanization, pp. 10–10. DOI: 10.18356/9789213585207c004.

Vozenilek V. (2009). Artificial intelligence and GIS: Mutual meeting and passing. In 2009 International Conference on Intelligent Networking and Collaborative Systems, pp. 279–284. IEEE.

Wang T., Li Y., and Zhou M. (2025). Scaling strategies in convolutional neural networks: Impacts on classification accuracy. Journal of Machine Learning Research, 26(4), 215–230.

Yang X., Wang Y., Lu K., Wu Y., and Zhao D. (2024). Artificial neural network modelling in GIS spatial analysis. Academic Journal of Computing and Information Science, 7(6), 32–37.

Yu D., Yoon J., and Lee Y. (2024, July). Detection and register of illegal garbage dumping action using the consecutive processing and Embedded-NAS. In 2024 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1–8. IEEE. DOI: 10.1109/avss61716.2024.10672606.

Zainun N. Y., Rahman I. A., and Rothman R. A. (2016, November). Mapping of construction waste illegal dumping using geographical information system (GIS). In IOP Conference Series: Materials Science and Engineering, Vol. 160, No. 1, p. 012049. IOP Publishing. DOI: 10.1088/1757-899X/160/1/012049.

Zeng W. (2024). Image data augmentation techniques based on deep learning: A survey. Mathematical Biosciences and Engineering, 21(6), 6190–6224.

Zorić M., Štula M., Markić I., and Braović M. (2024, September). Transfer learning in building neural network model case study. In 2024 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 1–6. IEEE.

Check for updates

cc BY

RESEARCH PAPER

THE CONCEPT OF HEALTH IN ATLAS MAPPING

Svetlana M. Malkhazova[1], Tatiana V. Kotova[1], Dmitry S. Orlov[1*], Nilufar K. Komilova[2], Nasiba I. Safarova[3], Li Wang[4]
[1]Lomonosov Moscow State University, Faculty of Geography, Leninskie Gory, 1, Moscow, 119991, Russia
[2]National University of Uzbekistan named after Mirzo Ulugbek, Department of Ecology, University street, 4, Tashkent, 100174, Uzbekistan
[3]Nizami Uzbek National Pedagogical University, Department of Geography and Methods of Teaching Geography, Bunyodkor, 27, Tashkent, 100185, Uzbekistan
[4]Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing, 100101, China
***Corresponding author:** orlovds@list.ru

Received: April 2nd 2025 / Accepted: February 26th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-4401

ABSTRACT. The concept of "health" is a multidimensional phenomenon reflecting the idea of the physical and mental state of the body. This is one of the most important characteristics of society, closely related to economic and social development, demographic structure, the state of the environment, as well as the country's public health management system. A special place in the information support for solving health problems is given to atlases, primarily those created based on health concepts.

Through the example of some domestic and foreign atlas works, the features of public health mapping are considered. A review of health atlases of various purposes, territorial coverage, and subject matters showed: 1) the concept of "health" is cartographically represented by a significant number of general comprehensive and sectoral atlases; 2) the primary sources of information for preparing atlases and illustrating the concept of "health" remain statistical data; 3) the sections "Demography", "The Natural and Social Environment", and "Lifestyle and Risk Factors" are included in virtually all comprehensive health atlases; 4) a comprehensive analysis of regional health determinants and their cause-and-effect relationships requires the integration of supplementary, region-specific data sources; 5) the thematic coverage of atlas-based health mapping is achieved through interdisciplinary collaboration.

The Medico-Geographical Atlas of Russia "Population Health", currently under development at the Faculty of Geography of Lomonosov Moscow State University, constitutes a continuation of a series of previously published medical-geographical atlases. The conceptual framework of the Atlas is based on new approaches in medical geography, moving "from the geography of disease to the geography of health".

KEYWORDS: public health, atlas mapping, concept of health, Medico-geographical atlas of Russia "Population Health"

CITATION: Malkhazova S. M., Kotova T. V., Orlov D. S., Komilova N. K., Safarova N. I., Wang L. (2026). The Concept Of Health In Atlas Mapping. Geography, Environment, Sustainability, 1 (19), 152-166
https://doi.org/10.24057/2071-9388-2026-4401

ACKNOWLEDGEMENTS: The study was carried out within the framework of the project of the Russian Geographical Society № 09/2025 "Medico-Geographical Atlas of Russia "Population Health" and the State Assignment of Lomonosov Moscow State University with support of the MSU Development Program (#1220).

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

The concept of "health" is a multidimensional phenomenon reflecting the physical and mental state of an organism (Petkau 2015). When applied to the population of a country, region, or specific administrative territory, the concept of "health" is interpreted as "public health" (Prokhorov 2001). This is one of the most crucial characteristics of society as a social organism: its potential for continuous harmonious growth and socioeconomic development. Public health serves as the most striking and comprehensive indicator of living conditions, as it largely determines the demographic and economic development of individual states and territories.

In recent years, public health has emerged as a strategic priority within state policy for numerous nations. As emphasized by the renowned cardiac surgeon Leo Bokeria at the "Health of the Nation – the Foundation of Russia's Prosperity" forum, *"Health is not just medicine. It is infrastructure, lifestyle, and even politics…"* Public health is closely linked to economic and social development, demographic structure, morbidity, environmental conditions, and the national public health management system (Lisitsyn and Sakhno 1989; Matros 1992; Lisitsyn and Ulumbekova 2011; Medik and Yuryev 2016). While the concept is multifaceted, several core principles are widely acknowledged:

- Public health is a multidimensional, dynamic state of society, a resource for productive life;
- Public health has many aspects (biological, demographic, socio-economic, historical, ethnographic, etc.) and is conditioned by many factors, which means it requires an integrated approach for its study and for decision-making;

- Elements of public health always have certain regional specifics;
- Currently, a universally accepted set of indicators and criteria for a composite health index remains elusive, despite numerous methodological approaches that have been tried.

In recent years, the public health system in our country has become a national priority. According to the Russian Minister of Health, creating a public health system has become "a task not only for federal, regional, and municipal structures, but also for public associations, volunteer movements, and the media. Such a system must become a kind of network operating at all levels and encompassing all areas of human life".

The national projects for 2025-2030 include the "Long and Active Life" project. This project is a successor to the "Healthcare" project and aims to develop and implement a comprehensive program to protect maternal and reproductive health and increase life expectancy in rural areas. The goal of the national project is to increase life expectancy to 78 years by 2030 and to 81 years by 2036, including accelerated growth in healthy life expectancy indicators.

Comprehensive public health information is crucial for both the development and evaluation of healthcare efforts, determining their effectiveness, and addressing population health issues across individual, group, and societal levels.

With various information sources available, including statistical publications, census and medical examination results, sociological surveys, expert assessments, databases, and annual reports from international organizations such as the WHO and World Bank, etc., comprehensive atlases, primarily those based on health concepts, play a pivotal role in providing information, taking into account the specific characteristics of the mapping object.

The Medico-Geographical Atlas of Russia "Population Health" is currently under development at the Faculty of Geography at Lomonosov Moscow State University. This atlas continues a series of previously published medico-geographical atlases of Russia (Natural Focal Diseases 2015, 2017; Healing Springs and Plants 2019; Risk Factors for Oncological Diseases 2024).

The purpose of this paper is to provide an overview of the main traditional atlases and to prepare the scientific and methodological foundation for the development of the Medico-Geographical Atlas of Russia "Population Health".

The objectives of this article are:

1. To review the atlases describing public health;
2. To analyze, through a review of select domestic and international atlases, the specific features and methodologies of public health mapping, with the aim of advancing its substantive and methodological foundations;
3. To present the preliminary conceptual framework of the Medico-Geographical Atlas of Russia "Population Health".

MATERIALS AND METHODS

In pursuit of the stated objectives, a systematic examination of domestic and international atlases was conducted. The selection criteria ensured a representative sample across various purposes, geographical scales, and thematic orientations to investigate the conceptualization of "health". Source materials, procured in both print and digital formats from the Russian State Library and online databases, were analyzed in conjunction with a critical review of the relevant academic literature.

RESULTS AND DISCUSSION

As a key societal characteristic, public health in Russian mapping is a necessary subject of study in comprehensive national and regional atlases.

The *National Atlas of Russia* (2007) contains health maps, primarily in the sections "Population and Social Development" (Population Reproduction; Birth Rate; Mortality; Life Expectancy; Population Health, including Public Health Status, Population Morbidity subsections) and "Social Sphere" (Healthcare: Organization of Medical and Preventive Care for the Population; Medical Personnel: Physicians; Nursing Personnel; Healthcare Financing subsections). Collectively, these provide a comprehensive assessment of population health across more than 90 indicators. Among these, the composite map of "Public Health Status" (Fig. 1) is of particular interest.

Fig. 1. Map "State of public health". Scale 1:60,000,000

A significant amount of information on public health is presented in the comprehensive Ecological Atlas of Russia (2017). This resource features cartographic representations of socially significant and natural focal diseases, mortality rates, and a composite "Public Health Index". The latter is accompanied by a textual analysis detailing population health status across the regions of the Russian Federation (Fig. 2).

Individual population health maps (Morbidity, Mortality, Healthcare System) can be found in general comprehensive regional atlases (Chistobaev and Semenova 2013). For illustrative purposes, a representative map from the *Atlas of the Baikal Region: Society and Nature* (2021) is provided (Fig. 3).

In domestic mapping, public health has emerged as a distinct thematic focus in a series of specialized and comprehensive national and regional atlases (Fig. 4). The following section provides a detailed examination of these cartographic works.

Of the national-level thematic atlases documenting the health status of the country's population, the *Atlas of Health of Russia* is particularly noteworthy. To date, it remains the only national-level publication that is unique in the comprehensiveness of its health indicators. Published on an almost annual basis since 2005, the most recent iteration, Issue 15, was released in 2019 (Atlas. Health of Russia 2019). The Atlas is based on extensive statistical information provided by the "League of the Nation's Health." According to its developers, it is intended for Russian citizens, with a primary focus on regional administrators at various levels, for whom it serves as a tool for situational monitoring and evidence-based decision-making. It is also intended for a broad spectrum of users engaged in professional and public activities related to disease prevention, environmental improvement, and the mitigation of adverse social phenomena, as well as for any individuals with an interest in the contemporary state of the national healthcare system.

The Atlas presents key indicators for the demographic and socioeconomic development of the Russian Federation, healthcare, and environmental impacts. Each annual edition significantly expands the scope of traditional sections (Demography, Morbidity, Resource Provision for Healthcare

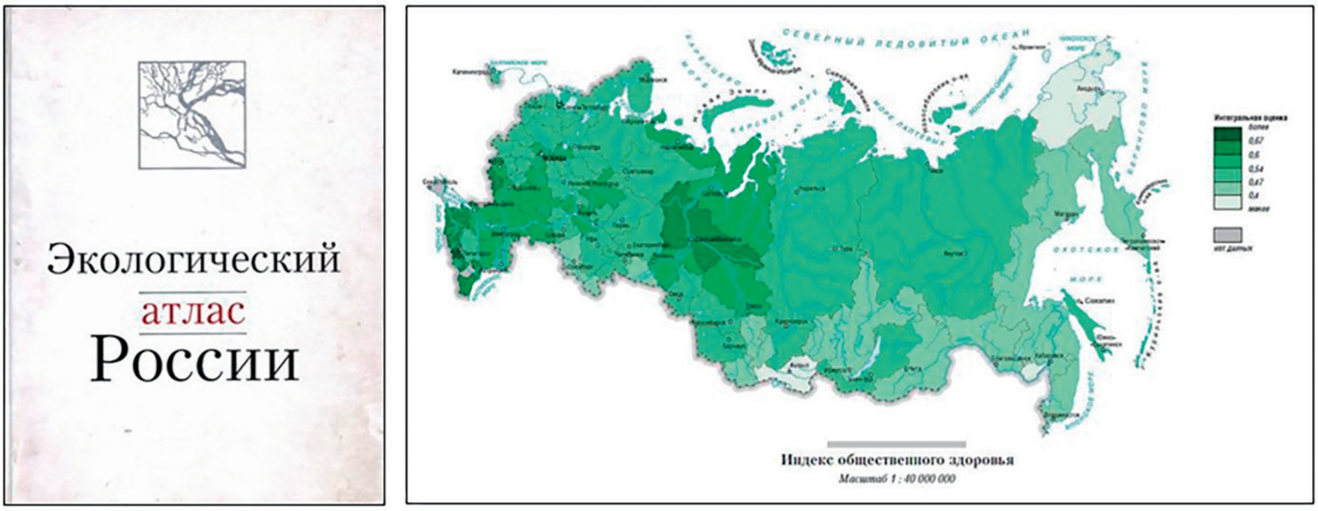


Fig. 2. Map "Public Health Index". Scale 1:40,000,000

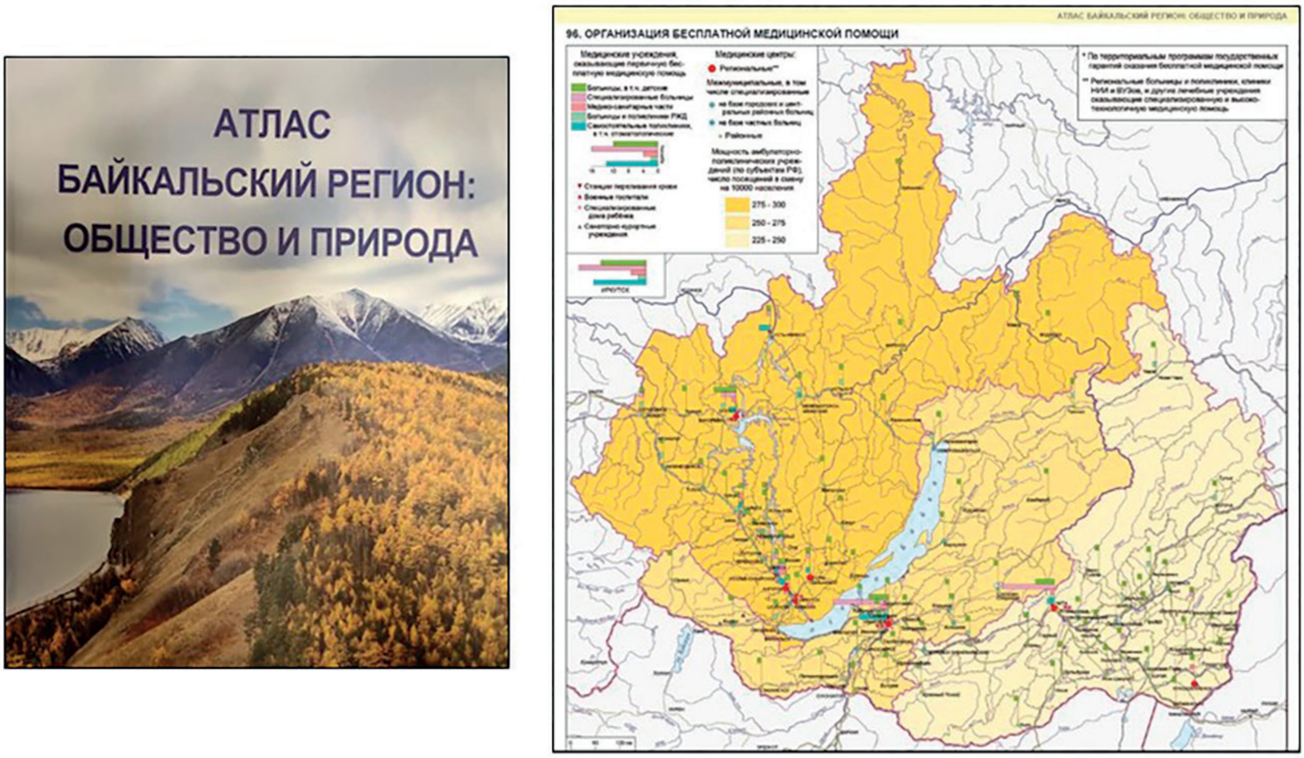


Fig. 3. Map "Organization of free medical care". Scale 1:6,000,000

Fig. 4. Examples of Health atlases

Organizations, Factors Affecting Population Health). For example, "Section 12. Socioeconomic Indicators in the Russian Federation" displays over 40 indicators, grouped into subsections: Living Standards, Employment, Pension Provision, Living Conditions, Agricultural Production, and Agricultural Product Consumption.

As a rule, new sections are added annually: Development of Physical Culture and Sports, Crime as a Threat to Health and National Security, Penitentiary Medicine, etc. Issue 15 contains over 230 maps, grouped into 16 sections; for comparison, the initial editions contained only 6 to 8 sections.

The data within the Atlas are conveyed through a combination of cartograms, constructed from multi-year statistical datasets, table data, line diagrams, and explanatory text. A notable example in this domain is Russia's first atlas on this subject, *"The Environment and Health of the Population of Russia"* (1995). This comprehensive work, produced in both print and digital formats in Russian and English, is structured into eight chapters: 1) Russia and its Population; 2) The Natural Environment and Biophysical Factors of Population Life Support; 3) The Economic and Technical Complex and its Impact on Living Conditions; 4) Social Processes and their Impact on Public Health; 5) Population Morbidity in Russia; 6) Disability; 7) Mortality; and 8) Life Expectancy, as well as a Conclusion section. Each chapter is further subdivided into two to seven thematic subchapters. An innovative cartographic representation from the subsection "4.1. The Structure and Stability of Society in the Regions of Russia" is illustrated in Fig. 5.

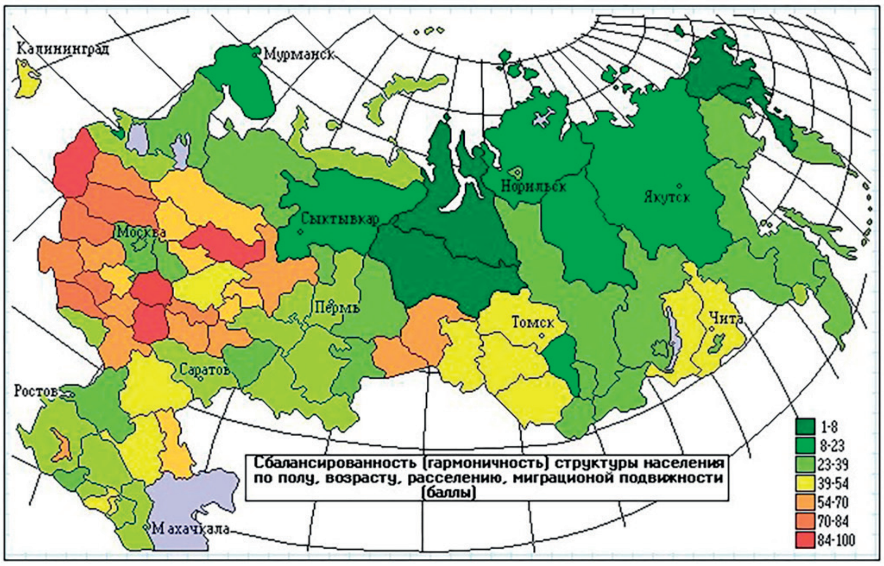


Fig. 5. Map "Balance (harmony) of the population structure by gender, age, migration mobility (points)"

The concluding section, "Development of a List of Population Health Indicators Related to the State of the Environment for Use in Databases for Assessing the Impact of the Environment on Health" delineates the core requirements for such indicators and offers specific recommendations for metrics related to population health, as well as the quality of atmospheric air, drinking water, and the water supply system.

The content is presented using cartograms at a scale of approximately 1:60,000,000, supplemented by diagrams, textual analyses, and tables. The temporal scope of the data varies, with the maximum coverage spanning the period from 1980 to 1996.

Among the regional health atlases, the medical and geographical atlases of Yakutia are of significant interest. In 2005, the *Medico-geographical Atlas of Population Health and Healthcare in the Republic of Sakha (Yakutia) at the Turn of the Century* was prepared, followed by the *Medico-geographical Atlas of Population Health Protection in the Republic of Sakha (Yakutia)* in 2012. The former contains more than 85 maps, distributed across the following sections: 1. Medical-Demographic Situation (6 indicators); 2. Population Morbidity and Disability (24 indicators); 3. Healthcare Resource Provision (8 indicators); 4. Healthcare Institution Activities (13 indicators); 5. Treatment and Preventive Institutions across 35 uluses (districts). Population susceptibility to various diseases is represented by two distinct metrics: prevalence and incidence (Fig. 6).

The second atlas contains 164 maps characterizing the socio-demographic situation, population morbidity, and the activities of medical and preventive care, and physical education and health facilities in the administrative-territorial units of the Republic of Sakha (Yakutia) – uluses (districts). These indicators are assessed and compared with similar indicators for the Russian Federation and the Far Eastern Federal District.

A separate publication, *The Atlas of Population Health of the Astrakhan Region* (2010), provides a focused analysis on maternal and child health, alongside region-specific environmental determinants. Its structure is organized into six thematic sections: 1) Demographic Situation; 2) Population Morbidity; 3) Children's Health; 4) Microelement and Vitamin Content in Children's Bioenvironments; 5) Women's Health; 6) Risk of Environmental Factors on Population Health. The final section is of particular relevance, as it quantifies critical regional health challenges, including the toxic and carcinogenic risks associated with atmospheric chemical exposure and the public health hazards of drinking water contamination (Fig. 7). The district cartograms are supplemented with diagrams and text descriptions.

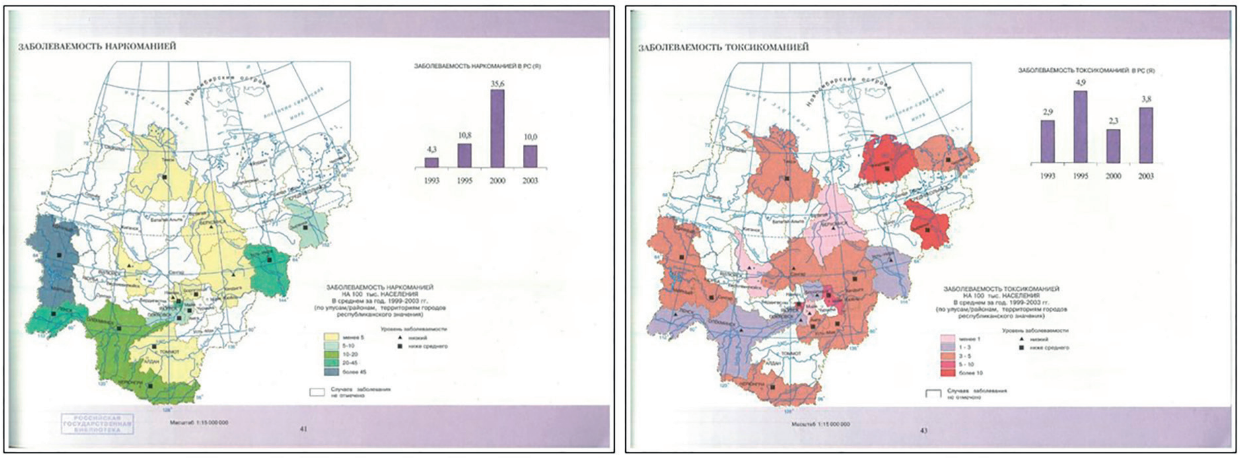


Fig. 6. Maps "Incidence of drug addiction" and "Incidence of toxicomania". Scale 1:15,000,000

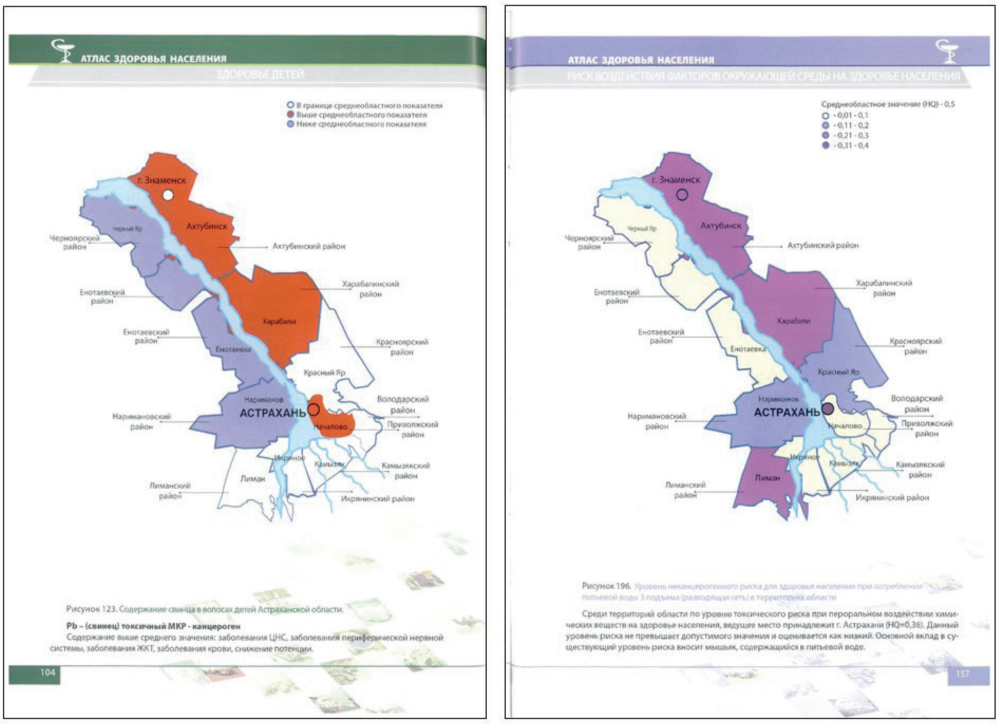


Рисунок 123. Содержание свинца в волосах детей Астраханской области.

Pb – (свинец) токсичный МКР - канцероген

Содержание выше среднего значения: заболевания ЦНС, заболевания периферической нервной системы, заболевания ЖКТ, заболевания крови, снижение потенции.

Рисунок 196. Уровень неканцерогенного риска для здоровья населения при потреблении питьевой воды 3 подъема (разводящая сеть) в территориях области

Среди территорий области по уровню токсического риска при пероральном воздействии химических веществ на здоровье населения, ведущее место принадлежит г. Астрахани (HQ=0,36). Данный уровень риска не превышает допустимого значения и оценивается как низкий. Основной вклад в существующий уровень риска вносит мышьяк, содержащийся в питьевой воде.

Fig. 7. MCartograms "Level of non-carcinogenic risk to public health when consuming drinking water of the 3rd rise of distribution network" and "Lead content in children's hair"

Significant attention in the study and regional mapping of population health is devoted to specific disease categories, such as mental health disorders (Fig. 8).

The International Association of Ethnopsychologists and Ethnopsychotherapists (IAEE), the concept of which belongs to RAMS Academician V. Ya. Semka, has been instrumental in advancing this field of study. A principal contribution of the IAEE has been the systematic development of a series of regional atlases, including those for Siberia and the Far East (1988), the Republic of Buryatia (2007), Tomsk (2008), Kemerovo (2011), Omsk (2012) regions, as well as Zabaikalsky Krai (2020).

The structure of these atlases is exemplified by the *Atlas of Mental Health of the Kemerovo Region*. This publication is organized into five chapters: 1) Historical, climatic, geographic, and socioeconomic characteristics of the region, including psychiatric services; 2) Prevalence of mental disorders and schizophrenia across administrative territories; 3) Prevalence of non-psychotic mental disorders and mental retardation; 4) Mental disorders in children and adolescents; and 5) The regional drug addiction situation.

Examples of cartograms for the Kemerovo Region and the Zabaikalsky Krai are shown in Fig. 9.

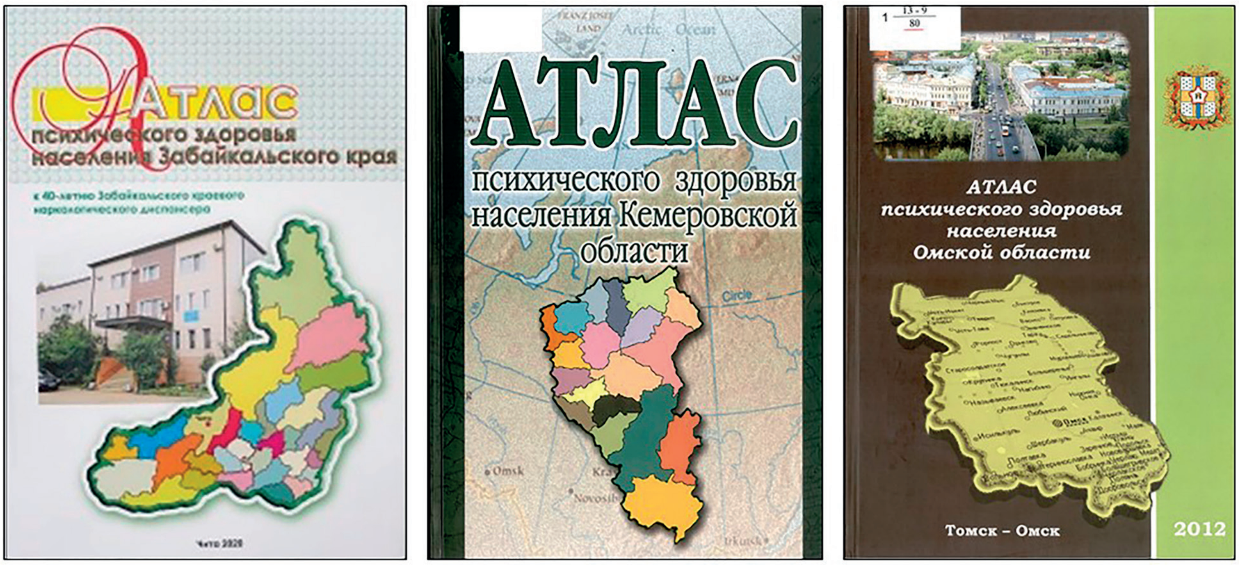


Fig. 8. Atlases of mental health of the population

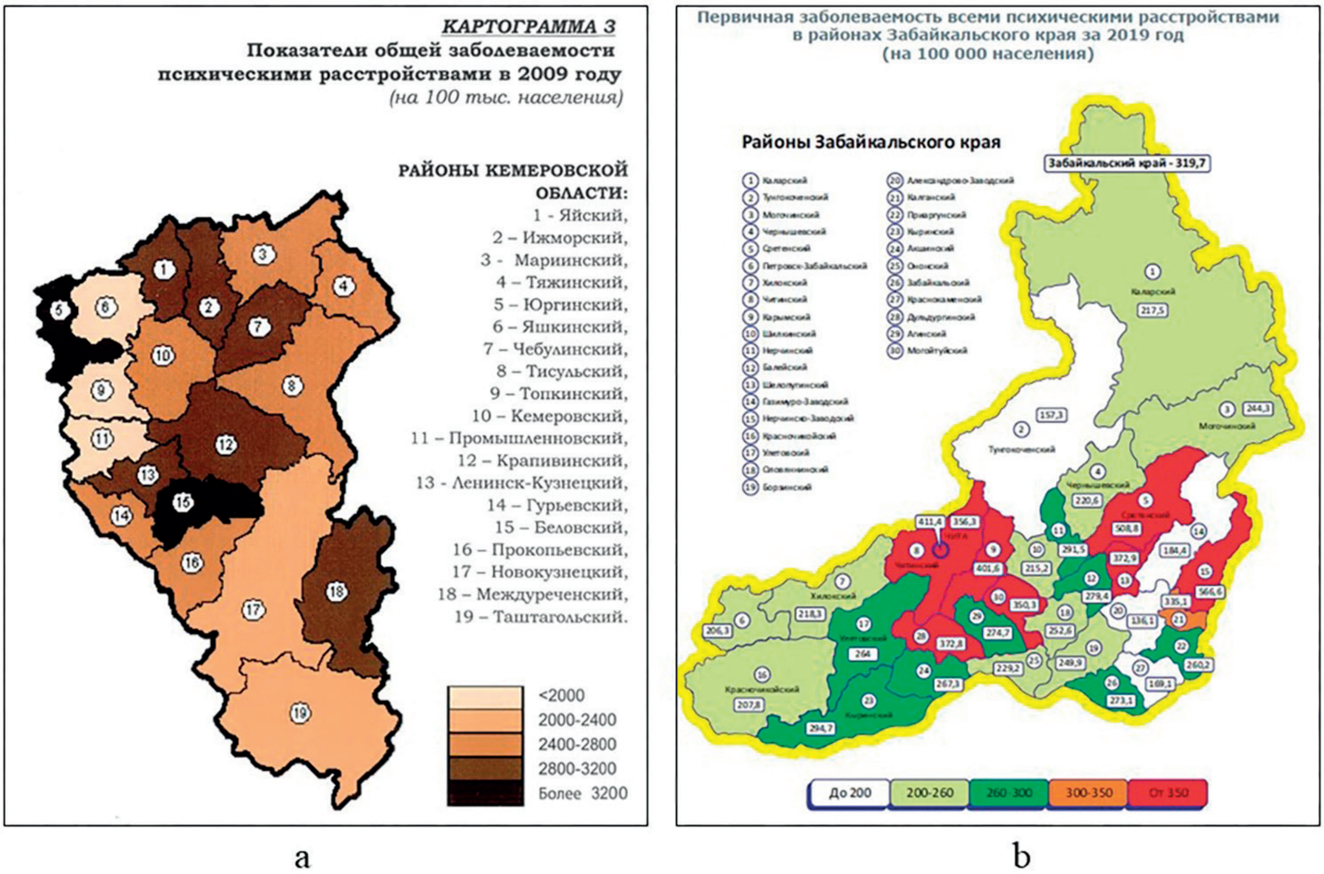


Fig. 9. a) Atlas of mental health of the population of the Kemerovo region. Cartogram "Indicators of the overall incidence of mental disorders in 2009 (per 100,000 population)"; b) Atlas of mental health of the population of Zabaikalsky Krai. Cartogram "Primary incidence of all mental disorders in the districts of Zabaikalsky Krai for 2019 (per 100,000 population)"

As previously noted, the study of disease prevalence necessitates the consideration of a multifactorial framework of environmental determinants. This includes not only sociodemographic and sociocultural factors but also ethnocultural and ethnoecological variables.

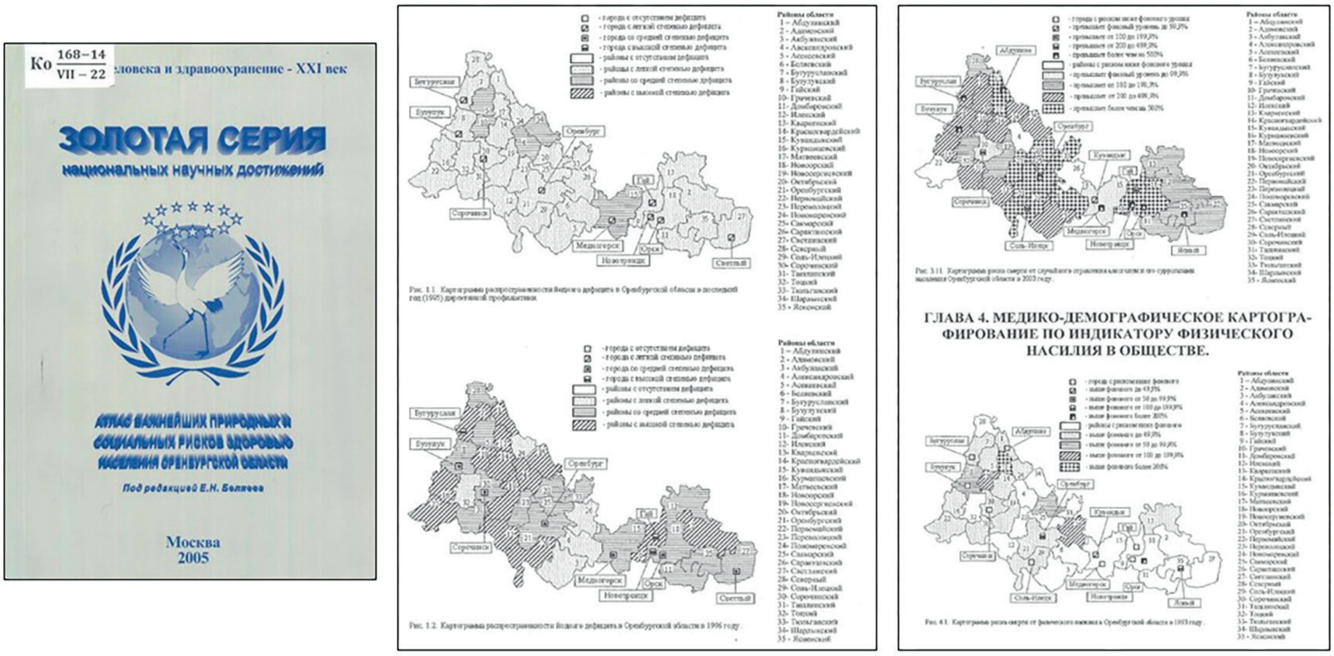
Natural focal diseases are a serious threat to public health. The researchers from the Faculty of Geography at Lomonosov Moscow State University have made significant contributions through the publication of the *Medical and Geographical Atlas of Russia "Natural Focal Diseases"* (2015, 2017). This atlas represents the first comprehensive cartographic compilation of data on diseases whose pathogens circulate independently of humans within natural ecosystems. The significance of this work is substantial, particularly given the vast, sparsely populated, and underdeveloped territories of Russia. An illustrative map from this atlas is provided in Fig. 10.

Another notable contribution is the *Atlas of the Most Important Natural and Social Health Risks in the Orenburg Region* (2005), designed to facilitate an in-depth analysis of population health determinants in relation to natural, geochemical, and anthropogenic risks. The atlas is divided into two parts. Part One comprises eight chapters: 1) Medical and demographic mapping by iodine deficiency parameters (Fig. 11a); 2) Medical and demographic mapping by iron deficiency in nutrition; 3) Medical and demographic mapping by parameters of the risk of premature death from poisoning by alcohol and surrogates; 4) Medical and demographic mapping by the indicator of physical violence in society (Fig. 11b); 5) Medical and demographic mapping by the indicator of mental health; 6) Medical and demographic mapping of fluorine deficiency; 7) Ecological and hygienic characteristics of trends in the dynamics of the most important natural and social health risks; 8) Ecological and hygienic ranking of risk based on average long-term data. Part Two contains four chapters providing methodological explanations for the mapping of demographic processes, primary morbidity, and leading causes of mortality, alongside a hygienic characterization and territorial ranking based on these indicators. The analysis covers the period from 1995 to 2004.

More recently, the scope of cartographic research on risk factors has been expanded by the *Medical and Geographical Atlas of Russia "Risk Factors for Oncological Diseases"* (2024), also created by the Faculty of Geography at Lomonosov Moscow State University. The relevance of this publication is underscored by the critical status of public health in Russia and the prominent role of cancer among the leading causes of mortality. The atlas presents a broad classification of risk factors, systematically organized into the following sections: 1) Prerequisites for Cancer (Demography, Health, Natural Conditions, Socioeconomic Determinants); 2) Key Risk Factors (Environment, Lifestyle, Living and working conditions, Occupational risks, Infectious agents); 3) Prevalence and Mortality from Malignant Neoplasms; and finally; 4) Risk Factor Mitigation and Cancer Prevention (Oncology service infrastructure, Use of natural resources for treatment, Public prevention initiatives, and Regional policy directions). A representative map from this atlas is shown in Fig. 12.

Some of the key public health indicators (primarily demographic, such as life expectancy at birth, mortality; also incidence, etc.) are reflected in a number of regional medical-demographic, medical-geographical, and medical-ecological atlases. These include the following atlases: *Medical-Demographic Atlas of the Kaliningrad Region* (2007);

Fig. 10. Map "Tick-borne encephalitis distribution". Scale 1:30,000,000



a b

Fig. 11. a) Cartogram of the prevalence of iodine deficiency in the Orenburg Region in the last year (1995) of directive prevention; cartogram of the prevalence of iodine deficiency in the Orenburg Region in 1996.
b) Cartogram of the risk of death from accidental poisoning by alcohol and its surrogates of the population of the Orenburg Region in 2003; cartogram of the risk of death from physical violence in the Orenburg Region in 1993

Fig. 12. Map "Smoking prevalence. 2021"

Medical-Demographic Atlas of the Moscow Region (2007), *Medical-Geographical Atlas of the Smolensk Region* (2012), and *Medical-Ecological Atlas of the Voronezh Region* (2010) (Fig. 13). The atlases are structured in four sections: 1) Introductory; 2) Population and Social Conditions; 3) Population Health; and 4) Health Care. The Environmental Quality section of the *Medical and Ecological Atlas of the Voronezh Region* contains approximately 20 maps that display medical and demographic processes, general population morbidity, incidence of specific disease classes, the frequency of reproductive pathologies, nutritionally induced diseases, and some particularly dangerous natural focal infections, etc.

Globally, the field of population health has been increasingly documented through a series of specialized atlases. A discernible trend in this domain is the shift towards electronic publications, which facilitate the recording, processing, and dynamic updating of extensive datasets.

Traditional atlas cartographic studies are an important tool in studying and assessing Chinese population health, consistently used to address pressing health-related issues. A substantial number of atlases have been developed to date, exhibiting considerable diversity in their thematic scope. This review will highlight several seminal atlases without an exhaustive structural analysis.

The severity of oncological disease problems became the basis for studying their distribution and preparing a number of atlases (Fig. 14).

The *Atlas of Cancer Mortality in the People's Republic of China* (1979) represents a foundational bilingual (Chinese-English) publication. It was constructed from a nationwide cause-of-death survey (1973-1975) encompassing 29 provincial-level administrative regions and a population of 850 million. Utilizing mathematical models and classification indices, the atlas delineated the geographical distribution of 15 malignant tumors across 2,392 county-level units. This work systematically identified, for the first time, high-risk geographical patterns for cancers such as oesophageal and nasopharyngeal carcinoma. The analytical frameworks established in this atlas provided a foundation for oncological research over the subsequent three decades.

In subsequent years, systematic work has been underway to create cancer atlases. In 2015, an updated

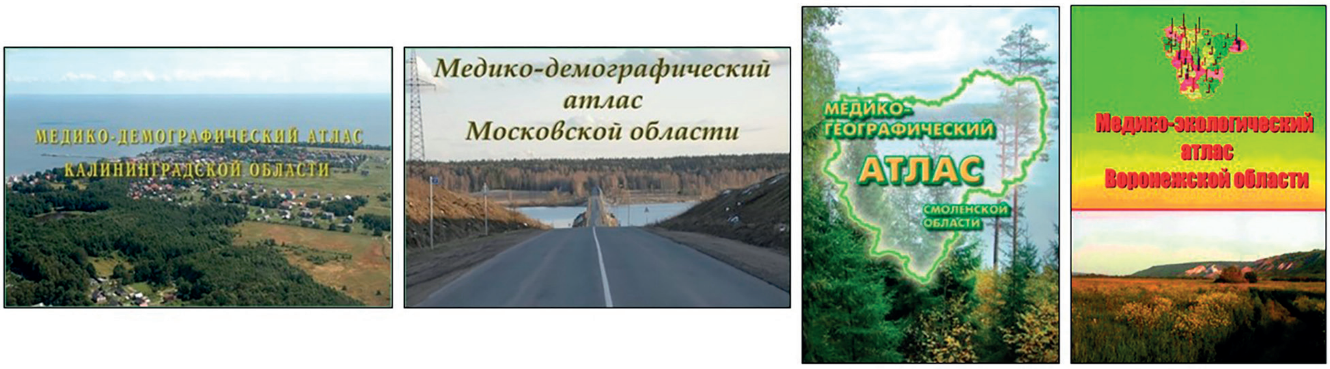


Fig. 13. Regional atlases

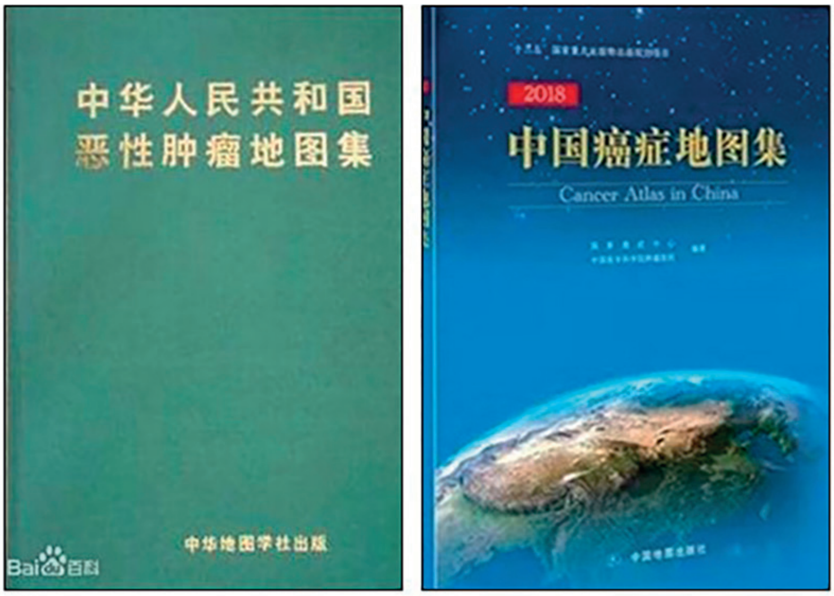


Fig. 14. The Atlases of Cancer Mortality in the People's Republic of China

county-level atlas (China Cancer Atlas 2015) was published, coordinated by the Cancer Institute of the Chinese Academy of Medical Sciences in collaboration with the Institute of Geographic Sciences and Natural Resources Research (IGSNRR CAS) and the National Center for Chronic and Noncommunicable Diseases Control and Prevention (CDC). It combines data from the national tumor registry, three rounds of mortality surveys, and demographic statistics. The atlas depicts the geographic distribution of major cancer types in China, providing reliable scientific data for interdisciplinary research, early diagnosis/treatment, and optimization of healthcare resource allocation. The atlas serves as a basis for medium- and long-term chronic disease prevention strategies in China and provides data to the WHO and the International Agency for Research on Cancer (IARC).

Much attention is paid to the study and mapping of endemic diseases (Fig. 15). The *Atlas of Endemic Diseases and the Environment of the People's Republic of China* (1989) was created as part of an interdisciplinary project (1980–1988) supervised by a joint committee of the Ministry of Health and the Chinese Academy of Sciences. It synthesizes geographic data on four major endemic diseases (Keshan disease, Kashin-Beck disease, endemic goitre / cretinism, and fluorosis) with environmental factors on 221 maps at the national, regional, and local levels. It constituted a milestone in medical geography by innovatively establishing linkages between geochemical imbalances and disease etiology, thereby informing subsequent policy and research directions.

Research on population-based diseases continued and was reflected in the *Atlas of Endemic Diseases of China*

Fig. 15. The Atlases of endemic diseases in the People's Republic of China

(2020). The atlas is a national publication project of the 14th Five-Year Plan (2021-2025) and documents China's public health campaigns against endemic diseases (including Keshan disease, Kashin-Beck disease, iodine deficiency/excess disorders, fluorosis, and arsenic poisoning) from 1949 to 2020, employing multi-scale mapping, dynamic surveillance data, and elimination assessments.

An example of a regional-level atlas of endemic diseases is the Atlas of *Endemic Diseases of Yunnan Province* (1989). It was compiled by a group of medical and geographic experts organized by the Yunnan Provincial Health Department. Published by the People's Medical Publishing House in 1989, the atlas includes introductory maps, 66 distribution maps for diseases such as Keshan disease, endemic goiter and cretinism, fluorosis, and brucellosis, as well as 18 supplementary maps. It comprehensively, systematically, and clearly depicts the distribution patterns, epidemic patterns, regional differences, and the relationship with the geochemical environment of four human malaria pathogens – *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae,* and *Plasmodium ovale*. The atlas holds both theoretical and practical significance.

The *Atlas of Major Endemic Diseases and Geoecology of the Tibetan Plateau was published* in 2025. It summarizes and reflects the relationships between the major geochemical endemic diseases of the Tibetan Plateau and the chemical, natural, and anthropogenic environments. It also examines the effectiveness of prevention and control measures. This atlas provides an overview of the spatiotemporal patterns of plague epizootics in the Himalayan marmot population on the Tibetan Plateau since the 1950s, revealing the ecological characteristics of natural plague foci and the extent to which they have changed due to human activity. The atlas consists of five sections (Introduction; Distribution of Kashin-Beck Disease; Distribution of Endemic Fluorosis; Environmental Arsenic Exposure and Risk; Plague Epidemic Patterns and Ecological Characteristics of Himalayan Marmots).

A consistent study of natural focal diseases served as the basis for the development of several epidemiological atlases of China (Fig. 16).

The sustained investigation of natural focal diseases has similarly yielded significant cartographic outputs, notably the three-volume *Epidemiological Atlas of Natural Focal Diseases of China* (2019) and the bilingual *Atlas of Plague and the Environment of the People's Republic of China* (2000). The former comprises over 1,600 maps detailing the distribution of 33 natural focal diseases, one vector-borne disease, and more than 1,100 host and vector species (encompassing 158 pathogens) across 2,900 counties. It provides a systematic categorization of disease-environment interactions, host/vector distributions, and complex epidemiological patterns. The latter atlas analyzes over two centuries of plague epidemiology in China through 284 maps and diagrams, covering spatiotemporal dynamics, natural foci, host/vector ecology (rodents and fleas), Yersinia pestis biotypes, and historical control measures.

The *Atlas of Schistosomiasis in China* (2012) delineates the prevalence of schistosomiasis, including human and bovine infection rates, and the distribution of intermediate host snails, detailing their geographic extent, density, and infection rates at national and key regional levels.

The relationship between public health and the environment is further expanded on in the *Atlas of Climate Risks to Health* (2021), which operates at the provincial scale. This publication monitors over 20 indicators of climate-related health impacts and tracks the progression of climate change adaptation initiatives throughout China (Fig. 17).

A broader contextualization of public health is presented in the *Atlas of Population, Environment, and Sustainable Development of the People's Republic of China* (2000). This work integrates health metrics with a comprehensive set of environmental and socioeconomic indicators. Its 184 maps, organized into nine thematic sections, combine demographic trends, environmental pressures (including land, water, and air pollution), and regional sustainable development strategies.

The *Atlas of the Elderly Population of the People's Republic of China* published in 1986. It is a 224-page compendium derived from national census data. It delineates age-group distributions across 46 categories (e.g., 60+, 65+, 80+, 90+, 100+) at the county level, complemented by foundational maps of socioeconomic and environmental conditions (Fig. 18).

Addressing a critical research gap, the *Atlas of Gender Equality and Women's Development in China* (2010) represents one of the first national atlases on this subject. Its eight sections of cartographic representations systematically assess gender equality across domains, including politics, employment, education, healthcare, marriage and family, and environmental contexts.

Fig. 16. Epidemiological Atlas of Natural Focal Diseases of China. Atlas of Plague and the Environment of the People's Republic of China. Atlas of Schistosomiasis in China

Fig. 17. Atlas of Climate Risks to Health

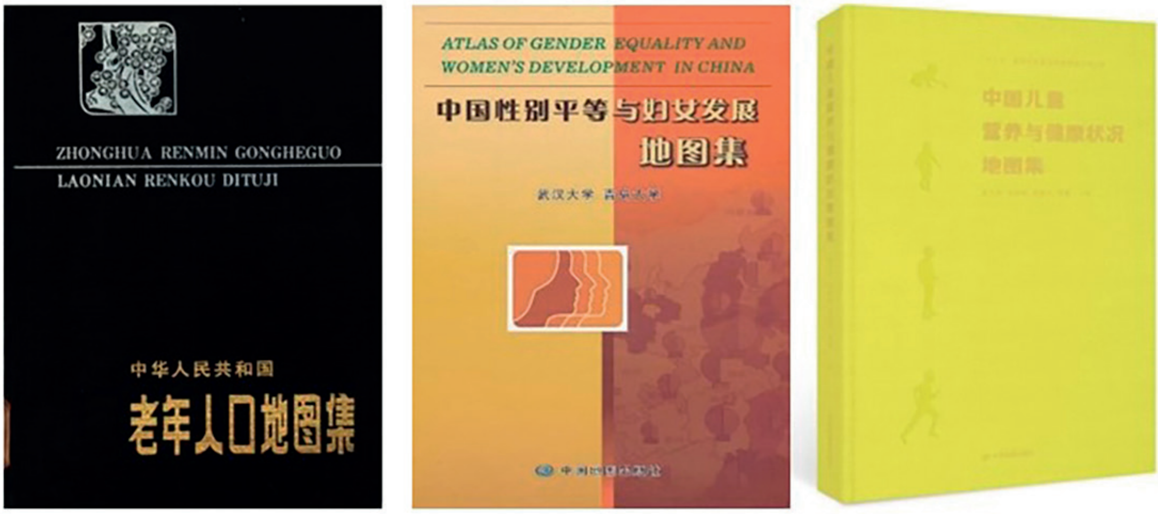


Fig. 18. Atlas of the Elderly Population of the People's Republic of China. Atlas of Gender Equality and Women's Development in China. Atlas of Child Nutrition and Health in China

A more recent and distinctive contribution is the *Atlas of Child Nutrition and Health in China* (2024). Its 144 maps are systematically divided into six sections (Overview, Dietary Nutrition, Growth and Development, Malnutrition, Physiology / Metabolism, and Knowledge / Behavior), providing detailed data on infant feeding practices, nutrient intake, physical and mental development, undernutrition, obesity, metabolic indicators (e.g., blood pressure, lipids, glucose), and associated lifestyle factors.

Multifaceted medico-geographical research is being conducted in the Republic of Uzbekistan. The results are presented in several atlases.

The *Medico-Geographical Atlas of the Republic of Uzbekistan* (O´zbekiston respublikasining…) includes three parts: 1) Syrdarya and Jizzakh Regions, 2022; 2) Tashkent and Tashkent Region, 2024; 3) Republic of Karakalpakstan and Khorezm region, 2025 (Fig. 19). Part 1, consisting of 87 maps, reflects the nosoecological situation at national and regional scales. For Uzbekistan as a whole, in addition to traditional introductory maps (Natural Conditions, Political and Administrative Structure, Population, etc.), there are demographic maps, mortality (4 maps), morbidity (21 maps), and healthcare system maps (for example, Number of employees in sanatorium-resort institutions, Outpatient and Polyclinic Institutions in Uzbekistan, Activities of Hospital Institutions in Uzbekistan, Activities of Obstetrics and Gynecology Departments, Service Network, etc.). For the regions, in addition to introductory maps displaying the main geographic factors influencing the health of the population, morbidity maps are provided. At the same time, attention is focused on diseases that are socially significant or specific to these regions due to a number of natural, ethnic, and other factors.

Part 2 presents nosogeographic maps of Tashkent and the Tashkent region and focuses on Tashkent and the Tashkent Region, with a thematic emphasis on "diseases of civilization" prevalent in urban settings: incidence of neoplasms, endocrine diseases, nutritional diseases, metabolic disorders, blood and hematopoietic diseases, diseases with certain immune system disorders, mental and behavioral diseases, diseases of the nervous system, circulatory system, and respiratory organs. Several environmental maps of Tashkent city and region have also been created: 1) Green spaces; 2) Air pollution (TVOC, total volatile organic compounds), noise; 3) Drinking water supply to apartments (houses); 4) Atmospheric air monitoring; 5) Air pollution (PM2.5; PM10); 6) Air pollution (CO; CO_2).

Part 3 is intended to reflect the geoecological situation of the region, as well as the dynamics of individual classes of diseases in the context of climate change, through nosogeographic maps of the Republic of Karakalpakstan and the Khorezm region, covering regional morbidity, the incidence of certain infectious and parasitic diseases, and the incidence of tuberculosis, viral hepatitis, scarlet fever, salmonella infections, measles, and bacterial dysentery.

Work is currently underway on Part 4 of the atlas. Part 4 focuses on the medical and geographical situation of the

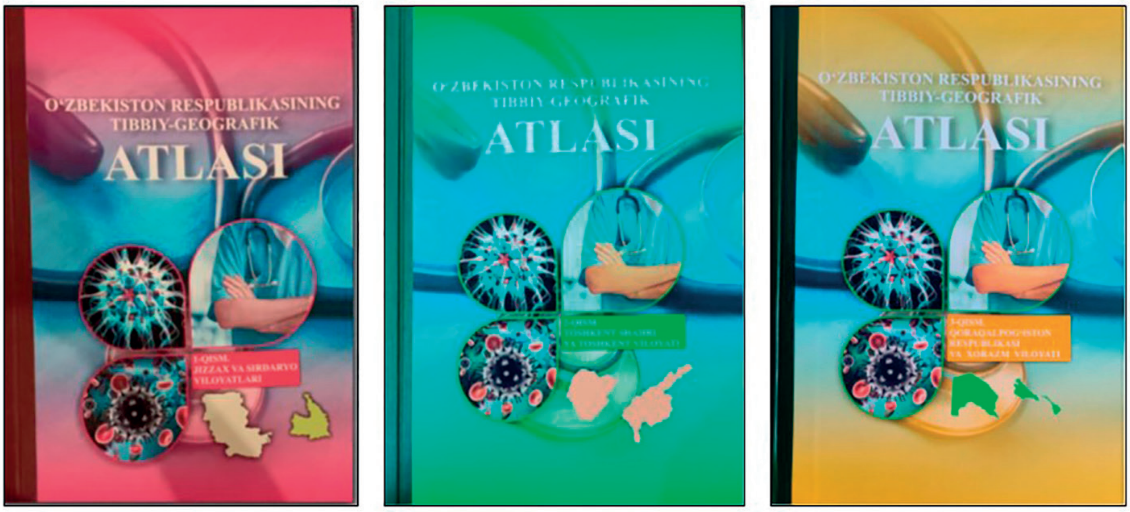


Fig. 19. Atlas of the Republic of Uzbekistan. Part 1. Syrdarya and Jizzakh regions; Part 2. City of Tashkent and Tashkent region; Part 3. Republic of Karakalpakstan and Khorezm region

Kashkadarya and Surkhandarya regions. This atlas primarily includes information on the epidemiological situation and the geographic distribution of infectious and parasitic diseases.

Active research on health and health systems in a wide variety of areas, with the results documented in atlases, is conducted at both the country and global levels, primarily under the auspices of the WHO and other international organizations. For the purposes of this paper, we will limit ourselves to examples of global atlases covering various topics.

The *Global Health Atlas* is a unified electronic platform that compiles standardized data and statistics on infectious diseases at the country, regional, and global levels for analysis and comparison. Data analysis and interpretation are complemented by information on demographics, socioeconomic conditions, and environmental factors. The Atlas highlights the wide range of factors influencing the spread of infectious diseases. This multi-year database focuses on key poverty-related diseases (malaria, HIV/AIDS, tuberculosis, etc.), as well as diseases on the path to eradication (dracunculiasis, leprosy, lymphatic filariasis), epidemic and emerging infections (meningitis, cholera, yellow fever), and drug resistance. In addition to epidemiological information, the system provides data on key support services, such as the network of Centers for Disease Control and Prevention, the Global Outbreak Alert and Response Network, and others. The database is constantly updated.

The *Surveillance Atlas of Infectious Diseases* was prepared by the European Centre for Disease Prevention and Control. The atlas contains data (for subpopulations: confirmed cases and data quality; indicators: registered cases, notification rate, age-standardized rate, hospitalized cases, number of deaths, proportion of cases associated with travel, etc.; and time) on more than 60 infectious diseases. The interface allows users to interact with the data, creating various maps and tables that provide an overview of the status of various infectious diseases in European countries.

Further advancing this field, the *Atlas of Population Health in the European Union Regions* (Atlas of Population Health in European… 2017) employs a composite metric, the Population Health Index (PHI), to provide a multidimensional assessment of European population health (Fig. 20). The atlas presents regional-level data across 17 health dimensions, such as employment and income, social protection, education, demographic change, health behaviors, environmental pollution, housing conditions, infrastructure, health system resources, expenditure, efficiency, and mortality and morbidity. The PHI itself is constructed from 39 indicators deemed relevant for a comprehensive assessment at the European level, supplemented by analyses of 22 additional indicators. The atlas is structured in three parts: 1) a foundation of theoretical and methodological considerations; 2) cartographic representations of the PHI across European regions; and 3) an analysis of regional inequalities in the determinants of population health. This final section is further divided, providing detailed data on the 39 PHI indicators and contextual information from the supplementary indicators to elucidate the geographic patterns of the population health index.

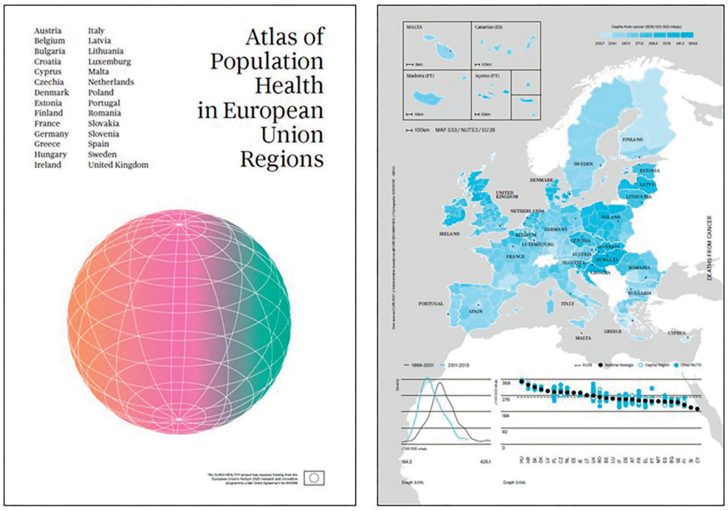


Fig. 20. Atlas of Population Health in the European Union Regions

Among specialized international atlases, the mental health atlases periodically produced by the World Health Organization (WHO) are noteworthy. They are considered the most comprehensive sources of information on mental health worldwide and an important tool for developing and planning mental health services in countries and regions. The first edition was published in 2001. The main goal of the atlases is to collect information on policies, plans, financing, human resources, and the provision of mental health services in different countries, as well as health monitoring and evaluation. Data is collected through a questionnaire sent to the health authorities of each country.

By providing governments and stakeholders with comprehensive data on mental health systems, services, and resources in different countries, the atlases help identify gaps and opportunities for improving mental health measures worldwide. The frequency of updates depends on the availability of data and resources, as well as the demand for new information. Typically, the update cycle is three years. For example, the *Mental Health Atlas* 2024 consists of five sections: 1) Mental Health Information Systems and Research; 2) Mental Health System Management; 3) Financial and Human Resources for Mental Health; 4) Access to and Use of Mental Health Services; and 5) Mental Health Promotion and Prevention. The content is presented through diagrams and text (Fig. 21). The publication is particularly important because it includes information and data on progress towards achieving the goals and targets of the Comprehensive Mental Health Action Plan for the period from 2013 to 2030.

The *Atlas of Health and Climate* (2012) was prepared by the World Health Organization and the World Meteorological Organization. It contains information on the links between weather, climate, and major health issues. The range of issues covers diseases caused by poverty, emergencies resulting from extreme weather events and epidemic outbreaks, as well as environmental problems. The influence of other factors on the relationship between health and climate is noted, such as psychology and individual behavior, environmental and socioeconomic conditions of the population, and the coverage and effectiveness of health programs. The atlas consists of three sections with subsections: 1) Infections (Malaria, Diarrhoea, Meningitis, Dengue Fever); 2) Emergencies (Floods and Cyclones, Drought, Airborne Hazards); 3) Emerging Environmental Issues (Heat Stress, UV Radiation, Pollen, Air Pollution). A representative map from this atlas is provided in Fig. 22.

As noted above, in addition to printed publications, online atlases dedicated to population health are currently being created. For example, in the United States, a significant number of atlases have been prepared, covering various geographic and thematic areas. They present many indicators that determine the structure of healthcare and population health. One of the most well-known projects is the Dartmouth Atlas (www.dartmouthatlas.org). The project was created by the Dartmouth Institute for Health Policy and Clinical Practice and operated from 1996 to 2024. The project provides comprehensive information and analysis at the national, regional, and local levels on the distribution and use of medical resources, identifying differences in the efficiency and effectiveness of healthcare across the country, including individual hospitals (capacity, utilization, costs, and outcomes) and their physicians. It uses

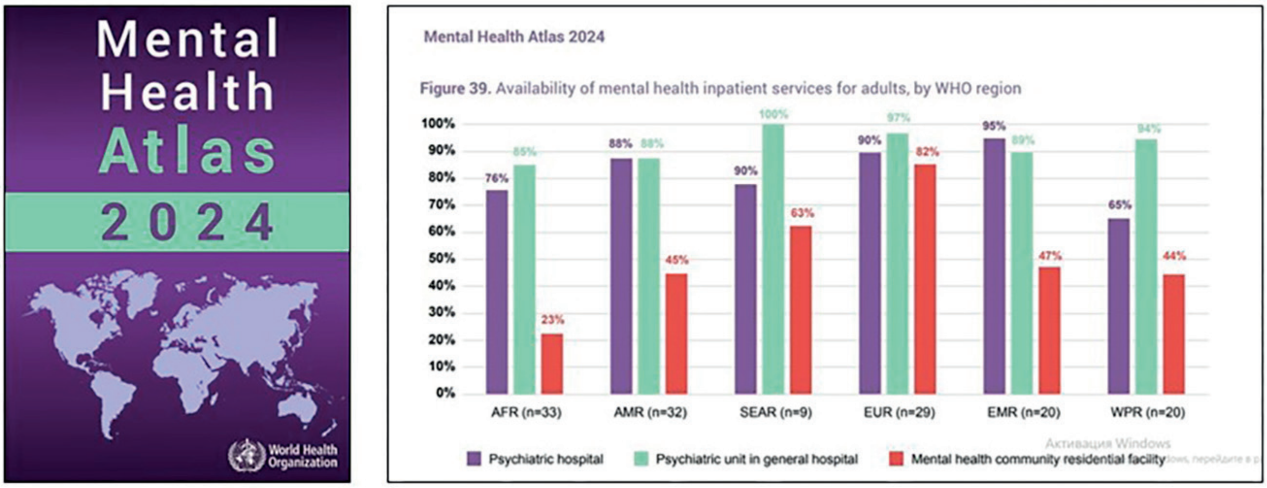
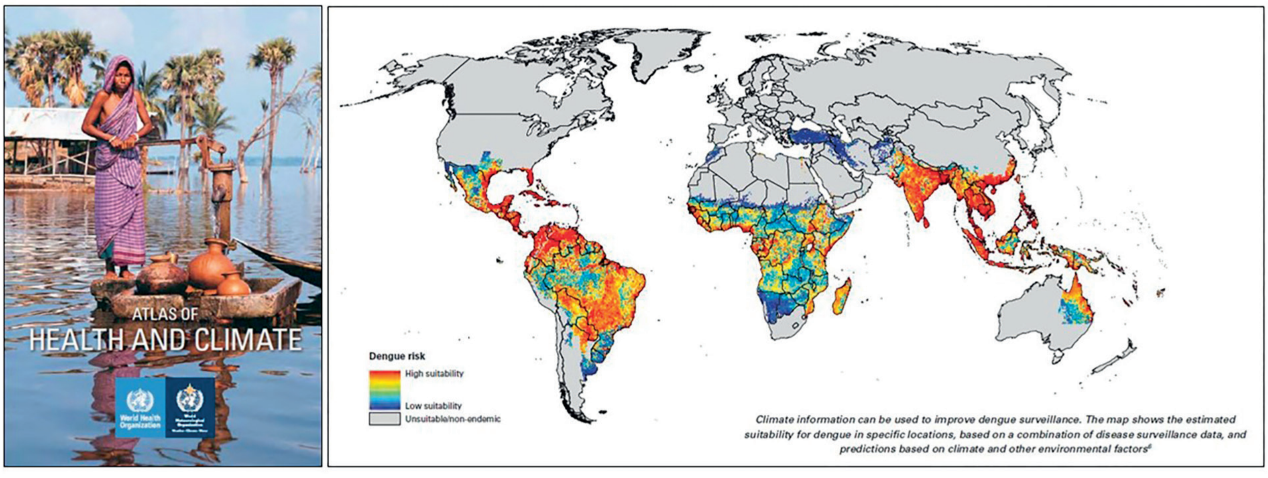


Fig. 21. Availability and reporting of mental health data in responding countries, by WHO regions



Climate information can be used to improve dengue surveillance. The map shows the estimated suitability for dengue in specific locations, based on a combination of disease surveillance data, and predictions based on climate and other environmental factors[6]

Fig. 22. Map of the spread of dengue fever depending on climatic conditions and other factors

data from various federal agencies, the American Hospital Association, and the American Medical Association. It includes maps, figures, and tables on topics such as physician personnel and the diagnosis and treatment of various diseases.

There are also other electronic health atlases, both for individual countries and regions. A detailed review of them requires a separate study. A number of atlases can be found on the Dartmouth Atlas website.

In a related development, the Faculty of Geography of Lomonosov Moscow University has initiated the development of the *Medico-Geographical Atlas of Russia "Population Health".* This project is founded upon a review of existing domestic and international health atlases and builds upon the faculty's prior experience in producing medical-geographical atlases (e.g., Natural Focal Diseases 2015, 2017; Healing Springs and Plants 2019; Risk Factors for Oncological Diseases 2024).

As noted above, the concept of the Atlas is based on new approaches in medical geography of moving "from the geography of diseases to the geography of health". This approach prioritizes the identification and cartographic representation of conditions and factors that contribute to enhanced public health, highlighting positive vital statistics such as high life expectancy, centenarian prevalence, and low rates of morbidity, mortality, and disability.

The preliminary structure of the atlas is organized into six sections: 1) Introduction; 2) Medical and Demographic Situation; 3) Socioeconomic Conditions and Lifestyle; 4) Health-improving Potential of the Environment; 5) Healthcare; 6) Tactics and Strategy of Actions to Improve Population Health.

CONCLUSIONS

A review of atlases with diverse scopes, territorial coverages, and thematic objectives yielded the following findings:

1) Being the most important characteristic of society, public health is a necessary subject of study in global, national, and regional atlas mapping.

2) The concept of "health" is represented by a significant number of general comprehensive and sectoral atlases. Along with traditional atlases, there is a clear trend toward electronic publications.

3) The set of indicators and the completeness of the "health" concept's representation vary depending on the purpose, concept of the atlas, and the completeness of the information base. Research continues to identify the dominant health factors and objectify its assessment.

4) Standardized sections, such as "Demography", "Natural and Social Environment", and "Lifestyle and Risk Factors" are included in virtually all health atlases.

5) The principal information sources for atlas compilation remain statistical data, which are primarily visualized through cartograms and schematic diagrams. These datasets form the foundation for subsequent spatial analysis and modeling to highlight cause-and-effect relationships. Along with traditional methods of representation, there is a widespread use of infographics. Cartographic representations are typically supplemented by explanatory text and table data.

6) A more comprehensive representation, analysis, and interpretation of cause-and-effect relationships in population health at the regional scale is facilitated by the integration of supplementary, region-specific information sources, alongside expert assessments and population survey data.

7) The thematic breadth, analytical depth, and conceptual focus of health mapping in atlases are achieved through an interdisciplinary methodology. This approach engages collaboration of medical geographers and specialists from other scientific fields at the intersection of medicine, ecology, and geography.

The creation of the Medico-Geographical Atlas of Russia "Population Health" at the Faculty of Geography of Lomonosov Moscow University, in a series of previously created medical-geographical atlases, should contribute to the information support for solving public health problems and improving the country's healthcare system. ■

REFERENCES

Atlas of Cancer Mortality in the People's Republic of China. (1979). Office of Cancer Prevention and Control Research, Ministry of Health, Chinese Cartography Society, 99 p.

Atlas of Child Nutrition and Health in China. (2024). Zhao Wenhua, Li Shujuan, Yang Zhenyu, Zhang Qian, China Map Press, 150 p.

Atlas of Health and Climate (2012). WHO, WMO. 64 p.

Atlas of health in Europe. 2nd edition, (2008), 128 p.

Atlas of health of the population of the Astrakhan region. (2010). Astrakhan: Medical expert center: Volga, 159 p. (in Russian)

Atlas of key endemic disease and geo-environment in Tibet Plateau. (2025). Yang Linsheng, Gong Hongqiang, Zhang Qiang., Sinomap press, Beijing.

Atlas of mental health of the population of Kemerovo region. (2011). Tomsk, Kemerovo, 100 p. (in Russian)

Atlas of mental health of the population of Omsk region. (2012). Tomsk, Omsk, 132 p. (in Russian)

Atlas of mental health of the population of the Republic of Buryatia. (2007). Tomsk, Ulan-Ude, 140 p. (in Russian)

Atlas of mental health of the population of the Zabaikalsky Krai. 2nd ed. revised and enlarged. (2020). Novosibirsk: DEALSIB, 128 p. (in Russian)

Atlas of mental health of the population of Tomsk region. (2021). Tomsk, 2008. 119 p. (in Russian).

Atlas of Population Health in European Union Regions. (2017). Santana P., Costa C., Freitas A., Stefanik I., Quintal C., Bana e Costa C., Borrell C. et al. Coimbra: Imprensa da Universidade, 267 p. DOI: 10.14195/978-989-26-1463-2.

Atlas of the Aged Population of the People's Republic of China. (1986). Institute of Geriatrics, Ministry of Health, Beijing Hospital, Ministry of Health, Map Press.

Atlas of the main mental illnesses in Siberia and the Far East. (1988). Tomsk, 94 p. (in Russian)

Atlas of the most important natural and social risks to the health of the population of the Orenburg region. (2005). Moscow.: Academy of Earth Sciences, 69 p. (in Russian)

Atlas. Baikal Region: Society and Nature. (2021). Moscow: Paulsen, 320 p. (in Russian)

Atlas. Health of Russia. (2019). Moscow: A.N. Bakulev NMRCCS, Iss. 15, 472 p. (in Russian)

China Cancer Atlas. (2018). National Cancer Center, Cancer Hospital, Chinese Academy of Medical Sciences, China Map Publishing House, 200 p.

Chistobaev A.I. and Semenova Z.A. (2013). Medical and geographical mapping in the former USSR and modern Russia. Bulletin of St. Petersburg State University, Ser. 7, Issue 4, 109-118. (in Russian)

Endemic Diseases Atlas of China. (2021). Sun Dianjun China Map Press, 236 p.

Environmental and health atlas of Russia. (1995). Moscow: PAIMS Publishing House, 448 p. (in Russian)

Epidemiological Atlas of Natural Focus Diseases in China. (2020). Cao Wuchun, Fang Liqun, Wang Juanle, Science Press, 908 p.

Gender Equality and Women's Development Atlas of China. (2010). Yu Dongmei, Huang Rentao, China Map Press, 118 p.

Lisitsyn Yu.P. and Sakhno A.V. (1989). Human health as a social value. Moscow: Mysl, 270 p. (in Russian)

Lisitsyn Yu.P. and Ulumbekova G.E. (2011). Public health and healthcare. 3rd ed. revised and enlarged. Moscow: GEOTAR-Media, 544 p. (in Russian)

Matros L.G. (1992). Social aspects of health problems. Novosibirsk: Science. Sib. publishing company, 158 p. (in Russian)

Medic V.A. and Yuriev V.K. (2021). Public health. 2nd ed., rev. and additional. Moscow: GEOTAR-Media, 420 p. (in Russian).

Medico-demographic Atlas of the Kaliningrad Region. (2007). Moscow: Lomonosov Moscow State University, Faculty of Geography, 86 p. (in Russian)

Medico-demographic Atlas of the Moscow Region. (2007). Moscow: Lomonosov Moscow State University, Faculty of Geography, 110 p. (in Russian)

Medico-ecological atlas of the Voronezh region. (2010). Voronezh: Istoki, 166 p. (in Russian)

Medico-geographical Atlas of Russia "Healing Springs and Plants" (2019). Edited by S.M. Malkhazova. Faculty of Geography, Lomonosov Moscow State University, Moscow, 304 p. (in Russian)

Medico-geographical Atlas of Russia "Natural Focal Diseases" (2015). Edited by S.M. Malkhazova. Moscow: Faculty of Geography, Lomonosov Moscow State University, 208 p. (in Russian)

Medico-geographical Atlas of Russia "Natural Focal Diseases" (2017). Edited by S.M. Malkhazova. 2nd revised. Moscow: Faculty of Geography, Lomonosov Moscow State University, 216 p. (in Russian)

Medico-geographical Atlas of Russia "Risk factors for oncological diseases" (2024). Edited by S.M. Malkhazova. Moscow: Faculty of Geography, Lomonosov Moscow State University, 254 p. (in Russian)

Medico-Geographical Atlas of the Republic of Uzbekistan (2022) (O, Zbekiston respublikasining tibbiiy-geografik atlasi). I-part. Syrdarya and Jizzakh Regions. (in Uzbek, in Russian, in English). 148 p.

Medico-Geographical Atlas of the Republic of Uzbekistan (2024) (O, Zbekiston respublikasining tibbiiy-geografik atlasi). II-part. City Tashkent and Tashkent Region. (in Uzbek, in Russian, in English). 199 p.

Medico-Geographical Atlas of the Republic of Uzbekistan (2025) (O, Zbekiston respublikasining tibbiiy-geografik atlasi). III-part. Republic of Karakalpakstan and Khorezm region. (in Uzbek, in Russian, in English). 197 p.

Medico-geographical Atlas of the Smolensk Region. (2012). Smolensk, 104 p. (in Russian)

National Atlas of Russia. (2007). Moscow: Roskartografiya, V. 3, 495 p. (in Russian)

Petkau A.Y. (2015). Modeling of the concept health in Soviet and post-Soviet information space in media texts. Journal of Language and Literature. 6, 1, 139-143, DOI: 10.7813/jll.2015/6-1/25.

Plague and Environment Atlas of the People's Republic of China. (2000). Liu Yunpeng, Tan Jian an, Shen Erli, Science Press, 201 p.

Population health and healthcare of the Republic of Sakha (Yakutia) at the turn of the century. (2005). Yakutsk: YakutAGE, 116 p. (in Russian)

Prokhorov B.B. (2001). Health of the population of Russia in the twentieth century. Moscow: Publishing House of MNEPU, 276 p. (in Russian)

Protecting public health in the Republic of Sakha (Yakutia). (2012). Yakutsk: Publishing house "Dani Almas Company", 211 p. (in Russian)

Schistosomiasis Atlas of China. (2012). Zhou Xiaonong, Chinese Cartography Society, 335 p.

The Atlas of Endemic Diseases and Their Environments in the People's Republic of China. (1989). Compilation Committee, 216 p.

The Atlas of Population, Environment and Sustainable Development of China. (2000). Edited by Liu Yue. Science Press, 251 p.

Ulumbekova G.E. (2010). Healthcare of Russia. What needs to be done: scientific substantiation of the "Strategy for the Development of Healthcare in the Russian Federation until 2020". Moscow: GEOTAR-Media, 592 p. (in Russian).

Ulumbekova G.E. (2018). Healthcare in Russia: 2018–2024. What needs to be done? ORGZDRAV: news, opinions, training. Bulletin of the Higher School of Health, No. 1, 9–16, (in Russian)

Yunnan Provincial Atlas of Endemic Diseases. (1989). Zheng Lingcai, The People's Health Press Co, 194 p.

RESEARCH PAPER

ANNUAL RUNOFF AND CLIMATE CHARACTERISTICS IN LARGE RIVER BASINS OF NORTHERN EURASIA UNDER GLOBAL WARMING

Alexander G. Georgiadi[1*], Elena A. Barabanova[1], Irina P. Milyukova[1], Polina A. Morozova, Aleksey N. Narykov[1]
[1]Institute of Geography Russian Academy of Sciences, Staromonetny lane, Moscow, 119017, Russia
***Corresponding author:** georgiadi@igras.ru

Received: October 21st 2025 / Accepted: January 20th 2026 / Published: March 31st 2026
https://doi.org/10.24057/2071-9388-2026-2423

ABSTRACT. The runoff and climatic characteristics averaged over the basins of major rivers of the East European Plain (Northern Dvina, Pechora, Volga, Don, and Dnieper) and Siberia (Ob, Yenisei, and Lena) under global warming were analyzed. This analysis was based on both observational data and estimates obtained from global climate models (GCMs) under a baseline period (1930s–1980), contemporary global warming conditions (1981–2010), and scenario-based anthropogenic global warming projections into the mid-21st century (2040–2069). The main results suggest that long-term averaged annual runoff estimates derived from an ensemble of 18 CMIP6 GCMs do not accurately reproduce the observed runoff during either the contemporary global warming or baseline periods. Nevertheless, these models generally capture the direction of runoff changes between these two periods, though they tend to underestimate the magnitude of these changes. Consequently, multi-model mean projections are useful in estimating the relative changes in runoff under different global warming scenarios. A key finding of this study is that selecting GCMs that can best reproduce the most reliably observed river runoff and climatic characteristics of river basins under contemporary global warming can improve ensemble model estimates of runoff. In addition, it was found that basin-averaged observed and modeled annual atmospheric precipitation over the Northern Dvina basin, obtained from the RIHMI-WDC archive (after applying all relevant corrections) and multi-model mean data, respectively, were closely aligned. This was in contrast to data of reanalysis and data that only included wetting corrections from the CRU archive.

KEYWORDS: large rivers, annual runoff, climate characteristics, global warming, global climate models, Northern Eurasia

CITATION: Georgiadi A. G., Barabanova E. A., Milyukova I. P., Morozova P. A., Narykov A. N. (2026). Annual Runoff and Climate Characteristics in large River Basins of Northern Eurasia under Global Warming. Geography, Environment, Sustainability, 1 (19), 167-180
https://doi.org/10.24057/2071-9388-2026-2423

ACKNOWLEDGEMENTS: The studies were supported by grant of the RSF 24-27-00143 project "River water flow in Russia during epochs of global warming."

Conflict of interests: The authors reported no potential conflict of interests.

INTRODUCTION

The impact of global warming on river flow is a central issue in modern hydrology. Research in this field has developed in two main directions, which differ in the type and source of data used, as well as in their methods.

The first direction focuses on analyzing changes that have occurred during periods for which long-term, regular observational data on runoff characteristics are available, along with climatic and anthropogenic factors that significantly influence river runoff. This branch of research utilizes a wide range of analytical techniques, including time series analysis, regression relationships, balance methods, and, to a much lesser extent, hydrological modeling. This research direction has been active for many years, with studies examining (1) long-term phases of increased/decreased water flow and other components of georunoff (Muraveisky 1960), which represent one of the major components of long-term hydrological variability (Bolgov et al. 2018; Frolova et al. 2022; Georgiadi and Kashutina 2016; Georgiadi et al. 2018; Georgievsky 2021; Sharma and Singh 2017; Shi et al. 2019; Shpakova and Wang 2023) and (2) the impact of contemporary global warming on runoff as well as its contribution to observed changes in runoff (Georgiadi et al. 2014). A key element of this research is the development and application of methods for naturalizing long-term series of natural runoff characteristics (Georgiadi and Milyukova 2023; Shiklomanov et al. 2011).

The second research direction involves the development and application of methods for assessing probable changes in river runoff under future climate scenarios. These methods rely on climate change projections obtained through numerical experiments with global climate models (GCMs). There are two main approaches used to develop models for runoff estimation. The first is the "hydrological" approach, which utilizes hydrological models with varying levels of detail to simulate river runoff formation using climatic data from GCM-based global warming projections as inputs. In contrast, the "climatological" approach estimates river runoff from

GCM outputs, either by recording changes in river runoff as calculated directly by GCMs or simulating changes with reference to the water balance equation and model estimates of precipitation and evapotranspiration.

In recent decades, a large number of studies have been conducted using these two approaches to assess potential changes in the flow of Russian rivers under anthropogenic warming scenarios projected for the 21st century (Dobrovolski et al. 2019; Döll et al. 2018; Gelfan et al. 2022; Georgiadi and Milyukova 2002; Georgiadi et al. 2011, 2014; Georgievsky et al. 1996; Guo et al. 2022; Gusev and Nasonova 2010; Kattsov and Govorkova 2013; Kislov et al. 2008; Mokhov et al. 2003; Motovilov and Gelfan 2019; Shiklomanov 2008; Yang and Kane 2021). However, estimates derived from GCM-based simulations exhibit a significant inter-model spread in both present-day climate characteristics (and consequently river runoff) and projected future changes.[1] This variability can be attributed to a variety of factors, including differences in the parameterization of processes used in GCMs, differences in major socio-economic and demographic characteristics (e.g., economic growth, technology, population, and management), changes in the optical properties of the atmosphere, and different levels of projected greenhouse gas and aerosol emissions into the atmosphere (Abramowitz et al. 2019; Anisimov and Kokorev 2013; Guo et al. 2022; Lehner et al. 2020; Motovilov and Gelfan 2019).

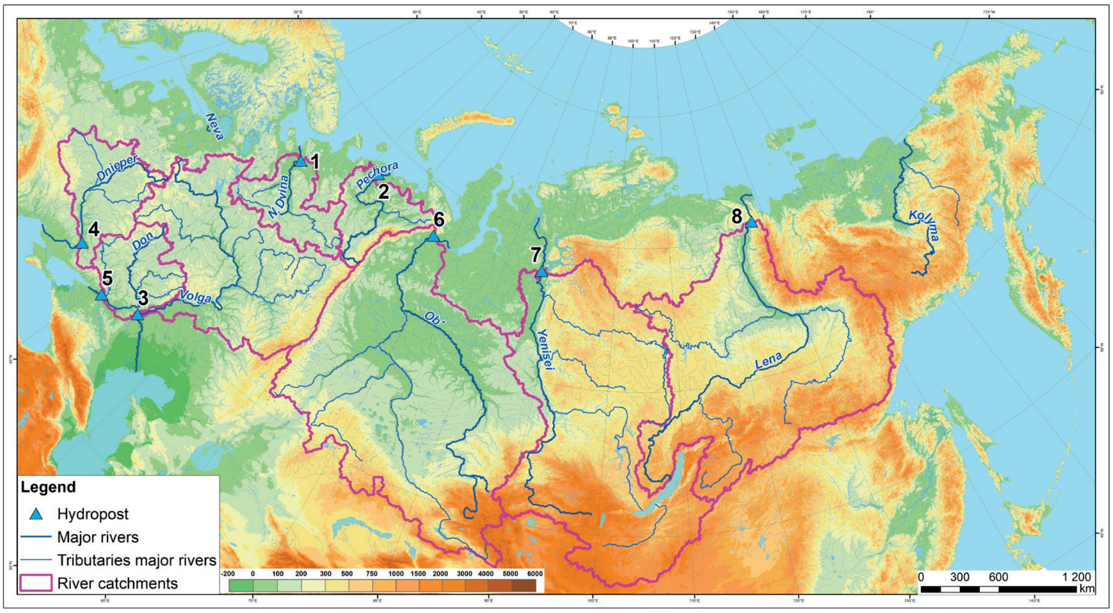
Consequences, studies on scenario-based changes in hydroclimatic characteristics often rely on averaging the results obtained from an ensemble of GCMs, especially since multi-model ensembles can help reduce the uncertainties associated with the climate system (Abramowitz et al. 2019; Guo et al. 2022; Lehner et al. 2020). Averaged climate characteristics derived from global data archives and electronic atlases[2,3] and scenario forecasts based on Coupled Model Intercomparison Project Phase 6 (CMIP6) global models can also be included to obtain higher resolution ensembles[4,5]. This ensemble-based approach can potentially be extended to the multi-model averaging of river flow estimates. However, researchers must consider whether all GCMs should be included in such ensembles to obtain more reliable runoff projections.

Considerable attention has also been given to the accuracy of observational climate data, which is crucial for addressing issues associated with the use of GCMs both in developing future scenario assessments (i.e., during the 21st century) and in analyzing the hydroclimatic impacts of contemporary global warming, including river runoff. This is particularly true for precipitation data collected in global and regional archives (Bogdanova et al. 2002; Fekete et al. 2004; Georgiadi et al. 2014; Groisman et al. 1996; Harding et al. 2011; Ye et al. 2004), where substantial discrepancies have been observed across different datasets (Harding et al. 2011).

This study provides a comparative analysis of river runoff in large Northern Eurasia River basins (on the East European Plain and in Siberia) as well as basin-averaged climatic characteristics that influence river runoff. This analysis uses observational data as well as GCM-derived outputs under present-day global warming conditions and for mid-21st century anthropogenic warming scenarios (2040–2069). The study utilizes long-term observational data of river runoff, precipitation, and air temperature from global and regional archives, reanalysis data, reconstructed natural runoff characteristics, and CMIP6-derived estimates of current and projected hydroclimatic changes. One of the main objectives of this study is to assess how accurately the averaged characteristics from multi-model ensembles reproduce the annual runoff and climatic characteristics in major Northern Eurasian River basins as observed during the period of contemporary global warming, as well as the preceding baseline interval.

OBJECTS OF THE STUDY

This study focused on large rivers with basins located in the East European Plain and Siberia (Fig. 1). These river basins exhibit noticeably different natural conditions (Table 1), as well as varying combinations and intensities of anthropogenic impacts on river runoff.



(1) Northern Dvina–Ust'-Pinega, (2) Pechora-Oksino, (3) Volga–Volgograd, (4) Dnieper–Lotsmanskaya Kamenka, (5) Don–Razdorskaya, (6) Ob'–Salekhard, (7) Yenisei–Igarka, (8) Lena–Kyusyur.

Fig. 1. Schematic map of the studied river basins

[1] AR6 Synthesis Report: Climate Change 2023. [online] Available at: https://www.ipcc.ch/report/sixth-assessment-report-cycle/ [Accessed 2 Oct. 2025]

[2] Interactive Atlas: Regional information (Advanced). [online] Available at: https://interactive-atlas.ipcc.ch / [Accessed 2 Oct. 2025]

[3] KNMI Climate Change Atlas. [online] Available at: https://climexp.knmi.nl/plot_atlas_form.py?id=someone@somewhere [Accessed 2 Oct. 2025]

[4] Roshydromet Climate Center. Official Website. [online] Available at: https://cc.voeikovmgo.ru/ru/klimat/cmip6hr [Accessed 2 Oct. 2025]

[5] Mapping and Data Analysis System for SESTRA project [online] Available at: https://sestra.unh.edu [Accessed 2 Oct. 2025]

Table 1. Main characteristics of the studied river basins

River–gauge	Basin area*, 10^3 km^2	Average elevation of the basin, in meters/ proportion of the basin with elevation ≥ 1000 m, % (SRTM 2025)	Share of main natural zones**, % of the basin area (ArcAtlas 1996)	Permafrost share (all types), % of the entire basin area (Brown et al. 2002)	Main reservoirs' total capacity, km^3 (Barabanova 2004)
Northern Dvina–Ust'-Pinega	348 (357)	143/0	T – 100	0	0
Pechora–Oksino	312 (322)	164/0.5	Tu –35 T – 65	44.3	0
Volga–Volgograd	1360 (1369)	183/0	T – 34 M&BLF – 40 FS – 15 S – 9	0	168
Dniepr–Lotsmanskaya Kamenka	459 (504)	168/0	M&BLF –77 FS – 15 S – 8	0	44.5
Don–Razdorskaya	378 (422)	153/0	M&BLF – 5 FS – 63 S – 32	0	28
Ob'–Salekhard	2450 (2990)	292/6.4	T – 29.7 M&BLF – 7 FS – 11 S – 25 SD&D – 6 I – 21	39.2	65.4
Yenisei–Igarka	2440 (2620)	753/27.3	Tu – 8 T – 58 FS – 17 S – 11	84.6	393
Lena–Kyusyur	2430 (2490)	594/17	Tu – 22 T – 71 I – 7	94	35.9

* In brackets – the entire basin area;

** TU – tundra, T – taiga, M&BLF – zone of mixed and broad-leaved forests, FS – forest-steppe, S – steppe, SD&D – semi-deserts and deserts, I – Intrazonal landscapes (swamps and others).

In the studied river basins, the majority of the annual runoff (45–65%) occurs during the spring–summer flood period due to melting snow cover (Kuzin and Babkin 1979). In the summer-autumn seasons, the main sources of river runoff include precipitation and underground runoff (accounting for 16–30% of annual runoff). The lowest runoff is observed during the winter low-water period, comprising only 10–20% of the annual total.

In recent decades, substantial changes have been observed in both annual runoff and its intra-annual distribution across basins in Northern Eurasia (Frolova et al. 2022; Georgiadi et al. 2014, 2024; Georgievsky 2021; Shiklomanov 2008).

DATA PREPARATION

Long-term observational data on runoff and climate characteristics

The study used annual hydroclimatic characteristics, including river runoff, total precipitation, and evapotranspiration (calculated based on water balance), as well as air temperature data obtained from global and regional archives. These values were averaged over two key periods: the period of modern global warming (1981–2010) and a baseline period (1930s–1980). The rationale for selecting these time intervals is discussed in the Methods section.

Annual river runoff

Long-term runoff records from the data archives of the Russian Hydrometeorological Service were used for large rivers with water regimes close to natural conditions (e.g., Northern Dvina, Pechora) as well as for rivers with naturalized runoff data. It should be noted that these reconstructions exclude anthropogenic changes caused by reservoirs and consumptive water withdrawal for use in industry. These observations were used as initial data to calculate average long-term annual runoff for the periods indicated above.

The long-term series were naturalized by transforming the annual hydrograph of daily water discharges using the Kalinin–Milyukov technique (Kalinin and Milyukov 1958). This approach yielded reconstructed long-term time series data on daily discharges for the Don, Ob, Yenisei, and Lena rivers (Georgiadi and Milyukova 2023; Yang and Kane 2021). The naturalization of annual and seasonal river flow for the Dnieper at Lotsmanskaya Kamenka and the Volga at Volgograd was based on regression relationships with the flow of tributaries and upper parts of the main rivers, which was not significantly changed by anthropogenic impact (Georgiadi et al. 2014).

Annual climatic characteristics

Long-term average annual precipitation values for the two key periods described above were calculated using data from two archives. Specifically, global gridded atmospheric precipitation and air temperature data (with a spatial resolution of 0.5°) were obtained from the Climatic Research Unit gridded Time Series (CRU TS) archive v. 3.10[6] prepared by the Climate Research Unit in the University of East Anglia (Norwich, UK; Harris et al. 2013). These gridded data were derived from the interpolation of meteorological measurements taken from a network of stations. In addition, atmospheric precipitation data were obtained from the RIHMI-WDC archive[7] for the Roshydromet observation network in Russia. This archive comprises two datasets: the first dataset (the RIHMI-WDC) contains wetting corrections starting from the mid-1960s, while the second dataset (the RIHMI-WDCcor), prepared at the Voeikov State Geophysical Observatory (SGO), incorporates corrections that account for key causes of measurement distortions, including aerodynamic effects (Bogdanova et al. 2002). Data from the first dataset were used in the preparation of the global CRU grid archive.

In addition, the average annual precipitation obtained from the reanalysis data was used to prepare data for the Northern Dvina basin. Specifically, data from the NCEP/NCAR (Kalnay et al. 1996), MERRA2 (Gelaro et al. 2017), and ERA5-Land (Muñoz Sabater 2019) climate reanalysis products were obtained from the Mapping and Data Analysis System for the SESTRA project[5], where they were aggregated based on the MERIT-Hydro-river network with a 5-minute spatial resolution (Yamazaki et al. 2019).

Model-based hydroclimatic data

Runoff, atmospheric precipitation, evapotranspiration, and air temperature data for the two study periods were obtained by averaging outputs obtained from the CMIP6 GCMs (Eyring et al. 2016; O'Neil et al. 2016). Monthly-resolution data from four experiments were used: historical, SSP 1-2.6, SSP 2-4.5, and SSP 5-8.5.

In the historical experiment, forcings were specified according to observational data for the period 1850–2014. In the scenario experiments (SSP 1-2.6, SSP 2-4.5, and SSP 5-8.5), forcings were specified according to different Shared Socioeconomic Pathways (SSP) for the period 2015–2100. The SSPs represent different levels of anthropogenic load: SSP 1-2.6 assumes a low level of greenhouse gas emissions and a reduction to zero values by 2075 (favorable scenario with insignificant warming expected); SSP 2-4.5 assumes that emissions remain at current levels until 2050 before gradually decreasing (moderate warming); and SSP 5-8.5 represents a high emissions scenario in which emissions double by 2050 (extreme scenario, strong warming). By 2100, the corresponding radiative anthropogenic forcing in these experiments reaches values of 2.6 W/m^2, 4.5 W/m^2, and 8.5 W/m^2, respectively.

Ensemble estimates were based on data obtained from 18 GCMs with relatively high spatial resolution used for scenario-based climate forecasts at the Climate Center of Roshydromet[8]. The model outputs were averaged annually and bilinearly interpolated to grid points with a resolution of 0.5°. The data were averaged over 1931–1980 (baseline period), 1981–2010 (contemporary global warming), and 2040–2069 (scenario projection period). In addition, all data were spatially averaged over the studied river basins.

METHODS

This study aimed to address several research problems. First, an assessment of the accuracy of the reproduction of observed annual runoff and annual climatic characteristics of large rivers of the East European Plain and Siberia from the outputs of a multi-model mean (MMM) derived from 18 CMIP6 GCMs. This was evaluated based on comparisons between (1) the model and observed average hydroclimatic characteristics during the periods of contemporary global warming and the baseline period, and (2) the differences between the model and observed data between these two periods.

Second, this study aimed to identify the spatial and temporal features of changes in annual river runoff, precipitation, air temperature, and evapotranspiration during the period of contemporary global warming and in the mid-21st century relative to the baseline period.

Finally, this study aims to compare the projected changes in annual runoff and climate characteristics under the different global warming scenarios for 2040–2069 relative to their model-derived baseline values based on CMIP6 MMM data.

Assessment of hydroclimatic characteristics and their changes during contemporary global warming

Estimates of runoff and climate characteristics (see Section 3 for more details on data preparation) and their changes under contemporary global warming were obtained by comparing the mean values obtained from two climatically distinct periods: modern global warming and the baseline period (Georgiadi and Kashutina 2016; Georgievsky 2021; Shiklomanov 2008). The modern period of global warming is generally considered to have started in 1981, while the baseline period (1930s–1980) has historically been used to calculate long-term river runoff characteristics in Russia[9]. For the CMIP6 ensemble and CRU archive data, as well as for river runoff data of the Pechora, Northern Dvina, Volga, Don, and Dnieper the base period 1931–1980 was used to calculate average hydroclimatic characteristics. For river runoff data of the Ob', Yenisei, and Lena and precipitation data from the RIHMI-WDC archive, the period 1936–1980 was used.

The transition between these two periods is thought to have occurred in the 1970s–1980s, where there is a noticeable increase in air temperature as well as changes to the long-term phases (lasting 10–15 years or more) of decreased/increased runoff in many rivers (Georgiadi and Kashutina 2016; Georgiadi and Milyukova 2023; Georgiadi et al. 2014, 2018; Georgievsky 2021; Shiklomanov 2008). However, it should be noted that the onset of warming was not simultaneous across river basins: based on the analysis of long-term phases in multi-year changes in annual and winter air temperatures averaged over the basins of large rivers (Volga, Don, Pechora, Yenisei, Lena), the transition to increased annual and winter air temperatures occurred gradually throughout the 1970s–1980s (Georgiadi and Kashutina 2016; Georgiadi and Milyukova 2023; Georgiadi

---

[6] CRU TS [online] Available at: https://crudata.uea.ac.uk/cru/data/hrg/) [Accessed 2 Oct. 2025]

[7] The All-Russian Scientific Research Institute of Hydrometeorological Information – World Data Center [online] Available at: http://meteo.ru/, http://meteo.ru/data/ [Accessed 2 Oct. 2025]

[8] Roshydromet Climate Center. Official Website. [online] Available at: https://cc.voeikovmgo.ru/ru/ [Accessed 2 Oct. 2025]

[9] State Water Cadaster. Surface and Groundwater Resources, Their Use and Quality (1982–2022). Leningrad/St. Petersburg: Rosgidromet. (In Russian)

et al. 2014). Indeed, in some basins such as the Northern Dvina, Lena and Don, this transition only became evident in the second half of the 1980s, especially for winter air temperatures (Georgiadi and Kashutina 2016; Georgiadi and Milyukova 2023; Georgiadi et al. 2014).

In terms of precipitation, the transition to a long-term phase of increased precipitation can be dated to the 1960s–1970s and even to the 1950s (Georgiadi and Kashutina 2016; Georgiadi and Milyukova 2023; Georgiadi et al. 2014). However, determining the exact onset of this change is complicated by the significant heterogeneity of long-term precipitation time series, especially due to the introduction of wetting corrections in the 1960s (Georgiadi and Kashutina 2016).

Finally, the timing of long-term phase shifts in annual and seasonal runoff varies among individual rivers and regions and often does not coincide with the onset of modern warming or the increase in air temperature (Frolova et al. 2022; Georgiadi and Kashutina 2016; Georgiadi and Milyukova 2023; Georgiadi et al. 2014).

Assessment of hydroclimatic characteristics and their changes under mid-21st century climate scenarios

As with the assessment of hydroclimatic changes under contemporary global warming, estimated changes in runoff and climate characteristics were evaluated by comparing the model-averaged values obtained for each projected global warming scenario and the baseline period. Ensemble-averaged data from the CMIP6 GCMs were used.

Results

Changes in hydroclimatic characteristics in large river basins based on long-term observations

Annual and seasonal runoff

During the period of contemporary global warming (1981–2010), most rivers flowing into the Arctic Ocean (including the Northern Dvina, Pechora, Yenisei, and Lena rivers), as well as the Volga basin, experienced increases in both annual and seasonal runoff (Fig. 2). In contrast, the annual runoff in the Dnieper increased due to increased summer–autumn and winter runoff, while flood runoff decreased. The annual runoff of the Don is decreasing due to a significant decrease in snowmelt flood runoff, while the runoff of low flow seasons are increasing. Annual and snowmelt flood runoff in the Ob' River exhibited little change during this period, though there has been a decrease in summer-autumn runoff and an increase in winter runoff.

The differences in runoff between the period of modern global warming and the baseline period are largely expressed in changes to the long-term runoff phases of increased/decreased annual runoff (lasting 10–15 years or more) as well as runoff during the main hydrological seasons. Specifically, these phases differ in terms of their duration, the timing of their phase transitions, and the magnitude of the differences between the average runoff during periods of increased and decreased runoff (Georgiadi et al. 2024).

Climatic characteristics

Based on data from the CRU global archive, an increase in annual runoff was observed during the period of contemporary global warming against the background of rising average annual air temperatures and precipitation (except for the Yenisei basin) (Table 2). The most pronounced increase in precipitation occurred in the Northern Dvina, Pechora, and Volga basins. The rise in mean air temperature averaged across all basins ranged from 0.7 to 1.1°C.

Comparison of annual hydroclimatic characteristics of large river basins based on long-term observations and CMIP6 MMM data

This section compares the annual runoff and climate characteristics obtained from both long-term observational time series and the mean values obtained from the MMMs of 18 CMIP6 GCMs over the period of contemporary global warming and the reference period.

Annual runoff

A comparison of the annual runoff derived from observed (Northern Dvina and Pechora) and naturalized (Volga, Dnieper, Don, Ob', Yenisei, Lena) time series data and the modeled annual runoff of these rivers obtained from CMIP6 MMM data over the study periods indicates that the modeled runoff underestimates the observed (naturalized) runoff of rivers flowing into the Arctic Ocean (Fig. 3), except for the Northern Dvina, where the modeled runoff slightly overestimates the observed runoff. In contrast, for rivers on the southern macroslope of the East European Plain (the Volga, Dnieper, and Don), the modeled runoff overestimates the observed runoff.

It should be noted that the greatest underestimations occurred in basins where a significant proportion of runoff is generated by mountainous areas or in regions where permafrost occupies a significant portion of the basin,

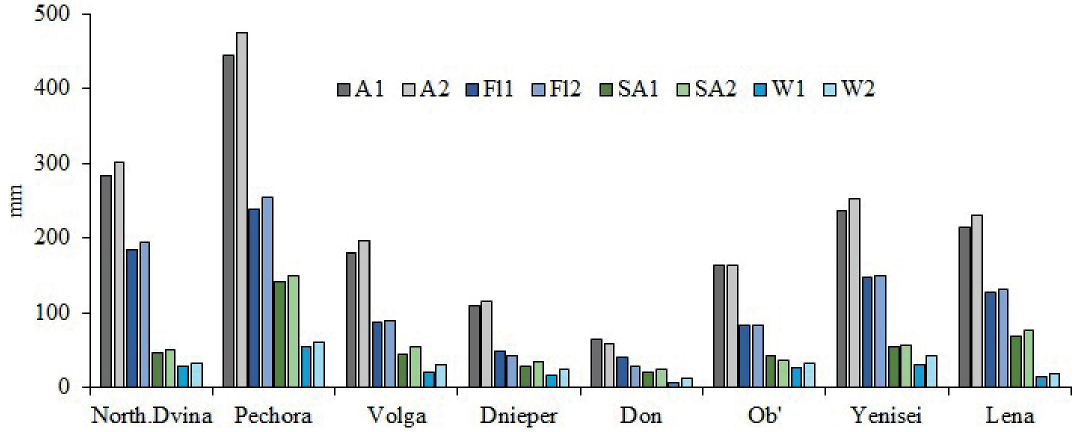


Fig. 2. Flood runoff (Fl), summer–autumn (SA), winter (W), and annual runoff (A) in (1) the baseline period and (2) the period of contemporary global warming

Table 2. Annual runoff of large rivers and annual precipitation and air temperature averaged for their basins according to the CRU archive data

River	Base period			Period of contemporary global warming		
	runoff, mm	precipitation, mm	air temperature, ^{0}C	runoff, mm	precipitation, mm	air temperature, ^{0}C
Northern Dvina[a]	283	584	0.9	301	612	1.6
Pechora[a]	445	500	-3.4	475	529	-2.8
Volga[b]	181	527	3.4	196	556	4.4
Dnieper[b]	109	559	6.9	116	569	7.7
Don[b]	64	456	7.0	59	466	7.8
Ob'[b]	164	426	-0.7	163	429	0.4
Yenisei[b]	237	438	-6.2	253	431	-5.2
Lena[b]	214	356	-10.5	231	357	-9.5

[a] observed runoff is close to natural conditions

[b] naturalized runoff (excluding anthropogenic changes)

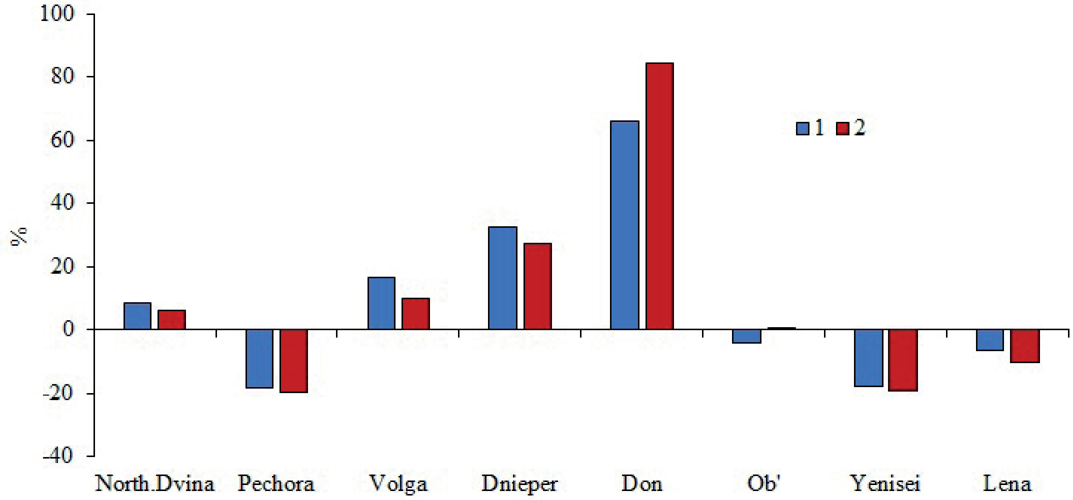


Fig. 3. The percentage difference between CMIP6 MMM data and observed runoff over (1) the baseline period and (2) the period of contemporary global warming

including the Pechora (runoff underestimated by 18–20%), the Yenisei (18–19%), and the Lena (7–10%).

In contrast, the greatest overestimations occurred in basins where forest–steppe and steppe zones account for a significant proportion of the region, such as the Don (runoff overestimated by 76–84%) and the Dnieper (27–33%). The discrepancies between observed and modeled runoff are generally smaller for Siberian rivers, than for rivers in the East European Plain.

Atmospheric precipitation

The differences between the observed (from the CRU archive) and simulated annual precipitation were quite significant, ranging from 20–49% for the period of contemporary global warming and 20–46% for the baseline period (Fig. 4a). The largest discrepancies were observed in the Lena River basin for both periods. In general, river basins of the Russian Plain exhibited smaller differences than river basins in Siberia.

Similarly, there were differences between the CMIP6-simulated annual precipitation data and the annual precipitation data obtained from the RIHMI-WDCcor archive (see Section 3) (Fig. 4b). In general, the modeled precipitation data overestimated the observed precipitation, though the differences between them were much smaller than when compared with data from the CRU archive (Fig. 4a). Specifically, in the Yenisei, Lena, and Volga basins, CMIP6 precipitation exceeds observations by 12–19% during the baseline period and 9–18% under contemporary global warming. In contrast, model estimates were lower than the corrected observed precipitation only in the Dnieper River basin, though these differences did not exceed 1%.

The long-term average annual precipitation over the Northern Dvina basin obtained from reanalysis data (NCEP, ERA5, and MERRA2) significantly exceeded both the CMIP6 model values as well as observed atmospheric precipitation from the global CRU, regional RIHMI-GDC, and RIHMI-GDCcor archives (Fig. 5). Notably, the CRU-based precipitation values and the RIHMI-GDC precipitation values with wetting corrections were practically identical for the Northern Dvina basin (Fig. 5). This is because the CRU dataset for Russia is primarily derived from data of meteorological station observations that already include wetting corrections.

Annual total evapotranspiration

This section compares annual evapotranspiration during the period of contemporary global warming and the baseline period using several estimation methods

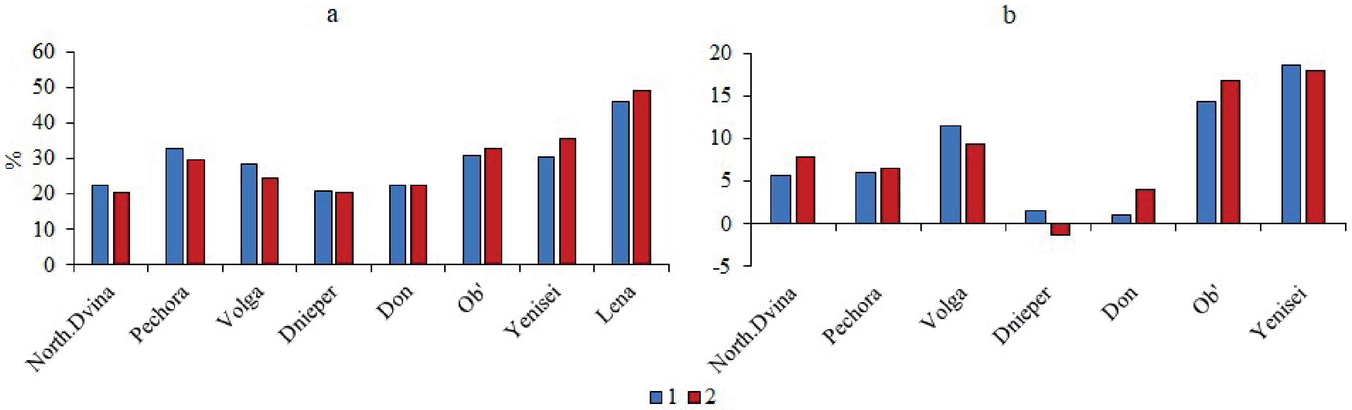


Fig. 4. The difference (%) between the average basin annual precipitation over (1) the baseline period and (2) the period of modern global warming, averaged over the ensemble of CMIP6 GCMs and determined based on data from (a) CRU and (b) RIHMI-WDCcor archives

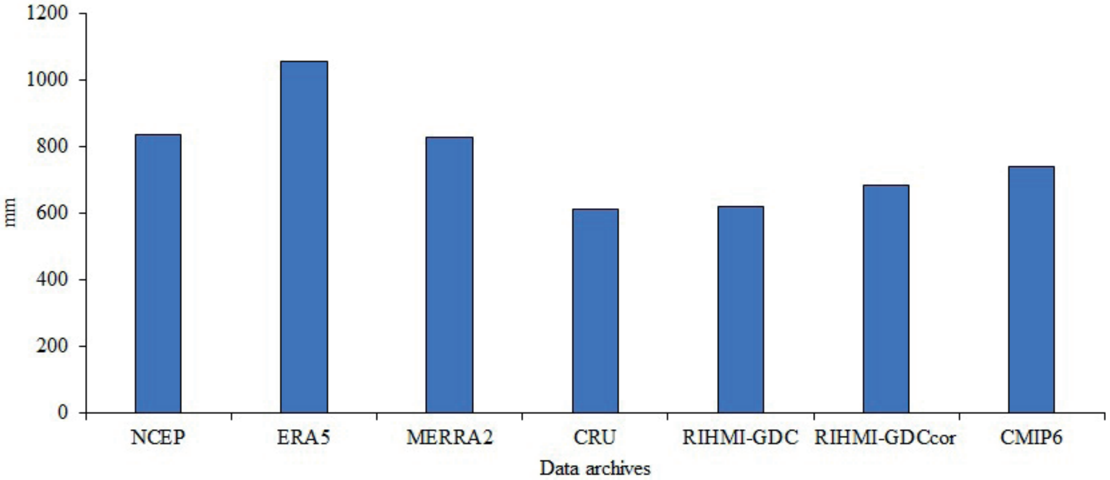


Fig. 5. Average annual atmospheric precipitation for period of modern global warming over the Northern Dvina River basin based on NCEP, ERA5, MERRA2, CRU, RIHMI-GDC, and RIHMI-WDCcor archives as well as CMIP6 ensemble averages

(Table 3). Specifically, total evapotranspiration values from CMIP6 ensemble means (CMIP6[d]) were compared with evapotranspiration calculated from the residual of the average long-term water balance equation for the study periods using annual precipitation and river runoff obtained from observational data. Inputs included annual precipitation from the CRU archive (CRU[a]), corrected data from the RIHMI-GDCcor (RIHMI-GDCcor[b]) archive, and using observed annual runoff and precipitation from CMIP6 data (CMIP6[c]).

It should be noted that the ensemble-averaged evapotranspiration calculated directly from the GCMs, the evapotranspiration obtained from the water balance equation using CMIP6 precipitation, and the observed runoff for some Siberian rivers, such as the Ob' and Lena, were quite similar. This similarity can be explained by the small difference between the modeled and observed runoff for these rivers: the smaller the difference between the modeled and observed runoff, the smaller the difference in evapotranspiration. Indeed, the difference in evapotranspiration was more significant for other rivers, such as the Pechora River basin, where modeled evapotranspiration can exceed the observed evapotranspiration by almost 50% due to the underestimation of runoff and precipitation in this basin.

Across all river basins, the lowest evapotranspiration estimates were obtained when using the water balance calculation with CRU precipitation data and observed runoff, primarily due to the underestimated precipitation in the CRU archive. The discrepancy is particularly noticeable for the Arctic rivers as well as the Volga basin. The largest discrepancy occurs in the Pechora basin, where evapotranspiration estimates based on CRU data are up to six times lower than other methods.

In contrast, evapotranspiration calculated using the water balance equation with corrected precipitation data from the RIHMI-GDCcor archive and observed (naturalized) runoff was much closer to the estimated evapotranspiration obtained from CMIP6 MMM data. However, in cases where model runoff or precipitation is significantly underestimated compared to their observed values, these evapotranspiration estimates turn out to be significantly lower than those obtained using the CMIP6 MMM data (e.g., Pechora, Yenisei, and Lena).

Finally, evapotranspiration estimates obtained from the use of atmospheric precipitation derived from reanalysis data and observed runoff in water balance calculations significantly exceeded the evapotranspiration estimates from CMIP6 MMM data (Fig. 6). For example, in the Northern Dvina basin, the evapotranspiration estimates obtained from reanalysis data were 527 mm (MERRA2), 534 mm (NCEP), and 753 mm (ERA5) under contemporary global warming.

Air temperature

The difference between the ensemble-averaged annual air temperature obtained from CMIP6 models and observed values from the CRU archive exhibited consistent signs for both the baseline period and the period of contemporary global warming (Table 2; Fig. 7). The sole exception is the Ob' river basin, where the sign of the difference between the modeled and observed annual air temperature was different across the two study periods,

Table 3. Total annual evapotranspiration (mm) obtained from ensemble-averaged data from CMIP6 GCMs, as well as average long-term observational data calculated using the water balance equation*

River	Evapotranspiration in the base period				Evapotranspiration during the modern global warming period			
	CRU^a	RIHMI-GDCcor^b	CMIP6^c	CMIP6^d	CRU^a	RIHMI-GDCcor^b	CMIP6^c	CMIP6^d
Northern Dvina	301	395	433	401	311	383	437	410
Pechora	55	181	219	293	54	169	211	301
Volga	346	426	496	465	360	437	496	477
Dnieper	450	n.d.a.	567	529	453	n.d.a.	570	537
Don	392	486	467	429	407	519	474	441
Ob'	262	388	394	398	266	385	407	406
Yenisei	201	263	334	364	178	248	332	371
Lena	142	216	306	306	126	221	302	312

*See the text for explanations of evapotranspiration calculation methods in this Section.

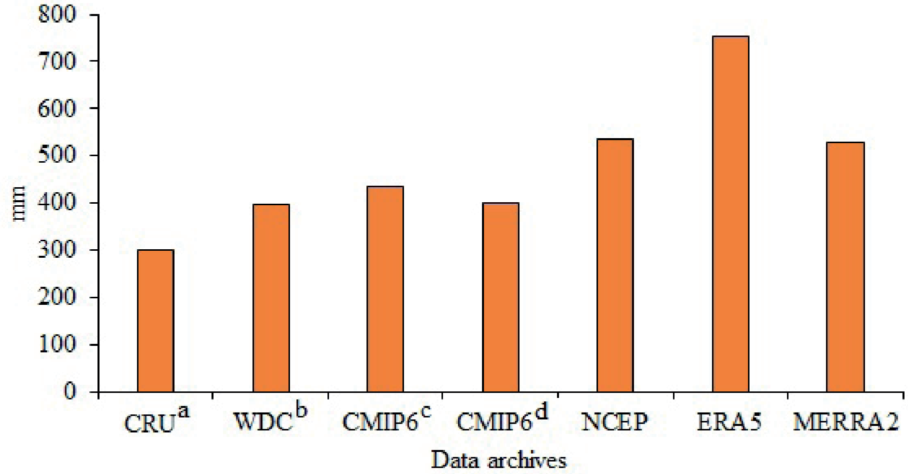


Fig. 6. Annual evapotranspiration in the Northern Dvina basin for period of modern global warming, calculated by averaging data from the CMIP6 GCMs ensemble and using the long-term average water balance equation. Further details on the evapotranspiration calculation methods used in this study can be found in Section 5.2.3

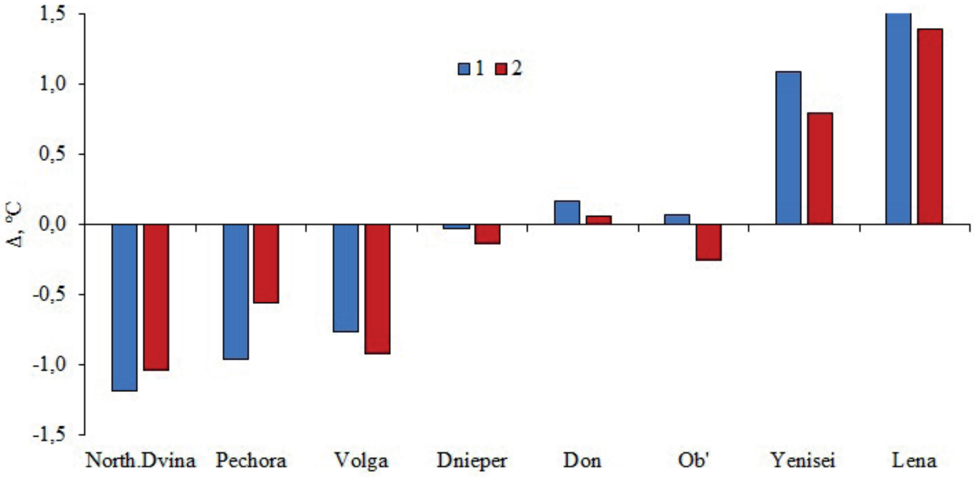


Fig. 7. The difference between the CMIP6 model-derived and CRU archive-derived air temperature data over (1) the baseline period and (2) the period of modern global warming

though the magnitude of the discrepancy was among the smallest across all basins. Model results underestimated the observed air temperature in the Northern Dvina, Pechora, Volga, and Dnieper basins, while temperatures were overestimated in the Yenisei, Lena, and Don basins.

Changes in annual river runoff and climate characteristics under contemporary global warming

This section compares the changes in annual hydroclimatic characteristics during the period of contemporary global warming relative to the baseline period based on observational data and CMIP6 MMM data.

Annual runoff

The direction of the relative changes in annual runoff obtained from ensemble-averaged CMIP6 GCMs and observational data under contemporary global warming relative to the baseline period (as a percentage of the annual runoff of the baseline period) was the same for most rivers, with the exception of the Don and Ob' (Fig. 3, 8, Table 2). For these two rivers, the modeled data indicated an increase in runoff, while the observed data revealed a decrease in runoff; this was particularly pronounced for the Don. The largest discrepancies between the modeled and observed changes in runoff were observed in the Volga, Dnieper, and Lena rivers. It should be noted that the difference in annual runoff between the two study periods did not exceed 10% and 5% for the observed and modeled data, respectively. In general, the ensemble-averaged model tended to underestimate the magnitude of changes in annual runoff compared to observational data.

Atmospheric precipitation

The direction of relative changes in annual precipitation according to CMIP6 MMM data and observational data from the CRU archive over the two study periods was the same in most river basins with the exception of the Yenisei basin (Fig. 9). Observed precipitation in the Yenisei basin decreased under contemporary global warming, while the model calculations indicated an increase. Changes in observed precipitation obtained from the CRU archive were significantly larger than the precipitation obtained from CMIP6 MMM data in several river basins of the Russian Plain, including the Northern Dvina, Pechora, and Volga. In contrast, the differences in observed and modeled precipitation were almost identical for the Dnieper and Don basins. Finally, the differences in observed precipitation changes were lower than the model estimates for the Siberian River basins (e.g., Ob' and Lena).

Precipitation data obtained from the RIHMI-GDCcor archive also exhibits increases during contemporary global warming, except for the Ob' basin. This increase was greater than the CMIP6 MMM estimates for the Volga, Don, and Lena River basins, but smaller than the model estimates for the Northern Dvina, Pechora, and Yenisei basins. Notably, in the Don basin, RIHMI-GDCcor data indicate an increase in precipitation that substantially exceeds both the CMIP6 ensemble estimate and the CRU observations.

Evapotranspiration

Under contemporary global warming, evapotranspiration increases in all river basins according to the CMIP6 MMM data (Fig. 10). Evapotranspiration also increases in most river basins when calculated using the water balance equation with modeled precipitation and runoff obtained from observational data; however, there are some river basins in which evapotranspiration does not change (the Volga basin), decreases slightly (Yenisei and Lena basins), or decreases significantly (Pechora basin). These differences arise from discrepancies between the modeled and observed runoff during the baseline and contemporary warming periods.

Evapotranspiration calculated with the water balance equation using precipitation data from the CRU archive and observed runoff is generally consistent with the estimates obtained from the CMIP6 ensemble; the only exceptions were the Pechora, Yenisei, and Lena basins, where evapotranspiration changes were found to have different signs. In contrast, when using precipitation data

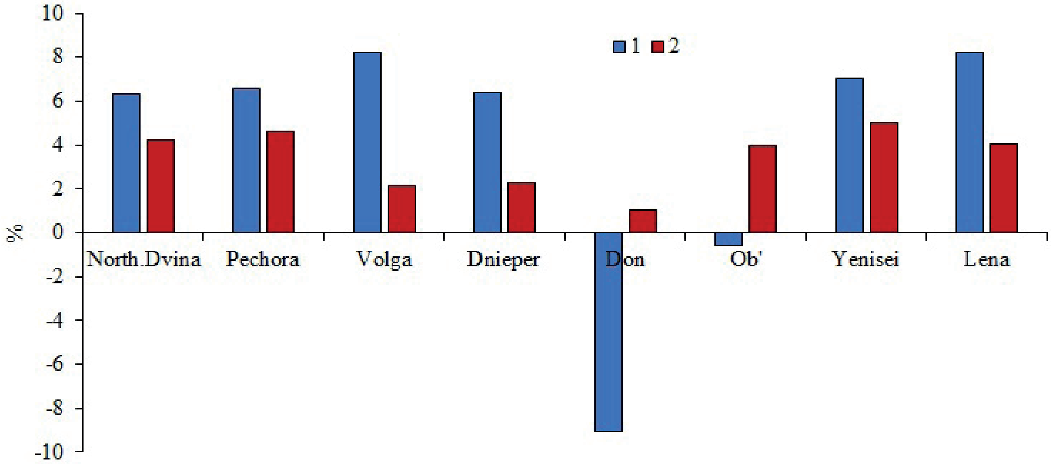


Fig. 8. Changes in annual runoff (Δ, %) during the period of contemporary global warming compared to the annual runoff of the baseline period, calculated using (1) long-term observational data and (2) CMIP6 MMM data

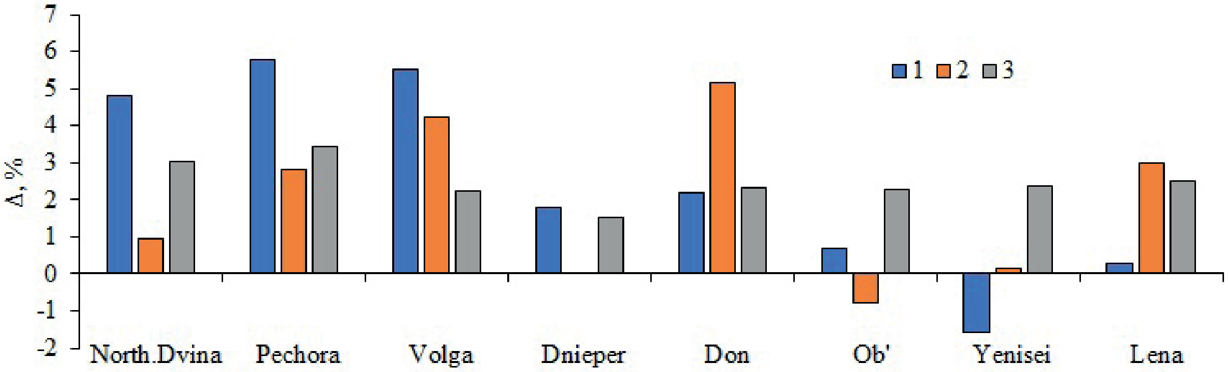


Fig. 9. Changes in annual precipitation (Δ, %) during the period of contemporary global warming compared to the baseline period, calculated using long-term data from the (1) CRU, and (2) RIHMI-GDCcor archives, and (3) CMIP6 MMM data

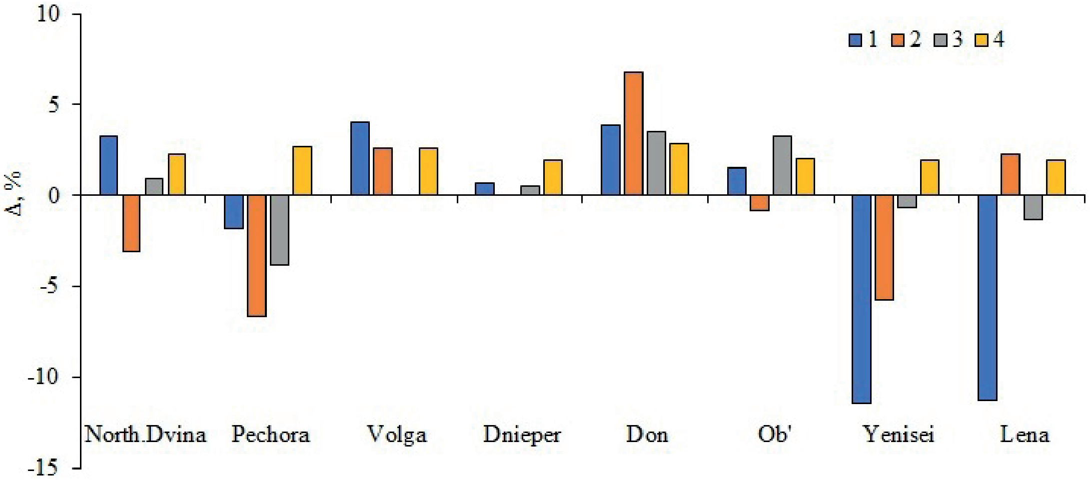


Fig. 10. Changes in annual evapotranspiration (Δ, %) during the period of contemporary global warming compared to the baseline period, calculated with the water balance equation using averaged long-term observational runoff and precipitation data from the (1) CRU and (2) RIHMI-GDCcor archives, (3) the water balance equation using observed runoff data and CMIP6 model estimates of atmospheric precipitation, and (4) the result of evapotranspiration calculations obtained from CMIP6 ensembles

from the RIHMI-GDCcor archive and observed runoff, the sign of evapotranspiration changes in Arctic river basins differed from those obtained from the CMIP6 ensemble (except for the Pechora and Lena rivers), while remaining the same in the Don basin. In the Volga basin, direct CMIP6 calculations indicated no change in evapotranspiration.

In general, evapotranspiration changes remained within ±7% of the baseline period, with only the Yenisei and Lena basins exceeding 11% when calculating using CRU precipitation data.

Air temperature

During the period of contemporary global warming, air temperature increased in all basins, both in the observed CRU data and in the CMIP6 GCM ensemble averages (Fig. 11). Observed increases ranged from 0.6°C in the Pechora basin to 1.1°C in the Ob' basin according to observed data and from 0.7°C in the Dnieper and Yenisei basin to 1°C in the Pechora basin according to ensemble data.

In most basins, the modeled temperature increase is slightly lower than the observed increase, though the opposite is true for the Northern Dvina and Pechora basins. The most noticeable discrepancies were observed in the Pechora, Yenisei, and Ob' basins.

Changes in annual runoff and climate characteristics for projected warming scenarios in the mid-21st century

The results show that estimates of runoff and other climate characteristics derived from the 18 CMIP6 GCM ensemble can differ significantly from their observed values in several river basins. This is consistent with previous studies using CMIP5 data (Georgiadi et al. 2024; Georgievsky and Golovanov 2015, 2019; Guo et al. 2022). Modeled trends in runoff and climate characteristics under modern global warming compared to the baseline period also differed from observed data.

Despite these issues, model estimates of projected global warming scenarios relative to the modeled values of their baseline periods remain of interest to researchers. These estimates can be used to calculate runoff and climate characteristics under different climate change scenarios as projected into the mid-21st century. Here, relative changes in the modeled runoff for each scenario (expressed as a percentage of the baseline runoff) can be used to calculate absolute runoff values for each scenario.

According to the CMIP6 ensemble average, all large rivers flowing into the Arctic Ocean are projected to experience increases in both runoff and climate characteristics under mid-21st-century anthropogenic warming compared to their baseline values (Table 4). The most significant increase in the annual runoff of these rivers is expected in the Pechora, Yenisei, and Lena basins (from 10.4% in the Pechora to 29.1% in the Lena), while the runoff of the Don and Dnieper may be lower than the model runoff for the baseline period by 3.4–17.1% and 4.8–14.7%, respectively depending on scenario. In contrast, the runoff of the Volga River may decrease under the moderate and extreme global warming scenarios by 1.7% and 2.5%, respectively.

Scenario projections indicate that climate characteristics will generally exceed baseline model values, except for slight decreases in annual precipitation in the Don River basin under SSP2-4.5 and SSP5-8.5 scenarios. The average annual air temperature across river basins is expected to increase across all basins, though the magnitude of this increase may differ. Increases in annual precipitation and total evapotranspiration will generally be more noticeable in Arctic river basins.

Percentage deviations in annual runoff for each warming scenario relative to the model runoff for the baseline period are closely linked to deviations in annual precipitation, while deviations in total evapotranspiration are closely linked to deviations in annual air temperature (Fig. 12).

It should be noted that the lowest correlation coefficients are observed when comparing the deltas between the model values calculated for the period of contemporary global warming and the baseline period. The closeness of the relationship increases with the intensity of global warming.

DISCUSSION

GCMs have improved significantly over the past decade. In particular, runoff calculation modules have begun to incorporate hydrological models, albeit in a simplified form (Georgievsky and Golovanov 2015, 2019).

This study shows that estimates of the long-term average annual runoff in several regions, including the southern and northern parts of the East European Plain and some basins of the largest Siberian rivers, obtained by averaging results from an ensemble of 18 CMIP6

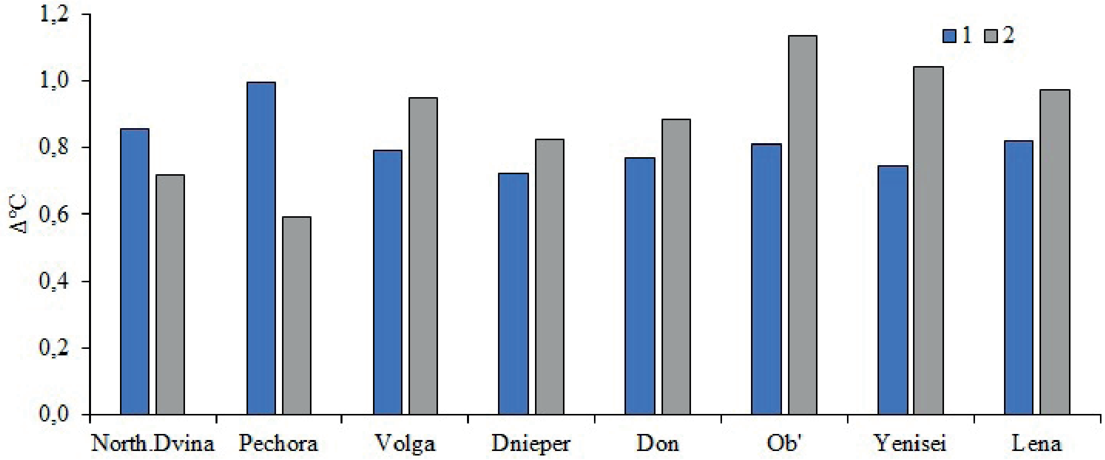


Fig. 11. Changes in annual air temperature (Δ, °C) during the period of modern global warming compared to the baseline period, calculated using (1) the CMIP6 MMM data and (2) long-term data from the CRU global archive

Table 4. Modeled changes in the components of the water balance equation ΔR, ΔPr, ΔE (%) and air temperature (ΔT, °C) in 2040–2069 relative to modeled values for the baseline period under different global warming scenarios

River	SSP1-2.6				SSP2-4.5				SSP5-8.5			
	ΔR	ΔPr	ΔE	ΔT	ΔR	ΔPr	ΔE	ΔT	ΔR	ΔPr	ΔE	ΔT
Northern Dvina	7.9	10.0	14.8	3.1	5.0	9.27	13.2	3.8	6.9	11.2	16.8	4.7
Pechora	10.4	3.2	18.8	3.6	11.0	12.40	17.3	4.3	14.1	12.8	22.0	5.3
Volga	2.6	7.8	11.5	3.1	-1.7	7.24	10.6	3.7	-2.5	8.0	12.6	4.5
Dnieper	-4.8	6.2	9.0	2.8	-14.0	4.66	8.2	3.4	-14.7	3.8	7.7	4.3
Don	-3.4	1.2	9.5	3.0	-11.8	-0.06	8.2	3.5	-17.1	-1.0	8.4	4.4
Ob	7.0	10.0	13.1	3.2	5.4	11.4	12.0	3.8	5.9	12.2	14.4	4.7
Yenisei	12.5	12.4	13.6	3.1	11.6	12.66	11.8	3.7	17.5	15.6	15.2	4.5
Lena	18.2	13.8	15.4	3.3	18.6	14.89	12.4	4.0	29.1	19.3	16.0	4.9

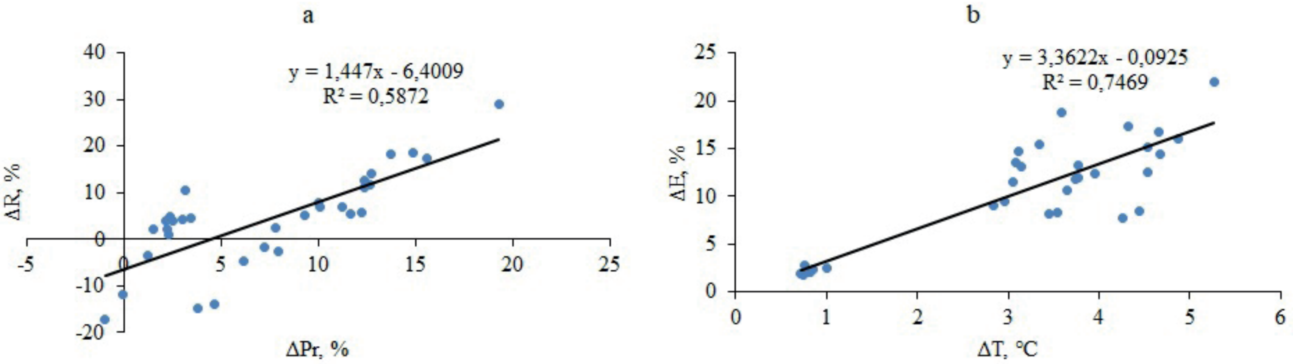


Fig. 12. Regression relationships between changes (%) in hydroclimatic characteristics for large river basins. Changes are shown between the contemporary global warming period, scenario period (2040–2069) under different intensities of global warming relative to the baseline period, calculated using CMIP6 MMM data. (a) annual runoff and precipitation; (b) annual evapotranspiration and air temperature

GCMs, do not adequately represent the average runoff observed under contemporary global warming relative to the baseline period; this is consistent with previous studies (Georgiadi et al. 2024; Georgievsky and Golovanov 2015, 2019; Guo et al. 2022). Nevertheless, these models generally reproduce the sign of the changes between the contemporary and baseline periods. In most rivers, these models underestimate the change in hydroclimatic characteristics between these periods.

A primary reason for this underestimation is that some GCMs runoff schemes do not sufficiently account for regional features of runoff formation. One potential solution is to select specific climate models within CMIP6 that best reproduce observed runoff and its climatic drivers. Similar approaches have been used in climate research (e.g., Anisimov and Kokorev 2013; Georgiadi et al. 2014; Kislov et al. 2008; Menzhulin et al. 2005), though river runoff is generally not included among the characteristics assessed. The results of this study demonstrate that, when selecting global climate models, it is useful to compare not only climate variables but also the observed and GCM-derived river runoff for the modern warming and baseline periods. Comparing long-term trends in reliably determined observed and modeled runoff changes further enhances the reliability of model selection (Motovilov and Gelfan 2019).

Another approach to improving the accuracy of GCM-based calculations involves bias correction of the model's climate variables using a specialized function, whose parameters are determined from the statistical analysis of long-term models and observed climatic fields (Haerter et al. 2011; Hoseini et al. 2024). Importantly, this method does not utilize observed runoff data for correction.

The further development of hydrological modules within GCMs, combined with improved methods for selecting the most suitable models based on both climatic variables and observed runoff, will enable GCMs to be more widely applied to address a broad spectrum of problems in the near future.

CONCLUSIONS

A comparative analysis of changes in the annual runoff, precipitation, total evapotranspiration, and air temperature of major river basins in the East European Plain and Siberia was conducted for both contemporary warming as well as projected global warming scenarios into the mid-21st century, relative to a baseline period. The main findings are as follows.

First, based on a comparison of annual runoff and climate characteristics between the period of contemporary global warming (1981–2010) and a baseline period (1930s–1980), calculated using observational data and the ensemble mean of 18 relatively high-resolution CMIP6 global climate models, it was found that:

– for both the period of contemporary global warming and the baseline period, the modeled annual runoff generally underestimates the observed (naturalized) runoff of rivers flowing into the Arctic Ocean with the exception of the Northern Dvina. In contrast, the modeled runoff overestimates the observed runoff in rivers on the southern macroslope of the East European Plain;

– modeled annual precipitation exceeded observed values from the CRU archive by 19–49% for the studied periods, with the difference being smaller for river basins in the East European Plain compared to river basins in Siberia. Modeled annual precipitation in most basins also exceeds the corrected observed precipitation values from the RIHMI-WDCcor archive, though this difference is much smaller and is hardly noticeable in some basins. Reanalysis-based precipitation data (NCEP, ERA5, MERRA2) for the Northern Dvina basin during 1981–2010 significantly exceed both the model-derived precipitation from CMIP6 and observed values from the CRU and RIHMI-GDC archives. Precipitation from the RIHMI-WDCcor archive (with set of corrections) was higher than the CRU data, lower than reanalysis data, and closely aligned with the CMIP6 model precipitation;

– the difference in annual air temperature calculated using CMIP6 data and observed values from the CRU archive for the two study periods has the same sign for most river basins, with the exception of the Ob' basin. Model estimates tend to underestimate air temperature in the Northern Dvina, Pechora, Volga, and Dnieper basins, while overestimating it in the Yenisei, Lena, and Don basins;

– for most rivers, both observed and modeled runoff increased during the contemporary warming period compared to the baseline. Exceptions include the Don and Ob' rivers, where observed runoff decreased, while modeled runoff increased. Differences in modeled runoff between the two study periods were generally smaller than the observed values, especially for the Volga, Dnieper, and Lena. The difference in observed runoff between the two study periods did not exceed 10%, while the differences in modeled runoff remained within 5%;

– similarly, the modeled and observed annual precipitation data (CRU and RIHMI-GDC (with set of corrections) archive) increased under contemporary global warming compared to the baseline period for almost all rivers except for the Yenisei basins (CRU) and the Ob' basins (RIHMI-GDC). Differences in precipitation between periods are typically smaller in model data than in CRU data for the East European Plain (except the Don basin, which exhibited similar differences), while the opposite is true for Siberian basins. According to the corrected RIHMI-GDC data, precipitation changes between the two periods exceed model values in the Volga, Don, and Lena basins, while remaining lower than the model values in other basins, with a typical range of 3–5%;

– both observed and modeled air temperatures exhibited increases under contemporary global warming compared to the baseline in all river basins. The modeled difference in air temperature between these periods was lower than the observed values in most basins; the only exceptions were the Northern Dvina and Pechora basins.

Second, when comparing projected global warming scenarios of the mid-21st century (2040–2069) with model data for the baseline period (1930s–1980), it was revealed that:

– both the runoff and climate characteristics of rivers flowing into the Arctic Ocean will be higher than the model values of their baseline periods based on CMIP6 ensembles. The only exception is for annual precipitation in the Don basin, which may slightly decrease under the SSP2-4.5 and SSP5-8.5 scenarios. The most pronounced increases in annual runoff are projected for the Pechora, Yenisei, and Lena rivers, ranging from 10.4% on the Pechora to 29.1% on the Lena. In contrast, the runoff of the Don and Dnieper may decrease by 3.4–17.1% and 4.8–14.7% depending on the scenario, respectively, while the Volga runoff may experience only minor decreases of 1.7–2.5% under moderate to maximum warming. The average annual air temperature over all river basins is expected to increase at approximately the same rate, but increases in annual precipitation and total evapotranspiration will generally be more noticeable in Arctic river basins;

– deviations in the annual runoff of the rivers in each global warming scenario relative to the baseline period are closely linked to deviations in annual precipitation, while deviations in total evapotranspiration are closely linked to deviations in annual air temperature. The lowest correlation coefficients are observed when comparing deltas between modeled values for contemporary warming and the baseline period, while correlations strengthen under more intense warming scenarios.

Finally, improving the accuracy of river runoff estimates, including the long-term averaged annual river runoff derived from GCM ensembles, may be achieved by selecting models that best reproduce the observed values of runoff and climatic characteristics of river basins, both during the period of contemporary global warming and the baseline period, as well as the differences in those characteristics between those two periods. ■

REFERENCES

Abramowitz G., Herger N., Gutmann E., Hammerling D., Knutti R., Leduc M., Lorenz R., Pincus R. and Schmidt G. A. (2019). ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing. Earth Syst. Dynam., 10, 91–105, DOI:10.5194/esd-10-91-2019, 2019.

Anisimov O.A. and Kokorev V.A. (2013). The optimal choice of hydrodynamic models to assess the impact of climate change on the cryosphere, Ice and Snow, 53(1), 83-92 (in Russian with English summary), DOI: 10.15356/2076-6734-2013-1-83-92

ArcAtlas: Our Earth. Environmental Systems Research Institute (1996). Inc. and DATA+ All Rights Reserved.

Barabanova E.A. (2004). Comparison of reservoirs by the complex of their positive and negative impacts on the environment and economy. Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya, 2, 72–82. (In Russian with English summary).

Bogdanova E. G., Ilyin B. M. and Dragomilova I. V. (2002). Application of a Comprehensive Bias-Correction Model to Precipitation Measured at Russian North Pole Drifting Stations, Meteorology and Hydrology, 3, 700–713. (In Russian).

Bolgov M.V., Filippova I.A., Osipova N.V., Korobkina E.A. and Trubetskova M.D. (2018). Modern features of the hydrological regime of the rivers of the Volga basin, Problems of Geography, 145, 206–218. (In Russian).

Dobrovolski S.G, Yushkov V.P. and Istomina M.N. (2019). Statistical Modeling of the Global River Runoff Using GCMs: Comparison with the Observational Data and Reanalysis Results // Water Resources, 46 (S2), 17–24, DOI: 10.1134/S0097807819080050

Döll P., Trautmann T., Gerten D., Müller Schmied H., Ostberg S., Saaed F. and Schleussner C.F. (2018). Risks for the global freshwater system at 1.5 °C and 2 °C global warming. Environ. Res. Lett., 13, 044038, DOI:10.1088/1748-9326/aab7.

Eyring V., Bony S., Meehl G. A., Senior C. A., Stevens B., Stouffer R. J. and Taylor K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev., 9, 1937–1958, DOI: 10.5194/gmd-9-1937-2016, 2016

Fekete B.M., Vörösmarty J.O., Roads J.O. and Willmott C.J. (2004). Uncertainties in Precipitation and Their Impacts on Runoff Estimates. Journal of Climate, 17, 294–304.

Frolova N.L., Magritskii D.V., Kireeva M.B., Grigor'ev V.Y., Gelfan A.N., Sazonov A.A. and Shevchenko A.I. (2022). Streamflow of the Russian rivers under current and forecasted climate changes. A review of publications. 1. Assessment of changes in the water regime of Russian rivers by observation data. Water Resour., 49(3), 333-350, DOI: 10.1134/S0097807822030046.

Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles C, Darmenov A, Bosilovich MG, Reichle R, Wargan K, Coy L, Cullather R, Draper C, Akella S, Buchard V, Conaty A, da Silva A, Gu W, Kim GK, Koster R, Lucchesi R, Merkova D, Nielsen JE, Partyka G, Pawson S, Putman W, Rienecker M, Schubert SD, Sienkiewicz M, and Zhao B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J Clim. 30(13),5419-5454, DOI: 10.1175/JCLI-D-16-0758.1.

Gelfan A.N., Gusev E.M., Kalugin A.S., Krylenko I.N., Motovilov Yu.G., Nasonova O.N., Millionshchikova T.D. and Frolova N.L. (2022). Runoff of Russian Rivers under Current and Projected Climate Change: a Review 2. Climate Change Impact on the Water Regime of Russian Rivers in the XXI Century. Water Resour., 49, 351–365, DOI:10.1134/s0097807822030058

Georgiadi A.G., Kashutina E.A. (2016) Long-term changes in the runoff of the largest Siberian rivers. Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya, (5), 70–81, DOI: 10.15356/0373-2444-2016-5-70-81(In Russian).

Georgiadi A.G. and Milyukova I.P. (2002). The scale of hydrological changes in the Volga River basin under anthropogenic warming, Meteorology and Hydrology, 2, 72–79. (In Russian).

Georgiadi A.G. and Milyukova I.P. (2023) Peculiarities of Long-term Phases of the Increased and Decreased Don and Lena Runoff in the 19th–21st Centuries. Russ. Meteorol. Hydrol., 48, 1066–1075, DOI: 10.3103/S1068373923120075

Georgiadi A.G., Kashutina E.A. and Milyukova I.P. (2018). Long-term Changes of Water Flow, Water Temperature and Heat Flux of the Largest Siberian Rivers. Polarforschung, 87 (2), 167–176, DOI:10.2312/polarforschung.87.2.167

Georgiadi A.G., Milyukova I.P., Borodin O.O. and Barabanova E.A. (2024). Water Flow of the Largest Russian Rivers in Modern and Scenario Global Warming. Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya, 88(6), 855-866, DOI: 10.31857/S2587556624060011. (in Russian with English summary).

Georgiadi A.G., Koronkevich N.I., Milyukova I.P., Kashutina E.A. and Barabanova E.A. (2014). Contemporary and Scenario Changes in River Runoff in the Basins of the Largest Russian Rivers. Part 2. Basins of the Volga and Don. Moscow: Maks Press Publ. (In Russian).

Georgievsky M.V. and Golovanov O.F. (2015). Assessment of probable changes in runoff in the Amur River basin for the period up to 2020 and 2050 based on data from atmospheric and ocean general circulation models. In: Georgievsky, V.Yu., ed., Extreme Floods in the Amur Basin: Hydrological Aspects. St. Petersburg: OOO EsPeKha Publ., 153–179. (in Russian).

Georgievsky M.V. and Golovanov O.F. (2019). Forecasting changes in river water resources of Russian Federation based on CMIP5 runooff data. Vestnik of Saint Petersburg University. Earth Sciences, 64 (2), 206–218, DOI: 10.21638/spbu07.2019.203 (In Russian with English summary)

Georgievsky V.Yu., Yezhov A.V., Shalygin A.L., Shiklomanov I.A. and Shiklomanov A.I. (1996). Assessment of the effect of possible climate changes on hydrological regime and water resources of rivers in the former USSR, Meteorology and Hydrology, 11, 89–99. (in Russian).

Georgievsky V.Yu. ed. (2021). Scientific and applied reference book: Long-term fluctuations and variability of water resources and the main characteristics of river flow in the Russian Federation. Saint-Petersburg: OOO «RIAL» Publ. (in Russian).

Groisman P. Ya., Easterling D. R., Quayle R. G. and Golubev V. S. (1996). Reducing biases in estimates of precipitation over the United States: Phase 3 adjustments. J. Geophys. Res., 101 (D3), 7185–7195.

Guo H., Zhan C., Ning L., Li X., Hu S. (2022). Evaluation and comparison of CMIP6 and CMIP5 model performance in simulating the runoff. Theor. Appl. Climatol., 149, 1451–1470, DOI: 10.1007/s00704-022-04118-0

Gusev E.M. and Nasonova O.N. (2010). Modelirovanie teplo- i vlagoobmena poverkhnosti sushi s atmosferoi [Modeling of Heat and Moisture Exchange Between the Land Surface and the Atmosphere. Moscow: Nauka Publ. (In Russian).

Haerter O.J., Hagemann S., Moseley C.and Piani C. (2011) Climate model bias correction and the role of timescales. Hydrol. Earth Syst. Sci., 15, 1065–1079.

Harding R., Best M., Blyth E., Hagemann S., Kabat P., Tallaksen L.M., Warnaars T., Wiberg D., Weedon G.P., van Lanen H.A. I., Ludwig F. and Haddeland I. (2011) WATCH: Current knowledge of the Terrestrial global water cycle. J. Hydrometeorol., 12, 1149–1156, DOI: 10.1175/ JHM–D-11–024.1

Harris I., Jones P.D., Osborna T.J. and Listera D.H. (2013). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset I. Int. J. Climatol. [published online in Wiley Online Library (wileyonlinelibrary.com)] DOI: 10.1002/joc.3711

Hoseini S.M., Soltanpour M. and Zolfaghari M. R. (2024). Projected changes in precipitation and air temperature over the Volga River Basin from bias-corrected CMIP6 outputs. Numerical Methods in Civil Engineering, 8(2), 36-47, DOI: 10.61186/NMCE.2303.1034

Kalinin G.P. and Milyukov P.I. (1958). Approximate Calculation of Unsteady Motion of Water Masses. Proceedings of the Central Institute of Forecasts. Issue 66. Leningrad: Gidrometeoizdat, 72 (In Russian).

Kalnay et al. (1996). The NCEP/NCAR 40-year reanalysis project. Bull. Amer. Meteor. Soc., 77, 437–470.

Kattsov V.M. and Govorkova V.A. (2013). Expected surface air temperature, precipitation and annual runoff changes over the territory of Russia: projections with an ensemble of global climate models (cmip5), Trudy GGO, (569), 75–97. (In Russian).

Kislov A.V., Evstigneev V.M., Malkhazova S.M., et al. (2008). Forecast of Climatic Resource Availability of the East European Plain in the Conditions of Warming of the 21st Century. Moscow: Max Press. (In Russian).

Kuzin P.S. and Babkin V.I. (1979). Geographical Regularities of the Hydrological Regime of Rivers. Leningrad: Gidrometeoizdat. (In Russian).

Lehner F., Deser C., Maher N., Marotzke J., Fischer E. M., Brunner L., Knutti R. and Hawkins E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. Earth Syst. Dynam., 11, 491–508, DOI:10.5194/esd-11-491-2020, 2020.

Menzhulin G.V., Shamshurin V.I., and Savvateev S.P. (2005). On the assessment of the accuracy of model climate change scenarios recommended by the IPCC commission for calculating the consequences of global warming. In: Modern Problems of Environmental Meteorology and Climatology. St. Petersburg: Nauka Publ., 55–85. (In Russian)

Mokhov I.I., Semenov V.A. and Khon V.K. (2003). Estimates of possible regional hydrologic regime changes in the 21st century based on global climate models, Izvestiya Rossiiskoi Akademii Nauk. Seriya Atmospheric and Oceanic Physics, (39), 130–144. (In Russian).

Motovilov Yu.G. and Gelfan A.N. (2019). Runoff Formation Models in Challengers of River Basin Hydrology. Moscow: Izdatel'stvo Rossiiskoi Akademii Nauk (In Russian)

Muñoz Sabater J. (2019). ERA5-Land monthly averaged data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.68d2bb30

Muraveisky S.D. (1960). Rivers and Lakes, Hydrobiology, Runoff. Moscow: Geografgiz Publ. (In Russian)

O'Neill B. C., Tebaldi C., van Vuuren D. P., Eyring V., Friedlingstein P., Hurtt G., Knutti R., Kriegler E., Lamarque J.-F., Lowe J., Meehl G. A., Moss R., Riahi K., and Sanderson B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci. Model Dev., [online] 9, 3461–3482. Available at: https://doi.org/10.5194/gmd-9-3461-2016 [Accessed 2 Oct. 2025]

Sharma S. and Singh P.K. (2017). Long Term Spatiotemporal Variability in Rainfall Trends over the State of Jharkhand, India. Climate, 5, 2–18, DOI:10.3390/cli5010018.Shi X., Qin T., Nie H., Weng B. and He S. (2019). Changes in major global river discharges directed into the ocean. Int. J. Environ. Res. Public Health, 16, 1469.

Shiklomanov A.I., Golovanov O., Lammers R.B., Tretjyakov M. and Yang D. (2011). Dam/Reservoir-Induced Hydrological Changes in Large Siberian Rivers. Amer. Geophys. Union, Fall Meeting 2011. Abstract C31B-06

Shiklomanov I.A. (Ed.) (2008). Water Resources of Russia and Their Use. Saint-Petersburg: State Hydrological Institute Publ. (in Russian).

Shpakova R.N., and Wang P. (2023). Change in Perennial Runoff Fluctuations of the Rivers in South Yakutia in the Context of Global Warming. Russian Arctic, 5(1), 33-44, DOI: 10.24412/2658-4255-2023-1-33-44 (In Russian)

Yamazaki D., Ikeshima D., Sosa J., Bates P.D., Allen G.H., Pavelsky T.M. (2019). MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets. Water Resources Research, 55, 5053–5073, DOI: 10.1029/2019WR024873

Yang D. and Kane D.L. (Eds.) (2021). Arctic Hydrology, Permafrost and Ecosystems. Cham: Springer, DOI: 10.1007/978-3-030-50930-9

Ye B. Aisheng, Daqing Yang D., Yongjian Ding Y., Tianding Han T., Toshio Koike T. (2004). A Bias-Corrected Precipitation Climatology for China. J. Hydrometeorol., 5, 1147-1160.

ges.rgo.ru/jour/

ISSN 2542-1565 (Online)