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.

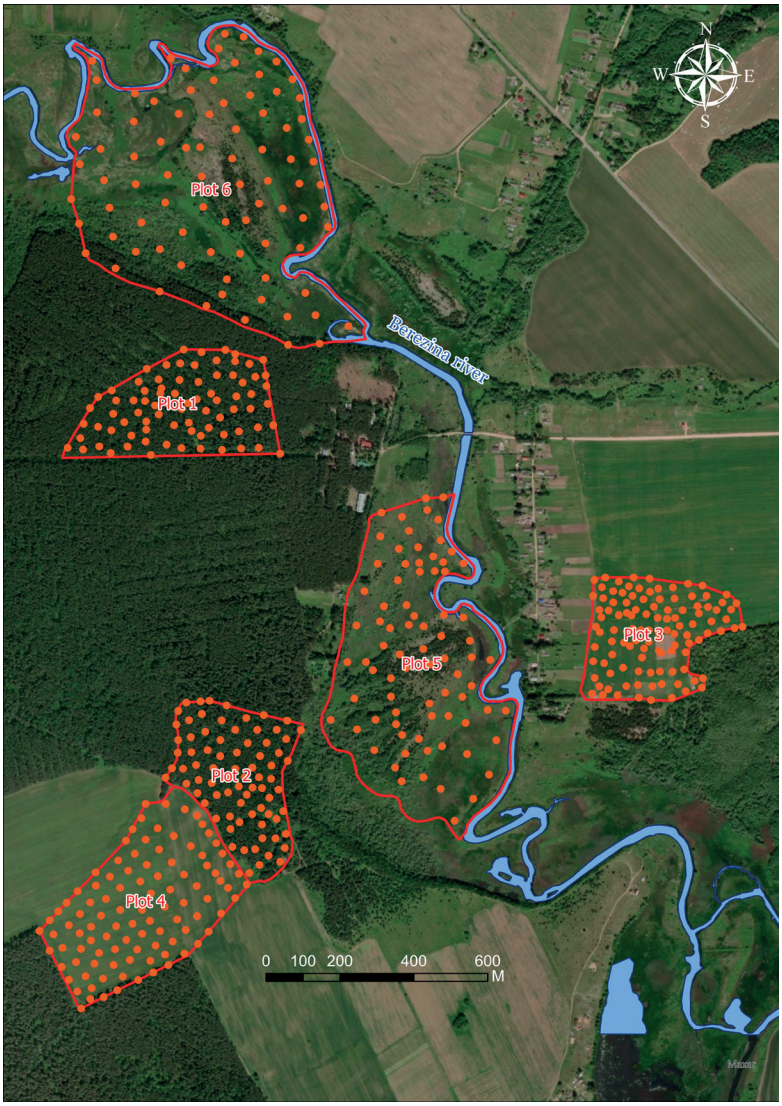

Fig. 1. Overview map of the study area location with sampling points

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

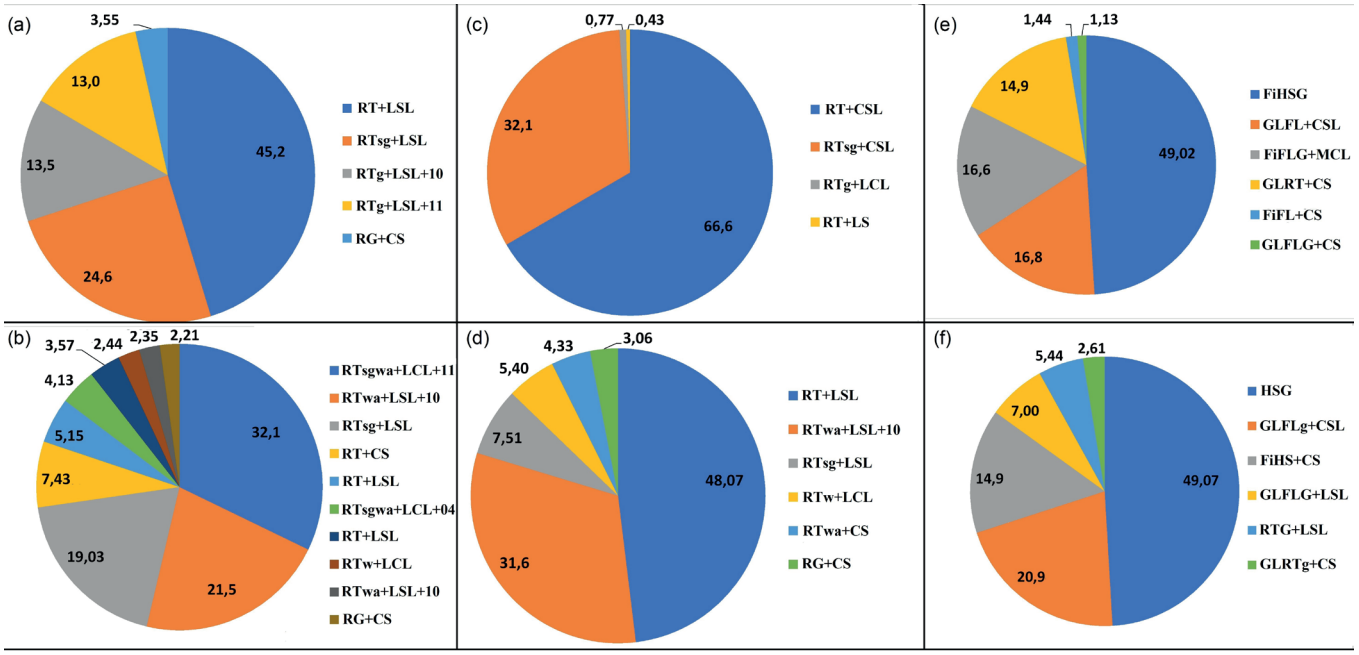

Fig. 2. Land cover composition of the study plots: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

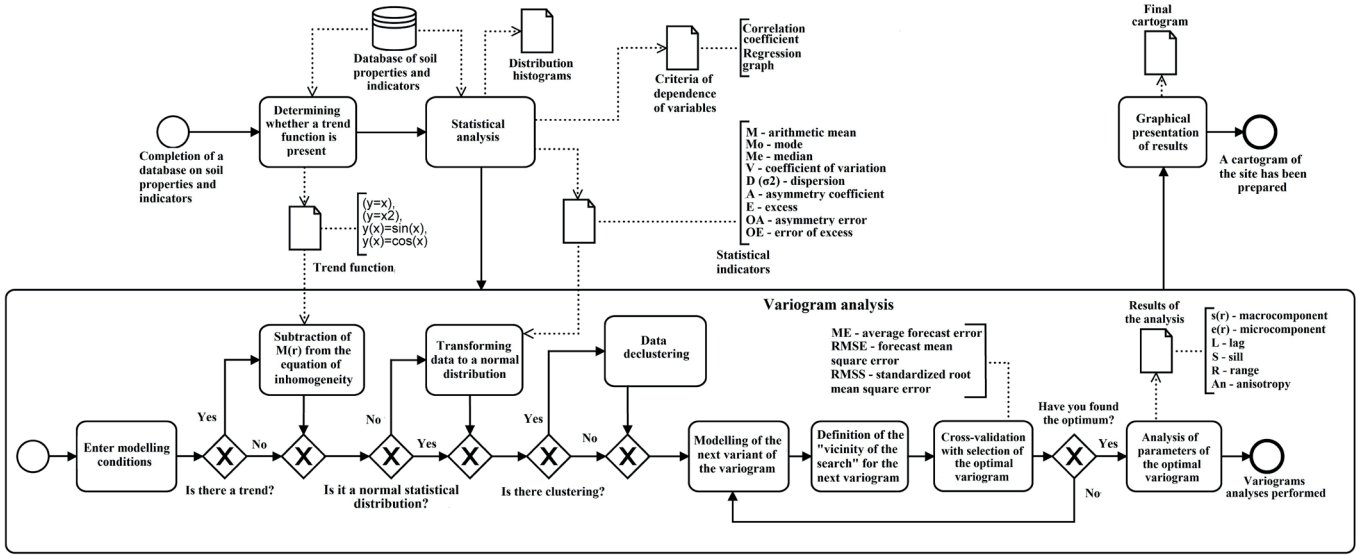

Fig. 3. Block diagram of the model of geostatistical diagnostics of soil properties variability

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.

$$O_A = 3 \times \sqrt{\frac{6 \times (n-1)}{(n+1) \times (n+3)}} \quad (1)$$

$$O_E = 5 \times \sqrt{\frac{24 \times n \times (n-2) \times (n-3)}{(n+1)^2 \times (n+3) \times (n+5)}} \quad (2)$$

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):

$$ME = \frac{\sum_{i=1}^{n} \left(\hat{Z}(s_i) - z(s_i) \right)}{n} \quad (3)$$

where ME – average forecast error;

$\hat{Z}(s_i) - z(s_i)$ the average difference between the measurement and the predicted value;

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{Z}(s_i) - z(s_i) \right)^2}{n}} \quad (4)$$

where *RMSE* – forecast mean square error.

$$RMSS = \sqrt{\frac{\sum_{i=1}^{n} \left[\left(\hat{Z}(s_i) - z(s_i) / \sigma(s_i) \right) \right]^2}{n}} \quad (5)$$

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

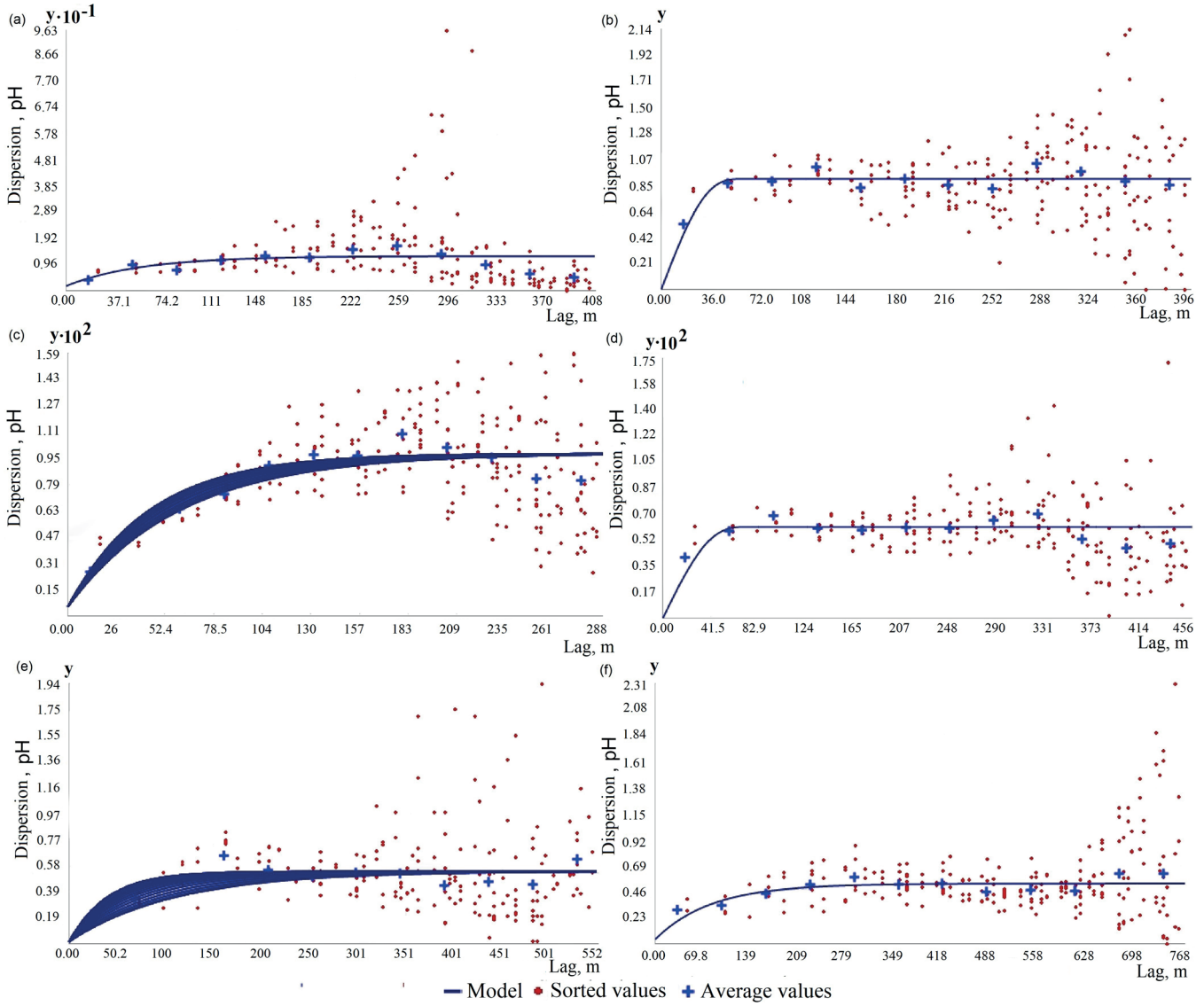

Fig. 4. Soil acidity variograms of the study sites: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

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).

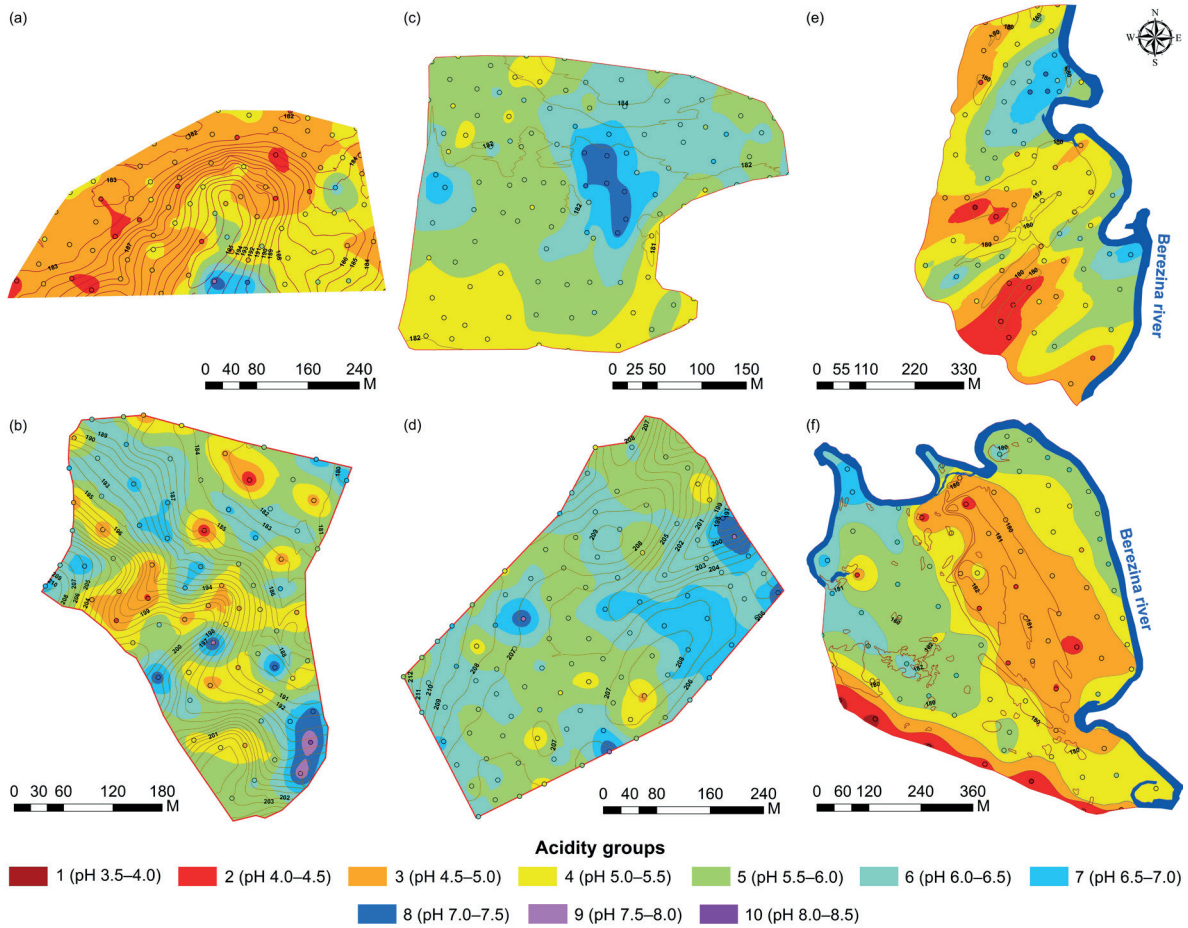

Fig. 5. Soil acidity cartograms of the study sites: (a) and (b) – forest plots 1 and 2; (c) and (d) – cropland plots 3 and 4; (e) and (f) – meadow plots 5 and 6

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 pHKCl 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 pHKCl >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 pHKCl 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)