ВЕСТНИК БЕЛОРУССКОЙ ГОСУДАРСТВЕННОЙ СЕЛЬСКОХОЗЯЙСТВЕННОЙ АКАДЕМИИ № 3 2020

МЕЛИОРАЦИЯ И ЗЕМЛЕУСТРОЙСТВО

UDC 631.474

USE GEOSTATISTICS TOOLS FOR EVALUATION THE SPATIAL DISTRIBUTION OF ACID-SOLUBLE ZINC IN THE SOIL

ИСПОЛЬЗОВАНИЕ МЕТОДОВ ГЕОСТАТИСТИКИ ДЛЯ ОЦЕНКИ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ КИСЛОТОРАСТВОРИМОГО ЦИНКА В ПОЧВЕ

T. N. MYSLYVA

Т. Н. МЫСЛЫВА

УО «Белорусская государственная сельскохозяйственная академия»,
г. Горки, Республика Беларусь, 213407, e-mail: byrty41@yahoo.com

(Поступила в редакцию 08.06.2020)

This article reflects on the problem of the possibility of using geospatial statistics methods to assess the spatial distribution of acid-soluble zinc in soil and applying geostatistical analysis methods to form management zones with different levels of acid-soluble zinc in the soil within the land use of an agricultural enterprise. The results of assessing the nature of the spatial distribution of data on the content of acid-soluble zinc in the soil by applying the functionality of the tools of the «Analysis of structural patterns» and «Calculation of clustering» modules of ArcGIS 10.5 are given. In particular, the data grouping analysis is performed using the k-means algorithm. The distance from which it is necessary to begin the analysis of spatial autocorrelation was 500 m, while the magnitude of the increment (lag), established empirically, is 50 m. The presence of reliable clustering of acid-soluble zinc in the soil was established (the actual value of the global Moran index is 0.360388; p-value> 2.58). Three management zones with different Zn contents were identified, with an area of 2074.64 ha (average acid-soluble copper content 3.10 mg/kg), 4684.91 (average acid-soluble copper content 3.99 mg/kg) and 1582.55 ha (average the content of acid-soluble copper 5.47 mg/kg). The presence of a steady trend in the west-east direction in increasing the zinc content in the central part of land for zone 1 and zone 3 and a steady trend in reducing the zinc content in the north-south direction for zone 3 while reliable trend for zone 2 was absent for any direction. The obtained information can be used to develop task maps for the differential application of micronutrient fertilizers during the introduction of precision farming.

Key words: *acid-soluble zinc, geospatial statistics, management zones, soil, clustering, Moran index, Getis-OrdGi, trend.*

В статье рассматривается проблема возможности использования методов геопространственной статистики для оценки пространственного распределения кислоторастворимого цинка в почве и применения методов геостатистического анализа для формирования менеджмент-зон с различным уровнем содержания в почве кислоторастворимого цинка в пределах землепользования сельскохозяйственного предприятия. Приведены результаты оценки характера пространственного распределения данных о содержании в почве кислоторастворимого цинка посредством применения функциональных возможностей инструментов модулей «Анализ структурных закономерностей» и «Расчет кластеризации» ArcGIS версии 10.5., в частности, выполнен анализ группирования данных с использованием алгоритма k-средних. Величина расстояния, с которого необходимо начать анализ пространственной автокорреляции, составила 500 м, тогда как величина приращения (лага), установленная эмпирическим путем, – 50 м. Величину лага рекомендуется использовать при создании мониторинговой сети наблюдений за содержанием Zn. Установлено наличие достоверной кластеризации содержания в почве кислоторастворимого цинка (фактическая величина глобального индекса Морана составляет 0,360388; p-значение > 2,58). Идентифицированы три менеджмент-зоны с различным содержанием Zn, площадью 2074,64 га (среднее содержание кислоторастворимого цинка 3,10 мг/кг), 4684,91 (среднее содержание кислоторастворимого цинка 3,99 мг/кг) и 1582,55 га (среднее содержание кислоторастворимого цинка 5,47 мг/кг). Установлено наличие достоверного тренда в направлении запад-восток в увеличении содержания цинка в центральной части землепользования для зон 1 и 3, а также устойчивого тренда в снижении содержания цинка в центральной части землепользования в направлении север-юг для зоны 3, тогда как для зоны 2 достоверные тренды в пространственном распределении данных отсутствовали. Полученная информация может использоваться для разработки карт-заданий по дифференцированному внесению микроудобрений при внедрении элементов точного земледелия.

Ключевые слова: *кислоторастворимый цинк, геопространственная статистика, менеджмент-зоны, почва, кластеризация, индекс Морана, Getis-OrdGi, тренд.*

Introduction

Zinc is an essential trace element that is part of enzymes and is involved in protein, carbohydrate, phosphorus metabolism, in the biosynthesis of vitamins, growth substances, RNA and chlorophyll. It is part of

30 enzymes (carbonic anhydrase, carboxypeptidase, glutamate dehydrogenase, aldolase, phospholipase, etc.) [12]. Its content in the soil primarily depends on the mineralogical and granulometric composition of the parent rocks, the type of soil-formed processes, the chemistry and level of groundwater, the quantity and quality of the organic matter of the soil, as well as the intensity of anthropogenic activity [2].

The main parent rocks of Belarus are traditionally poor in zinc especially fluvioglacial and ancient alluvial sand deposits, as well as weathering products of crystalline rocks. In this regard, soils formed on such rocks have low reserves of gross and mobile zinc. Over the past 15 years, the weighted average content of acid-soluble zinc in Belarusian soils has increased from 3.02 mg/kg to 3.06 mg/kg, and the share of arable land with a Zn content of less than 3.0 mg/kg ranges from 52.6 % to 73.8 % depending on the region. Arable soils in the Mogilev region are characterized by a predominantly low availability of acid-soluble zinc (weighted average content is 3.4 mg/kg). Moreover, 86 % of them are deficient in this trace element: the first group of Zn provides 52.7 % of the area of arable land, the second – 33.4 % and only 11.8% of the area of arable land has a high zinc content. In the Goretsky district, 56.9 % of arable land has a zinc content of less than 3.0 mg/kg; 31.4 % – from 3 to 5 mg/kg; 10.1 % – from 5 to 10 mg/kg and only 1.6 % contain more than 10 mg/kg of this element [4].

Zinc belongs to pedochemical active substances that create acid-base and redox conditions in the soil and thus affect the general soil-geochemical environment. Based on the fact that on the one hand it is one of the most important trace elements, and on the other it refers to heavy metals, the geoecological features of its distribution in natural and agricultural landscapes should also be considered from two perspectives:

1) assess the level of deficiency of this element for plants;

2) determine the degree of danger of soil pollution and vegetation.

The application of the capabilities of GIS analysis is the most optimal for the search for spatial patterns in the distribution of certain soil indicators in particular, acid soluble zinc in soil, and the relationships between them. However, in modern practice of agrochemical monitoring carried out both in the Republic of Belarus and in neighboring countries, soil surveys are provided without precise positioning, therefore, it is difficult to say with certainty that the samples were taken at the same place during repeated observation. This practice makes it impossible to reflect the real dynamics of soil indicators within land use and contrary to the concept of «monitoring», which subsequently leads to incorrect results when calculating the doses of fertilizers and chemical reclamants, and directly affects both the economic activity of the agricultural enterprise and the environmental situation within the agricultural landscape [5].

Main part

The purpose of the study is to analyze the possibility of using geospatial statistics methods to assess the spatial distribution of acid-soluble zinc in the soil of arable land of the Republican Unitary Enterprise “Educational Experimental Farm of BSAA” for the formation of management zones when introducing elements of the precision farming system.

The objectives of the study include the following:

1) to perform data grouping analysis using the k-means algorithm;

2) to determine the minimum and maximum distances of the neighborhood of the search for the nearest neighborhood, making it possible to choose the optimal value of the neighborhood of the search when modeling the spatial distribution of acid-soluble copper;

3) to calculate the global Moran index, which allows to determine whether there is a clustering phenomenon in relation to attributive data on the content of acid-soluble zinc in the soil;

4) to determine the overall Getis-OrdG index for assessing the overall structure and trend of geodata, as well as the degree of clustering of high and/or low sample values of acid-soluble zinc;

5) to calculate the Getis-OrdG * index, which allows to establish the presence of data clustering with high and low values;

6) to form management zones for the content of acid-soluble zinc in the soil.

The studies were carried out on the territory of Gorky district of Mogilev region within the land use of RUE “Educational Experimental Farm of BSAA” on an area of 8342.10 thousand hectares. The data about the content of acid-soluble zinc obtained from the agrochemical survey of the territory of RUE “Educational Experimental Farm of BSAA”, executed in 2018 by the Mogilev Regional Design and Exploration Station of Agrochemicalization, were used for the analysis. The soil cover of the study area is represented by Sod-podzolic, Umbric Retisols (WRB, 2016); Eutric Podzoluvisol (FAO, 1988) [6].

The spatial distribution analysis was performed using the functionality of the Spatial Statistics Tools of ArcGIS version 10.5. Statistical characteristics of a sample of data on the content of acid-soluble zinc used to perform geostatistical analysis are presented in the Table 1.

Table 1. **Statistical characteristics of a sample of data on the content of acid-soluble zinc in the soil used to perform geostatistical analysis**

Indicator name and sample size	Indicator value			Sd	Cv, %	Med	Kurtosis	Skewness	Interquartile range
	min	max	mid						
Zn, mg/kg, n=1638	1.85	10.01	4.06	1.26	31.0	3.87	9.32	1.93	1.24

Note: Sd is the standard deviation; Cv is the coefficient of variation; Med is the median.

The global Moran (I) index was calculated by the formula (1) [7]:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=i}^{n} w_{ij}(y_i - \bar{y})\cdot(y_j - \bar{y})}{\left[\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right]\cdot\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]} \qquad (1);$$

where n denotes the number of units in the sample; w_{ji} denotes the weight of the spatial relationship between the i-th and j-th sampling units; y_i denotes the attribute value for the i-th sample unit; $\bar{y}$ denotes the sample mean value of the attribute.

The Getis-OrdGi * index value was counted using the formula (2) [7]:

$$GetisOrdGi^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \qquad (2);$$

where x_j denotes the attributive value of the object of observation; $w_{i,j}$ denotes spatial weight between objects i and j; n denotes the total number of objects.

The first step in the study of geospatial data that characterize certain soil indicators is the analysis of their grouping. The main purpose of grouping is to search for the presence of natural clusters in the data. With its help, data on soil parameters are distributed on a given number of groups in which all indicators are most similar to each other, while the groups themselves are as different as possible from each other. Using the analysis of grouping it is possible to establish the presence within the land use of homogeneous zones with a specific set of parameters. In our case, a “set of parameters” means the intervals of the content of copper in the soil according to the gradation given in the guidelines for conducting large-scale agrochemical and radiological surveys of the soils of agricultural lands of the Republic of Belarus [8]. Since the minimum zinc content in the soil was 1.85 and the maximum was 10.01 mg/kg, four groups of clusters were identified in the analysis of grouping. The localization of the selected clusters is shown in Figure 1, and their main characteristics are described in Table 2.

Table 2. **Statistical characteristics of identified cluster groups according to the content of acid-soluble zinc in the soil**

Group of clusters	Mean	Sd	Minimum value	Maximum value	R^2	Group area, hectares
1	2.92	0.38	1.85	3.44	0.20	2590.61
2	3.98	0.32	3.45	4.61	0.15	3800.74
3	5.25	0.54	4.62	6.87	0.37	1683.75
4	8.66	1.30	6.98	10.01	0.28	267.00

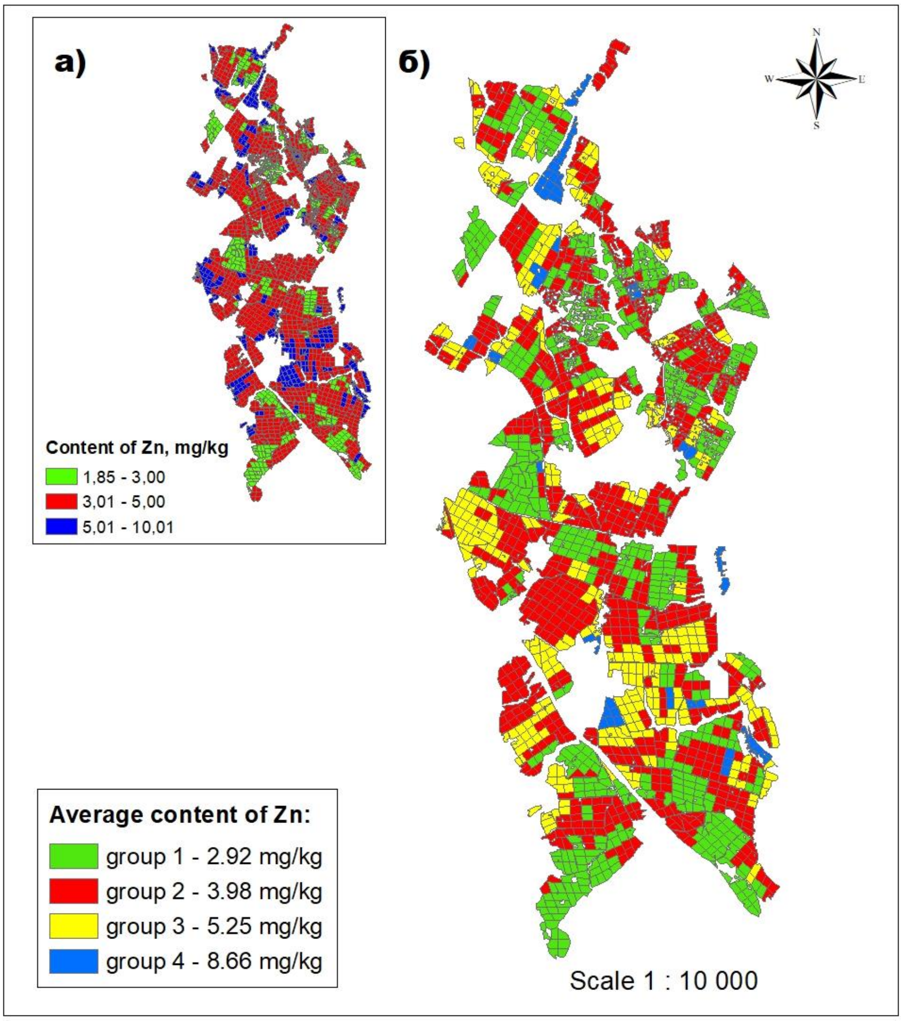

Fig. 1. Spatial localization of identified groups of acid-soluble zinc content clusters within the land use of the RUE «Educational Experimental Farm of BSAA»
(a – copper content according to the generally accepted classification; b – identified cluster groups of content of zinc)

It has been determined that group 2 has a maximum area of allocated clusters of 3800.74 ha, and group 4 has a minimum area, which is 267.00 ha. In the first group, the average value is outside the global lower quartile, in the group two it is equal to the median value, in the group three it is outside the global upper quartile, and in the group four, it lies outside the global upper quartile, but within the range of variation of the values in the group. It is necessary to add that the value of R^2 reflects the extent to which the variation in the source data was saved during the grouping process, respectively, the more R^2 is for a certain variable, the better this variable distinguishes between values.

In general, we can state the following:

1) the analysis of grouping allows to establish the presence of homogeneous zones with a certain content of acid-soluble zinc within the land use;

2) the selected groups of clusters give a certain idea of the nature of the spatial distribution of zinc within the study area, but are unsuitable for establishing the clear boundaries of management zones.

3) the performing grouping analysis can be used as an alternative to other, more complex methods of geospatial analysis, if it is necessary to perform zoning of the territory on one feature.

Performing cluster analysis, in contrast to grouping analysis, allows not only to establish the presence of clusters and to assess the reliability of clustering, but also to analyze clusters, identify outliers of high and low values and establish the boundaries of management zones with different content of zinc in soil. To determine the value of a fixed distance or the minimum distance of a neighborhood searching for a neighborhood between the values of the content of acid-soluble zinc in the soil, the tool «Incremental Spatial Autocorrelation» was used [9]. The value of the initial (distance at which it is necessary to start the analysis of spatial autocorrelation) and incremental (distance by which it is necessary to increase the initial distance at each subsequent iteration) distances were set in the dialog box of this tool. As a result of the calculations, the distance at which it is necessary to begin the analysis of spatial autocorrelation was 500 m, while the magnitude of the increment (lag) established empirically is 50 m. Ten distance intervals evenly distributed throughout the extent were highlighted by performing incremental spatial autocorrelation. The global Moran index was calculated for each interval and the interval for which this index would be the largest was recommended as the optimal distance for the search neighborhood. As a result, we got a graph on which the minimum and maximum distances of the neighborhood of the search for the nearest neighborhood are marked. It has been also established that the minimum distance to search for the closest neighborhood between the contents in the soil Zn is 700 m (Figure 2).

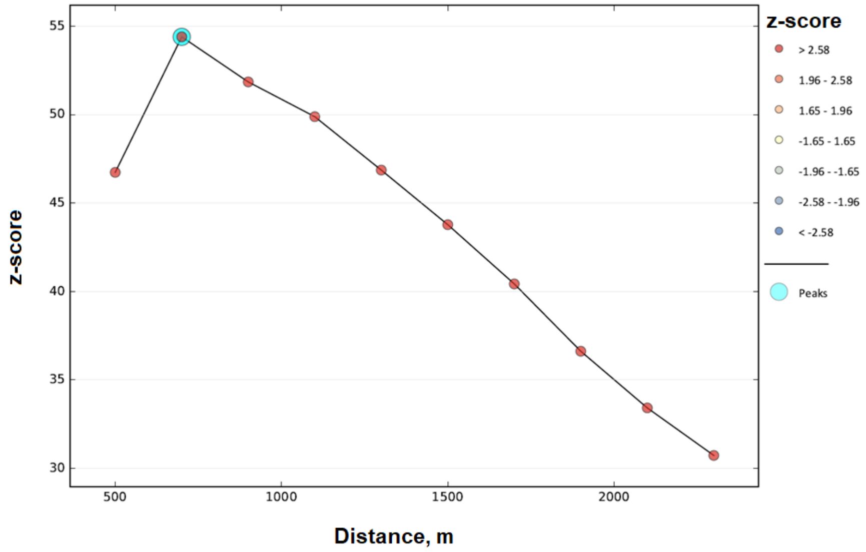

Fig. 2. Graphical interpretation of the minimum distance of the neighborhood search neighborhood between the values of the content of acid-soluble zinc in the soil

The value of the global Moran index was calculated in order to determine whether there is a clustering phenomenon in relation to attribute data. This index is a measure of spatial autocorrelation and characterizes the presence or absence of spatial autocorrelation of geodata. The results of determining the magnitude of the global Moran index, calculated for the sample by the attribute values, as well as the z-score value, which allow judging the nature of the data distribution, are presented in the Table 3.

Table 3. **The results of determining the magnitude of the global Moran index and the general Getis-OrdG index**

Indicator name and sample size	Actual value	Expected value	p-value	z-score
Global Moran index				
Zn, mg/kg, n=1638	0.360388	-0.000611	0.000058	47.335097
General Getis-OrdGi index				
Zn, mg/kg, n=1638	0.012800	0.012590	3.935778	0.000083

The actual value of the global Moran index is 0.360388; therefore, data on the content of acid-soluble zinc in the soil within the study area are not randomly distributed and clustered. Since the value of the z-score exceeds 2.58, it can be argued with a probability of 99 % that the clustered type of data distribution is not random. The degree of clustering of values (searching for unexpected bursts of high or low values in space) was determined by calculating the overall Getis-OrdG index, which was used to evaluate the overall structure and trend of geodata. Since the actual value of the overall Getis-OrdG index is larger than expected, there is a clustering of data with high attribute values.

As a result of the analysis of hot spots, statistically significant spatial clusters of high values (hot spots) and low values (cold spots) for the content of acid-soluble zinc in the soil were determined and visualization of the obtained data was performed. It should be noted that the purpose of this analysis is to determine whether the neighborhood of the object has statistically significant differences between the studied attribute and the entire range of values. The analysis of clusters and outliers allows to identify the concentrations of high and low values and helps to establish where the clearest boundaries between the contours with high and low zinc content in the soil are located (Figure 3).

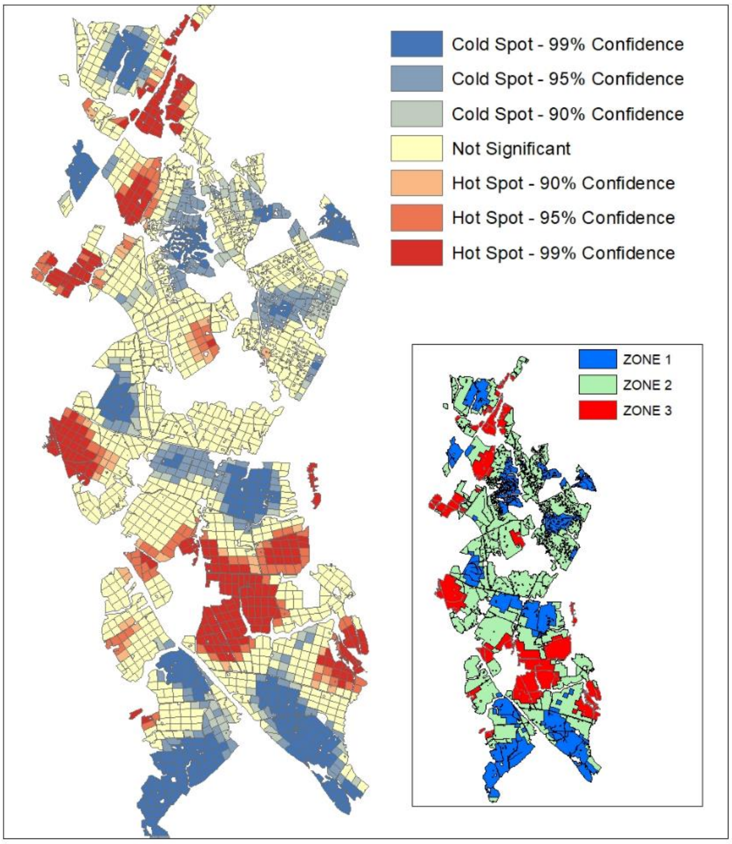

Fig. 3. «Hot Spot Analysis» results within the land use of the RUE «Educational Experimental Farm of BSAA» *(a – identified reliable localization of high and low values; b) dedicated management zones)*

Based on the results of the analysis of the hot spots, three management zones with different contents of acid-soluble zinc were identified: zone 1 – average content of acid-soluble zinc 3.10 mg/kg, area – 2074.64 hectares; zone 2 – the average content of acid-soluble zinc is 3.99 mg/kg, the area is 4684.91 hectares; zone 3 – the average content of acid-soluble zinc is 5.47 mg/kg, the area is 1582.55 hectares. The statistical characteristics of the content of acid-soluble copper within the identified zones are shown in Table 3.

Table 3. **Statistical characteristics and square of a sample of data on the content of acid-soluble zinc (mg/kg) within the limits of identified zones**

Identified zones and sample size	Indicator value			Sd	Cv, %	Med	Iinterquartile range	Area, hectares
	min	max	mid					
Zone 1, n=408	1.85	5.60	3.10	0.62	20.0	2.99	0.87	2074.64
Zone 2, n=915	1.85	7.46	3.99	0.74	18.9	3.95	0.88	4684.91
Zone 3, n=314	2.41	10.01	5.47	1.69	30.8	5.06	1.71	1582.55

Note: Sd is the standard deviation; Cv is the coefficient of variation; Med is the median.

It was also found that there is a steady trend in the west-east direction in increasing the zinc content in the central part of land for zone 1 and zone 3 and a steady trend in reducing the zinc content in the north-south direction for zone 3 while reliable trend for zone 2 was absent for any direction.

It should also be noted that the selected management zones through the use of GIS functionality can be divided into work parcels formed according to the working width of the used high-precision agricultural equipment used for the differential application of mineral fertilizers when introducing precision farming system, and the resulting cartographic images can be used as task maps to ensure the effective operation of the equipment.

Conclusion

Using a geostatistical analysis of data on the content of acid-soluble zinc in the soil allows you:

1) to identify and mathematically evaluate the spatial distribution of this essential trace element;

2) to study spatial autocorrelation of data and determine the lag value that should be taken into account when selecting a step in the process of creating a monitoring network for monitoring the zinc content in soil for precision farming;

3) to evaluate the clustering of data on the zinc content in the soil and determine the location of clusters in space;

4) to visualize clusters by constructing cartographic images;

5) to determine the boundaries and areas of management zones for precision farming, within which it is possible to apply the differential application of micronutrient fertilizers.

Further research should be concentrated in the direction of studying the mutual influence of the spatial distribution of humus and the pH of the soil solution on the spatial differentiation of acid-soluble zinc in the soil.

REFERENCES

1. Школьник, М. Я. Микроэлементы в жизни растений / М. Я. Школьник. – М.: Наука, 1978. – 250 с.

2. Атлас агроекологічного стану ґрунтового покриву Житомирської області / В. А. Трембіцький, Т. М. Мислива, Г. М. Мартенюк, Ю. А. Білявський. – Житомир: ПТЦ «Облдержродючість», 2011. – 56 с.

3. Мислива, Т. М. Цинк у ґрунтах Житомирського Полісся / Т. М. Мислива // Вісник ЖНАЕУ. – 2011. – № 1(28). – С. 30–45.

4. Агрохимическая характеристика почв сельскохозяйственных земель Республики Беларусь (2013–2016 гг.) / И. М. Богдевич [и др.]; под общ. ред. И. М. Богдевича. – Минск: Ин-т почвоведения и агрохимии, 2017. – 275 с.

5. Якушев, В. В. Точное земледелие: теория и практика / В. В. Якушев. – СПб.: ФГБНУ АФИ, 2016. – 364 с.

6. Мыслыва, Т. Н. Геостатистический анализ пространственного распределения агрохимических свойств почв земель сельскохозяйственного назначения / Т. Н. Мыслыва, Ю. А. Белявский / Приемы повышения плодородия почв и эффективности удобрения : материалы Международной научно-практической конференции, посвященной памяти ученых: А. И. Горбылевой, Ю. П. Сиротина и В. И. Тюльпанова. – Горки: БГСХА, 2019. – С. 101–103.

7. Mitchell, A. The ESRI Guide to GIS Analysis /A. Mitchell. – Esri Press, 2005. – Volume 2. – 252 p.

8. Крупномасштабное агрохимическое и радиологическое обследование почв сельскохозяйственных земель Республики Беларусь: методические указания / И. М. Богдевич [и др.]. – Минск, 2006. – 63 с.

9. Myslyva, T. N. Use geostatistics tools for evaluation the spatial distribution of acid-soluble copper in the soil / T. N. Myslyva // Bulletin of BSAA. – 2020. – №2. – P. 170–175.