

# Research Article Geostatistical Analysis Methods for Estimation of Environmental Data Homogeneity

# Aleksandr Danilov,<sup>1</sup> Inna Pivovarova,<sup>2</sup> and Svetlana Krotova<sup>2</sup>

<sup>1</sup>Geoecology Department, Saint-Petersburg Mining University, St. Petersburg 199106, Russia <sup>2</sup>Department of Informatics and Computer Technology, Saint-Petersburg Mining University, St. Petersburg 199106, Russia

Correspondence should be addressed to Aleksandr Danilov; aleksandrsdanilov@gmail.com

Received 20 January 2018; Accepted 30 April 2018; Published 3 June 2018

Academic Editor: Sunil Nautiyal

Copyright © 2018 Aleksandr Danilov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The methodology for assessing the spatial homogeneity of ecosystems with the possibility of subsequent zoning of territories in terms of the degree of disturbance of the environment is considered in the study. The degree of pollution of the water body was reconstructed on the basis of hydrochemical monitoring data and information on the level of the technogenic load in one year. As a result, the greatest environmental stress zones were isolated and correct zoning using geostatistical analysis techniques was proved. Mathematical algorithm computing system was implemented in an object-oriented programming C #. A software application has been obtained that allows quickly assessing the scale and spatial localization of pollution during the initial analysis of the environmental situation.

### 1. Introduction

The construction of functional models of geosystems and prediction of the geosystems behavior are mandatory and necessary in conditions of increasing anthropogenic load in the present period of environmental studies. Optimization of nature management is an urgent need. To optimize the environmental management it is necessary to know how geosystem exists or existed in the absence of anthropogenic impact, which components of geosystems are most susceptible to anthropogenic impact, and ecosystems specificity of functioning under anthropogenic load. Classification of objects is necessary by the degree of environmental disturbance and environmental safety. Obviously, that it is wrong to apply uniform methods and standards to all geosystems. However, it is impossible to develop methods and standards for each individual geosystems because of the unacceptable time and material cost. Therefore, the necessity to integrate geosystem into taxa occurs. Application of unified methods for determining the permissible limits of anthropogenic impact and predicting their evolution is fully justified within the taxa. Thus, the task of evaluating the spatial homogeneity

of ecosystems with the possibility of subsequent zoning occurs.

Environmental scientists of various countries have long paid attention to the concept of ecological zoning. In 1967, Crowley first presented the concept of ecoregion, which refers to the land and water areas with similar ecosystem or being supposed to play similar functions [1]. Basing on this concept, the purpose of ecological regionalization is to provide suitable spatial units for studying, evaluating, restoring, and managing the ecosystem [2]. The concept of aquatic ecoregion originated from America. It refers to the freshwater ecosystem or living organism and the interrelated land units [3]. The aquatic ecosystems regionalization is one of the most important fields of ecological regionalization, and it is also the field most successfully studied [4].

The "quality" of regionalization or the correspondence of allocated area to the set goals depends largely on the choice of research method. The most widely used methods of regression [5] and cluster [6, 7] analysis in many examples reveal a high degree of subjectivity. So, applying different data sets, overly careful accounting or vice versa, neglect of the influence of constantly changing anthropogenic factors can lead to different zoning schemes. Applying the methods of geostatistical analysis and cartographic visualization is most suitable for environmental problems studying (that have a pronounced spatial aspect) relying on the experience of previous research in this field [8, 9].

## 2. Materials and Methods

An attempt to reconstruct the contamination of a water body and to isolate the zones of the greatest environmental stress based on the values measured at a limited number of points was the main objective of this study. Proof of correctness zoning with certain statistical algorithms also has been a major purpose. The whole course of research was conditionally divided into two stages. The first stage is the construction of a general view, analysis, and visualization of primary data. The second stage is the use of statistical calculation methods for estimating the spatial homogeneity of environmental characteristics and finalizing the model with further software implementation.

2.1. The First Stage Is the Construction of a Probability Model for the Distribution of the Characteristics. This problem was considered as a problem of interpolation in mathematics. In the standard approach, an unknown function is approximated by a parametric function whose form is given either explicitly (polynomial) or implicitly (the minimum curvature condition). The parameters are chosen to optimize some criterion of best approximation values at the points. The criterion can be statistical (least squares) or deterministic (exact coincidence at measurement points). Most of the existing interpolation methods are built into modern GIS packages. The main ones are as follows [10]:

- (i) IDW method is inversely weighted distances (average values of neighboring pixels by a predetermined number of neighbors or within a specified radius);
- (ii) Kriging is multistage selection of a mathematical function for a given number of points or for points within a given radius for propagation of dependencies on all points;
- (iii) Natural Neighbor finds the closest subset of input samples to the requested point and applies a weighted value based on proportionate areas to interpolate a value;
- (iv) Bilinear is bilinear interpolation, when the point value in the new image is calculated by linear interpolation between the values of the four nearest points;
- (v) TIN is a method when all the starting points are connected by triangles, resulting in an irregular triangulation network.

The Mapinfo-GIS package served as a tool for building the base map in this research project.

The initial data for modeling were materials of ecological and hydrochemical monitoring of the state of surface water bodies and information on the level of anthropogenic impact within one year (January-December 2016) on the territory of the Khibiny mountain massif (Figure 1), located in the central part of the Kola Peninsula of the Russian Federation (Apatity mining agglomeration). The values of the content of sulfate ion  $(SO_4^{2-})$  in surface waters were used to estimate the intensity of contamination. The probable source of sulfate ion entering the surface waters was the intensification of the extraction of the apatite-nepheline ore of the Rasvumchor field [11].

The method of inversely weighted distances (IDW) was chosen as the method of interpolation.

In the inverse distance weighted method (IDW), which can be assigned to a group of kriging methods, estimated points are determined on the basis of source points, found in its surroundings. The result is affected by several parameters such as range searches, the number of points involved in the analysis, and power factor. The process of IDW interpolation can be divided into the following steps [12]:

- (1) Searching for points that meet the criterion of neighborhood (the amount or the distance).
- (2) Allocating weights to each typed point. At this step, it is possible to determine the power factor (*p*); the bigger it is, the points which are farther will have a greater impact on the result.
- (3) Calculating the value of estimating points [13].

$$ww_{nn} = \left[ \left( x_n - x_0 \right)^2 + \left( y_n - y_0 \right)^2 \right]^{(-p/2)}$$
  
$$zz = \sum \left( ww_{nn} * \frac{zz_{nn}}{\sum ww_{nn}} \right),$$
 (1)

where  $w_n$  is weight of the points used to interpolate,  $z_n$  is value of the points used in interpolation,  $x_0$ ,  $y_0$  are the coordinates of estimating point, p is power factor, and z is value of the estimated point.

The method worked well with a large amount of initial data and showed the result in a convenient form for perception (Figure 2).

As a result, it was clarified that pollution is absent at points 1.1, 1.2, 1.3, 3.1, 3.2, and 3.5. Values at points 2.2, 2.11, 3.3, and 3.4 are excluded from further calculations due to data being uninformative. This is caused by close proximity. A single impregnation was detected at point 3.6, which can be caused by the infiltration of polluting components from the tailing dump of the mining enterprise located in the source of the stream. The site limited by points 2.1 and 2.12 is contaminated. The site is conditionally divided by the degree of pollution into districts I and II. The method of statistical estimation of data homogeneity was used to verify the actual presence of a spatial trend.

2.2. The Second Stage Is the Statistical Evaluation of the Spatial Homogeneity of Environmental Characteristics. There are a number of criteria for verifying spatial data for homogeneity. These criteria allow us to determine whether two samples (data on two different objects) are related to one general population or not [14]. If the samples belong to the same population, then the difference between the samples is within



FIGURE 1: Region of research and sampling points.



FIGURE 2: GIS-distribution of pollutants.

the limits of random variations of the quantities and there are no fundamental differences between the objects. In this case, parametric criteria require that the distribution of the sample is subject to a specific distribution law. Thus, the classical criteria of Student and Fisher require that the law of distribution of samples be sufficiently close to the normal law [15]. Parametric criteria allow us to directly estimate the level of the main parameters of the general populations, the difference in the means and the difference in variances. The criteria can identify trends in data changes and evaluate the interaction of two or more factors. Recently, the Cramer and Welch criteria [16, 17] have also been used to estimate the homogeneity of data. An additional advantage of these criteria is the optional equality of the variances of the compared samples. Parametric criteria are considered to be more powerful than nonparametric ones, provided that

the characteristics are measured in an interval scale and are normally distributed.

*Nonparametric* criteria do not have the above limitations. The term "nonparametric method" means that it is not necessary to assume that the distribution functions of the results of observations belong to any particular parametric group while it is used. Nonparametric criteria do not impose conditions for the recognition of the distribution law. However, criteria of this type do not allow a direct assessment of the level of such important parameters as the average or variance. Using nonparametric criteria is impossible to estimate the interaction of two or more conditions or factors affecting the change in characteristics. Many nonparametric methods have been developed, Smirnov's criteria [18], such as the omega-square (Leman-Rosenblatt) [19, 20], Wilcoxon (Mann–Whitney) [21, 22], van der Waerden [23], Savage, etc.

In addition, the affinity between the variables is usually investigated using correlation functions [24]. In this study, the calculation technique was reduced to the construction and further analysis of the homogeneity of the spacecorrelation function. The analysis of the function homogeneity was carried out based on the principle of assessing the significance of the difference between the actual correlation coefficient and the assumed coefficient in the total population. The Z-Fisher distribution was used as the evaluation criterion. The value of statistics obtained for the compared data groups was compared with the theoretical value at the accepted level of significance. Mathematical algorithm was as follows.

The auxiliary values were determined by the Fisher method from the values of the empirical  $\tilde{r}(\alpha_{jk})$  and theoretical  $r(\alpha_{ik})$  correlation functions

$$z_{jk} = \frac{1}{2} \ln \frac{1 + r(\alpha_{jk})}{1 - r(\alpha_{jk})}$$

$$\tilde{z}_{jk} = \tilde{z}(\alpha_{jk}) = \frac{1}{2} \ln \frac{1 + \tilde{r}(\alpha_{jk})}{1 - \tilde{r}(\alpha_{jk})} + \frac{r(\alpha_{jk})}{2(N_{jk} - 1)}$$
(2)

and the deviation or difference  $z_{jk} - \tilde{z}(\alpha_{jk})$  was calculated for all  $c_l^2 = l(l-1)/2$  pair wise distances  $\alpha_{jk}$  between the observation points.

Standard deviations  $\sigma_{zjk}$  of auxiliary variables  $z_{jk}$  from their conditional average values  $\tilde{z}(\alpha_{jk})$  were determined from the formula

$$\sigma_{zjk} = \frac{1}{\sqrt{N_{jk} - 1}} \tag{3}$$

According to the law of normal distribution of the normalized deviations from the average value in the confidence limits

$$\widetilde{z}\left(\alpha_{jk}\right) - t_{\sigma jk} < z_{jk} < \widetilde{z}\left(\alpha_{jk}\right) + t_{\sigma z jk} \tag{4}$$

of all  $c_l^2 = l(l-1)/2$  empirical values  $z_{jk}$  should fall P(I) = 0.683 = 68.3% for t = 1 or P(II) = 0.954 = 95.4% for t = 2.

Therefore, a necessary and practically sufficient condition for the homogeneity of the correlation function within the region under consideration is the fulfillment of inequalities

$$\left|z_{jk} - \tilde{z}\left(\alpha_{jk}\right) \ge \sigma_{zjk}\right| \text{ or } \ge 2\sigma_{zjk}$$
 (5)

approximately 31.7% or 4.6% of the total number  $c_l^2 = l(l - 1)/2$  of empirical values  $z_{jk}$ . In other words, for t = 1 and t = 2, the total empirical number of excesses

$$K_{\mathfrak{s}}(1) = K_{\mathfrak{s}} \left[ z_{jk} - \tilde{z}_{jk} \right] > \sigma_{zjk}$$

$$K_{\mathfrak{s}}(1) = K_{\mathfrak{s}} \left[ z_{jk} - \tilde{z}_{jk} \right] > 2\sigma_{zjk}$$
(6)



FIGURE 3: Spatial-correlation function.

should be approximately equal to the theoretically possible according to the normal distribution law, the number of excesses, i.e.,

$$K_{9}(1) \approx 0.317C_{l}^{2} = 0.317\frac{l(l-1)}{2}$$

$$K_{9}(2) \approx 0.317C_{l}^{2} = 0.046\frac{l(l-1)}{2}$$
(7)

Thus, the  $C_{11}^2 = (11 * 10)/2 = 55$  pairs of correlation ratios between the arrays of initial data at the site limited by points 2.1 and 2.12 were calculated. A graph of correlation dependence was constructed and the equations of theoretical and empirical correlation functions were obtained (Figure 3).

The auxiliary values of  $z_{jk}$  and  $\tilde{z}(\alpha_{jk})$  were calculated from them. The standard deviation  $\sigma_{zjk}$  of the auxiliary values  $z_{jk}$  from their conditional average values  $\tilde{z}(\alpha_{jk})$  was determined (Table 1).

As a result, it was concluded that the spatially correlation function of the region under study is inhomogeneous, since the total empirical number of exceedances is greater than the theoretically possible.

When calculating the RMS, the number of exceedances was 20, and according to the law of normal distribution, there should be 17 by  $2\sigma_{ik}$ : 7 and 2, respectively.

#### 3. Results and Discussion

The heterogeneity of spatially distributed data within the initial study area is caused by an increased level of technogenic impact along the line of sampling points 2.1–2.12. Discharges of sewage from a mining enterprise are located in the catchment basin of small rivers in this area.

The remaining sampling points are located on small rivers and streams, the catchment basins of which lie at the base of the Khibiny mountain massif and the main power source is the melting of snow in the summer season and precipitation; this causes a low content of polluting components, including sulfate.

In order to obtain a conclusion on the zoning of ecological components, this region is divided into two subareas: 1, 2 (determined by the results of the previously constructed GIS project) for each of which similar calculations were

·/~~	Case, when a rejection exceeds a doubled MSE $(\Delta > 2\delta)$	1											-
סזוו סווו זד המווס זווס	Case, when a rejection exceeds a MSE $(\Delta > \delta)$	1	1	1	1	1		1	1	1	1		-
0	Doubled mean square error (28)	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0,603023	0.603023
0	Mean square error (ð)	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0,301511	0.301511
menty of cross-correlation function of the investigated descriptions (a fi	Rejection (Δ)	1,203298	0,526303	0,427799	0,362274	0,330208	0,218134	0,597625	0,492953	0,391579	0,351679	0,22768	0.759216
	$\widetilde{Z}_{jk}$	1,373064	1,221999	1,167063	1,119238	1,091341	0,958347	1,252727	1,205259	1,141991	1,110392	0,972574	1.303627
	$Z_{jk}$	2,576362	1,748302	1,594861	1,481512	1,421548	1,176481	1,850351	1,698212	1,53357	1,462071	1,200254	2.062843
	Value of theoretical cross- correlation function	0,988499	0,941182	0,920891	0,901751	0,889921	0,826339	0,951779	0,935185	0,911034	0,898054	0,833732	0.968209
	Value of empiric cross- correlation function	0,868792	0,827205	0,809371	0,792549	0,782152	0,726269	0,836519	0,821934	0,800707	0,7893	0,732767	0.850959
0	Coefficient of pair wise correlation	0,986	0,984	0,5	0,7	0,99	0,74	0,85	0,7	0,4	0,8	0,99	0.5
	Distance between sampling points	0,14	0,716	0,963	1,196	1,34	2,114	0,587	0,789	1,083	1,241	2,024	0.387
	Pairs	2,1-2,3	2, 1-2, 4	2,1-2,5	2,1-2,6	2,1-2,7	2, 1-2, 8	2, 3-2, 4	2,3-2,5	2,3-2,6	2,3-2,7	2,3-2,8	2.4-2.5

	Eo
ç	Ę
	airs
	2
	ň
	t o
-	Igua
-	bro
	S 15
	Suo
-	lati
-	ICU
ç	t ca
	L O
	ove
	ent
	a m
ç	tras
`	g
	ons
:	pti
•	SCTI
-	б Ф
	ted
	llga
	vest
	III
7	the
ر	lot
	tior.
	unct
ç	ם tu
	tioi
-	rela
	COL
	SS-C
	cro
ç	of
	ŝιtγ
	ene
	gon
	nor
ç	ot
	on
	lati
•	stun.
F	э Ц
,	E

📃 Ho	mogeneity calculatio	ı						
Data	entry field							
2,1	24,2775 22,2025	17,845	17,2225	16,6	19,7125	20,75	21,995	19,2975
2,3	43,2549 39,5579	31,7942 31 4245	30,6851	29,576	35,1215	36,97	39,1882	34,3821
2,4	64,6191 59,0961 53,5731 49,707	47,4978	45,8409	44,184	52,4685	55,23	<b>58,5</b> 438	51,3639
2,5	16,0173 14,6483	11,7734	11,3627	10,952	13,0055	13,69	14,5114	12,7317
2,6	57,3651 52,4621	42,1658	40,6949	39,224	46,5785	49,03	51,9718	45,5979
2,7	38,5398 35,2458 31,9518 29,646	28,3284	27,3402	26,352	31,293	32,94	34,9164	30,6342
2,8	31,7187 29,0077	23,3146	22,5013	21,688	25,7545	27, <b>1</b> 1	28,7366	25,2123
								Calculate
Unev	ven							

FIGURE 4: Example of calculations in the software.

made, and a conclusion was made about the homogeneity of environmental characteristics in each of them.

The preliminary carried out reconnaissance studies indirectly point to correctness of the proposed results. [25]. Sampling points 1.1-1.3 are located in the zone of the botanical garden of the Kola Branch of the Russian Academy of Sciences and are not affected by any technogenic impact. In the upper course of the river, in the area of point 2.4, a discharge of wastewater from a mining and processing plant with an extremely high content of sulfate ion in water has been detected. At the same time, there are no large tributaries in the investigated area. For this reason, the sewage hardly changes its composition and the uniformity of ecological parameters throughout the site (region 1). After sampling point 2.8, there is a mixture of pure natural waters from the Khibin foothills with industrial wastewater. Also, a large volume of water flows of meltwater enters the river system. Thus, in region 2, there is a general decrease in the concentration of polluting components due to dilution of sewage waters of the mining enterprise with clean natural waters, and because of the absence of inflows downstream this area is defined as ecologically homogeneous.

3.1. Finalization of the Methodology and Automation of Calculations. Numerous calculations and a large amount of input data have revealed the necessity to develop a software to solve the task, despite the good results of using the methodology. It was decided to replace the segment of the graphic finding of the parameters of the equation of the empirical and theoretical correlation functions by the construction of approximating dependencies. In the future, the approximation parameters were found from the condition of a minimum of the total quadratic error (least squares

method) in order to fully automate the whole process of calculations from the introduction of the initial data to obtaining a response about the homogeneity of the characteristics studied. Thus, the parameters of the equation y = ax + b form were found by the formulas

$$a = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$
(8)  
$$b = \frac{\sum_{i=1}^{n} y_i - a \sum_{i=1}^{n} x_i}{n},$$

where *n* is the number of terms in the series,  $x_i$  is the distance between the observation points, and  $y_i$  is the pairwise correlation coefficient. Further calculations were made according to the above algorithm.

The result was the implementation of a mathematical computation algorithm in the system of object-oriented programming C # (Figure 4). Simplicity of use, full compatibility with Windows and all office applications, loading of initial data from MS Excel, and being undemanding to a certain format of initial data make the developed software solution a convenient tool for the user.

The main result is that the software allow any ecologist to quickly assess the scale and spatial localization of pollution at the stage of the initial analysis of the environmental situation.

#### 4. Conclusion

The process of constructing a model of the spatial structure of various natural systems is quite complex and requires the joint consideration of a large number of very diverse factors. This heterogeneity itself has both a thematic and a spatial nature. The spatial heterogeneity of information is expressed in the fact that statistical and descriptive data are often correlated with different spatial objects that differ in nature and in scale, which creates additional difficulties in the joint processing and analysis of information. Therefore, in problems of this kind, the role of coordinate data binding is great, without which spatial analysis does not make sense. Pollution zones are geographically related to sources of environmental hazard. The strength of the hazardous effect and possible damage depend on the proximity of the risk element to the source of contamination, and the risk depends on the frequency of the dangerous manifestations. Thus, when allocating zones of adverse impact, the use of a geographic coordinate space is necessary to assess the area and intensity of environmental damage. That was done in our work at the first stage of research. Moreover, since the basic sample map showed the possible presence of a spatial trend in the data, this fact was verified by statistical methods. For this purpose, the relationship between the values of the investigated variable and coordinates in a two-dimensional space is distinguished using various indicators of the correlation relationship. The result of the work is the development and successful use of a certain mathematical algorithm with its further software solution for estimating the uniformity of spatially distributed data. Creation of the information model of the investigated territory is reflecting the spatial structure and location of the zones of environmental pollution.

#### **Data Availability**

The data used to support the findings of this study are available from the corresponding author upon request.

#### **Conflicts of Interest**

The authors declare that they have no conflicts of interest.

#### References

- M. Crowly J, "Canadian Geographer," *Biogeography*, vol. 11, no. 4, pp. 312–326, 1967.
- [2] J. M. Omernik and R. G. Bailey, "Distinguishing between watersheds and ecoregions," *JAWRA Journal of the American Water Resources Association*, vol. 33, no. 5, pp. 935–949, 1997.
- [3] J. M. Omernik, "Ecoregions of the Conterminous United States," Annals of the Association of American Geographers, vol. 77, no. 1, pp. 118–125, 1987.
- [4] I. Brewer, The conceptual development and use of ecoregion classifications. Masters Thesis [Master, thesis], Oregon State University, Corvallis, 1999.
- [5] G. D. Tasker and J. R. Stedinger, "Regional skew with weighted ls regression," *Journal of Water Resources Planning and Management*, vol. 112, no. 2, pp. 225–237, 1986.
- [6] M. C. Acreman and C. D. Sinclair, "Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland," *Journal of Hydrology*, vol. 84, no. 3-4, pp. 365–380, 1986.
- [7] R. H. Jongman, C. J. Ter Braak, and O. F. van Tongeren, *Data Analysis in Community and Landscape Ecology*, Cambridge University Press, Cambridge, 1995.

- [8] I. Pivovarova and A. Makhovikov, "Statistical methods of ecological zoning," *Research Journal of Applied Sciences*, vol. 11, no. 6, pp. 321–326, 2016.
- [9] I. Pivovarova and A. Makhovikov, "Ecological regionalization methods of oil producing areas," *Journal of Ecological Engineering*, vol. 18, no. 1, pp. 35–42, 2017.
- [10] M. A. Pashkevich and T. A. Petrova, "Assessment of Widespread air Pollution in the Megacity Using Geographic Information Systems," *Zapiski Gornogo instituta*, vol. 228, pp. 738–742, 2017.
- [11] A. E. Isakov and V. A. Matveeva, "OAO «Kovdorsky MCC» manganese-containing waste water purification study," *Obo-gashchenie Rud*, no. 2, pp. 44–48, 2016.
- [12] R. Rozpondek, K. Wancisiewicz, and M. Kacprzak, "GIS in the studies of soil and water environment," *Journal of Ecological Engineering*, vol. 17, no. 3, pp. 134–142, 2016.
- [13] G. Y. Lu and D. W. Wong, "An adaptive inverse-distance weighting spatial interpolation technique," *Computers & Geosciences*, vol. 34, no. 9, pp. 1044–1055, 2008.
- [14] D. H. Parks, G. W. Tyson, P. Hugenholtz, and R. G. Beiko, "STAMP: statistical analysis of taxonomic and functional profiles," *Bioinformatics*, vol. 30, no. 21, pp. 3123-3124, 2014.
- [15] R. A. Fisher, "On a distribution yielding the error functions of several well known statistics," *Toronto*, vol. 2, p. 805, 1928.
- [16] H. Cramer, Mathematical Methods of Statistic, University of Stockholm, 1946, p. 648.
- [17] B. L. Welch, "The Significance of the Difference Between Two Means when the Population Variances are Unequal," *Biometrika*, vol. 29, no. 3/4, p. 350, 1938.
- [18] L. N. Bolshev and N. V. Smirnov, Tables of Mathematical Statistics, Nauka publ., Moscow, Russia, 3ed edition, 1983, p. 474.
- [19] E. L. Lehmann, "Consistency and unbiasedness of certain nonparametric tests," *Annals of Mathematical Statistics*, vol. 22, pp. 165–179, 1951.
- [20] M. Rosenblatt, "Limit theorems associated with variants of the von Mises statistic," *Annals of Mathematical Statistics*, vol. 23, pp. 617–623, 1952.
- [21] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals* of *Mathematical Statistics*, vol. 18, pp. 50–60, 1947.
- [22] M. Hollender and D. Wolf, *Methods of nonparametric statistics*, Finance and Statistics Publ, Moscow, Russia, 1983, p. 518.
- [23] Bartel van der Waerden, *Mathematische statistik*, Springer-Verlag, Berlin, Germany, 1957, p. 436.
- [24] A. I. Orlov, "Key stages of statistical methods development Scientific," *Scientific Journal of KubSAU*, vol. 97, pp. 73–85, 2014.
- [25] A. E. Isakov and M. A. Chukaeva, "Ecological and geochemical peculiarities of surface water transformation in the area of the enterprise JSC "Apatit" impact," *International Journal of Ecology* and Development, vol. 31, no. 2, pp. 90–98, 2016.

















International Journal of Biodiversity



Submit your manuscripts at www.hindawi.com









Applied & Environmental Soil Science





Forestry Research







Journal of Marine Biology



Advances in Meteorology



Advances in Agriculture





Journal of Geological Research