GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 11 • Number 3 • 2017
http://dx.doi.org/10.7494/geom.2017.11.3.49

Tomasz Gonet*, Kamil Gonet**

Alternative Approach to Evaluating Interpolation Methods of Small and Imbalanced Data Sets

Abstract: The research concerns an alternative approach to the evaluation of interpolation methods for mapping small and imbalanced data sets. A basic statistical analysis of the standard cross-validation procedure is not always conclusive.
In the case of the investigated data set (which is inconsistent with normal distribution), three interpolation methods have been selected as the most reliable (according to standard cross-validation). However, maps resulting from the aforementioned methods clearly differ from each other. This is the reason why a comprehensive statistical analysis of the studied data is a necessity.
We propose an alternative approach that evaluates a broadened scope of parameters describing the data distribution. The general idea of the methodology is to compare not only the standard deviation of the estimator but also three additional parameters to make the final assessment much more accurate.
The analysis has been carried out with the use of Golden Software Surfer. It provides a wide range of interpolation methods and numerous adjustable parameters.

Keywords: interpolation, Surfer, cross-validation, small data set

* University of Warsaw, Faculty of Geology, Warsaw, Poland
** AGH University of Science and Technology, Faculty of Drilling, Oil and Gas, Krakow, Poland

1. Introduction

The research concerns the evaluation of the influence of the Stanisław Siedlecki Polish Polar Station (Hornsund, Svalbard – Fig. 1) on the environment with the use of magnetic methods.

The study area is located in the southern part of Spitsbergen and covers approximately 17.5 ha (500 × 350 meters). There is mostly tundra vegetation growing on initial lithosols and frost-deformed regosols [6].

It is essential to emphasize that this area is placed within a territory of South Spitsbergen National Park and has been identified as an Important Bird Area [1]. Hence, legal regulations have strongly limited the number of samples, so only 73 topsoil specimens were collected from the vicinity of the station (Fig. 2).

Consequently, a grave problem arose; i.e., how to visualize, interpolate, and finally interpret such a small data set? In this paper, we propose a methodology for coping with the evaluation of interpolation methods for mapping small and imbalanced data sets.

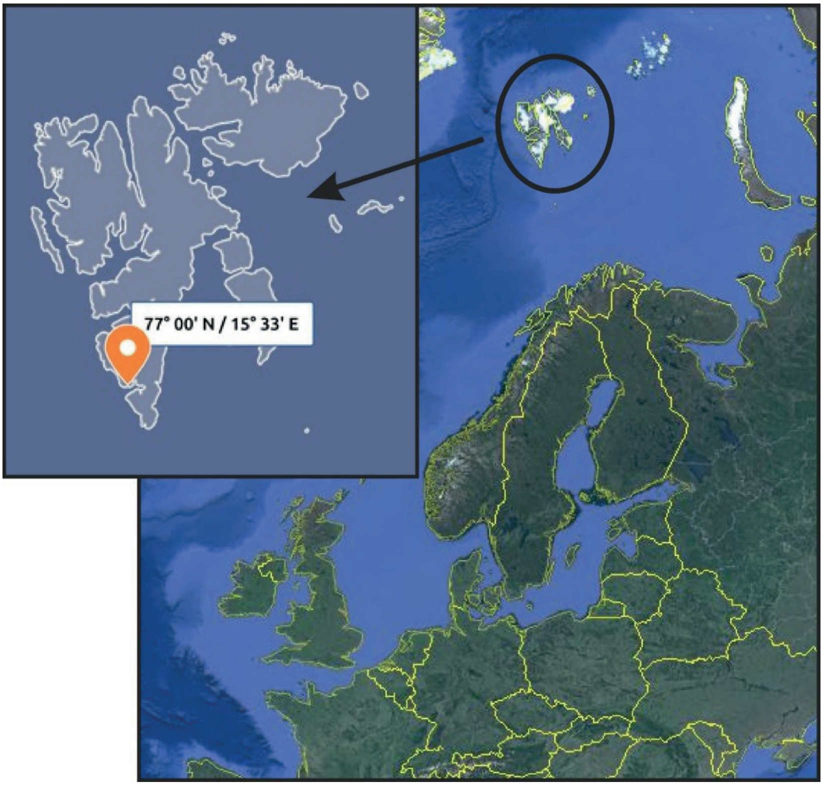

Fig. 1. Position of Stanisław Siedlecki Polish Polar Station

Source: [5, 11]

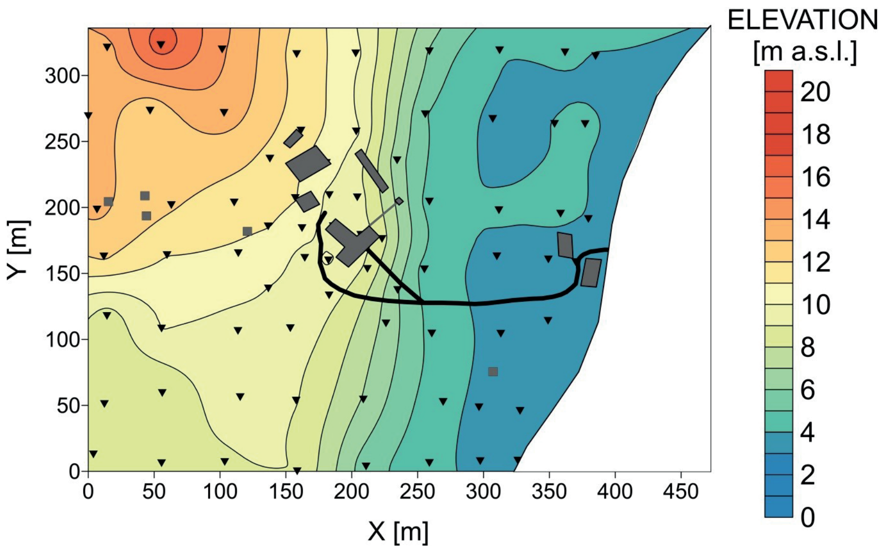

Fig. 2. Hypsometric map of study area (grey polygons – buildings of PPS; black triangles – sampling sites)

2. Magnetic Susceptibility, Descriptive Statistics, and Chi-Squared Test

For each sample, magnetic susceptibility was measured using Multi-Function Kappabridge (AGICO, Czech Republic). The susceptibility values were normalized by a mass unit (kg), and the specific magnetic susceptibility (χ) was calculated. This parameter corresponds to the magnetic mineral content that is usually accompanied by heavy metals and polycyclic aromatic hydrocarbons [8], [13]. Thus, the area where magnetic susceptibility is enhanced can be considered as having been influenced by the activity of the Polish Polar Station (PPS).

After our measurements, a basic statistical analysis was carried out. Figure 3 presents a decile plot and histogram. Deciles are values that divide the sorted data into ten equal parts; e.g., the 3rd decile is the value below which there are 30% of the values [12]. Consequently, the 5th decile is equal to the median value.

Magnetic susceptibility values vary in wide extremes; i.e., from $7.98 \cdot 10^{-8}$ m$^3 \cdot$kg^{-1} to $138.86 \cdot 10^{-8}$ m$^3 \cdot$kg^{-1}. The mean is $19.43 \cdot 10^{-8}$ m$^3 \cdot$kg^{-1} (dashed line in Figure 4a), and the median is $13.20 \cdot 10^{-8}$ m$^3 \cdot$kg^{-1}. The mean value is higher than the 75th percentile, which means that more than 75% of the samples feature magnetic susceptibility lower than the mean. The high level of discrepancy between the mean and median values indicates the strong asymmetry of the data distribution. This fact is also proven by the high positive value of skewness (i.e., equalling 4.24). Moreover, the standard deviation is very high – $18.68 \cdot 10^{-8}$ m$^3 \cdot$kg^{-1} (over 96% of the mean). The data distribution is leptokurtic – kurtosis equals 23.35 [7].

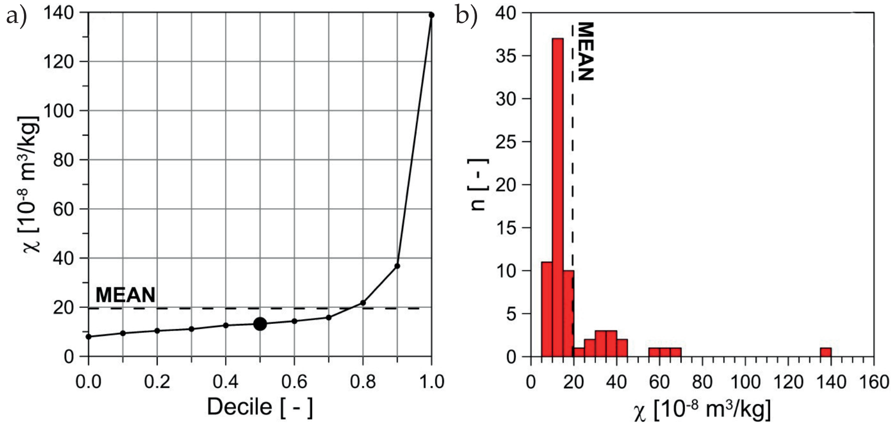

Fig. 3. Decile plot (a) and histogram (b) of studied data

Subsequently, we carried out the chi-squared test with a significance level of 5%. It was proven that investigated data is not consistent with normal distribution [15].

This kind of data requires exceptional caution during both the interpolation and interpretation processes. Furthermore, the literature does not recommend one particular interpolation method in such cases in order to create a reliable map of the studied variable. Hence, we have decided to examine all of the interpolation methods available in Golden Software Surfer.

3. Interpolation Simulations Using Surfer

Golden Software Surfer provides a wide range of interpolation methods with numerous adjustable parameters [3]:

1. *Inverse Distance to a Power* is a weighted average interpolator. Weights of the data are determined in accordance with the inverse of the distance from a grid node. Moreover, a weighting power can be adjusted. This parameter controls how quickly weighting factors drop with distance from a grid node. If the weighting power is high, the influence of the points far from the grid node is insignificant. Thus, this method can be either an exact or smoothing interpolator. In our case, power values of 1, 2, 3, and 4 were used.
2. *Modified Shepard's Method* is similar to the previous one, and it uses the same type of interpolator. Nonetheless, it eliminates (or at least reduces) the bull's-eye effect by using the local least-squares method.
3. *Kriging* is a geostatistical gridding method that is most-frequently recommended by the literature [3, 4, 10, 14]. This method is based on the theory of regionalized variables. Variograms, anisotropy, and other options help to find trends in the analyzed data.

4. *Minimum Curvature* is one of the smoothing interpolators. It helps to generate a smooth surface that honors the data as closely as possible.
5. *The Natural Neighbor* method is based on Delaunay triangulation, where weights are proportional to the areas of particular Voronoi cells.
6. *Nearest Neighbor* is a simple interpolation method assigning the value of the nearest point to each grid node. This method is useful in cases when the data is already evenly spaced.
7. *Polynomial Regression* defines large-scale trends and patterns in the analyzed data using various types of polynomials.
8. *Radial Basis Function* is a set of exact interpolation methods. The basis kernel functions are analogous to the variograms in *Kriging*. Additionally, a smoothing factor can be introduced in order to produce a smoother surface.
9. *The Triangulation with Linear Interpolation* method uses optimal Delaunay triangulation. In result, no triangle edges are intersected by another triangle. Each triangle determines a plane over the grid nodes lying within the triangle, with the tilt and elevation of the triangle defined by the three original data points.
10. *Local Polynomial* assigns grid node values with the use of a weighted least-squares fit within an adjustable search ellipse. Local polynomials can be of degrees 1, 2, or 3.
11. *The Moving Average* interpolation method is based on averaging the data within a user-defined search ellipse for each grid node [3].

Altogether, a series of 83 simulations was carried out using aforementioned 11 interpolation methods (Tab. 1). Each simulation had different values of parameters characteristic for the particular method.

Table 1. Number of simulations conducted using particular interpolation method

Interpolation method	Number of simulations
Inverse Distance to a Power	8
Minimum Curvature	7
Kriging	6
Local Polynomial	14
Modified Shepard's Method	14
Moving Average	9
Natural Neighbor	1
Nearest Neighbor	4
Polynomial Regression	4
Radial Basis Function	15
Triangulation with Linear Interpolation	1
Total	83

4. Preliminary Interpolation Assessment – Standard Cross-Validation

In order to be sure that the presentation of the data in the form of a map is reliable, an evaluation of the quality of the interpolation was needed. For this purpose, standard cross-validation is routinely used. During this procedure, Surfer removes one point from the data set; then, using the remaining data and a particular algorithm, it interpolates a new value at this point – it is called the *Estimated* value (E_i).

The difference between the *Estimated* value and the measured one (M_i) at the particular point is called the *Residual* (R_i) (Equation (1)) [3, 12].

$$R_i = E_i - M_i \tag{1}$$

Afterwards, the program repeats the process for the rest of the points from the data set. After calculating the *Residual* for each point, various types of statistical analysis could be carried out and used as a quantitative objective proxy for the interpolation method quality [2]. The literature [3] suggests focusing on a comparison between the standard deviation and standard error of the residuals during the selection of the best method.

The standard error is defined by the following equation:

$$D_R = \frac{S_R}{\sqrt{n}} \tag{2}$$

where:

S_R – standard deviation of the residuals,
n – sample size.

Being aware of the proportionality between the standard error (D_R) and standard deviation (S_R); we analyzed only the latter parameter.

The results of the cross-validation procedure for the studied data set are shown in Table 2. The last column in Table 2 contains R_E – calculated as the relative difference between the analyzed parameter value for a particular interpolation method and the parameter value for the best one. The conducted analysis has revealed that *Minimum Curvature* is the best interpolation method (having the smallest value of standard deviation). The second is *Polynomial Regression*; the third – *Local Polynomial*; and the fourth – *Radial Basis Function.*

Despite the fact that there is only a 9.6% difference between the first four methods (in compliance with the cross-validation algorithm), their graphical representations are significantly disparate (Fig. 4). What is more, these kinds of discrepancies could be observed in almost half of the utilized methods (Figs. 5, 6, Tabs. 3–5).

Table 2. Results of cross-validation procedure for studied data

Rank	Interpolation method	S_R	R_E [%]
1	Minimum Curvature	15.31	0.0
2	Polynomial Regression	16.40	7.1
3	Local Polynomial	16.43	7.3
4	Radial Basis Function	16.78	9.6
5	Inverse Distance to a Power	16.86	10.1
6	Modified Shepard's Method	17.07	11.5
7	Moving Average	17.25	12.7
8	Kriging	18.13	18.4
9	Nearest Neighbor	19.53	27.6
10	Natural Neighbor	19.66	28.4
11	Triangulation with Linear Interpolation	21.19	38.4

Table 3. Results of second parameter (Eq. 3) for studied data

Rank	Interpolation method	R_R	R_E [%]
1	Minimum Curvature	0.35	0.0
2	Local Polynomial	0.38	8.6
3	Radial Basis Function	0.41	17.1
4	Kriging	0.43	22.9
5	Natural Neighbor	0.47	34.3
6	Nearest Neighbor	0.47	34.3
7	Moving Average	0.51	45.7
8	Modified Shepard's Method	0.53	51.4
9	Triangulation with Linear Interpolation	0.54	54.3
10	Inverse Distance to a Power	0.55	57.1
11	Polynomial Regression	0.55	57.1

Table 4. Results of third parameter (Eq. 4) for studied data

Rank	Interpolation method	Δ_M	R_E [%]
1	Polynomial Regression	0.00	–
2	Kriging	0.14	0.0
3	Radial Basis Function	0.21	50.0
4	Modified Shepard's Method	0.40	185.7
5	Nearest Neighbor	0.43	207.1
6	Local Polynomial	0.60	328.6
7	Minimum Curvature	0.75	435.7
8	Moving Average	1.15	721.4
9	Natural Neighbor	1.56	1014.3
10	Inverse Distance to a Power	2.05	1364.3
11	Triangulation with Linear Interpolation	2.42	1628.6

Table 5. Results of fourth parameter (Eq. 5) for studied data

Rank	Interpolation method	Δ_S	R_E [%]
1	Triangulation with Linear Interpolation	2.62	0.0
2	Nearest Neighbor	3.80	45.0
3	Minimum Curvature	4.81	83.6
4	Natural Neighbor	4.90	87.0
5	Kriging	5.10	94.7
6	Modified Shepard's Method	7.02	167.9
7	Polynomial Regression	9.75	272.1
8	Moving Average	10.13	286.6
9	Local Polynomial	10.29	292.8
10	Radial Basis Function	11.34	332.8
11	Inverse Distance to a Power	12.08	361.1

It is important to emphasize that each interpolation method has slightly different characteristics. For instance, *Moving Average* is most-reliable in cases with a large volume of data; *Local Polynomial* is most-applicable to data sets that are locally smooth; and *Polynomial Regression* shows only the underlying large-scale trends and patterns [9]. Hence, some of the methods are not suitable for interpolation of the studied data. However, we have considered all of the available methods, because the proposed methodology should be as universal as possible and applicable to each data set.

The other important issue is the use of the *Kriging* method. It is often recommended by the literature [2, 3, 14] as one of the most-flexible and most-useful method. Nevertheless, it only achieved eighth place in the case of the investigated data set (Tab. 2). The reason is that *Kriging* is most-applicable to data that is consistent with normal distribution. Otherwise, there is no clear suggestion on which interpolation method is most-reliable.

After critical visual assessment by the researchers, it becomes quite obvious that only *Minimum Curvature* is reliable among the best three methods (Tab. 1, Fig. 4a–c). However, such an evaluation would not be clear if the other methods were taken into consideration. For instance, the choice between *Minimum Curvature* and *Radial Basis Function* seems to be impossible without any statistical analysis (Fig. 4a, d).

Furthermore, the conducted simulations have proven that neither the standard cross-validation procedure nor the subjective assessment of the researcher may be reliable in assessing the quality of interpolation, especially in the case of data that is not consistent with normal distribution.

The literature recommends the Univariate and Willmott statistical methods in such situations [16]. They are partially based on a linear regression analysis and absolute difference measures. However, Willmott [16] emphasizes that the above-mentioned parameters can be misleading.

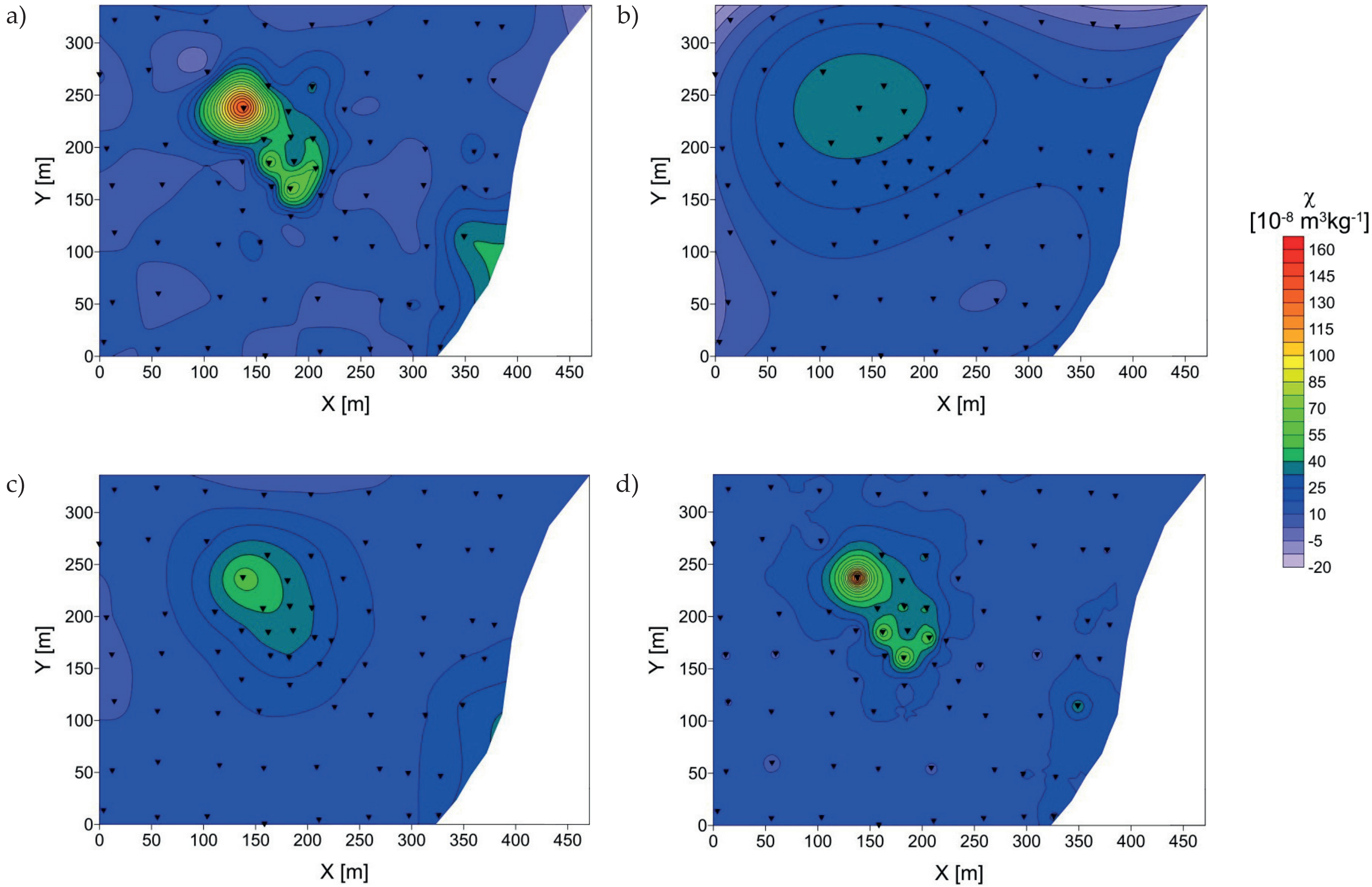

Fig. 4. Maps created using: Minimum Curvature (a); Polynomial Regression (b); Local Polynomial (c); and Radial Basis Function (d)

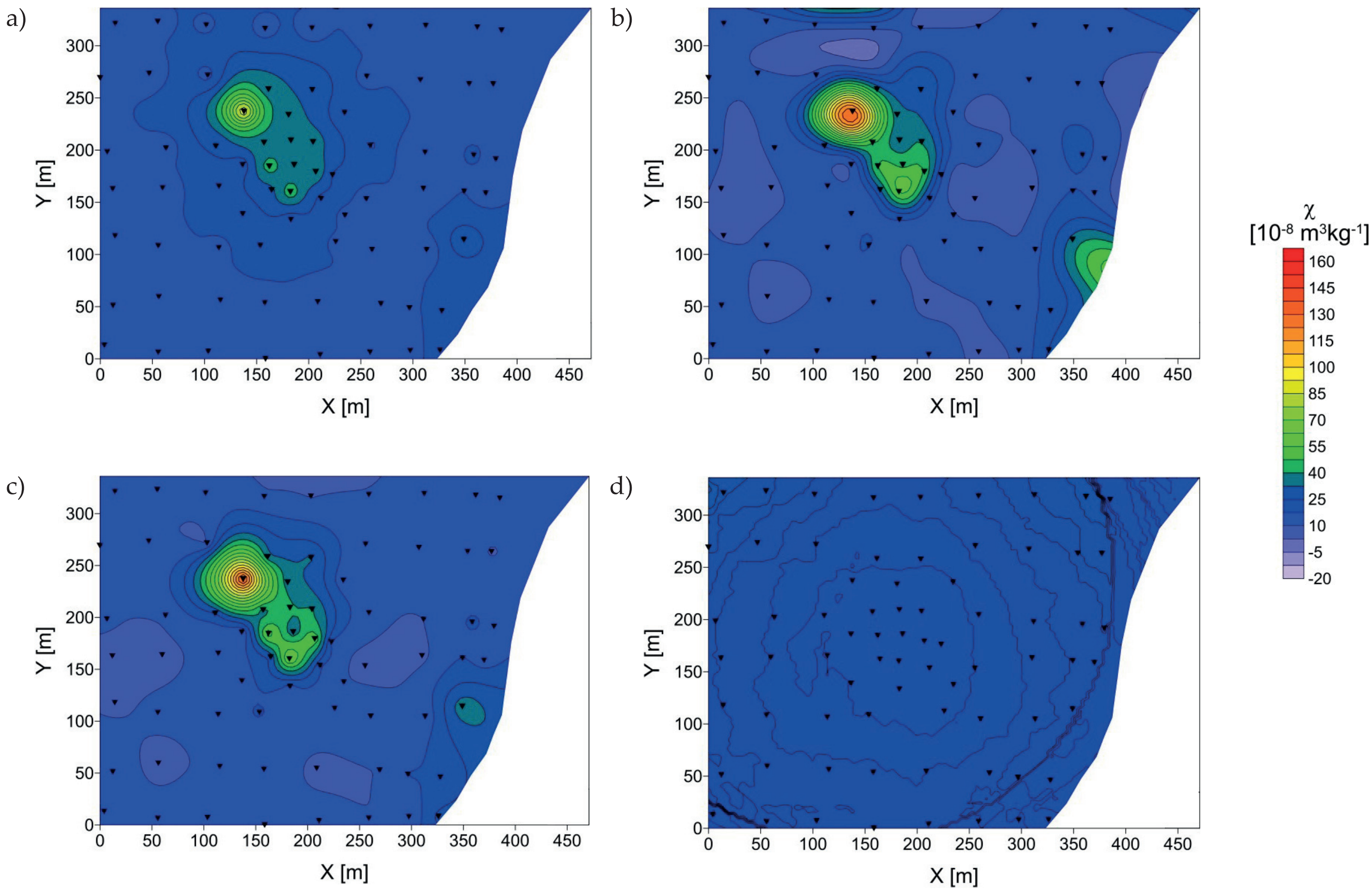

Fig. 5. Maps created using: Inverse Distance to a Power (a); Modified Shepard's Method (b); Kriging (c); and Moving Average (d)

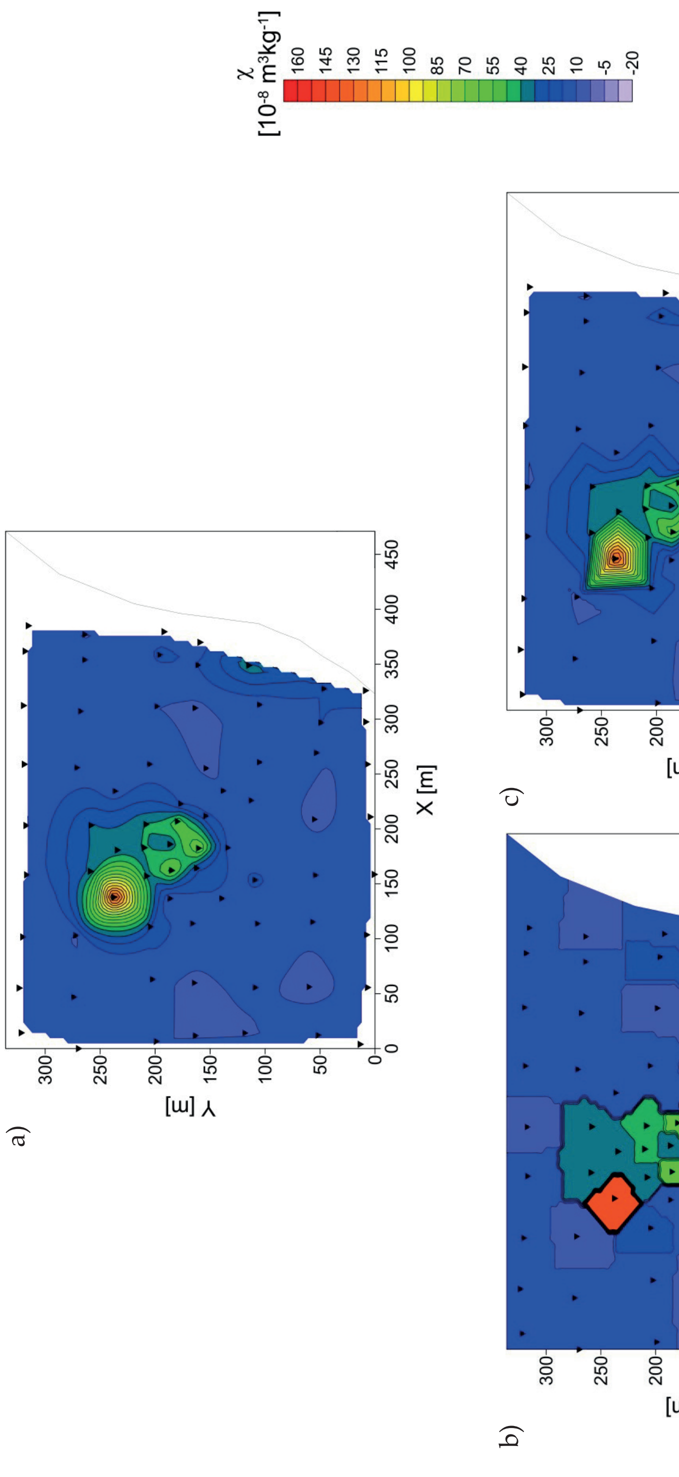

Fig. 6. Maps created using: Natural Neighbor (a); Nearest Neighbor (b); and Triangulation with Linear Interpolation (c)

The data set analyzed in this paper contains a number of outliers (values even several times higher than the mean; cf. Fig. 3); consequently, it is inconsistent with normal distribution. The aforementioned facts increase the risk of an unreliable evaluation of interpolation quality; this is the reason why we propose an alternative approach that evaluates a broadened scope of parameters describing the data distribution.

5. Alternative Approach – Error Index

The general idea of the proposed methodology is to compare not only the standard deviation of the estimator (Step 1) but also three additional parameters in order to make a final assessment much more accurate. Besides the standard deviation of the *Residuals*, our analysis shall consider:

- Mean of absolute relative residual (Step 2)

$$R_R = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{R_i}{M_i}\right| \tag{3}$$

- Mean residual (Step 3)

$$\Delta_M = \left|\sum_{i=1}^{n}\frac{R_i}{n}\right| \tag{4}$$

- Difference between standard deviations of the measured and estimated values (Step 4)

$$\Delta_S = \left|S_M - S_E\right| \tag{5}$$

where:

R_i – residual of the particular data point,
E_i – estimated value of the particular data point,
M_i – measured value of the particular data point,
n – number of points,
S_M – standard deviation of the measured values,
S_E – standard deviation of the estimated values.

It is important to emphasize that our goal is not to assess each parameter separately but to draw one final conclusion from all of them, making the assessment more accurate and thorough.

After calculating all of the parameters, we have normalized them from 1 to 10, where 1 is the best mark and 10 is the worst. The sum of these normalized values for each method gives a new parameter, which we call the *Error Index*. Therefore, the best interpolation method is the one with the lowest value of *Error Index*.

6. Interpolation Reassessment

The results of cross-validation (Step 1) are presented in Table 2. The procedure has revealed that *Minimum Curvature* is the most-reliable interpolation method (the smallest value of standard deviation – S_R). The second one is *Polynomial Regression*, and the third is *Local Polynomial*. Both methods have similar relative errors (R_E) – 7.1% and 7.3%, respectively. The final method (considered the worst) is *Triangulation with Linear Interpolation*, with a relative error of 38.4%.

The second parameter (Equation (3)) calculated for each interpolation method (Step 2) is shown in Table 5. It has indicated the same conclusion as in Step 1 – the *Minimum Curvature* method is the best. Second place was taken by *Local Polynomial*, followed by *Radial Basis Function* (with relative errors of 8.6% and 17.1%, respectively). On the basis of the second parameter, the worst method is *Polynomial Regression* (with relative error of 57.1%), which has achieved second place according to Step 1 (Tab. 2). All values of the relative error are noticeably higher than in the previous step.

Calculations during Step 3 have revealed that the best interpolation method is *Polynomial Regression*, followed by *Kriging* (R_E = 0.0%) and *Radial Basis Function* (R_E = 50.0%) (Tab. 4). Values of relative error were calculated in relation to the Δ_M for *Kriging* (which is in second place). The value of Δ_M for *Polynomial Regression* is infinitesimally close to 0; consequently, it was impractical to relate the errors to that value. Such a low Δ_M is the result of the nature of the *Polynomial Regression* algorithm – it is constructed in such a way that it minimizes the difference between the means of the *Measured* and *Estimated* values. The last method is *Triangulation with Linear Interpolation*, with a relative error of 1628.6%. In Step 3, there is a substantial increase in the R_E values as compared to Steps 1 and 2.

Surprisingly, Step 4 has indicated that the most-reliable interpolation method is *Triangulation with Linear Interpolation*, which had substantially worse results in the previous steps (cf. Tabs. 2–4). Second place is taken by *Nearest Neighbor*, followed by *Minimum Curvature* (with relative errors of 45.0% and 83.6%, respectively – Table 5). Tenth place was taken by *Radial Basis Function*, which was noticeably higher in Steps 1–3 (cf. Tabs. 2–4). The worst method is *Inverse Distance to a Power*, with relative error of 361.1%.

In both Steps 3 and 4, the relative error values are considerably higher than in Steps 1 and 2. This is the reason why normalization of the obtained values is a necessity. Therefore, *Error Index* – calculated as the sum of four obtained parameters – shall have an equal contribution from each step (Tab. 6, Fig. 7). The values of *Error Index* have indicated that *Minimum Curvature* is the best interpolation method (the lowest value of E_I in Table 6) for the investigated data set. The map created using the *Minimum Curvature* algorithm is presented in Figure 4a. Second place was taken by *Kriging*, followed by *Local Polynomial*. It is noticeable that the difference between the best *Minimum Curvature* and the other methods is now much more evident (68.2%) than in the case of standard cross-validation (Fig. 7).

Table 6. Final results of all normalized parameters and *Error Index*

Rank	Interpolation method	S_R	R_R	Δ_M	Δ_S	E_I	R_E [%]
1	Minimum Curvature	1.00	1.00	3.80	3.08	8.88	0.0
2	Kriging	5.31	4.77	1.51	3.36	14.94	68.2
3	Local Polynomial	2.72	2.51	3.24	8.30	16.77	88.9
4	Radial Basis Function	3.25	3.91	1.79	9.30	18.25	105.5
5	Nearest Neighbor	7.47	6.63	2.60	2.12	18.82	111.9
6	Modified Shepard's Method	3.70	8.96	2.48	5.18	20.33	128.9
7	Polynomial Regression	2.67	10.00	1.00	7.78	21.45	141.6
8	Natural Neighbor	7.66	6.43	6.82	3.17	24.09	171.3
9	Moving Average	3.97	8.35	5.30	8.15	25.77	190.2
10	Triangulation with Linear Interpolation	10.00	9.68	10.00	1.00	30.68	245.5
11	Inverse Distance to a Power	3.37	9.99	8.65	10.00	32.01	260.5

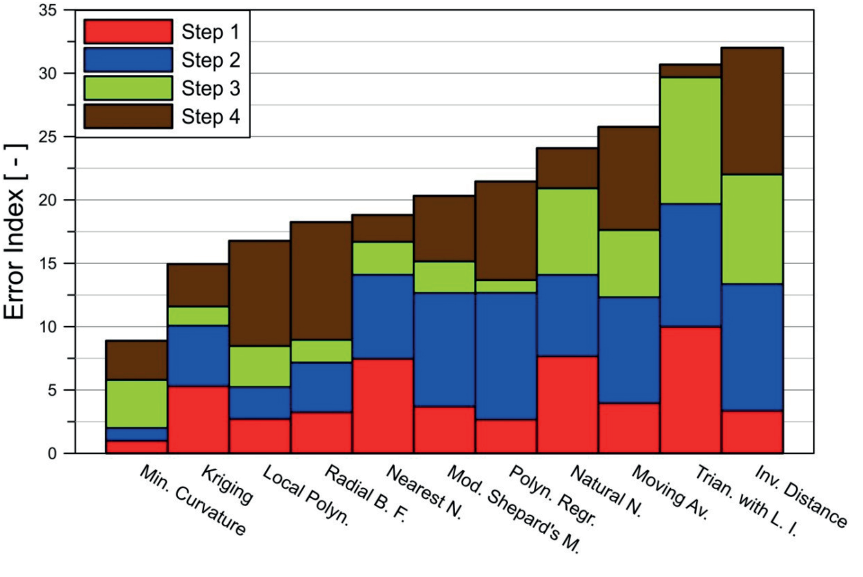

Fig. 7. Bar chart of *Error Index*

Admittedly, the conclusion is the same as that according to standard cross-validation, but it will not always be like this (especially in the cases of small and imbalanced data sets).

The conducted research has proven that, in the cases of small and imbalanced data sets, selection of the best interpolation method could be problematic.

Firstly, the researcher should be highly aware of the basics of each method. Each algorithm is constructed for specific sets of data; for instance, *Moving Average* is the most-applicable to large and very large data sets (more than 1000 data points); *Nearest Neighbor* is the most-useful when it comes to regularly (or almost regularly) spaced data points; methods based on polynomials are suitable for defining local/global trends and patterns in data; *Kriging* is based on the theory of regionalized variables, which works best with data consistent with normal distribution, etc. [2, 3, 9, 14].

Secondly, there are some methods that result in relatively similar maps (e.g., *Minimum Curvature* and *Radial Basis Function*; Fig. 4a, c) and, simultaneously, have considerably different statistical results. A basic statistical analysis of the standard cross-validation procedure used during the evaluation of the gridding method quality is not always conclusive, especially in the cases of small and imbalanced data sets.

Finally, as shown in the conducted analysis, nearly every single method takes a different place depending on which statistical parameter has been taken into consideration in a particular step. Hence, it is vital to remember that the statistics are not related to the physical basis of the investigated phenomena; therefore, the researcher's reliable assessment of the final map is crucial to the proper interpretation of the studied variable.

7. Conclusions

– The final results have revealed that *Minimum Curvature* is the best interpolation method for the investigated data set.
– Small and imbalanced data sets require careful consideration during interpolation method selection, because some algorithms might lead to unrealistic distortions and, consequently, a misinterpretation of the studied data.
– Both the researcher's knowledge about the physical basis of the studied parameter and a critical assessment of the final map are essential to the reliable interpretation of the investigated variable.
– Standard cross-validation analysis is not always conclusive. The use of *Error Index* allows us to evaluate a broadened scope of parameters describing the data distribution, which makes the final assessment more accurate and thorough.
– The proposed approach has been applied to only one data set. Hence, in order to assess its functionality, a further analysis of different data sets is strongly recommended.

References

[1] Birdlife International, [on-line:] http://www.birdlife.org (access: October 29, 2016).

[2] Davis J.C.: *Statistics and Data Analysis in Geology*. John Wiley and Sons, Canada 1973.

[3] Galon Z.: *Surfer. Podręcznik użytkowania*. Gambit Centrum Oprogramowania i Szkoleń Sp. z o. o., Kraków 2009.

[4] Goldsztejn P., Skrzypek G.: *Wykorzystanie metod interpolacji do numerycznego kreślenia map powierzchni geologicznych na podstawie nieregularnie rozmieszczonych danych.* Przegląd Geologiczny, vol. 52, no. 3, 2004, pp. 233–236.

[5] Google Maps, [online:] https://www.google.pl/maps [access: October 29, 2016].

[6] IUSS Working Group WRB: *World Reference Base for Soil Resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps*. World Soil Resources Reports No. 106. FAO, Rome.

[7] Lane D.M.: *Introduction to Statistics: An Interactive eBook*, [on-line:] https://trello.com/c/4nr4HFQj/80-introduction-to-statistics-an-interactive-e-book-by-david-m-lane [access: October 29, 2016].

[8] Magiera T.: *Wykorzystanie magnetometrii do oceny zanieczyszczenia gleb i osadów jeziornych*. IPIŚ PAN, Zabrze 2004.

[9] Magnuszewski A.: *GIS w geografii fizycznej*. Wydawnictwo Naukowe PWN, Warszawa 2012.

[10] Oliver M.A., Webster R.: *Kriging: a method of interpolation for geographical information systems.* International Journal of Geographical Information Systems, vol. 4, issue 3, 1990, pp. 313–332.

[11] Polska Stacja Polarna Hornsund im. Stanisława Siedleckiego, [online:] http://hornsund.igf.edu.pl/ [access: October 29, 2016].

[12] Sheskin D.J.: *Handbook of parametric and nonparametric statistical procedures.* 3rd ed. Chapman and Hall/CRC, London, New York, Washington, DC. 2004.

[13] Strzyszcz Z., Rachwał M.: *Zastosowanie magnetometrii do monitoringu i oceny ekologicznej gleb na obszarach objętych wpływem emisji przemysłowych*. IPIŚ PAN, Zabrze 2010.

[14] *Surfer. User's Guide. Contouring and 3D Surface Mapping for Scientists and Engineers*. Golden Software, 1993–2002.

[15] Taylor J.R.: *Wstęp do analizy błędu pomiarowego*. Wydawnictwo Naukowe PWN, Warszawa 1999.

[16] Willmott C.J.: *On the evaluation of model performance in physical geography*. [in:] Gaile G.L., Willmott C.J. (eds.), *Spatial Statistics and Models*, Springer Science+Business Media, Dordrecht, 1984, pp. 443–460.

Nowe podejście do oceny jakości metod interpolacji niewielkich i zróżnicowanych zestawów danych

Streszczenie: Badania dotyczą alternatywnego podejścia do oceny jakości metod interpolacji niewielkich i zróżnicowanych zestawów danych. Podstawowa analiza statystyczna oparta na klasycznej walidacji krzyżowej nie zawsze daje jednoznaczne wnioski.

W przypadku analizowanego zestawu danych (niezgodnego z rozkładem normalnym) trzy metody interpolacji zostały wybrane jako najlepsze (zgodnie z procedurą klasycznej walidacji krzyżowej). Niemniej jednak mapy powstałe na podstawie tych metod wyraźnie się od siebie różnią. To jest powód, dla którego dogłębna analiza statystyczna była konieczna.

Zaproponowano alternatywne podejście do tego zagadnienia, które uwzględnia szersze spektrum parametrów opisujących badany zestaw danych. Głównym założeniem tej metodyki jest porównanie nie tylko odchylenia standardowego estymatora, ale również trzech dodatkowych parametrów. To powoduje, iż końcowa ocena jest znacznie dokładniejsza.

Analizę wykonano za pomocą programu Surfer (Golden Software). Zapewnia on możliwość wykorzystania wielu metod interpolacji wraz z różnorakimi, regulowanymi parametrami.

Słowa kluczowe: interpolacja, Surfer, walidacja krzyżowa, mały zestaw danych