

Article

Application of Modified Shepard's Method (MSM) case study with interpolation of Neogene reservoirs variables in the Northern Croatia

Tomislav Malvić ^{1a}, Josip Ivšinić ^{2*}, Josipa Velić ^{1b}, Jasenka Sremac ³ and Uroš Barudžija ^{1c}

¹ University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering (^aFull Prof., ^bProf. Emer., ^cAssist. Prof.)

² INA Plc., Zagreb (DSc.)

³ University of Zagreb, Faculty of Science (Full Prof.)

* Correspondence: josip.ivsinovic@ina.hr

Abstract: Interpolation is procedure that depends on spatial and/or statistical properties of analysed variable(s). It is special challenging task for data that included low number of samples, like dataset with less than 20 data. This problem is especially emphasized in the subsurface geological mapping, i.e. in the cases where data are taken solely from wells. Successful solutions of such mapping problems ask for knowledge about interpolation methods designed primarily for small datasets and dataset itself. Here are compared two methods, namely Inverse Distance Weighting and Modified Shepard's Method, applied for three variables (porosity, permeability, thickness) measured in the Neogene sandstone hydrocarbon reservoirs (Northern Croatia). The results showed that pure cross-validation is not enough condition for appropriate map selection, but also geometrical features need to be considered, for datasets with less than 20 points.

Keywords: Modified Shepard's Method (MSM); Inverse Distance Weighting (IDW); sandstone; neogene; Northern Croatia

1. Introduction

Geostatistics is a branch of applied statistics that applies theories of deterministic estimation, then stochastic processes, that is, statistical inference to different spatial phenomena in numerous sciences. The most important geostatistical term applied in geology and petroleum geology is the regionalized variable [1]. It means a measured or estimated magnitude whose behaviour is described in part by a random and partly by a deterministic variable. Each such variable shows the property of spatial continuity and as such is modelled in space. A qualitative description of the property of a regionalized variable leads to a more accurate assessment of the regionalized variable.

The interpolation methods used in the Croatian part of the Pannonian Basin System (CPBS) are described in [2], while the last applied specific interpolation method is presented in [3]. The IDW method is widely used in the CPBS for mapping geological variables, and its application is described in [4-8].

The SMS method has been applied all over the world, and the following papers are selected as appropriate examples. In [9] authors have applied the MSM method to the digital terrain model region of the Chen-Yu-Lan River (Taiwan). Authors in [10] have applied MSM for multicomponent aerosol-cloud parameterization (global climate modelling), and the same method was used for modelling of geological reserve of hydrocarbons calculation [11]. The authors in [12] applied the MSM method on selected parameters of subsurface waters in the area of approximately 10 ha, located in the Ciemięga Valley (Poland). The MSM was used by the authors [13] to determine the boundary of the smooth particle hydrodynamics (SPH) method. In selected area, the IDW and SMS methods were applied in the field "B", Late Miocene (Lower Pontian) reservoir "K" on three geological variables: porosity, permeability and thickness.

2. Geological settings of the analysed field “B” (Sava Depression, Northern Croatia)

The oil and gas field “B” is located in the CPBS, i.e. in the western part of the Sava Depression. The location of the field “B” in the depression is shown in Figure 1.

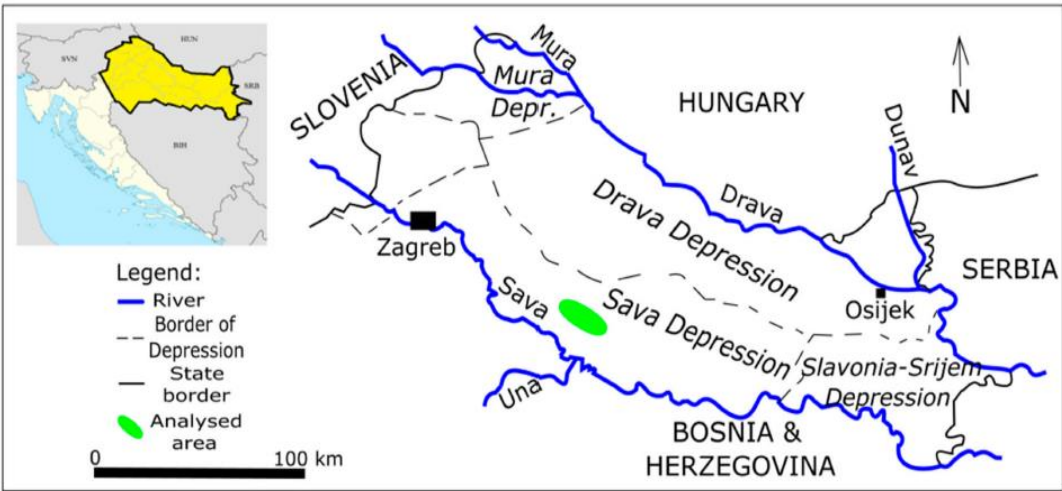


Figure 1. Location of field “B” in Sava Depression (from [3]).

Most of the wells drilled only sediments to the base of the Kloštar Ivanic Formation, since hydrocarbon reservoirs were proven within it. The formation is described as the alteration of marls and well-sorted sandstones. In the oldest part sandstones are hard, which according to younger part become fine-grained, even weakly consolidated sands. Marls are grey to greyish-brown, with medium hardness, whose compaction decreases upwards as well as the proportion of clay component increases. Marls represent impermeable rocks on a top of each sandstone reservoir. Inside the formation marl’s thickness is 30-150 m, while sandstone’s thickness is 20-150 m.

Structurally, tectonic blocks are considered to represent in most cases separate structural traps, where hydrocarbon accumulated in the highest parts. Therefore, these tectonic blocks represent separate hydrodynamic units with their reservoir characteristics. The structural map of the analysed Lower Pontian reservoir "K" is shown Figure 2.

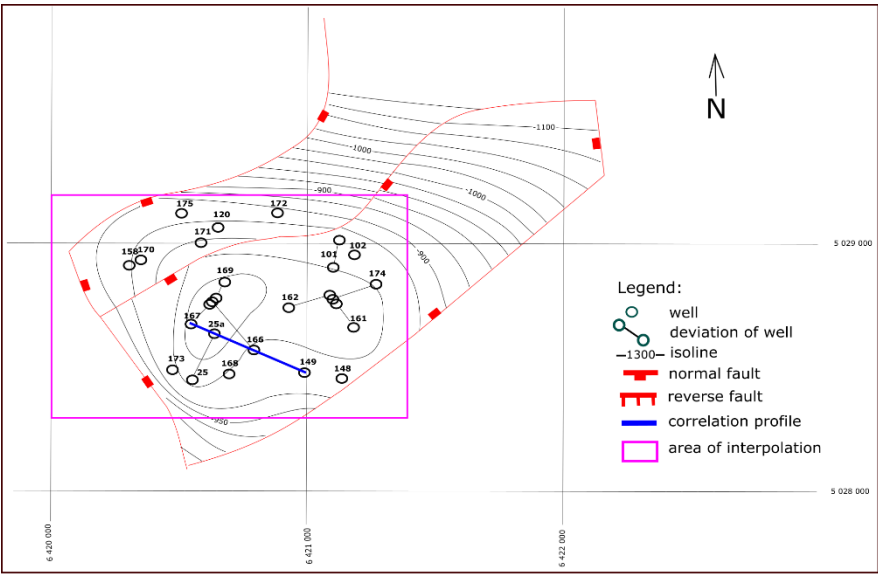


Figure 2. Structural map of the Lower Pontian reservoir "K".

A typical geologic section is derived from the well data, and clearly outline the sandstone member (blue borders) where are discovered hydrocarbon reservoirs (Figure 3). The section also reveals partial structural inversion within the formation, where monocline of deeper members is changed in the syncline in upper part.

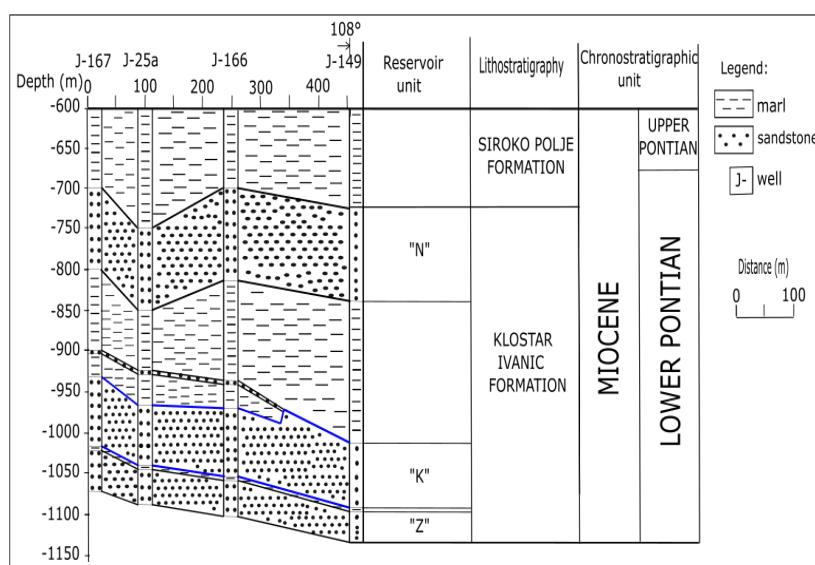


Figure 3. Correlation section through wells J-167 – J-25alfa – J-166 – J-149 shown on Figure 2 with blue line (from [14]).

The mapped sandstone “K” could be described as one of two typical sandstone types in the Northern Croatian Neogene. The younger are well-sorted, medium-grained, quartz-mica dominated sandstones, deposited in the lacustrine environment of the Pannonian Lake, mostly brackish. The detritus had been transported with sequential turbidites during Upper Pannonian and, especially, Lower Pontian (e.g., [15]). The dominant, primary porosity is intra-granular. The second one type are poorly sorted and coarse-grained sandstones, containing some larger (ruditic) clasts, are bound together with sparritic calcite cement. They contain various lithoclasts derived from older sedimentary and metamorphic rocks: dolomites, quartzites, gneisses and various schists, as well as individual quartz grains originated from metamorphic rocks. Such calcarenaceous sandstones (e.g., [16-17]) are deposited during Middle Badenian, in the in the shallow marine back-reef environments of the Paratethys Sea, with a significant input of siliciclastic material from the nearby land (Figure 4). The dominant, primary porosity is intra-skeletal (Figure 5).

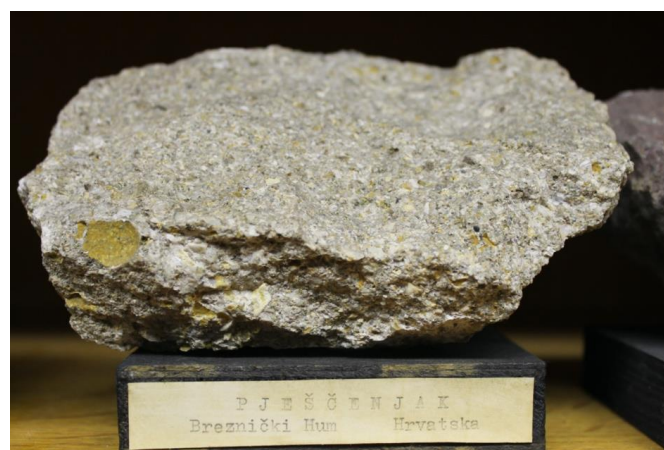


Figure 4. Macro sample of typical Badenian sandstone in the Northern Croatia (Breznički Hum locality).

89



90

91 **Figure 5.** Micro sample of clasts in typical Badenian sandstone in the Northern Croatia (Breznički
92 Hum locality).

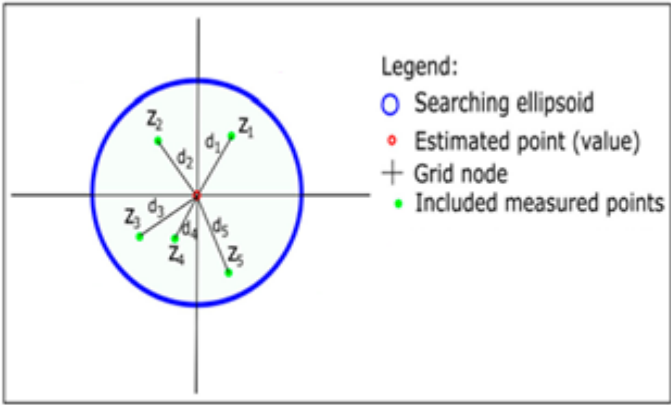
93 The analysis presented in this paper, i.e. interpolation of porosity, permeability and thickness
94 of sandstone units in the Neogene deposits in the Northern Croatia, is valid for both cases. The
95 applied methods, interpretations and conclusion could be applied both in the Badenian as well as
96 Pannonian/Pontian sandstone, although their origin is partially different.

97 **3. Basics of the applied interpolation methods**

98 The interpolation means the estimation among the values of variables in places where no values
99 have been measured. When interpolating, one (primary) or multiple (primary + secondary) variables
100 can be used. The interpolation methods applied in the reservoir "K" are Inverse Distance Weighting
101 (IDW) and Modified Shepard's Method (MSM). Only primary variables (porosity, permeability,
102 thickness) had been used.

103 **3.1. Inverse Distance Weighting (IDW)**

104 The Inverse Distance Weighting method is a mathematically simple interpolation method,
105 where the unknown value of a variable is estimated according to the known values inside the
106 searching radius. Generally, each of them is weighted by distance (Figure 6). Single data are inversely
107 proportional dependent on the distance between the unknown and measured values of variable.
108



109

110 **Figure 6.** Points included within the IDW interpolation area (from [3]).

The estimation (e.g., [18-19]) in IDW is done using (Equation 1):

$$z_{IW} = \frac{\frac{z_1}{d_1^p} + \frac{z_2}{d_2^p} + \dots + \frac{z_n}{d_n^p}}{\frac{1}{d_1^p} + \frac{1}{d_2^p} + \dots + \frac{1}{d_n^p}} \quad (1)$$

where:

z_{iw} —interpolated (unknown) value;

d_i —distance to “i-th” location;

z_i — interpolating (known) value at “i-th” location;

p —power of distance.

3.2. Modified Shepard's Method (MSM)

The MSM interpolation is modification of the IDW method, with the aim of reducing the expressive local values (outliers, extremes) that could cause “bull-eyeing” or “butterfly shapes” effects. The method was developed by [20], sometimes named as Shepard's method, and modification of the method was carried out in the works of, for example, [21] and [22]. The estimation (e.g., [22]) using MSM function is done by (Equation 2):

$$F(x, y) = \frac{\sum_{k=1}^n W_k(x, y) \cdot Q_k(x, y)}{\sum_{i=1}^n W_i(x, y)} \quad (2)$$

where:

F - MSM function;

W - relative weights;

Q_k - bivariate quadratic function;

x, y -data coordinates;

n - number of data.

MSM used so called relative weights determined (e.g., [22]) with (Equation 3):

$$W_k(x, y) = \left[\frac{(R_\omega - d_k)_+}{R_\omega \cdot d_k} \right]^2 \quad (3)$$

where:

W - relative weights;

d_k - Euclidean distance between points at locations (x, y) and (x_k, y_k) ;

R_ω - radius of influence about node (x_k, y_k) .

The Figure 2 shows the searching radius of the MSM within the measured and estimated values (grid). Here could be note each unknown point is surrounded with local searching radius, instead of global like in the IDW.

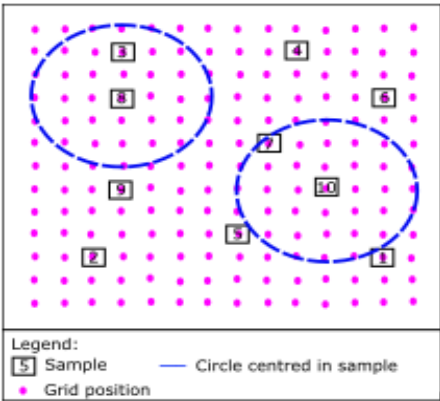


Figure 7. Radii of coverage of the MSM (modified from [23]).

4. Interpolation results in reservoir “K” for variables porosity, thickness and permeability

Here are interpolated three geological variables, measured in the reservoir "K", namely: porosity, permeability and thickness. Table 1 shows part of their descriptive statistics, calculated from values measured from the laboratory measurement and/or interpretation of the logging curves (resistivity, density and neutron).

Table 1. Geological variables porosity, permeability and thickness data in the "K" reservoir

Description	No data	Min	Max	Mean
Porosity	19	0.217	0.315	0.232
Permeability (10 ⁻⁹ m ²)	18	29.6	121.2	85.7
Thickness (m)	14	1	24	7.7

According to Table 1, analysed sets belong to a small geological input data set [24; 8]. So, the interpolation has been done with mathematically simpler methods. The porosity maps obtained by the IDW and SMS methods are given in Figure 8. The value cross-validation (porosity) for IDW is 0.00119 and SMS is 0.00345.

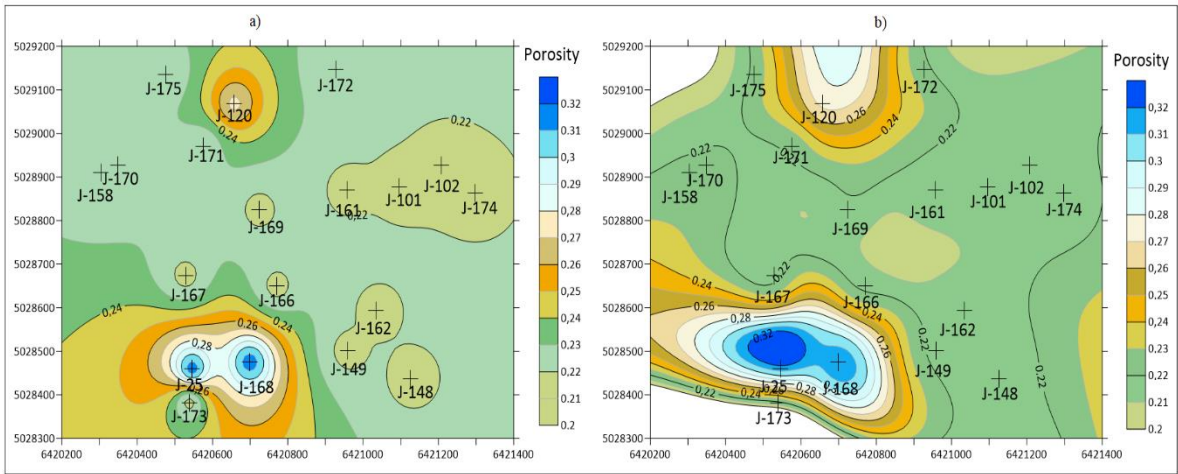


Figure 8. Porosity of the reservoirs "K": a) IDW b) SMS.

A permeability maps obtained by the IDW and SMS methods are shown in Figure 9. The accompanied cross-validation (permeability) for IDW is 480.8 and SMS 516.1.

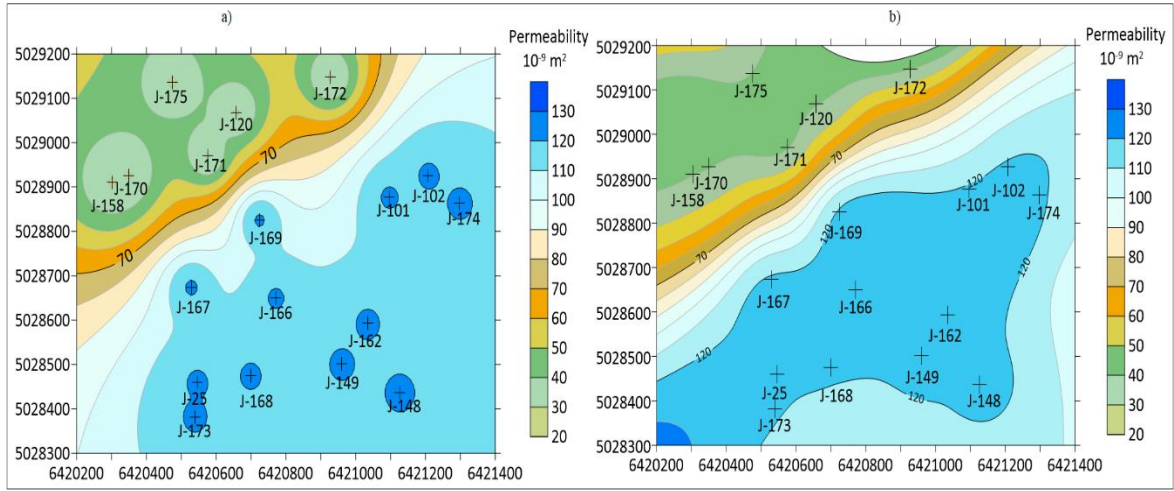


Figure 9. Permeability of the reservoirs "K": a) IDW b) SMS.

A thickness maps has been interpolated by the IDW and SMS methods are shown in Figure 10. The value cross-validation (thickness) for IDW is 40.7 and SMS is 60.5.

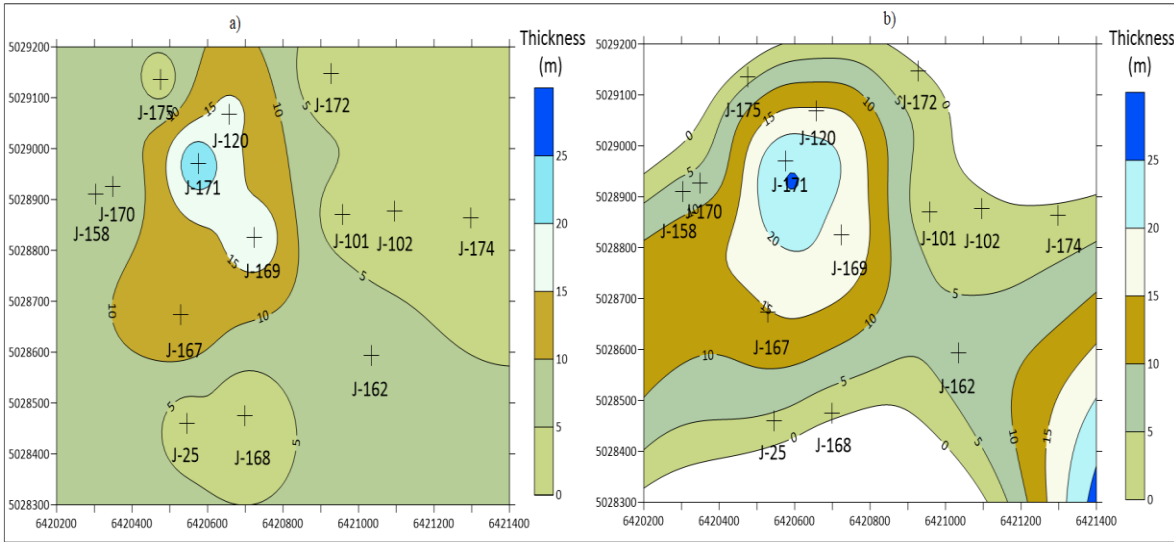


Figure 10. Thickness reservoirs "K": a) IDW b) SMS.

The maps obtained by the IDW and SMS methods are evaluated in quick-look, according the most observable feature of highly expressed local value, i.e. bull-eye or butterfly shape effects. The SMS is clearly advantageous because it mitigates the mentioned effects, however the numerical check using the cross-validation value strongly favours the IDW method.

5. Discussion and conclusion

To consider the differences between results obtained with IDW and MSM methods it is necessary to describe the main differences in their equations. The IDW does not use weighting coefficient, i.e.

each value is “weighted” by simple (powered) inversely proportional distance from the measured point. General form of the IDW is (Equation 4):

$$u(x) = \frac{\sum_{i=1}^N w_i(x) \cdot u_i}{\sum_{i=1}^N w_i(x)}, d(x, x_i) \neq 0 \quad (4)$$

Where:

$u(x)$ - interpolated value in the location “x”;

u_i - i -th known value inside searching radius centred in the location “x”;

i - no. of samples (known/measured value);

w_i - weight for the “ i -th” sample with known value;

N - number of samples with known value.

So, the weighting coefficients for location “x” could be generally defined as (Equation 5):

$$w_i(x) = \frac{1}{d(x, x_i)^p} \quad (5)$$

Where are:

w_i - weight for location “x”;

$d(x, x_i)$ - is a given distance (metric operator) from the known point x_i to the unknown point x ,

p - power factor (positive real number).

This is basic IDW function as defined by Shepard ([20], where weights decrease as distances to known values increase. Oppositely, greater values of power (“p”) favour larger influence of measured points closer to interpolated point, eventually finished in zonal interpolation (like Voronoi diagram), where inside zone is used constant value of single known point centred into such polygon and no isolines are constructed. On contrary, for 2D interpolation, if power is 2 or less, more distanced measured points will have greater influence. In the “M” dimensions this statement is valid for $p \leq M$. In all cases the IDW minimize function that measure deviation between unknown and known points using condition (Equation 6):

$$\frac{\delta(x, u)}{\delta u} = 0 \quad (6)$$

On contrary, the MSM is more advanced technique of the IDW, including weighting coefficient (Equation 8), and calculating interpolated value using only nearest known points within R-sphere (instead of using full set of measured samples). Weights are modified in (Equation 7):

$$w_k(x) = \left(\frac{\max(0, R - d(x, x_k))}{Rd(x, x_k)} \right)^2 \quad (7)$$

Where:

$Rd(x, x_k)$ - is local searching radius, so called “R sphere”, centred in point “x” and including “k” measured points;

$w_k(x)$ - weighting coefficient for known point “k”, at distance “x” from interpolated point.

By default, the IDW takes into account all measured points or work with general searching radius (or radii for ellipsoid). However, the MSM, by default, works with local searching (both can

be changed in the most software, but those are theoretical defaults). So, "R sphere" modification would need to reduce anomalous local values, i.e. outliers or extremes. The maps on **Figures 7-9** showed that it is really so.

However, there is problem of calculated cross-validation values. They were significantly lower for the IDW then MSM (Table 2). The differences are not negligible. For porosity MSM cross-validation is 289 % larger, permeability 7 %, and thickness 49 %.

Table 2. Summary results of application of the IDW and SMS methods in reservoir "K"

Description	No data	Cross-validation	
		Inverse Distance (IDW)	Modified Shepard's Method (MSM)
Porosity	19	0.00119	0.00345
Permeability	18	480.8	516.1
Thickness	14	40.7	60.5

So, let once more time analyse the **Equations 2** (MSM function), **3** (weighting coefficient in the MSM searching radius) or **8** (the more general form of Eq. 3). It is not surprisingly that the MSM will mostly give larger cross-validation values, because it works by default with local search and, consequently, smaller number of known values in each cross-validation iteration. If "one point out" method is applied with small number of remaining known values, the difference between real and estimated value in such point will be mostly higher than if all known measurements had been used. So, generally the MSM will tend to give larger cross-validation value for the same dataset, compared with the IDW.

Secondly, the MSM approximate the unknown value with quadratic function, i.e. polynomial of degree 2, for one or more variables in which is the 2nd degree the highest one. Any quadratic polynomial with two variables may be given as (Equation 8):

$$f(x, y) = ax^2 + by^2 + cxy + dx + ey + f \quad (8)$$

Where:

x, y - variables;

a, b, c, d, e, f - coefficients.

Consequently, the MSM will harder force gradual interpolation for outliers (extremes), i.e. in such ways it will done smoothing instead of bull-eyeing. At the end, let's look the selection criterium of interpolation algorithm for the small datasets in the same spatial area as it was selected for this work (the Sava Depression, Northern Croatia; [3]):

Table 3. Recommended interpolation methods for small input data set [3].

Number of data	Applicability of interpolation method		
	Inverse Distance Weighting	Nearest Neighbourhood	Natural Neighbour
1-5	Yes	Yes	No
6-10	Yes	Yes	Yes
11-19	Yes	Yes	Yes

Authors in [3] the small datasets divided into three classes regarding included points: (a) 1-5, (b) 6-10 and (c) 11-19 points. It is concluded, if the main selection criteria could be cross-validation, the

IDW would be the best choice for interpolation. However, they also warned that this method sometimes created numerous “bull’s-eyes” or “half-butterfly” features. So, we can summarize our results in conclusion and recommendation given in the following statements and Table 4:

1. The MSM could be recommended for geological subsurface mapping of Neogene in the Northern Croatia;
2. That is valid for datasets smaller than 20 measured values;
3. The selected variables could be porosity, permeability and thickness measured in Neogene lithostratigraphic units by laboratory analysed well data (cores) or interpreted logs;
4. The interpreter need to be aware of: (a) the IDW resulted with the lowest cross-validation of all applied methods without spatial model (variogram), i.e. namely Nearest neighbourhood and Natural neighbour, (b) the MSM interpolation eliminated all unwanted geological features and did not result in the cross-validation 250 % higher than in the IDW, for the same variable.
5. Based on visual (geometric) criteria, the sets with less than 5 measured points need to be consider as too small for interpolation (the IDW or MSM), and they are appropriate only for zonal estimation.

Table 4. Recommended choice between the IDW and MSM for small input data set (less than 20 data.

Number of data	Applicability of interpolation method	
	Inverse Distance Weighting	Modified Shepard’s Method
Criterion	(1) If Nearest Neighbourhood, Natural Neighbour or any available zonal method is tested and the IDW has the smallest cross-validation. (2) If the cross-validation is single criteria for quick-look.	(1) If the IDW has the smallest cross-validation. (2) If the outliers are strongly and unnaturally emphasised on the map (“bull-eye”, “butterfly” features). (3) If expert geologist, with knowledge about regional geology of mapped volume, can evaluate map(s).
5-20	Yes	Yes

Author Contributions: Tomislav Malvić led the research and selected the fields for analyses. Josip Ivšinić collected data, selected Sheppard method and interpolated maps. Josipa Velić and Jasenka Sremac prepared regional geology presentation and Uroš Barudžija visualised samples and connect with regional geology.

Funding: This research (analyses) was partially supported with the project “Mathematical methods in geology IV” (led by T. Malvić). Funds were given from the University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, for the 2019 year.

Acknowledgments: some data from technical documentation of INA Plc. had been used and presented in some maps and tables. That is collected during doctoral researching J. Ivšinić.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Mesić Kiš, I.; Malvić, T. Zonal estimation and interpolation as simultaneous approaches in the case of small input data set (Šandrovac field, northern Croatia). *The Mining-Geology-Petroleum Engineering Bulletin* **2014**, *29*, 9-16.
2. Malvić, T. History of geostatistical analyses performed in the Croatian part of the Pannonian Basin System. *Nafta* **2012**, *63*, 223–235.
3. Malvić, T.; Ivšinović, J.; Velić, J.; Rajić, R. Interpolation of Small Datasets in the Sandstone Hydrocarbon Reservoirs, Case Study of the Sava Depression, Croatia. *Geosciences* **2019**, *9*, 201.
4. Malvić, T.; Đureković, M. Application of methods: Inverse distance weighting, ordinary kriging and collocated cokriging in porosity evaluation, and comparison of results on the Beničanci and Stari Gradac fields in Croatia. *Nafta* **2003**, *54*, 331–340.
5. Smoljanović, S.; Malvić, T. Improvements in reservoir characterization applying geostatistical modelling (estimation & stochastic simulations vs. standard interpolation methods), Case study from Croatia. *Nafta* **2005**, *56*, 57–63.
6. Balić, D.; Velić, J.; Malvić, T. Selection of the most appropriate interpolation method for sandstone reservoirs in the Kloštar oil and gas field. *Geologia Croatica* **2008**, *61*, 27-35.
7. Ivšinović, J. Deep mapping of hydrocarbon reservoirs in the case of a small number of data on the example of the Lower Pontian reservoirs of the western part of Sava Depression. Proceedings of the 2nd Croatian congress on geomathematics and geological terminology, 2018, Zagreb, Croatia, 6 October 2018, Malvić, T. (ed.); Velić, J. (ed.); Rajić, R. (ed.), University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering; pp. 59-65.
8. Ivšinović, J. Odabir i geomatematička obradba varijabli za skupove manje od 50 podataka pri kreiranju poboljšanoga dubinskogeološkoga modela na primjeru iz zapadnoga dijela Savske depresije (Selection and geomathematical calculation of variables for sets with less than 50 data regarding the creation of an improved subsurface model, case study from the western part of the Sava Depression), Doctoral thesis, University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, 2019.
9. Yang, C.S.; Kao, S. P.; Lee, F.B.; Hung, P.S. Twelve different interpolation methods: A case study of Surfer 8.0. Proceedings of the XXth ISPRS Congress, Instabul, Turkey, 12-23 July 2004, Altan, M. O. (ed.), Photogrammetry and Remote Sensing; pp. 778-785.
10. Rap, A.; Ghosh, S.; Smith, M. Shepard and Hardy Multiquadric Interpolation Methods for Multicomponent Aerosol-Cloud Parameterization. *Journal of The Atmospheric Sciences* **2009**, *66*, 105-115.
11. Pál, L.; Oláh-Gál, R.; Mako, Z. Shepard interpolation with stationary points. *Acta Univ. Sapientiae, Informatica* **2009**, *1*, 5-13.
12. Bronowicka-Mielniczuk, U.; Mielniczuk, J.; Obroślak, R.; Przystupa, W. A. Comparison of Some Interpolation Techniques for Determining Spatial Distribution of Nitrogen Compounds in Groundwater. *International Journal of Environmental Research* **2019**, *13*, 679–687.
13. Huang, X.; Chen, W.; Hu, Z.; Zheng, X.; Jin, S.; Zhang, X. Application Research of an Efficient and Stable Boundary Processing Method for the SPH Method. *Water* **2019**, *11*, 1110.
14. Malvić, T.; Ivšinović, J.; Velić, J.; Rajić, R. Kriging with a Small Number of Data Points Supported by Jack-Knifing, a Case Study in the Sava Depression (Northern Croatia). *Geosciences* **2019**, *9*, 36.
15. Malvić, T.; Velić, J. Neogene Tectonics in Croatian Part of the Pannonian Basin and Reflectance in Hydrocarbon Accumulations. In *New Frontiers in Tectonic Research: At the Midst of Plate Convergence*; Schattner, U. Ed.; InTech, Rijeka, Croatia, 2011; pp. 215-238.
16. Šimunić, A.; Pikija, M.; Hećimović, I.; Šimunić, Al. Basic Geological Map of SFR Yugoslavia; 1:100.000; Geology for Varaždin Sheet L33-69.-Geological survey Zagreb, 1981; Federal geological survey Belgrade. (in Croatian).
17. Pettijohn, F.J.; Potter, P.E.; Siever, R. Sand and Sandstones; Springer: Berlin, Germany, 1972.
18. Malvić, T. Primjena Geostatistike u Analizi geoloških Podataka (Application of Geostatistics in Geological Data Analysis); INA-Industrija nafte d.d.: Zagreb, Croatia, 2008.

352 19. Ly, S.; Charles, C.; Degré, A. Geostatistical interpolation of daily rainfall at catchment scale: the use of several
353 variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences*
354 **2011**, *15*, 2259-2274.

355 20. Shepard, D. A two-dimensional interpolation for irregularly spaced data function. Proceedings of the 1968
356 ACM national conference, New York, USA, 27-29 August 1968, Blue, R. B. (ed.); Rosenberg, A. M. (ed.),
357 Association for Computing Machinery; pp. 517–523.

358 21. Franke, R.; Nielson, G. Smooth Interpolation of Large Sets of Scattered Data. *International Journal for*
359 *Numerical Methods in Engineering*, **1980**, *15*, 1691-1704.

360 22. Renka, R. J. Multivariate Interpolation of Large Sets of Scattered Data. *ACM Transaction on Mathematical*
361 *Software*, **1988**, *14*, 139-148.

362 23. Basso, K.; Zingano, P.; Freitas, C. (1999). Interpolation of Scattered Data: Investigating Alternatives for the
363 Modified Shepard Method. Proceedings of the XII Brazilian Symposium on Computer Graphics and Image
364 Processing (SIBGRAPI '99), Campinas, Brazil, 17-20 October, 1999, Stolfi, J. (ed.); Tozzi, C. L. (ed.),
365 Universidade Estadual de Campinas (UNICAMP) & Sociedade Brasileira de Computação (SBC); pp. 39-48.

366 24. Husanović, E.; Malvić, T. Review of deterministic geostatistical mapping methods in Croatian hydrocarbon
367 reservoirs and advantages of such approach. *Nafta* **2014**, *65*, 57-63.