5th Croatian geomathematical congress
UDC: 622.1; category: original scientific paper

Statistical analysis and mapping of imaginary coal layer using inverse distance interpolation, simple case study of recommended methods

Ana Ljubić[1]; Mirta Petrović[1]
[1]University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, 10000 Zagreb, Croatia
ORCID: Ana Ljubić: 0009-0007-7753-4608; Mirta Petrović: 0009-0006-6131-5905
Corresponding author: mirta.petrovic@rgn.unizg.hr

Abstract

In order to correctly describe the minerals within a deposit for the purpose of exploitation, a large amount of geological data from the selected deposit is required. In the case of underground exploitation, the amount of necessary data increases. This paper describes an imaginary coal seam using both simulated data and primary research findings. The selected coal seam is located underground, making it difficult to get the necessary data. That is why proper spatial analysis of known data is necessary. Statistical methods are applied to measure spatial diversity with the measured samples. Results are described according to the total variant of estimated points or blocks. The methods used in order to observe the distribution of the data are the Shapiro-Wilk, Kolmogorov-Smirnov tests and the QQ diagram. The inverse distance method was used to create the maps. The goal was to obtain reliable and spatial results and show set of methods for fast and reliable representation of 2D maps of deposits, like coal one. All conducted research provided a large amount of data on the mentioned layer, which was successfully described using statistical methods.

Key words: coal seam; normal distribution; null hypothesis; data analysis

1. Introduction

In many geosciences, but specific to mining industry, it is often necessary to collect a large amount of information about the composition and stability of the subsoil. These informations are not easy to obtain because each specific type of work has a need for unique equipment and large number of workforce. Also, underground researches come with high risks and costs. In these conditions, statistical data analysis is of great help, which simplifies research and greatly reduces the time and costs.

This paper describes an imaginary coal seam, with imagined data on depth and size intended only for the purpose of displaying statistical analyses. The described coal seam is located in the settlement of Siverić, which is located 4 kilometers northeast of the town of Drniš, in the municipality of Promina (**Figure 1-1**). The Promina deposits are groups of sedimentary rocks that are named after the mountain in which they are located. They consist of conglomerate, limestone, marl and coal of Eocene-Oligocene age with a high degree of carbonization. It covers an area of 187 km² and is located at a depth of 500 m (**Zupanič & Babić, 2011**).

Precisely this depth of the deposit represents the mentioned obstacles in the research that would be carried out for the purpose of exploitation. That is why it is necessary to reduce the influence of these obstacles. Specifically, by examining the normal distribution of data on the depth of the top and the bottom area of the deposit, it is possible to visualize the shape of the deposit itself. Since that examination will only give informations whether our data is normally distributed, and it is difficult to conclude with the obtained numerical indicators which data in the data set causes a potential deviation from the normal distribution, statistical tools such as variance and standard deviation will also be used for better interpretation. In order to present the analytical results of the mentioned tests in graphic form, the easiest way is to use the mapping method. The chosen method is inverse distance weighting, which is also used for interpolation of a limited number of known data. The objectives of this paper is to determine how much data can be obtained about the mentioned imaginary coal seam based on limited data, and what methods are best to use for our purposes.

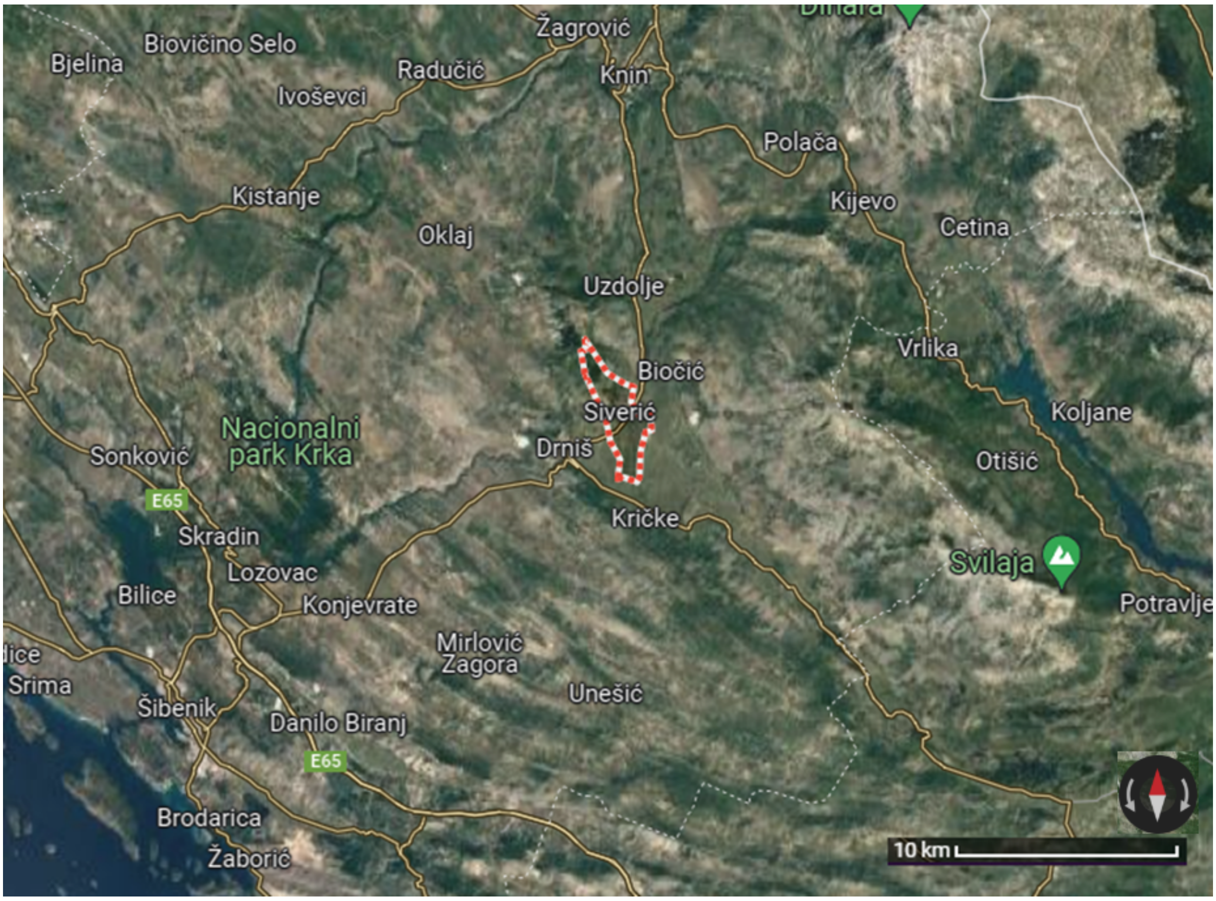

Figure 1-1: Geographical location of the settlement Siverić (**Google Maps**, accessed: 18th December 2023)

2. Materials and methods

To approximate the data of the imaginary coal seam, and to find out whether they are normally distributed, it is necessary to use statistical analysis. In this paper, the collected data were analyzed by Shapiro-Wilk, Kolmogorov-Smirnov test and QQ diagram. For easier data interpretation, they were also described with simple statistical tools.

2.1. Case study

General rule for coal deposition environments is that coal layers usually have a form of a syncline. Based on that, imaginary coal seam in the Siverić settlement is described in the same model. The data on the thickness of the layer i.e., the depth of the top and bottom surfaces are defined by imaginary coal data set, which is shown in **tables 1** and **2**. In said tables it is shown how the thickness of the layers differs and gradually increases approaching the middle of the syncline, where the layer is up to 50 meters thick, while moving away from the centre, towards the edges, the layer becomes thinner. These data sets are imagined based on previous investigations of similar coal deposits in the area, so the x and y variables can be considered as points in the imaginary coordinate system in the area of the Siverić settlement. The entire coal deposit is located approximately 500 meters below the surface, and precisely because of this great depth it is considered that research could be expensive, risky and time-consuming.

2.2. Statistical tools

Some of the statistical tools that form the basis of correct data analysis are the mean value, the minimum and maximum of the measured values, and the median and centiles, which are independent of extreme values and thus provide more detailed information.

Also useful statistical tools are variance and standard deviation which are closely related to each other. Variance can be defined as a measure of the dispersion of measured or random variables, i.e. the average sum of the squares of the deviations of the characteristic values from the arithmetic mean, and was described by **Equation 1**:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \quad (1)$$

where are:
σ^2 - variance;
x - one of the values from series;
$\bar{x}$ - arithmetic mean;
N - number of values in the series (**Pfaff, 2012**),

while the standard deviation indicates the positive square root of the variance, i.e. the measure of deviation (**URL 1**). Due to this mathematical connection, it can be emphasized that these two statistical tools are closely related.

For data processing, a deep study of the data distribution and statistical parameters are very helpful. The most common and well-known distribution function in statistics is the normal distribution. It is bell-shaped and symmetrical around the ordinate axis, which is shown in **Figure 2-1**. It was described by the function (**Equation 2**):

$$f(x) = \frac{1}{s\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\bar{x}}{s^2}\right)^2\right] \text{ for } -\infty < x < \infty \quad (2)$$

where for any „x“ are :
s - standard deviation;
x - measured value;
$\bar{x}$ - mean value.

The function reaches its maximum at a value of 0.4/s, which was valid in the case x=$\bar{x}$, and it was precisely this case that explains the symmetry around the ordinate axis and the bell shape. Due to the central limit axiom, which describes the behavior of a large number of independent measurements and the property that most of the data will be located exactly around x=$\bar{x}$, the use of this type of distribution was frequent. Due to the poor availability of data, the normal distribution approximation was a very useful method in all phases of research (e.g. **Malvić et al., 2015**).

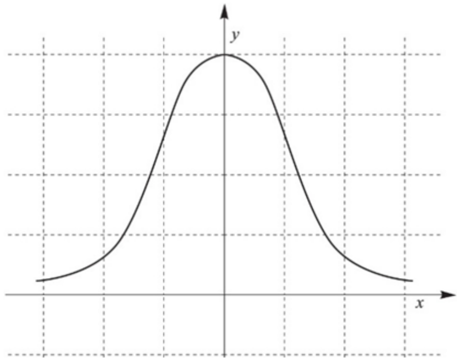

Figure 2-1: Example of normal distribution diagram (**URL 2**)

Three tests, Shapiro-Wilk, Kolmogorov-Smirnov and QQ plot, were used to test normal distribution.

2.3. Examination of the normal distribution

The Shapiro-Wilk test is used for testing normal distribution. It was published by Samuel Sanford Shapiro and Martin Wilk in 1965 (**URL 3**), and was described with **Equation 3**:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (3)$$

where are:
W - test statistic
α - level of significance
$x_{(i)}$ - measured values ordered from smallest to largest;
$\bar{x}$ - arithmetic mean.

The null hypothesis of this test was that the data is normally distributed. If the p-values of the test are less than alpha, then we reject the hypothesis and the data are not normally distributed. When the p-value is greater than alpha (α), we do not reject the hypothesis and conclude that the data are normally distributed. In general, there is a possibility that this test will show errors in case of too much data. Therefore, it is necessary to use several different tests for the same data set in order to compare the results. That's how we can consider them more reliable.

An example of the second type of test is the Kolmogorov-Smirnov test (K-S test), which can be defined as a non-parametric test of the equality of one-dimensional probability distributions, and with it we can find out whether a sample comes from a given probability distribution and whether two samples come from the same distribution. It differs from parametric tests because it does not require a defined variance and mean value for data analysis (**URL 4**). Because of this, it is less precise so it is required to do additional tests. K-S test can be calculated using the equation (**equation 4**):

$$D_n = \frac{sup}{x} \left| F_n(x) - F(x) \right| \qquad (4)$$

where are:
n - number of values in series;
sup - distance supremum;
F_n - empirical distribution function;
F - theoretical cumulative distribution function (**Bralić et al., 2022**).

While the empirical distribution function F_n was defined by the equation (**Equation 5**):

$$F_n(x)=\frac{number\ of\ values\ in\ series \leq x}{n}=\frac{1}{n}\sum_{i=1}^{n} 1_{<-\infty,x]}(x_i) \quad (5)$$

where are:

$1_{<-\infty,x]}(x_i)$ - an indicator function whose value equals 1 when $x_i \leq x$, while it equals 0 in other cases.

In addition to qualitative tests, the quantitative QQ diagram (quantile-quantile diagram) is also useful for data analysis, which was defined as a graphical method for comparing two probability distributions (**URL5**). With the QQ diagram, we can check whether the assumption that the data set is normally distributed is correct. A scatterplot is obtained by plotting two sets of quantiles, and a 45-degree reference line is also plotted. If the data set is normally distributed, the points should fall approximately along this reference line. The further the data set is from the normal distribution, the greater the deviation from the reference line. The QQ diagram is only a visual representation of the data, and the assessment itself is subjective, which is why this solution is only an aid to analysis (**URL6**).

Data obtained from previous research in the observed area should be interpolated in order to best understand the details of the exploitation field. The method used for this purpose is called the inverse distance weighting method, which is also the simplest interpolation method, and was often used in statistical procedures. This method assigns corresponding weighting coefficients to the control points depending on their distance from the points of the regular spatial mesh. The distance exponent (Power), that is shown in **tables 5** and **8**, is a weighting parameter that determines how quickly the weight will decrease depending on the distance from the nodes of the spatial mesh. The closer the parameter is to zero, the more similar the resulting display is to a horizontal surface, while larger parameter, results in more elaborate topography of the surface (**Medved et al., 2010**).

networks. With the help of a mathematical expression, the value at the selected points was estimated (**6**):

$$z_{IU} = \frac{\frac{z_1}{d_1^p}+\frac{z_2}{d_2^p}+\cdots+\frac{z_n}{d_n^p}}{\frac{1}{d_1^p}+\frac{1}{d_2^p}+\cdots+\frac{1}{d_n^p}} \quad (6)$$

where are:

z_{IU} - estimated value;

$d_1 \ldots d_n$ - distances of locations 1...n from the assessment site z_{IU};

p - distance exponent;

$z_1 \ldots z_n$ - measured values at locations 1...n.

The number of points is determined by the radius of the circle placed around the test location. The influence of each point is inversely proportional to its distance from the location where the value is estimated, which means that the values that are closer to the points will have a greater influence on the interpolation process. (**Malvić, 2008**).

Table 1: Artificial data of coal layer (in meters), a)

x	1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	4	4	4	4
y	10	6	11	4	5	6	7	8	9	10	11	3	4	5	6	7	8	9	10	11	12	3	4	5	6
Thickness of layers	1	0	0	2	3	5	4	2	1	6	5	1	3	12	8	10	6	8	12	8	0	3	10	14	10
Depth of the roof surface	497	493	498	492	494	500	495	496	496	497	499	492	493	494	501	500	495	495	500	497	501	492	496	500	502
Depth of the bottom surface	498	493	498	494	497	505	499	498	497	503	504	493	496	506	509	510	501	503	512	505	501	495	506	514	512

x	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5	6	6	6	6	6	6	6	6	6
y	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11
Thickness of layers	18	20	18	15	11	4	4	12	25	18	22	27	25	27	15	6	2	10	28	20	28	32	42	30	16
Depth of the roof surface	500	500	499	501	501	506	492	499	500	503	501	501	500	500	502	502	493	499	500	504	503	501	501	501	503
Depth of the bottom surface	518	520	517	516	512	510	496	511	525	521	523	528	525	527	517	508	495	509	528	524	531	533	543	531	519

x	6	6	7	7	7	7	7	7	7	7	7	7	7	8	8	8	8	8	8	8	8	8	8	9	9
y	12	13	3	4	5	6	7	8	9	10	11	12	13	4	5	6	7	8	9	10	11	12	13	4	5
Thickness of layers	7	2	0	9	28	32	35	45	47	28	25	7	1	6	15	30	42	48	50	49	32	12	1	0	18
Depth of the roof surface	500	497	496	497	500	502	504	505	502	505	501	501	492	496	500	503	506	503	506	505	502	498	498	491	497
Depth of the bottom surface	507	499	496	506	528	534	539	550	549	533	526	508	493	502	515	533	548	551	556	554	534	510	499	491	515

Table 2: Artificial data of coal layer (in meters), b)

x	9	9	9	9	9	9	9	9	10	10	10	10	10	10	10	10	10	11	11	11	11	11	11	11	11
y	6	7	8	9	10	11	12	13	5	6	7	8	9	10	11	12	13	4	5	6	7	8	9	10	11
Thickness of layers	29	38	50	50	22	8	4	2	5	17	42	50	47	43	11	2	1	0	6	27	50	50	50	42	13
Depth of the roof surface	503	507	503	506	504	500	498	500	496	502	502	502	501	503	502	500	501	493	496	501	501	500	501	502	499
Depth of the bottom surface	532	545	553	556	526	508	502	502	501	519	544	552	548	546	513	502	502	493	502	528	551	550	551	544	512

x	11	11	12	12	12	12	12	12	12	12	12	12	13	13	13	13	13	13	13	13	13	13	14	14	14
y	12	13	4	5	6	7	8	9	10	11	12	13	3	4	5	6	7	8	9	10	11	12	3	4	5
Thickness of layers	3	1	10	28	32	48	50	40	38	12	3	0	0	8	7	16	30	24	22	25	22	4	0	5	7
Depth of the roof surface	499	500	497	502	500	500	500	501	498	497	502	497	496	500	501	497	497	499	499	499	496	497	497	500	500
Depth of the bottom surface	502	501	507	530	532	548	550	541	536	509	505	497	496	508	508	513	527	523	521	524	518	501	497	505	507

x	14	14	14	14	14	14	14	15	15	15	15	15	15	15	15	15	16	16	16	16	16	17	17	17	17
y	6	7	8	9	10	11	12	3	4	5	6	7	8	9	10	11	4	5	6	7	8	4	5	6	7
Thickness of layers	8	12	8	9	15	8	2	0	3	6	7	10	5	4	4	2	1	3	3	2	2	1	5	4	1
Depth of the roof surface	496	494	499	500	496	496	496	497	497	496	494	493	499	497	496	494	496	497	493	492	496	495	495	492	492
Depth of the bottom surface	504	506	507	509	511	504	498	497	500	502	501	503	504	501	500	496	497	500	496	494	498	496	500	496	493

3. Results

By performing the S-W and K-S test, total of 750 samples were entered that contain values of depths of top and bottom area, thickness of the deposit and artificial positions of known points. (**Table 1** and **2**) Those tests also showed histograms that describe distribution of values of depths. **Figure 3-1** decribes the data for the top area of the deposit, while **Figure 3-2** displays bottom area. In the histograms, the x axis shows the ranges of the depths, while the y axis represents count of samples that fall within each range.

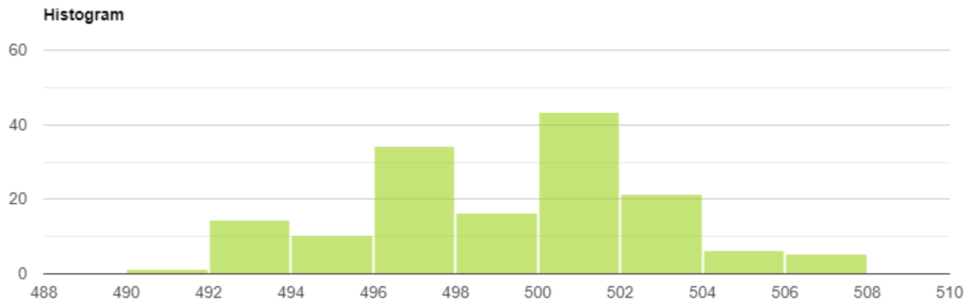

Figure 3-1: Values of depths of top area (**URL7**)

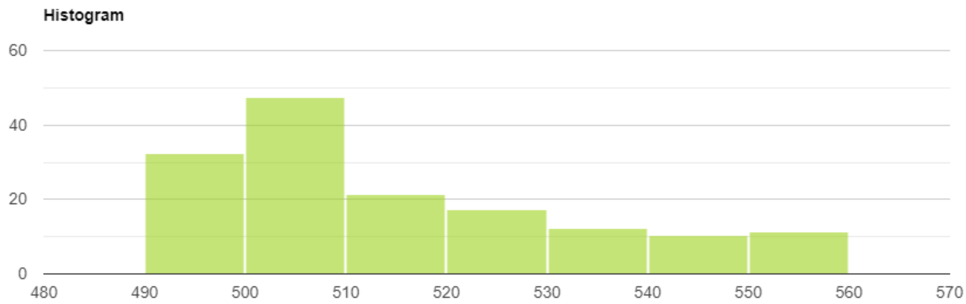

Figure 3-2: Values of depths of bottom area (**URL7**)

Table 3 shows the results of the Shapiro-Wilk test. The initial null hypothesis was that the data are normally distributed. The obtained p-value was 0, also visible from the table (**table 3**), which is a smaller value than alpha (α=0.5). For that reason the null hypothesis was rejected, and it could be concluded that the data were not normally distributed.

Table 3: Presentation of Shapiro-Wilk test results

P-value	<0.001
W	0.66
Number of samples	750
Average	290.33
Median	15
Sample standard deviation (S)	243.41
Skewness	0.41

Due to the amount of alpha (α=0.05), the S-W test allows 5% deviation of the data, and since the p-value was less than α, the data significantly deviate from a normal distribution. The same could be seen with observing the test statistic (W), which was valued at 0.6602. In S-W test, the test statistic (W) close to 1 is desirable, so its obtained value confirmed deviation from the normal distribution.

In the Kolmogorov-Smirnov test, the null hypothesis was also rejected because the obtained p-value was lower than alpha. The test statistic D representing the deviations between the empirical and the theoretical relative cumulative frequency was equal to 0.34. Similar to the S-W test, the test statistic D serves as a confirmation for deviation from the normal distribution, which was also seen with the derived value of 0.34. (**Table 4.**). Also, in **Figure 3-3**, the curve obtained by K-S test is shown, which can be compared with **Figure 2-1** for a better understanding of the results, and how far they deviate from the normal distribution.

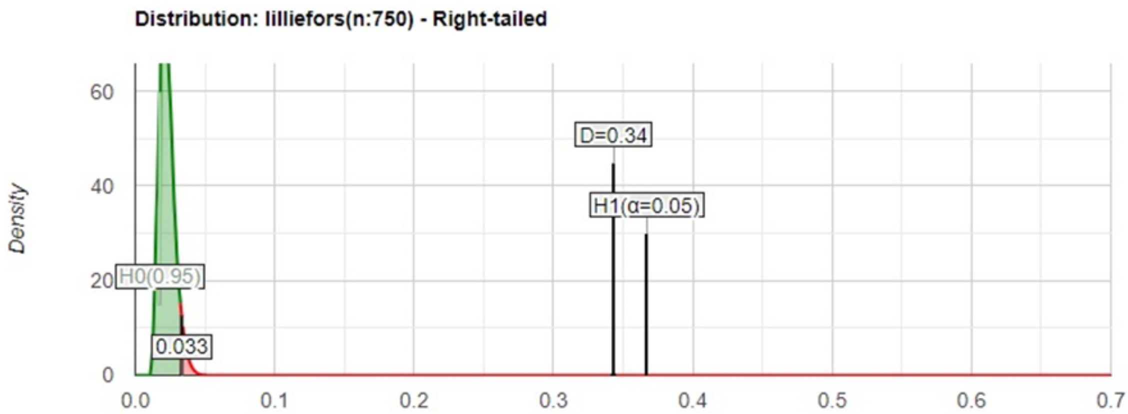

Figure 3-3: Display of the curve obtained by the K-S test (**URL7**)

Table 4: Displaying the results of the Kolmogorov-Smirnov test

P-value	<0.001
D	0.34
Number of samples	750
Average	209.23
Median	15
Sample standard deviation (S)	243.41
Skewness	0.41
K	9.41

Also, a QQ plot can graphically display the dispersion of results. This graphic representation can be obtained with the formula for calculating the quantiles of the entered data (from the **Tables 1** and **2**). Since the points do not follow the blue reference line, it could be concluded that the data do not come from the same distribution and that the assumption of a normal data distribution is wrong (**Figure 3-4**)

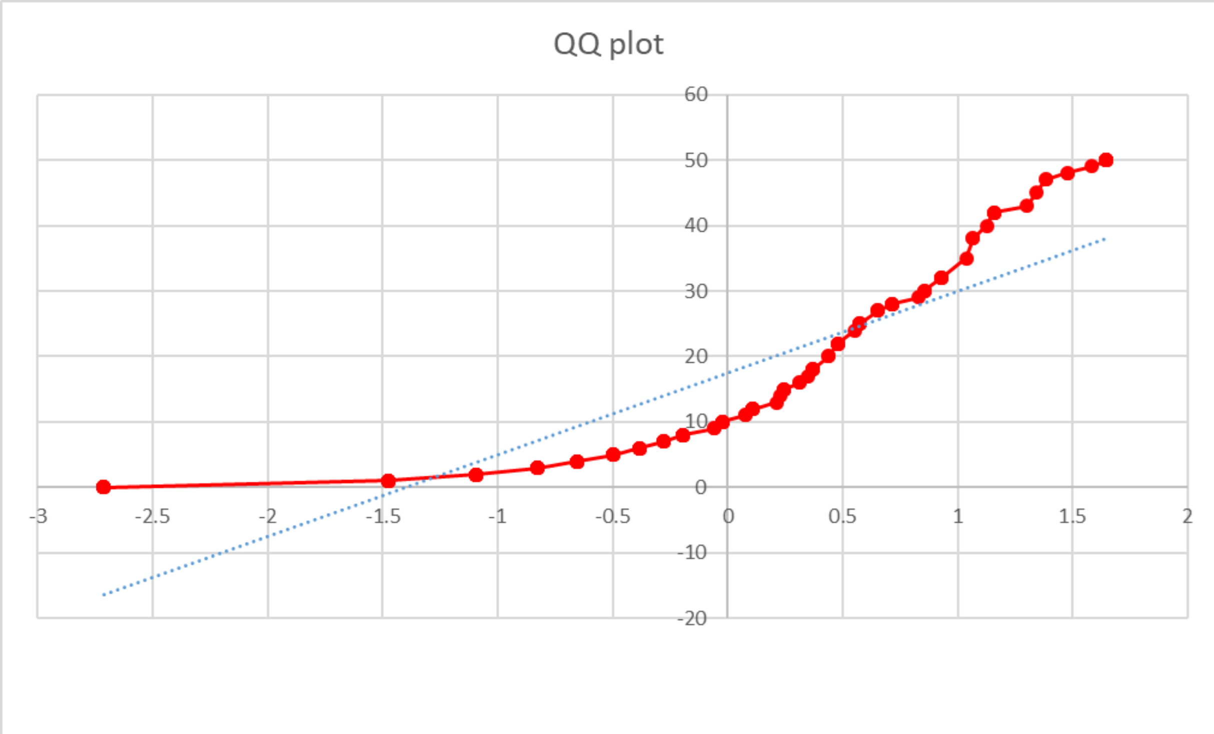

Figure 3-4: Display of QQ diagram (**URL7**)

Using the Surfer 8 program, x and y data, depth of the top and bottom, were entered. The inverse distance method was used for interpolation. The obtained results provide a basic statistical analysis, the size of the network and the number of nodes, which can be seen in **Tables 5**, **6**, **7** and **8**. Limits of the different depth values must be clearly seen, which is why in this paper the distance exponent value equals 2, which shows the simplest form of the map due to properly rounded lines. The map was displayed through 63 rows and 100 columns, which, due to the ratio of the map, shows all the given values of the depths.

Table 5: Data for the top area

Weighting power	2
Smoothing power	0
Anisotropy ratio	1
Anisotropy angle	0
Grid size	63 rows and 100 columns
Total nodes	6300

Table 6: Basic statistical analysis for top area

	x	y	z
Minimum	1	3	491
25%-tile	5	5	496
Median	9	8	500
75%-tile	13	10	501
Maximum	17	13	507
Average	8.83	7.83	498.77
Standard deviation	4.38	2.87	3.54
Variance	19.18	8.26	12.54

Table 7: Basic statistical analysis for bottom area

	x	y	z
Minimum	1	3	491
25%-tile	5	5	501
Median	9	8	509
75%-tile	13	10	527
Maximum	17	13	556
Average	8.83	7.83	514.75
Standard deviation	4.38	2.87	17.95
Variance	19.18	8.26	322.25

Table 8: Data for bottom area

Weighting power	2
Smoothing factor	0
Anisotropy ratio	1
Anisotropy angle	0
Grid size	63 rows and 100 columns
Total nodes	6300

In the Surfer 8 program, maps of the depth of the top and bottom were created. A darker color represents smaller depth, while a lighter color represents deeper surfaces. **Figure 3-5** shows a map of the depth of the top, while **Figure 3-6** shows a map of the depth of the bottom.

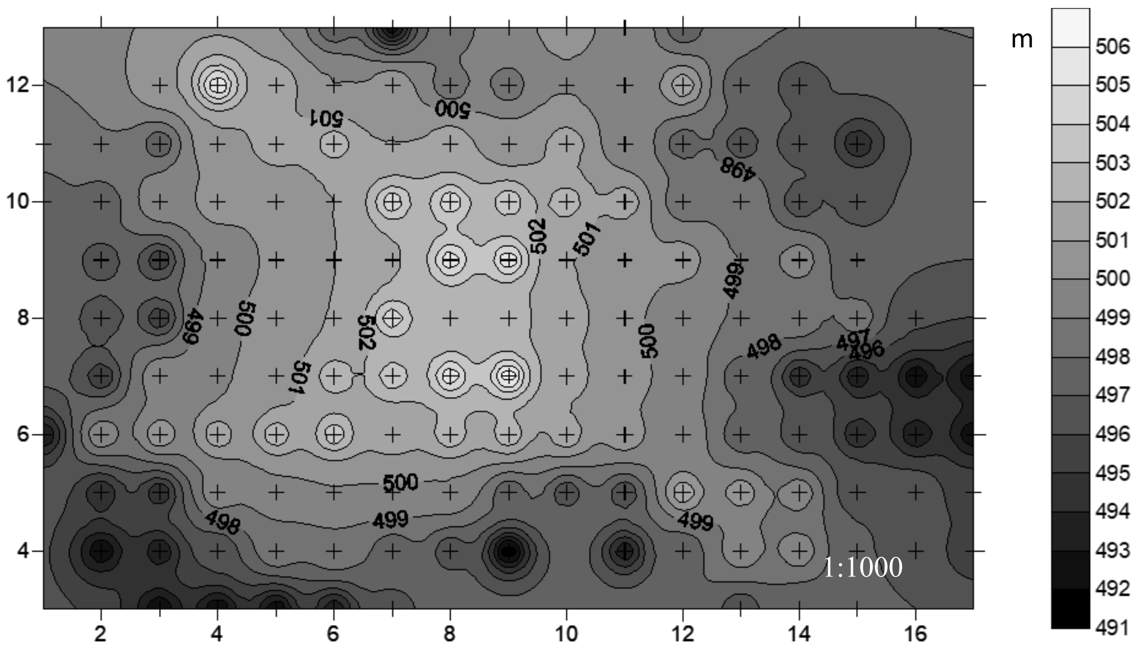

Figure 3-5: Map of the relative depth of the top area

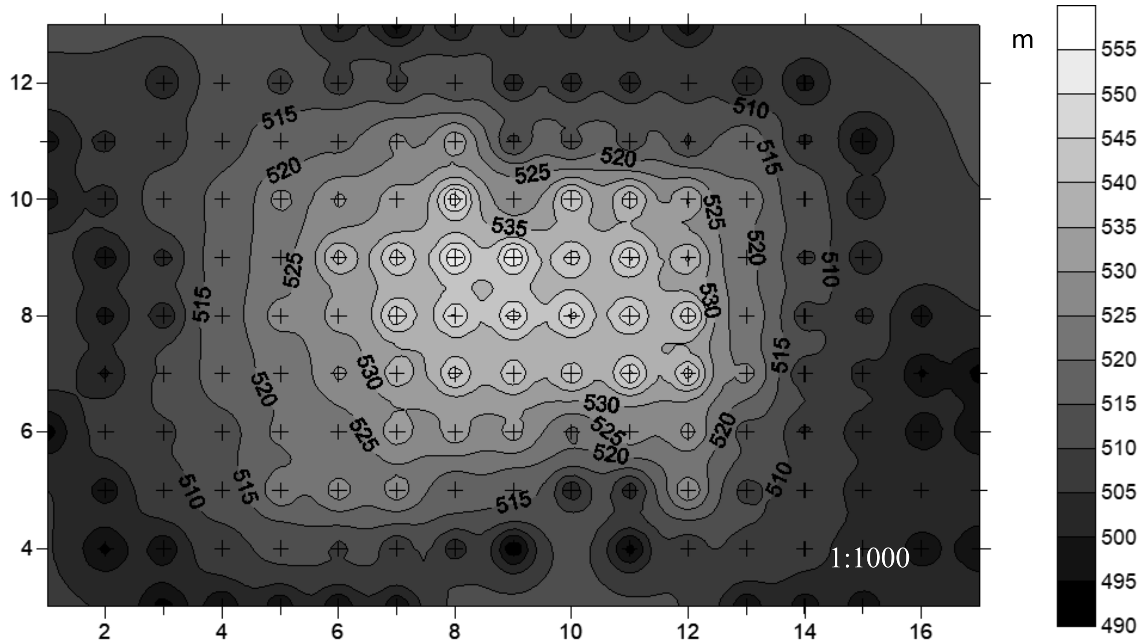

Figure 3-6: Map of the relative depth of the bottom area

4. DISCUSSION AND CONCLUSIONS

Testing for normal distribution was performed using the Shapiro-Wilk, Kolmogorov-Smirnov and QQ plot tests. The inverse distance interpolation method was used when creating the depth map of the top and bottom. The results of the Shapiro-Wilk test showed that the data were not normally distributed because the obtained p-value was less than alpha, which rejects the initial null hypothesis. The coefficient of skewness is 0.4109 and shows an asymmetric curve where most of the data is on the left side, which does not match the Gaussian normal distribution curve. The results of the Kolmogorov-Smirnov test show a p-value less than alpha, which also rejects the null hypothesis that the data are normally distributed, and the coefficient of skewness is equal to the coefficient of skewness as in the Shapiro-Wilk test. Graphical data analysis was performed with the help of QQ diagram. The data were analyzed according to a reference line with a slope of 45°, the trend of which the data should follow, however, the obtained QQ diagram shows a deviation, so it can be concluded that it is not a normal distribution.

The inverse distance method was used because it is cosidered to be one of the most simple method of interpolation as it is originally an exact interpolator, which means that it keeps the input data values fixed even during interpolation. Due to the relatively simple algorithm, speed in the calculation is significant, so for smaller data sets, all available inputs can be used when interpolating (**Medved et al., 2010**).

The reason for using three different methods for the same data set is to eliminate the flaws of each separate test and thus we can get a more realistic interpretation of that data set.

The inverse distance method was used to obtain maps which resulted in variance values, and for the top it was 12.5, and for the bottom 322.3. This confirms the shape of the syncline due to the large mutual difference. Also, taking into account the large value of the variance of the subsurface, it can be concluded that the depths on the bottom differ quite a bit, and that is precisely why the data did not pass the normal distribution tests. It is precisely this dispersion of data on the bottom area and the shape of the syncline of the coal layer that can be observed on the resulting maps of the depth of the top and bottom areas, where the lighter color shows greater and the darker color smaller depths. The map was made according to the ratio of anisotropy of 1 and the angle of anisotropy 0. These coefficients define the shape that encloses the boundaries of the depths. In our maps these boundaries are in the shape of the circle, rather than ellipse because of the proper distribution of data around the chosen points.

All conducted research provided a large amount of data on the mentioned layer, which was successfully described using statistical methods and mapping. If the normal distribution of the measured data is proven by the implemented statistical methods, the behavior of the layer on its unexplored parts can be easily concluded, and mapping results are more reliable. In the case of imaginary coal seam in Siverić, where normal distribution is not proven, with auxiliary statistical analyses it was still possible to obtain a large amount of data which can greatly help with reducing costs and high time consumption of research. So, it can be concluded that all analysis of the data carried out in this paper could be very helpful.

5. References

1. BRALIĆ, N., MALVIĆ, T. (2022): Interpretation of chemical analyses and cement modules in flysch by (geo)statistical methods, example from the southern Croatia. Processes 2022, 10(5), 813, pg. 2, 8-9 p.
2. MALVIĆ, T. (2008): Primjena geostatistike u analizi geoloških podataka (*Application of geostatistics in the analysis of geological data*). Udžbenici Sveučilišta u Zagrebu. Zagreb (*in Croatian, without English abstract*), 16-18 p
3. MALVIĆ, T., MEDUNIĆ, G. (2015): Statistika u geologiji *(Statistics in geology)*. Udžbenici Sveučilišta u Zagrebu. Zagreb *(in Croatian, without English abstract),* 12 p
4. MEDVED, I., PRIBIČEVIĆ B., MEDAK, D., KUZMANIĆ, I. (2010.) Usporedba metoda interpolacije batimetrijskih mjerenja za praæenje promjena volumena jezera (*Comparison of bathymetric interpolation methods measurements to monitor lake volume changes)* Geod. list 2010, 2, 71–86, Zagreb *(in Croatian, with English abstract),* 74-75 p
5. PFAFF, S., (2012): Osnove statistike (*Basics of statistics*). Zagreb: Element. (*in Croatian, without English abstract*), 23 p
6. ZUPANIČ, J., BABIĆ, LJ., (2011): Sedimentary evolution of an inner foreland basin margin: Palaeogene Promina Beds of the type area, Mt. Promina *(Dinarides, Croatia)* 102 p

Internet sources:

1. URL 1: https://www.enciklopedija.hr/clanak/63913 (accessed 18th January 2024)
2. URL 2: https://www.enciklopedija.hr/natuknica.aspx?ID=21410 (accessed 9th January 2024)
3. URL 3: https://builtin.com/data-science/shapiro-wilk-test (accessed 12th January 2024)
4. URL 4: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35g.htm (accessed 20th January 2024)
5. URL 5: https://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm (accessed 18th January 2024)
6. URL 6: https://library.virginia.edu/data/articles/understanding-q-q-plots (accessed 18th January 2024)
7. URL 7: https://www.statskingdom.com/shapiro-wilk-test-calculator.html (accessed 12th January 2024)

SAŽETAK

Statistička analiza i kartiranje metodom inverzne udaljenosti zamišljenog sloja ugljena, jednostavna studija slučaja preporučenih metoda

Za ispravno opisivanje ležišta mineralnih sirovina u svrhu eksploatacije potrebna je velika količina geoloških podataka o navedenom ležištu. U slučaju podzemne eksploatacije povećava se količina potrebnih podataka, a sama istraživanja su složenija. Ovaj rad opisuje zamišljeni ugljeni sloj koristeći simulirane podatke i rezultate primarnog istraživanja. Navedeni sloj ugljena nalazi se duboko ispod površine, što otežava dobivanje potrebnih podataka, zbog čega je potrebna pravilna prostorna analiza. Za mjerenje prostorne raznolikosti mjerenim uzorcima, primjenjuju se statističke metode istraživanja, a rezultati se opisuju prema nesigurnosti procijenjenih točaka ili blokova. Također je potrebno promotriti vrstu distribucije podataka, za što su korišteni Shapiro-Wilkov, Kolmogorov-Smirnovljev test i QQ dijagram, koji se međusobno nadopunjavaju. Za izradu karata i interpolaciju podataka korištena je metoda inverzne udaljenosti, a sam cilj rada je dobiti pouzdane analitičke rezultate te pokazati skup metoda za brz i pouzdan prikaz 2D karata ležišta. Postizanjem tog cilja, moguće je smanjiti troškove te vrijeme istraživanja. Sva provedena istraživanja dala su veliku količinu podataka o navedenom sloju koji je uspješno opisan statističkim metodama.

Ključne riječi: ugljeni sloj; normalna distribucija; nul-hipoteza; analiza podataka

Acknowledgment

This research is made under motivation of Tomislav Malvić, tenure Full Professor on University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, whose knowledge on geostatistics has made vast contribution.

Author's contribution

Mirta Petrović (graduate student) has contributed with a description of research goal mentioned in abstract as well as in introduction. In her involment in this research, she described methods that were used for the interpretation. **Ana Ljubić** (graduate student) provided and edited results obtained from Surfer 8 program, in written and graphic form, while designing the maps, as well as conveying discussion and conclusions, were done by both authors.