INTERPOLTION OF SULFUR DIOXIDE (SO_2) AND NITROGEN DIOXIDE (NO_2) IN YOGYAKARTA SPECIAL PROVINCE USING THE COKRIGING METHOD

Andina Aulia Pramesti[a,1,*], Nur'eni [b,2], Iman Setiawan [b,3]

[a b]Tadulako University, Soekarno Hatta St KM 10, Palu City, Indonesia
[a.,1] andinaaulia31@gmail.com;

ARTICLE INFO

Article history
Received June 7, 2023
Revised November 12, 2023
Accepted December 22, 2023

Keywords
Cokringing
Spherical Model
Nitrogen Dioxide
Sulfur Dioxide

ABSTRACT

Air is a mixture of gases found in the layer that surrounds the Earth. The kriging method is the development of the kriging method to estimate a variable that minimizes estimation errors by utilizing cross-correlation between several variables. This research was conducted using data from 45 coordinate locations for air monitoring by the Department of Environment and Forestry of the Special Region of Yogyakarta in 2017. The results showed that the best model for estimating sulfur dioxide and nitrogen dioxide air pollution was the spherical model with the smallest Mean Square Error (MSE) value, which is 317,527. Interpolation of sulfur dioxide and nitrogen dioxide content values using cokringing resulted in 100,390 new points. The value of air pollution is in the range of 0-40, which according to ISPU, means that at that point, good air quality has no impact on humans but causes a certain smell and injury to some plant species.

1. Introduction

Air is a mechanical mixture of various gases. The composition of normal air consists of 78.1% nitrogen, 20.93% oxygen, and 0.03% carbon dioxide, while the rest is argon, neon, krypton, xenon, and helium. Ambient air is free air on the surface of the Earth in the troposphere, which is within the jurisdiction of the Republic of Indonesia which is needed and affects the health of humans, living things, and other elements of the environment.[1] Research on aerosol pollution requires sufficient data, but in practice, the available data is limited due to tools, and obtaining large-scale data is expensive. Therefore it is necessary to do interpolation to collect the minimum amount of data points and estimate data at nonsampled locations without compromising data accuracy [2]. Cokriging is a method for estimating a variable that minimizes estimation errors by utilizing cross-correlation between several variables. [3]. The Kriging method has the advantage of providing more accurate estimates, and the spatial correlation between variables can be identified. Interpolation is a method of estimating an unknown value by using known values around it. The dots around it can be in regular or irregular shapes. The accuracy of the results of the interpolation depends on the number and distribution of points whose values are known and the mathematical functions used in estimating the model. [4] The Objectives of this research are to determine which model is the best to use in interpolating the content of sulfur dioxide (SO_2) and nitrogen dioxide (NO_2) air pollution in the province of the Special Region of Yogyakarta using the *kriging* method. To Determine the interpolation of sulfur dioxide (SO_2).

2. Literature Review

2.1 Spatial Data

Spatial data is data obtained from measurement results that contain information about locations, generally based on maps that contain interpretations and projections of all phenomena that are on Earth [5].

2.2 Interpolation

Interpolation is a numerical procedure that forms a calculation of the height of a certain location based on the dependency function of the location surrounding it [6].

2.3 Kriging

$$\hat{z}(u_0) = \sum_{i=1}^{n} w_i z(u_i)$$

2.4 Stationarity

The main requirement in analyzing geostatistical data using kriging is to fulfill the assumption of stationarity because it can minimize model errors. Data is said to be stationary if there are no drastic changes in the data, thus making the pattern random and not containing a trend.

$$(Z_t) = \frac{(Z_t^{\lambda} - 1)}{\lambda}, \quad \lambda \neq 0$$

2.5 Experimental Semivariogram

An experimental variogram is the estimated value obtained from sampling and is made based on the spatial correlation value between two variables separated by a certain distance of ÿ. The experimental Variogram is formulated as follows [7]

$$\hat{\gamma}(\text{h}) = \frac{1}{2|N(h)|} \sum_{i=1}^{N(h)} \left(z(s_i) - z(s_i + h)\right)^2$$

2.6 Experimental Semivariogram

After calculating the experimental semivariogram values, next is the calculation of the parameters used in the theoretical semivariograms. The following is an experimental semivariogram pattern:

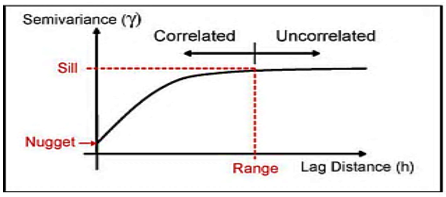

Fig 1. Experimental Semivariogram

2.7 Theoritical Semivariogram

A theoretical variogram is a variogram that is arranged according to a particular function or model. Therefore, further analysis will be carried out in structural analysis, namely comparing the Mean Square Error (MSE) values between models on theoretical semivariograms. This aims to determine the best model, which is then used to assist the estimation calculation results. The following are three theoretical semivariogram models used as comparisons:

a. Exponential Model

The semivariogram form is formulated as follows :

$$\gamma(h) = c\left[1 - exp\left[-\frac{h}{a}\right]\right]$$

b. Spherical models

The form of this semivariogram is formulated as follows:

$$\gamma(h) = \begin{cases} c\left[\frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3\right] & ,jika\ h \leq a \\ c & ,jika\ h > a \end{cases}$$

c. Gaussian models

The form of this semivariogram is formulated as follows:

$$\gamma(h) = c\left[1 - exp\left(-\left(\frac{h}{a}\right)^2\right)\right]$$

d. *Coriging*

Coriging is an interpolation method that minimizes the variance of the estimation error by utilizing cross-correlation between several variables obtained from the main variable and covariate variables. Appropriate weighting will minimize the error variance, and the estimation results will be unbiased.[7]

$$\hat{u}_0 = \sum_{i=1}^{n} a_i . u_i + \sum_{j=1}^{m} b_j . v_j$$

e. Air Pollution

Environmental Pollution is the entry or inclusion of living things, energy substances, and/or other components into the environment by human activities so that the quality decreases to a certain level which causes the environment to not function according to its form [10]. The Table 1 shows the National Ambient Air Quality Standards. The ISPU value is used to categorize air quality conditions in a place. The following are air quality categories based on ISPU values.

Table 2. ISPU Value

No	***ISPU value***	***Category***
1	0-50	Good
2	51-100	Currently
3	101-199	Not healthy
4	200-299	Very Unhealthy
5	>300	Dangerous

g. Sulfur Dioxide (SO_2)

Sulfur Dioxide (SO_2) is a group of gases in the atmosphere with an unpleasant odor which is often found in industrial areas that use coal and cork as their main energy sources. Sulfur dioxide is a form of gas resulting from volcanic activity, volcanic eruptions, natural sulfur gas sources, hot springs, and natural hot steam [11].

h. Nitrogen Dioxide (NO_2)

Nitrogen Dioxide (NO and NO_2, referred to together as NOx) is a highly reactive gas formed when oxygen and nitrogen react at high temperatures during combustion or in the event of a lightning strike [12].

3. Research Methods

3.1 Location and Place of Research

The research location was the Department of Environment and Forestry of the Special Province of Yogyakarta, and the place for the research was the Applied Statistics Laboratory, Statistics Study Program, Mathematics Department, Faculty of Mathematics and Natural Sciences, Tadulako University.

3.2 Population and Research Sample

The population in this study is all air monitoring location points in the province of the Special Region of Yogyakarta. Meanwhile, the sample for this research is data from 45 coordinate locations for air monitoring by the Department of Environment and Forestry of the Special Province of Yogyakarta in 2017.

3.3 Data Collection Procedures

The data used in this study is secondary data obtained from the results of the publication of the Yogyakarta Special Province Environment and Forestry Service in 2017. The variables used in this study are the coordinates in latitude (X), longitude (Y), Sulfur Dioxide (A)) and Nitrogen Dioxide (B).

3.4 Data analysis

a. Gatherexploration data from the Department of Environment and Forestry of the Special Province of Yogyakarta
b. Perform data stationarity testing. Stationary testing is carried out to see whether there is a trend in the data through whether the stationary assumptions are fulfilled or not. If the scatter plot on both data does not show a particular trend, then the spatial data meets the stationary assumption.
c. Perform experimental semivariogram calculations and experimental cross-variograms for each variable.
d. Calculating theoretical semivariogram and theoretical cross variogram using three models, namely spherical, gaussian, and exponential.
e. Perform co-kriging interpolation using selected theoretical cross variograms and produce interpolation plots. Making conclusions from research results.

4. Results and Discussion

4.1 Data Exploration

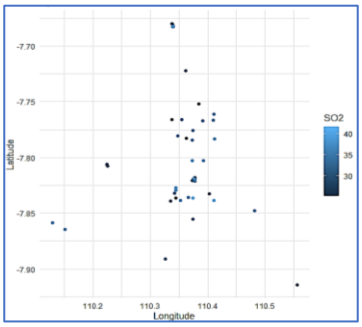

Fig 2. Plot of Sampled SO_2 Data

Figure 2. shows the value of SO_2 air pollution content From each coordinate point at 45 air monitoring points, where the blue color indicates a fairly large SO_2 air pollution content, the darker the blue color, the smaller the pollution content.

 Tadulako Science and Technology Journal
Vol 4, No. 1, December 2023, pp. 22-34
ISSN 2776-4893

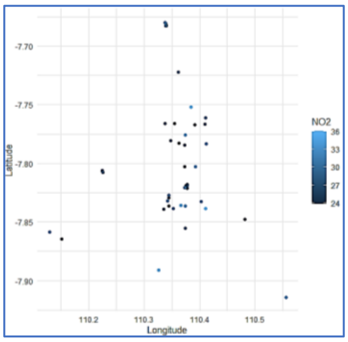

Fig 3. Sampled NO_2 Data Plot

Figure 3. shows the value of NO_2 air pollution content. From each coordinate point, where the blue color indicates a fairly large NO_2 air pollution content, the closer to the darker blue, the smaller the pollution content. Furthermore, descriptive data from sampled air pollution points SO_2 and NO_2. Based on the calculation results shown in Table 3, it can be seen that the value of the minimum SO_2 air pollution content is 25.35 µg/, which is located at one of the coordinate points (-7.836222222, 110.3446944), the minimum NO_2 air pollution content value is 23.86 µg/ which is located at one of the coordinate point (-7.766277778, 110.3552222). The maximum SO_2 air pollution content value is 41.69 µg/ located at coordinates (-7.836222222, 110.3743333), the maximum NO_2 air pollution content value is 36.05 µg/ located at coordinates (-7.838472222, 110.41125), the average SO_2 air pollution content value is $m\ ^3 m^3 m^3 m^3$34,136µg/, the average value of NO_2 air pollution content is $m^3$28.84µg/ with the variance value for $m^3$$SO_2$ of 16,287 for NO_2 of 11,215, and the standard deviation value for SO_2 is 4,036, for NO_2 is 3,349.

4.2 Stationary Assumption

In kriging, it is necessary to assume that the data is stationary for the SO2 and NO2 pollution content variables. The stationarity of the data can be seen from the graph if it makes the pattern random, does not contain a trend, and if it appears to have a constant mean and variance value, as shown in Figure 4 and Figure 5 below:

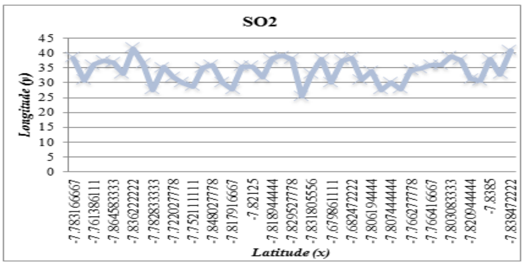

Figure 4. Plot of Sampled SO_2 Pollution Point Data

The plot above is a sample data plot of SO_2 air pollution points; from the plot above, it can be seen that the data is spread out on an average line, the pattern is random, and the line does not contain a particular trend, so it can be said that the data meets the assumption of stationarity.

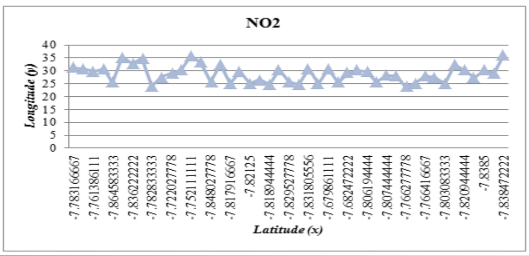

Figure 5. Data Plot of Sampled NO_2 Pollution Points

Figure 5 is a sample data plot of NO_2 air pollution points; from the plot above, it can be seen that the data is spread on an average line, the pattern is random, and the line does not contain a particular trend, indicating that the data meets the assumption of stationarity and the next steps can be taken.

4.3 Experimental Semivariogram and Cross Variogram

The experimental semivariogram value is obtained based on the distance between the same points. Equation (2.7) The experimental semivariogram is symbolized by (gamma). This value can indicate a spatial correlation between observation points. Experimental semivariogram values obtained from each variable are experimental semivariogram values of SO_2, NO_2, and cross-variogram air pollution.γ Based on Table 4, there are 15 experimental semivariogram distances. Where the value is the number of adjacent data distance pairs, the value is the average distance based on the number of data pairs, and the value is the experimental semivariogram value calculated based on equation (2.7). In the first distance, the number of pairs of adjacent distances is obtained, namely 45 data distances with an average of 0.004, so based on the calculation results, the experimental semivariogram value is 19,678.

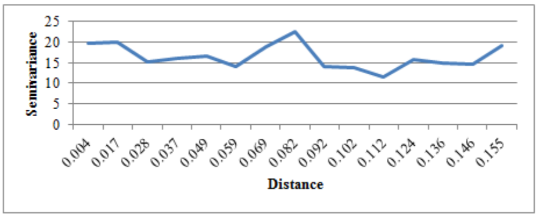

Figure 6. SO_2 Experimental Semivariogram Plot

From Figure 6, it can be seen that there are 15 distances to get the nugget, sill, and range values that will be used to calculate the theoretical semivariogram. The semivariogram value of NO_2 air pollution, the following is the resulting experimental semivariogram value. Based on Table 5, there are 15 experimental semivariogram distances. Where the value is the number of adjacent data distance pairs, the value is the average distance based on the number of data pairs, and the value is the experimental semivariogram value calculated based on equation (2.7).$npDistGamma$

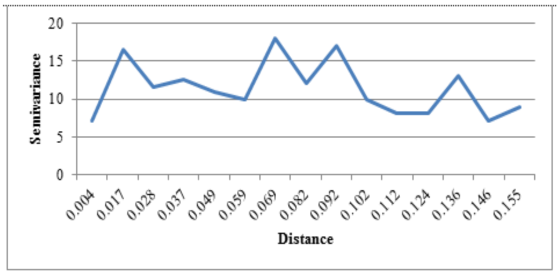

Figure 7. Experimental Semivariogram Plot NO_2

Based on the NO2 experimental semivariogram values in Table 5, when depicted in plot form, it will look like in Figure 7. The cross-variogram value is the observed value derived from two variables. Following are the cross-variogram values of the two variables, namely SO_2 air pollution and NO_2 air pollution. Based on Table 6, there are 15 experimental semivariogram distances. Where the value is the number of adjacent data distance pairs, the value is the average distance based on the number of data pairs, and the value is the experimental semivariogram value calculated based on equation (2.7).$npDistGamma$

Tadulako Science and Technology Journal
Vol 4, No. 1, December 2023, pp. 22-34
ISSN 2776-4893

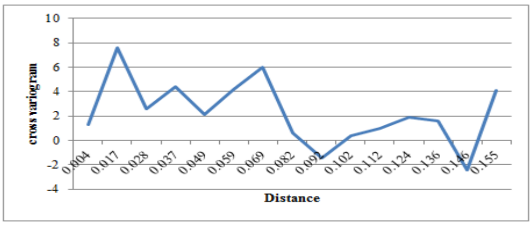

Figure 8. Experimental Cross-Variogram Plot

Based on the cross variogram values in Table 6, if depicted in plot form, it will look like in Figure 8. From Figure 8, it can be seen that there are 15 distances to get the nugget, sill, and range values that will be used to calculate the theoretical semivariogram.

4.4 Theoretical Semivariogram and Cross Variogram

The theoretical semivariogram and cross-variogram have the shape of the curve closest to the experimental semivariogram and cross-variogram model. In this study, there are three theoretical semivariogram, and cross-variogram models used, namely the Spherical, Gaussian, and Exponential models. The following is a theoretical SO2 semivariogram model.

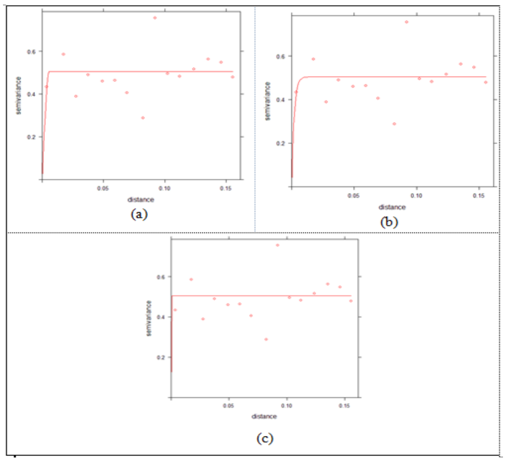

Figure 9. Theoretical SO2 semivariogram (a) Spherical Model, (b) Exponential Model, (c) Gaussian Model

Figure 9. shows the three SO_2 theoretical semivariogram models from the calculation results. The dot pattern indicates the experimental semivariogram model, while the lines indicate the theoretical semivariogram model. The spherical model has a nugget value of 0, a sill value of 0.276, and a range value of 0.006. The exponential model has a nugget value of 0, a sill value of 0.342, and a range value of 0.002, and the Gaussian model has a nugget value of 0, a sill value of 0.206, and a range value of 0.004. Furthermore, The following is a theoretical NO_2 semivariogram model.

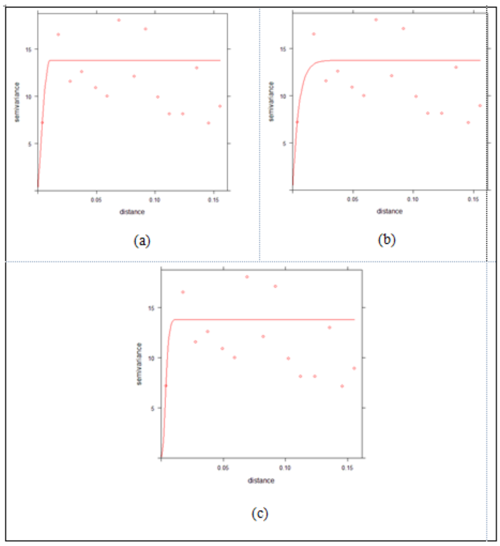

Figure 10. Theoretical NO_2 semivariogram (a) Spherical Model, (b) Exponential Model, (c) Gaussian Model

The spherical model has a nugget value of 0, a sill value of 13,822, and a range value of 0.01. The exponential model has a nugget value of 0, a sill value of 13,774, and a range value of 0.005, and the Gaussian model has a nugget value of 0, a sill value of 13,822, and a range value of 0.004. After getting the values of the nuggets, sills, and ranges from the theoretical SO_2 and NO_2 semivariograms, then look at the theoretical cross-variogram models. The following are the three theoretical cross-variogram models:

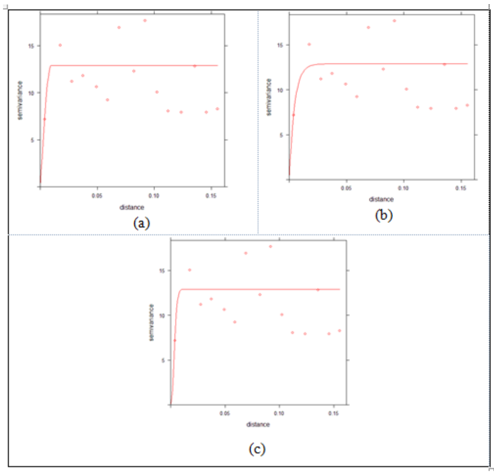

Figure 11. Theoretical NO_2 semivariogram (a) Spherical Model, (b) Exponential Model, (c) Gaussian Model

Figure 11. shows the three models of the cross-variogram from the calculation results. The spherical model has a nugget value of 0, a sill value of 12.91, and a range value of 0.01. The exponential model has a nugget value of 0, a sill value of 12,889, and a range value of 0.005. Meanwhile, the Gaussian model has a nugget value of 0, a sill value of 12.91, and a range value of 0.004.

4.5 Best Model Determination

In this study, value *means square error* (MSE) is used as a comparison to determine the best theoretical model. The following are the MSE values for each semivariogram and theoretical model. Table 7. shows the MSE values of the theoretical semivariogram and cross-variogram models. In Table 7, it is known that the smallest MSE value in SO_2 is the model *Gaussian* with a value of

Tadulako Science and Technology Journal
Vol 4, No. 1, December 2023, pp. 22-34
ISSN 2776-4893

1132,851. The smallest MSE value for NO_2 is the spherical model, with a value of 254,294. The MSE value on the smallest cross variogram is on the spherical model, namely 317,527.

4.6 Interpolation of SO_2 and NO_2 Air Pollution Content

The value resulting from interpolation with the cokringing method is calculated based on equation (2.13); the following is an interpolation of SO_2 and NO_2 air pollution content. Table 8. is the result of interpolation showing the point and value of SO_2 and NO_2 air pollution content.

4.7 Interpolation Plot of SO_2 and NO_2 Air Pollution Content

After getting the results of the interpolation that displays the point and value of SO_2 and NO_2 air pollution content. The next step is to look at the area plots at points with SO_2 and NO_2 air pollution content values marked with various colors. The following is an interpolation plot of SO_2 air pollution content:

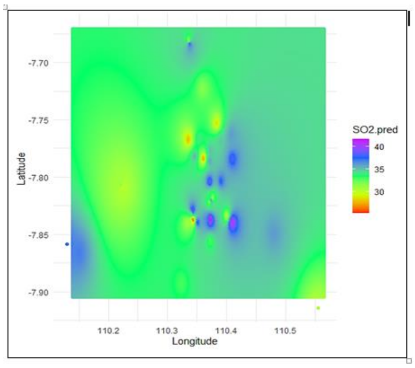

Figure 12. Interpolation Plot of SO2 Air Pollution Content

Figure 12. is the result of the interpolation of SO2 air pollution content which shows the points and values of air pollution from 45 initial points to 100,390 new points. It can be seen that the color degradation in the plot starts from orange, which indicates a very small point of air pollution value; from green to blue indicates the value of air pollution at a moderate point, while dark blue to purple indicates a high air pollution value. Figure 6.10 has an air pollution value with a range of 0-40 according to the Air Pollution Standard Index (ISPU) value (Table 2). Air quality at that point is good and has no impact on humans, but exposure to SO2 with a range of 0-40 causes injury to several plant species. Furthermore, the interpolation plot of NO2 air pollution content is as follows:

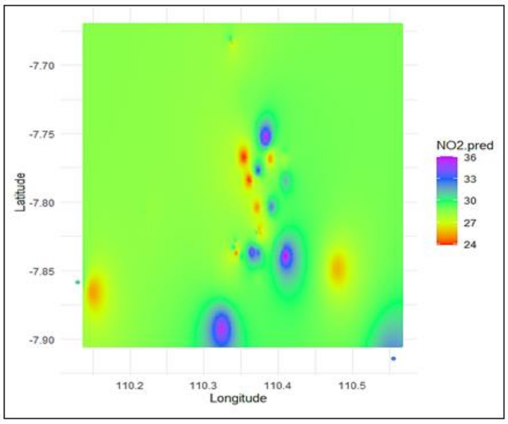

Figure 13. Interpolation Plot of NO_2 Air Pollution Content

Figure 13. is the result of the interpolation of NO_2 air pollution content which shows the point and value of air pollution from 45 initial points to 100,390 new points. It can be seen that the color

degradation in the plot starts from orange, which indicates a very small point of air pollution value; from green to blue indicates the value of air pollution at a moderate point, while dark blue to purple indicates a high air pollution value. Figure 13 has air pollution values in the range of 0-40, according to the Air Pollution Standard Index (ISPU) value (Table 2) air quality at that point is good and has no impact on humans, but exposure to NO_2 in the range 0-40 causes a slight odor.

5. Conclusion

Based on the results and discussion that has been described previously, the conclusions that can be drawn are as follows:

1. The best model for interpolating the content of SO2 and NO2 air pollution in the Special Region of Yogyakarta using the cokringing method is *spherical*, with the smallest MSE value of 317,527.
2. Interpolation of SO2 and NO2 air pollution content using the coking method yields 100,390 new points from 45 starting points. Air quality prediction with color degradation in plots ranging from orange to purple indicates a very small level of pollution to a high level of air pollution. The SO2 and NO2 air pollution values at that point have a range of 0-40; according to the Air Pollution Standard Index (ISPU) value, the air quality at that point is good, with no impact on humans, but exposure to SO2 with a range of 0-40 causes injury to several plant species and exposure to NO2 in the range 0-40 causes a slight specific odor.

References

[1] The Republic of Indonesia Government Regulation Number 41 of 1999 concerning Air Pollution Control, Article 1 paragraph (4)
[2] Memarsadeghi, N. (2004). Coriging Interpolation, Masters degree, Computer Science Department of the University of Maryland, Maryland.
[3] Isaacs, EH, and Srivastava, RM (1989). Applied Geostatistics. New York: Oxford University Press.
[4] Jaya, AJ (2002). Spatial Estimation of Altitude Distribution Ordinary Kriging Interpolation Model to Obtain Continuous Distribution of Surface Temperature, Thesis, Bogor Agricultural University, Bogor.
[5] Girsang, U. (2016). Robust Kriging (Case Study: Estimation of Nickel Content in the Morowali area). Thesis. Statistics Study Program. Mathematics Department. FMIPA. Gadjah Mada University. Yogyakarta.
[6] Pertiwi, A. (2011). Inverse Distance Interpolation Method for Elevation (Contour) Maps. Thesis. Informatics Engineering Study Program. FMIPA. Dian Nurwantoro University. Semarang.
[7] Cressie, NAC (1993). Statistics For Spatial Data. John Wiley and Sons: New York.
[8] Makridakis. (1995). *Geostatistics for Engineers and Earth Scientists*. Kluwer Academic Publishers. United States of America.
[9] Armstrong, M. (1998). Basic Linear Geostatistics. New York: Springer
[10] Decree of the Head of the Environmental Impact Management Agency No. 107 of 1997 concerning Calculation and Reporting and Information on Air Pollution Standard Index.
[11] Siregar, EM (2006). Air Pollution, Plant Responses and Their Effects on Humans. Medan: University of North Sumatra, Faculty of Agriculture, Forestry Study Program.
[12] Treatise, N. (2011). The Relationship between Air Pollutants and Land Surface Temperature and Their Distribution in DKI Jakarta, Thesis, University of Indonesia, Jakarta.

Appendix :

Table 1. National ambient air quality standards

Tadulako Science and Technology Journal
Vol 4, No. 1, December 2023, pp. 22-34
ISSN 2776-4893

No	*Parameter*	*Time*	*Quality standards*
1	Aerosol (PM10)	24 hours	150 µg/m^3
2	Carbon monoxide (CO)	1 hour	30000 µg/m^3
		24 hours	10000 µg/m^3
3	Ozone (O3)	1 hour	235 µg/m^3
		One year	50 µg/m^3
4	Sulfur Dioxide (SO2)	24 hours	365 µg/m^3
		One year	80 µg/m^3
5	Nitrogen Dioxide (NO2)	1 hour	0.25µg/m^3
		One year	100µg/m^3

Table 3. Descriptive Data

Variational	*Min*	*Max*	*Means*	*Var*	*std. Dev*
SO2	25.35	41.69	34.14	16.29	4,036
NO2	23.86	36.05	28.84	11.22	3,349

Table 4. SO2 Experimental Semivariogram Values

No	*np*	*Dist*	*Gamma*	*id*
1	45	0.004	19,678	SO2
2	44	0.017	19,844	SO2
3	46	0.028	15.207	SO2
4	91	0.037	16.148	SO2
5	66	0.049	16,496	SO2
6	66	0.059	14,094	SO2
7	63	0.069	18,928	SO2
8	45	0.082	22,574	SO2
9	35	0.092	14.144	SO2
10	48	0.102	13.78	SO2
11	37	0.112	11,478	SO2
12	42	0.124	15,627	SO2
13	31	0.136	14,949	SO2
14	65	0.146	14,699	SO2
15	63	0.155	19,251	SO2

Table 5. Experimental NO2 Semivariogram Values

No	*np*	*Dist*	*Gamma*	*id*
1	45	0.004	7.11	NO2
2	44	0.017	16,543	NO2
3	46	0.028	11,594	NO2
4	91	0.037	12,599	NO2
5	66	0.049	10,897	NO2
6	66	0.059	10006	NO2
7	63	0.069	18046	NO2
8	45	0.082	12,098	NO2
9	35	0.092	17.113	NO2
10	48	0.102	9,942	NO2
11	37	0.112	8.115	NO2
12	42	0.124	8.149	NO2
13	31	0.136	12.99	NO2
14	65	0.146	7.136	NO2
15	63	0.155	8,937	NO2

Table 6. Experimental Cross Variogram Values

No	*np*	*Dist*	*Gamma*	*id*
1	90	0.004	1,278	NO2. SO2
2	88	0.017	7,571	NO2. SO2
3	92	0.028	2,556	NO2. SO2
4	182	0.037	4,421	NO2. SO2
5	132	0.049	2,099	NO2. SO2
6	132	0.059	4.17	NO2. SO2
7	126	0.069	6.01	NO2. SO2
8	90	0.082	0.572	NO2. SO2
9	70	0.092	-1,478	NO2. SO2
10	96	0.102	0.385	NO2. SO2
11	74	0.112	0.994	NO2. SO2
12	84	0.124	1876	NO2. SO2
13	62	0.136	1617	NO2. SO2
14	130	0.146	-2,427	NO2. SO2
15	126	0.155	4,095	NO2. SO2

Table 7. MSE Semivariogram and Cross Variogram SO2 and NO2 Values

Model	*MSE Value*		
	SO2	NO2	Cross Variogram
Spherical	1139,119	253,864	317,527
Exponential	1141553	254,294	318,123
Gaussian	1132851	317,527	339,319

Table 8. Interpolation of SO_2 and NO_2 Pollution Content

Tadulako Science and Technology Journal
Vol 4, No. 1, December 2023, pp. 22-34
ISSN 2776-4893

Point	*Latitudes*	*Longitude*	*NO2*	*SO2*
1	-7,783	110,412	31.14	38.34
2	-7,833	110,402	30.76	30.81
3	-7,761	110,410	29.69	36.14
...	...	...	...	...
...	...	...	...	...
100388	-7,671	110,563	28.87	34,354
100389	-7,671	110,564	28,871	34,354
100390	-7,671	110,565	28,871	34,353