ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 3, Issue 1, April 2023, pp. 48-58

Prediction of PM2.5 in DKI Jakarta Using Ordinary Kriging Method

Shafira Salsabilla [a,1] , Amadea Fitri Syaharni [a,2], Nur Chamidah [a,3,*]

[a] Statistics Study Program, Department of Mathematical, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia 60115
[1] shafira.salsabilla-2019@fst.unair.ac.id; [2] amadea.fitri.syaharani-2019@fst.unair.ac.id; [3] nur-c@fst.unair.ac.id *
* Corresponding author

ARTICLE INFO

Keywords
Ordinary kriging
Predict on
PM2.5 levels
DKI Jakarta

ABSTRACT

Air pollution is a serious matter that must be addressed promptly and quickly. One of the most dangerous pollutants in the air is PM2.5. This pollutant is particulates dust measuring 2.5 micrometers. PM2.5 can cause environmental and health problems such as acute respiratory infections, lung cancer, cardiovascular cancer, and premature death. Air pollution occurs in big cities such as the capital city of Indonesia, DKI Jakarta, which is the city with the highest PM2.5 levels in Indonesia. There are 6 six stations in DKI Jakarta that measure PM.2.5 level at 6 areas. The ordinary kriging is one of spatial methods that can be used to predict PM2.5 level in outside the existing stations, for example in the Pulogadung industrial area. This area was selected because there are many factories in this area that can increase levels of PM2.5 in the air. To predict the concentration of PM2.5 in one area could be done by calculating the surrounding PM2.5 concentrations that were not available to measure air quality. In study, we use mean an absolute percentage error (MAPE) value to evaluate Ordinary Kriging performance for predicting PM2.5 level in DKI Jakarta.

1. Introduction

Clean air is the right of every human being, but the low quality of air has become a concern for several countries in the world in the midst of the climate crisis. Based on the iQAir 2022 New Delhi city, India ranks first for the place with the worst air quality in the world. Indonesia is ranked 17th in the world with the worst air pollution and 1st in Southeast Asia. The three cities in Indonesia that have the lowest quality are Jakarta, Surabaya, and Bandung where all three are big cities in Indonesia. The iQAir website assigns a country's rating to healthy air quality by calculating the levels of PM2.5 contained in the air. The capital city of Indonesia, DKI Jakarta ranks first with the highest PM2.5 levels in Indonesia. Air pollution is caused by movable sources and immovable sources which include the transportation, industrial, and domestic sectors. Other factors that indirectly affect the occurrence of air pollution are population growth, the rate of urbanization. high, unbalanced spatial development and low level of public awareness regarding air pollution [1].

p-ISSN 2798-253X
e-ISSN 2798-3153

One of the parameters of air pollution according to Government Regulation of the Republic of Indonesia No. 41 of 1999 is particulate matter (PM) or particulate dust, which is distinguished based on its size, namely PM2.5 and PM10. PM2.5 is airborne particles smaller than 2.5 micrometers (micrometers). Various materials contained in PM2.5 can cause respiratory tract disorders such as acute respiratory infections (ARI), lung cancer, cardiovascular disease, premature death, and chronic obstructive pulmonary disease [2]. Particulates can settle in the respiratory tract through various physical mechanisms, including sedimentation, impaction, diffusion, interception, and electronic precipitation [3]. With the adverse effects caused by PM2.5, the World Health Organization (WHO) has set a limit for PM2.5 levels. There is a limit for PM2.5 levels set by WHO, areas with PM2.5 levels that exceed the normal limit need to be handled to improve air quality.

One of the methods used to solve the problem of regionalized variables is the kriging method. The kriging method was developed by Georges Matheron which was applied to geostatistical data [4] . Solving problems related to geostatistics uses many different kriging methods, distinguished based on the presence or absence of a population average between simple kriging, ordinary kriging and universal kriging with different functions [5]. To determine the estimation of PM2.5 levels in an area, geostatistical spatial data can be used, namely the Ordinary Kriging method. Ordinary Kriging is a geostatistical method used to predict spatial data with an unknown population mean and assumed to be stationary at a certain location [6]. This method can provide an estimate that is used to estimate the value at a measurement sample point using a semivariogram model that aims to visualize, model and exploit the spatial autocorrelation of regionalized variables. This method compares the experimental semivariogram values with several theoretical semivariogram models, namely spherical, exponential and gaussian.

The purpose of this study was to estimate the best model formed from the three semivariogram models used. In addition, this study also aims to determine the estimated value of PM2.5 levels in the crocodile hole area and find out the right categories and recommendations to minimize air pollution due to PM2.5 levels in DKI Jakarta. There are some previous studies that have examined using the same method and object in different places. The reason for choosing the place is because it wants to know the estimated value of PM2.5 levels in other industrially dense areas in the city with the highest pollution levels in Indonesia. The results of the estimation of PM2.5 levels in DKI Jakarta using the ordinary kriging method are expected to explain the categories of air pollution and provide appropriate recommendations to the local government to minimize air pollution due to PM2.5 levels. This can be seen by comparing and selecting the best model from the three existing semivariogram models.

2. Related Works and Literature Review

2.1. PM2.5

Particulate matter (PM) are particles found in the air including dust, soot, dirt, smoke, and water droplets [7] . PM2.5 is a very fine particle below 2.5µm which is so small that it can only be seen with an electric microscope. In general, the particles contained in PM2.5 contain higher heavy metals than the particles contained in the coarse filter [8]. Sources of PM2.5 originating from human activities include engine combustion, industrial processes, fireworks smoke, and cigarette smoke. Meanwhile, natural sources of PM2.5 include dust, exhaled salt, plant spores, pollen, and smoke from fires. The main sources of PM2.5 may vary by season, weather, climate, level of urbanization, country and region.

PM2.5 particles can be formed if there is a reaction between certain gases such as sulfur dioxide, ammonia, nitrogen oxides, carbon black, and volatile organic compounds. The very small size of PM2.5 particles causes these particles to enter the tissue in the lungs so that they can cause health problems, such as ARI, symptoms of anemia, heart disease, growth retardation, weak immune system, symptoms of autism, lung cancer. lung disease, even premature death. In addition to having a negative impact on human health, PM2.5 particles can also harm the environment if excessive. Some of the consequences include causing damage to buildings, increasing levels of ozone damage, formation of acid deposition.

p-ISSN 2798-253X
e-ISSN 2798-3153

The safe level of PM2.5 according to WHO after being updated on 22 September 2021 is 15 gram/m ^{3} in a day. Meanwhile, in a year the safe level of PM2.5 particles is 5 gram/m ^{3} . PM2.5 particles can be reduced in several ways, including slowing the growth of mold by controlling humidity, not smoking in the house, reducing burning of candles and incense, increasing electricity and fuel efficiency, limiting the burning of wood or garbage, using a mop/wet cloth. /HEPA cleaner, does not use a generator or indoor barbecue, and does not use external heating [7]. BMKG Indonesia (2015) stipulates the category of PM2.5 levels in the air and is a danger to human health. Threshold Limit Value (NAV) is the air pollution concentration limit that is allowed to be in the ambient air. NAV for PM2.5 pollutant is 65 grams/m^3 . The following are categories of PM2.5 levels according to BMKG in the Air Quality Index (AQI).

Table 1. Level Categories of PM2.5

Level	Concentration Level Range
Well	0-15 gram/m ^{3}
Currently	16-65 gram/m ^{3}
Not healthy	66-150 gram/m ^{3}
Very Unhealthy	151-250 gram/m ^{3}
Dangerous	>250 gram/m ^{3}

2.2. Kriging

The kriging method is a method that is often used in the field of geostatistics to estimate the alleged unsampled location points based on the surrounding sampled points [9]. The Kriging method is divided into three types of kriging, namely, Simple Kriging, Ordinary Kriging, and Kriging with a trend/Universal Kriging. In the Simple kriging method, it is assumed that the average is known and has a constant value. The ordinary kriging method is the simplest method in geostatistics. This method is assumed to have a population mean of unknown but constant. The data used in this method must not contain trends and contain no outliers. The Universal Kriging method is a general form of simple kriging and has a trend tendency. This method can solve the stationary problem from the data taken.

Ordinary Kriging is the interpolation of a variable value at a certain point by observing similar data in other locations . Ordinary Kriging predicts a variable at a certain point by observing similar data in an area. The Ordinary Kriging method is a Kriging method that produces an estimator that is BLUE (Best Linear Unbiased Estimator). It means that it has the smallest variance compared to other estimators. The data used in the Ordinary Kriging method is spatial data with an unknown population mean and is assumed to be stationary [6]. Ordinary kriging weight satisfies the unbiased property with where n is the number of known samples [10]. In using Ordinary Kriging , the following steps are required:

a. Determining the empirical semivariogram
b. Determining the numeric semivariogram
c. Calculating semi variance
d. Counting predictions

2.3. Related literature

There are several other studies that used the previous Kriging method. Rachmawati et al conducted research in 2021 with the title Estimation of dissolved oxygen using spatial analysis based on ordinary kriging method as effort to improve the quality of Surabaya's river water [11] . This study used the Ordinary kriging method to predict dissolved oxygen levels as an effort to improve river water quality in Surabaya. In the process of estimating dissolved oxygen levels, the researcher used the weights obtained from the spherical semivariogram model. The results of the comparison of the estimated results and the actual value of dissolved oxygen levels were not significantly different. This shows that the ordinary kriging method used in this study has high accuracy. The results of this study also provide data that can be used as a reference for the Surabaya government in evaluating the quality of river water in Surabaya to make it even better. Based on good estimation results, it would be better if this was also done with other observational data to estimate dissolved oxygen levels in

p-ISSN 2798-253X
e-ISSN 2798-3153

rivers in other areas so that the estimation results can be used as evaluation material for the local government.

Another study with the same method was conducted by Wulansari *et al.* in 2021 with the title *Prediction concentration of PM2.5 in Surabaya using ordinary Kriging method* [12]. This study used the Ordinary Kriging method to estimate PM2.5 concentrations in industrial areas of Surabaya and several other areas in Surabaya. In the process of estimating dissolved oxygen levels, the researcher used the weight values obtained from the spherical semivariogram model. The results of the study stated that the accuracy of the Ordinary kriging method in estimating PM2.5 concentration levels in the industrial area of Surabaya was very high as evidenced by the Mean absolute percentage error value which was in the high or very good category. The results of the estimated PM2.5 concentration level also showed good results in accordance with the conditions when the data was taken, namely during the COVID-19 pandemic. With these satisfactory results, it would be nice if the same estimation process could be carried out with different data, for example estimating the levels of other types of pollutants produced from industrial processes and endangering the health of the surrounding environment.

Another study was conducted using the Co-Kriging method by Safitri *et al* in 2021 with the title *Co-Kriging method performance in estimating the number of COVID-19 positive confirmed cases in East Java Province* [13]. This study uses the Co-Kriging method to estimate the number of positive cases of COVID-19 in East Java with the second variable, namely the latest data on the positive number of COVID-19 on July 21, 2020. The estimation process produces a Mean Absolute Percentage Error (MAPE) value of 18.17% which is in the range of 10%-20%. This shows that the Co-Kriging method has a good performance for estimating COVID-19 cases in East Java. The steps taken in the Co-Kriging method in estimating an estimated value are similar to the Ordinary Kriging method. The difference between the two methods is the variables used in the estimation process. The accuracy and performance of each method can always be tested using different data and seeing the resulting MAPE value.

3. Materials & Methodology

This section describes the data, testing steps, and methods used in the research.

3.1. Ordinary Kriging

Ordinary kriging is the simplest method of geostatistical data. This method has the assumption that the population mean (mean) is unknown but has a constant value and the spatial data used does not contain trends and outlier [14].

Ordinary kriging interpolation can be obtained by (1):

$$\hat{Z}(s_0) = \lambda^T . Z \quad (1)$$

To form a distance matrix between observation locations, it can be obtained by calculating the Euclidian distance which is located in the latitude and longitude coordinates. Obtained by using (2):

$$h_{ab} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \quad (2)$$

where $a, b = 1,2, \dots, n$ and $U_a = (x_a, y_a)$ are the spatial coordinates (Latitude, Longitude) at the th location U_a, where $a = 1,2, \dots, n$.

Meanwhile, to calculate the distance between the observation location and the estimation location, (3) can be used:

$$h_{a0} = \sqrt{(x_a - x_0)^2 + (y_a - y_0)^2} \quad (3)$$

where $a, b = 1,2, \dots, n$

 p-ISSN 2798-253X
e-ISSN 2798-3153

3.2. Variogram and Semivariogram

The variogram is a measure of the spatial variance. The variogram is used to determine the distance at which the observed data values become independent or have no correlation [15]. The semivariogram is half of the variogram with the symbol. Semivariogram is used to measure spatial correlation in the form of error variance at location u and location $u + h$.

Experimental semivariogram is a semivariogram obtained from measurement data or sample [10] which is formulated as in (4).

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(s_i + h) - Z(s_i)]^2 \tag{4}$$

Where $\gamma(h)$ denotes variogram value with distance h, $N(h)$ denotes the number of data pairs at distance h, $Z(s_i)$ denotes the value of the distance of the observation location to the i-th location, and $Z(s_i + h)$ denotes the value of the observation at the i-th location with respect to the additional distance h.

After obtaining the experimental semivariogram value, further analysis was carried out to calculate the theoretical semivariogram value. There are several theoretical semivariogram models that are used as a comparison with experimental semivariogram values. These models include:

1. Exponential Model

   The function of the exponential model is expressed by:

$$\gamma(h) = \begin{cases} 0 & , for\ h = 0 \\ C_0 + C\left(1 - exp^{\frac{h}{a}}\right) & , for\ h \neq 0 \end{cases} \tag{5}$$

2. Gaussian Model

   The function of the Gaussian model is expressed by:

$$\gamma(h) = \begin{cases} 0 & , for\ h = 0 \\ C\left(1 - exp\left(-\frac{h^2}{a^2}\right)\right) & , for\ h \neq 0 \end{cases} \tag{6}$$

3. Spherical Model

   The function of the exponential model is expressed by:

$$\gamma(h) = \begin{cases} C\left[\left(\frac{3}{2}\right)\left(\frac{h}{a}\right) - \left(\frac{1}{2}\right)\left(\frac{h}{a}\right)^3\right] & , for\ h \leq a \\ C & , for\ h > a \end{cases} \tag{7}$$

Where h is the distance between locations, $C_0 + C$ is the value of spatial data variance (sill), and a is the distance when the auto covariance and cross covariance values reach the maximum point.

3.3. Calculation Accuracy

Every result of the estimation does not have a perfect level of accuracy. There will always be errors in the estimation results obtained. Therefore, a method is needed to calculate the level of accuracy in a model.

One of them is to look for Absolute Percentage Error (APE). is an absolute value with (8) [12].

$$APE = \frac{|Approximate\ value - Precise\ value|}{Precise\ value} \times 100 \tag{8}$$

After calculating the APE value, it can be continued by calculating the Mean Absolute Percentage Error (MAPE) with (9) [16]:

$$MAPE = \frac{\sum_{t=1}^{n}\frac{|x_t - y_t|}{x_t}}{n} 100\% \tag{9}$$

p-ISSN 2798-253X
e-ISSN 2798-3153

Where n is the amount of data, t is the forecasting period, x_t is the actual data or original data in period t, y_t is the t period forecasting data.

MAPE values are divided into several categories that interpret the level of model accuracy. These categories include [12]:

Table 2 . MAPE Value Interpretation Category

MAPE Value	Prediction Accuracy
MAPE < 10%	High forecasting model capability
10% < MAPE < 20%	Good forecasting model ability
20% < MAPE < 50%	The ability of the forecasting model is feasible
MAPE > 50%	Poor forecasting model ability

3.4. Spatial Outlier Detection

A spatial outlier can be defined as an observation location value that is inconsistent or highly deviant (extreme) from the value of other observation locations [14]. One method used to detect outliers is the spatial statistics Z test. Spatial statistics Z test is obtained in equation (9) [17].

$$Z_{s(x)} = \left|\frac{s(x)-\mu_s}{\sigma_s}\right| > \theta \tag{9}$$

Where $s(x)$ is the difference between the observed value of location x and the average observed value of the location close to x, μs is the mean of $s(x)$, σs is the deviation standard of $s(x)$, θ is the Z table value for a certain level of significance. If *Zs(x)* $> \theta$, then x is detected as an outlier.

3.5. Data and step of analysis

The data used in this study comes from the IQAir website [18] which provides news and data on world air quality. One of the Indonesian government agencies that contributes to this website is the Meteorology, Climatology and Geophysics Agency (BMKG). The unit of observation in this study is all areas in DKI Jakarta in 2022 which consists of 6 air quality monitoring stations. There are 3 parameters, namely (X, Y) as the coordinates of the longitude and latitude of the measurement location and the content of PM2.5 in the air of DKI Jakarta

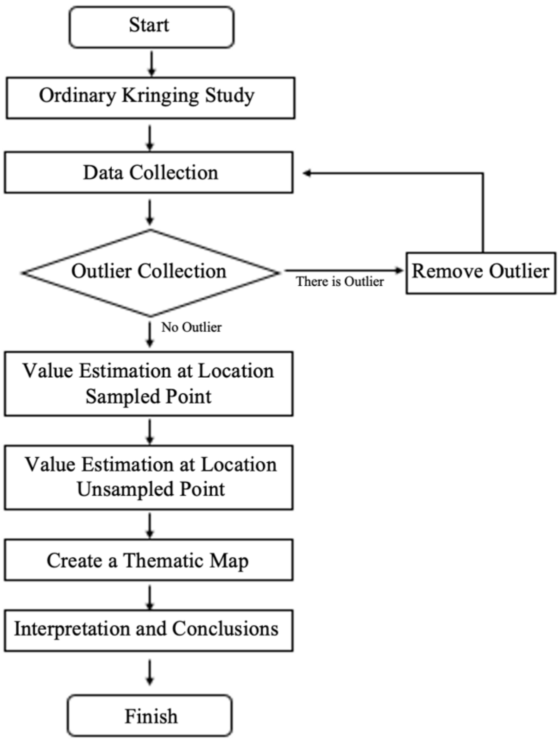

Fig. 1 Flowchart of the research.

p-ISSN 2798-253X
e-ISSN 2798-3153

Fig. 1 is a flowchart of the steps that would be taken in this research. Several stages in this research include:

1. Entering CO gas pollutant data on Microsoft excel and then analyzed descriptively.
2. Entering the latitude and longitude coordinates of the observation location $u_a = (x_a, y_a)$as well as the latitude and longitude coordinates for the estimated location $u_0 = (x_0, y_0)$, as well as the variable $Z(u_a)$ with a = 1,2,...,n.
3. Forming a distance matrix between observation locations by calculating the Euclidian distance between observation locations based on geographic position to form a distance matrix between observation locations. The Euclidian distance between the a-th location which is located on latitude and longitude coordinates $(x_a, y_a)$to the b-th location which is located on the latitide and longitude coordinates (x_b, y_b) based on (4).
4. Calculates the distance between the observation locations which are located in latitude and longitude coordinates $(x_a, y_a)$to the estimation results located in latitude and longitude coordinates (x_0, y_0) based on (3).
5. Calculating experimental semivariogram based on (4).
6. Calculating the value of the theoretical semivariogram model with the three models, namely the Exponential model based on (5), Spherical model based on (6), and using the Gaussian model based on (7).
7. Calculating the MSE value between the experimental semivariogram value and the theoretical variogram value to determine the most appropriate model used in kriging analysis.
8. Calculating the matrix value P and S.
9. Estimating PM2.5 levels in each location by calculating the weighted value.
10. Estimating PM2.5 levels at the alleged location, namely at the crocodile hole point in the same way as before.
11. Calculates the value of Absolute Percentage Error (APE) for each location based on (8).
12. Calculates the Mean Absolute Percentage Error (MAPE) value based on (9).

4. Results and Discussion

4.1. Statistics Descriptive

The data are secondary data from six observation stations in Jakarta which were obtained from the IQAir website and presented in a descriptive analysis as follows.

Table 3. Statistics Descriptive from PM2.5 Data Data

Criteria	Score
Minimum	26.3
Maximum	44.3
Average (*Mean*)	34.884
Standard Deviation	7.427629
Variant	55.1697

Based on the **Table 3**, it is known that the highest PM2.5 concentration was observed in the AHP Capital Place area with a PM2.5 concentration of 44.3 gram/m^3 then followed by the Kemayoran area with a PM2.5 concentration of 40.5 gram/m^3. Meanwhile, the lowest PM2.5 concentration was in the West Pejanten area, which was 26.3 gram/m^3. From the 6 monitoring stations, it is known that the average PM2.5 level is 34.884 gram/m^3 and the standard deviation is 7.427 gram/m^3. Based on the PM2.5 concentration index as shown in Table 2 it was concluded that at all points PM2.5 levels were still in the range of 16-65 gram/m^3, which was a moderate level.

4.2. Process Analysis

The next step is to determine the Euclidian distance from the coordinates of the observation data. There are 15 data pair points that would be used in the estimation process. After determining the Euclidian distance, the experimental Semivariogram calculation is carried out as in the formula

p-ISSN 2798-253X
e-ISSN 2798-3153

number (4) . Furthermore, plotting is carried out between the Euclidian distance and the experimental semivariogram value as shown in Fig. 2.

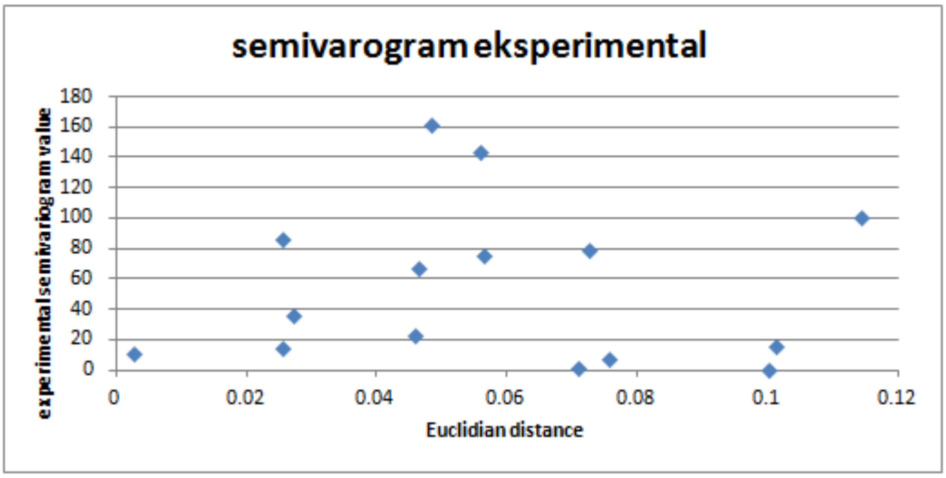

Fig. 2 Plot of the experimental semivariogram.

Based on the plot it can be seen that the value of Sill C is 55.16967, obtained from the variance of the spatial data. Meanwhile, the Range value is 0.048297, which is the highest Euclidean distance from the experimental semivariogram value.

After calculating the experimental semivariogram value, the next step is to calculate the theoretical semivariogram value. There are three models for the theoretical semivariogram, namely Spherical, Exponential, and Gaussian. From the three models, the MSE value obtained by comparing the experimental semivariogram value with the theoretical semivariogram value would be calculated. The best model would be determined by choosing the smallest MSE value from the three models. The results of the calculation of the theoretical semivariogram values are presented in the table below.

Table 4. Calculation Results for Exponential, Spherical, and Gaussian Models

Observation data	Exponential	Spherical	Gaussian
1	-2.948229281	4.304336	0.149323
2	-38.2040859	39.52619	13.34317
3	-38.34596211	39.61695	13.41005
4	-41.41700877	41.49974	14,85229
5	-86.53281632	54.9088	32.51043
6	-88.36256818	55.01284	33.05596
7	-94.79591238	55.16967	34,87358
8	-119.7741024	55.16967	40.60529
9	-121.7272748	55.16967	40.97552
10	-182.956002	55.16967	48.6694
11	-191.6633777	55.16967	49.32509
12	-207.6874017	55.16967	50.34821
13	-382,4363807	55.16967	54.41255
14	-392.73277	55.16967	54.48245
15	-531.7600071	55.16967	54.96378

Based on the three models, the Mean Square Error (MSE) value is calculated to determine the best model among the three models by choosing the lowest MSE value. After doing the calculations, it was found that the Spherical model has the lowest MSE value so that the Spherical semivariogram

p-ISSN 2798-253X
e-ISSN 2798-3153

model is used in the process of estimating the PM2.5 concentration level in Jakarta in this study. Wulansari *et al.*(2021).

4.3. Results

In the estimation process, the next step is to determine the P matrix and S matrix which would be used to find the weights in the process of estimating PM2.5 concentration levels. Matrix P is a matrix in which each element contains a semivariogram value of the distance between monitoring station locations. The S matrix is a matrix whose elements contain the semivariogram value of the distance between each station and the estimated area.

$$\underbrace{\begin{bmatrix} \gamma(h_{11}) & \cdots & \gamma(h_{1n}) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(h_{n1}) & \cdots & \gamma(h_{nn}) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ p \end{bmatrix}}_{Q} = \underbrace{\begin{bmatrix} \gamma(h_{10}) \\ \vdots \\ \gamma(h_{n0}) \\ 1 \end{bmatrix}}_{S}$$

$$Q = S$$

$$Q = P^{-1} S$$

The above is an illustration of the P matrix and the S matrix which would be used to find the weights in the process of estimating PM2.5 concentration levels.

After calculating the weighting matrix which is the product of the inverse matrix P and matrix S, an estimate is made for each point of observation location. The estimation results and Absolute Percentage Error (APE) values from each observation location are presented in Fig. 3.

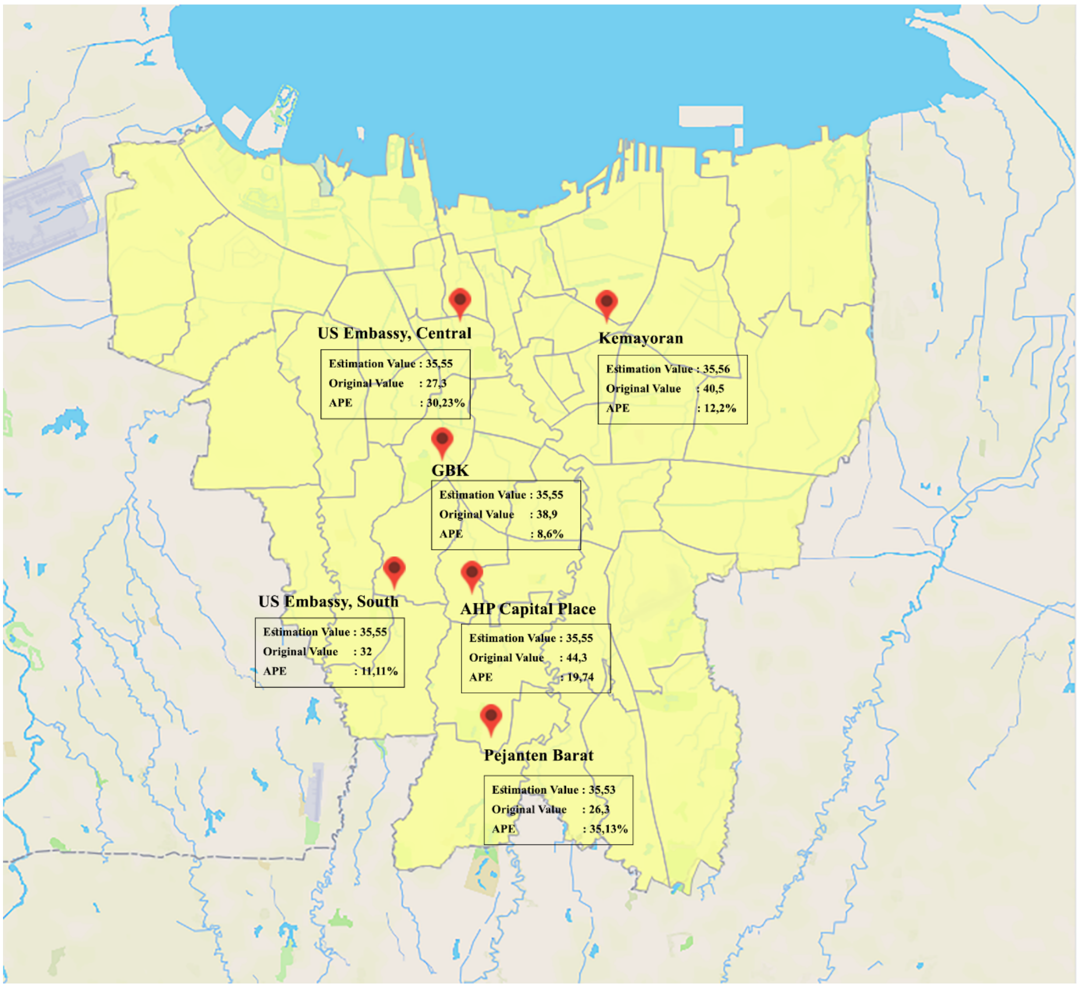

Fig. 3 Map of estimation results, actual value, and APE value for each observation location.

Based on the APE value of each observation location calculated using the Spherical semivariogram model, then the Mean Absolute Percentage Error (MAPE) value is calculated to determine the accuracy of the estimate. The MAPE value obtained is 19.50% . The MAPE value is in the *Good category* so it can be seen that the accuracy of the estimation process results is in a good level. Based on the map, it is known that the estimated value of the Jakarta GBK area has the lowest APE value of 8.6% with an estimated value of 35.55561 gram/m^3 . The area with the highest APE value is in the West Pejanten area of 35.13% with an estimated value of 35.53957 gram/m^3 .

4.4. Discussion

Based on the results from MAPE which has a good category, the researchers decided to predict the concentration of PM2.5 in the Lubang Buaya area, East Jakarta. Researchers chose this area because the area is located around an industrial area that produces smoke and air waste produced by industry and wants to know whether the smoke and air waste produced affects the health of the surrounding environment. To predict PM2.5 concentrations in the Lubang Buaya area, the researchers used the same steps to estimate PM2.5 concentrations in the previous area. Six sample points were used with known Sill values and range values previously in this study, which were 55.169 and 0.0482, respectively. Based on this value, the lowest MSE value was determined, namely from the Spherical semivariogram model among the three theoretical semivariogram models. The spherical model was used to predict PM2.5 concentrations in the Lubang Buaya area, East Jakarta. The results of the estimated PM2.5 concentration in the Lubang Buaya area are 68.69841 gram/m^3 with an Absolute Percentage Error (APE) value of 0.23% . Based on the PM2.5 indicator table from the Indonesian BMKG, the PM2.5 concentration level includes the unhealthy level of health because it is in the range of 66-150 gram/m^3 .

5. Conclusion

The Ordinary Kriging method has a high enough accuracy to predict PM2.5 concentrations in Jakarta with a MAPE value of 19.5%. Prediction result of PM2.5 concentration in Lubang Buaya area using ordinary kriging method is 68.69841 gram/m^3 and is categorized as Unhealthy because it is in the range of 66-150 gram/m^3 . This is also the impact of the return to normal activities of the people of Jakarta after going through the PPKM phase during the high period of the COVID-19 pandemic that occurred in Indonesia, causing the concentration of air in Jakarta to increase again.

This research was conducted with data on May 23, 2022, which includes the period after PPKM which occurred during the COVID-19 Pandemic. The accuracy of the ordinary kriging method to predict PM2.5 concentration levels in Jakarta can be retested by using data before, during and after PPKM. By doing this, researchers will find out the causes and impacts generated by the industrial production process during the three phases. So that researchers can provide advice to the Jakarta government or industry players in minimizing the high levels of concentration in Jakarta.

References

[1] A.G. Simandjuntak ,"Pencemaran Udara," *Buletin Limbah*, vol. 11, no. 1, pp. 34–40, 2007.

[2] R. Novirsa, U. Achmadi, and Fahmi, "Analisis Risiko Pajanan PM2,5 di Udara Ambien Siang Hari terhadap Masyarakat di Kawasan Industri Semen" *Kesmas National Public Health Journal*, vol. 7, no. 4, pp. 173–179, 2012, doi: http://dx.doi.org/10.21109/kesmas.v7i4.52.

[3] J.S. Brown, "Deposition of Particles," in *Comparative Biology of the Normal Lung*, R.A. Parent, Ed., 2nd ed. MA, USA: Academic Press, 2015, ch. 27, pp. 513-536, doi: https://doi.org/10.1016/C2012-0-01154-4.

[4] R. Yendra and R. Risman, "Penerapan Metode Ordinary Kriging pada Pendugaan Kriminalitas di Kota Pekanbaru Riau," *Jurnal Sains Matematika dan Statistika,* vol. 5, no. 1, pp. 13–22, 2019, doi: http://dx.doi.org/10.24014/jsms.v4i1.6701.

[5] I. Fauzi and E.S. Hariyadi, "Analisis Geostatistik dalam Menentukan Keseragaman Nilai Kepadatan Tanah Dasar," *Jurnal Teknik Sipil*, vol. 25, no. 3, pp. 195–202, 2018, doi: https://doi.org/10.5614/jts.2018.25.3.4.

[6] N.A. Alfiana, "*Metode Ordinary Kriging pada Geostatistika*," Bachelor thesis, Universitas Negeri Yogyakarta, Sleman, Indonesia, 2010.

[7] IQAir Staff Writers, “PM2.5,” IQAir.com. https://www.iqair.com/us/blog/air-quality/pm2-5 (accessed May 16, 2022).

[8] R. Mukhtar, E.H. Panjaitan, H. Wahyudi, M. Santoso, and S. Kurniawati, “Komponen Kimia PM2,5 dan PM10 di Udara Ambien di Serpong – Tangerang,” *Jurnal Ecolab* . Vol.7 , No. 1, pp. 1–7, 2013, doi: https://doi.org/10.20886/jklh.2013.7.1.1-7.

[9] A.A. Awali, H. Yasin, and R. Rahmawati, “Estimasi Kandungan Hasil Tambang Menggunakan Ordinary Indicator Kriging,” *Jurnal Gaussian,* vol. 2, no. 1, pp. 1–10, 2013, doi: https://doi.org/10.14710/j.gauss.2.1.1%20-%2010

[10] A.D.R. Bahtiyar, A. Hoyyi, and H. Yasin, “Ordinary Kriging Dalam Estimasi Curah Hujan di Kota Semaran,” *Jurnal Gaussian*, vol. 3, no. 2, pp. 152–153, 2014, doi: https://doi.org/10.14710/j.gauss.3.2.151%20-%20159.

[11] K.B. Rachmawati, I.A. Sukma, A. Triamartha, K.P. Dela, N. Chamidah, “Estimation of Dissolved Oxygen Using Spatial Analysis Based on Ordinary Kriging Method as Effort to Improve the Quality of Surabaya’s River Water,” *Ecology, Environment and Conservation*, vol. 25, pp. 62–66, 2019.

[12] D.W. Fitri, N. Afifah, S.M.D. Anggarani, N. Chamidah, “Prediction Concentration of PM2.5 in Surabaya Using Ordinary Kriging Method,” *AIP Conference Proceedings*, 2021, pp. 1–6, doi: https://doi.org/10.1063/5.0042284.

[13] S.A.D. Safitri, F.A. Putri, A.A. Belindha, N. Chamidah, “Co-Kriging Method Performance in Estimating the Number of COVID-19 Positive Confirmed Cases in East Java Province,” *AIP Conference Proceedings,* 2021, pp. 1–10, doi: https://doi.org/10.1063/5.0042286.

[14] G. Rozalia, H. Yasin, and D. Ispriyanti , “Penerapan Metode Ordinary Kriging pada Pendugaan Kadar No2 di Udara (Studi Kasus: Pencemaran Udara di Kota Semarang),” *Jurnal Gaussian*, vol. 5, no. 1, pp. 113–121, 2016, doi: https://doi.org/10.14710/j.gauss.5.1.113-121.

[15] S. Munadi, *Pengantar Geostatistik*. Jakarta, Indonesia: Universitas Indonesia, 2005.

[16] R. Hayami, Sunanto, and I. Oktaviandi, “Penerapan Metode Single Exponential Smoothing pada Prediksi Penjualan Bed Sheet,” *Jurnal Computer Science and Information Technology (CoSciTech)*, vol. 2, no. 1, pp. 32–29, 2021, doi: https://doi.org/10.37859/coscitech.v2i1.2184.

[17] N.M.S. Fridayani, P.E.N. Kencana, and K.G. Sukarsa, “Perbandingan Interpolasi Spasial dengan Metode Ordinary dan Robust Kriging pada Data Spasial Berpencilan (Studi Kasus: Curah Hujan di Kabupaten Karangasem),” *E-Jurnal Matematika*, vol. 1, no. 1, pp. 68–74, 2012, doi: https://doi.org/10.24843/MTK.2012.v01.i01.p012.

[18] iQAir, https://www.iqair.com/id/ (accessed May 9, 2022).