Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 143 (2020) 155-169 EISSN 2392-2192

Uncertain Semivariogram Model using *Robust* Optimization for Application of Lead Pollutant Data

Annisa Azizah*, Budi Nurani Ruchjana, Diah Chaerani

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang – 45363, Indonesia

*E-mail address: amarannisa@gmail.com

ABSTRACT

Semivariogram is a half variance diagram of the difference between observations at the location s_i with another location that is as far as h units of distance. Semivariogram is used to describe the correlation of observation sorted by location. This research discusses the theoretical Semivariogram for the *Spherical*, *Gaussian*, and *Exponential* Semivariogram models through the Linear Programming approach. Next, the Semivariogram parameter estimation is studied with the assumption that there are data uncertainties, called the Uncertain Semivariogram. The method used to overcome the uncertainty data is *Robust* Optimization. The Uncertain Semivariogram using *Robust* Optimization are solved using the *box* and *ellipsoidal* uncertainty set approach. The calculation of the application of the model was carried out using the R software. For the case study, the application of the model used secondary data of Lead pollutant data in the Meuse River floodplains on the borders of France and the Netherlands at 164 locations. Based on the calculation results, the Exponential theoretical Semivariogram model is obtained as the best Semivariogram model, because it has a minimum SSE. Furthermore, the application of the Uncertain Semivariogram model using *Robust* Optimization on the Semivariogram Exponential model of Lead pollutant data is carried out using the box and ellipsoidal uncertainty set approach which is to obtain computationally tractable results.

***Keywords*:** Semivariogram, Linear Programming, *Robust* Optimization, Software R, Lead Pollutan

(Received 16 February 2020; Accepted 05 March 2020; Date of Publication 06 March 2020)

1. INTRODUCTION

In daily life, a lot of data in a sequence of locations is used called spatial data which is part of a stochastic process. The function that describes the spatial correlation in the form of error variance in spatial data is expressed by a Semivariogram or a half variance difference diagram from observations at the location s_i over a distance h unit. In the spatial analysis, the Semivariogram is used to find the parameters that determine the weights in Kriging, which is the prediction method at the unsampled location.

Theoretical semivariogram has the curve shape that is closest to the Experimental Semivariogram so that for the purposes of further analysis many theoretical Semivariogram models can be used, including *Spherical* Model, *Gaussian* Model, and *Exponential* Model (Armstrong, 1998).

Before using the Kriging method, the Semivariogram parameters are estimated first. One of the method that can be used to estimate parameters in the Semivariogram model is the Linear Programming method (Chen and Jiao, 2001). In this study, the researcher assumes there is a problem in estimating the parameters of the Semivariogram model which involves the uncertainty data.

The uncertainty data can be caused by measurement errors, modeling errors, or unavailability of required information. *Robust* optimization is a method for solving problems that are affected by data uncertainty and there is no probability distribution available for parameter uncertainty (Gabrel et al, 2011). The purpose of *Robust* Optimization is to obtain robust solutions, which help decision-makers to avoid losses from uncertainty. The urgency of this research lies in estimating the parameters of the semivariogram model using linear programming but involves the uncertainty of the data. Therefore, the Semivariogram model needs to be transformed first into the form of the Linear Programming problem and then solved using the *Robust* Optimization method with the *box* and *ellipsoidal* uncertainty set to obtain *computationally tractable* results.

2. MATERIALS AND METHODS

2. 1. Experimental and Theoretical Semivariogram

Estimated experimental semivariogram at distance h, can be written as follows (Youkuo et al, 2015):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(x_i + h) - Z(x_i))^2 \quad (1)$$

where:

$\hat{\gamma}(h)$: Experimental semivariogram value with distance h
$Z(x_i)$: Observation value in location x_i
$Z(x_i + h)$: Observation value in location $x_i + h$
$N(h)$: The number of point pairs within h

There are 3 Theoretical Semivariogram model that are commonly used, that is *Spherical* model, *Gaussian* model, and *Exponential* model (Youkuo et al., 2015).

1. *Spherical* Model

The *Spherical* model is formulated as follows:

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

2. *Gaussian* Model

The *Gaussian* model is a quadratic form of *Exponential* so that it obtained a parabolic form at close range. The *Gaussian* model is formulated as follows:

$$\gamma(h) = c\left[1 - exp\left(-\frac{h}{a}\right)^2\right] \tag{3}$$

3. *Exponential* Model

The *Exponential* Model curve there was a very steep increase and reached the value of the *sill* asymptotically, formulated as follows:

$$\gamma(h) = c\left[1 - exp\left(-\frac{h}{a}\right)\right] \tag{4}$$

2. 3. Linear Programming

Optimization is an effort to get the best results from a situation. The aim is to minimize the effort required or to get the maximum results from the benefits obtained (Rao, 2009).

According to Rao (2009), linear programming is an optimization method to finding solutions problems with objective functions and constraint functions that arise in the form of linear functions of decisions.

Linear programming problems are generally stated in the following standard form:

$$minimize \quad \mathrm{c^T x} \tag{5}$$

$$s.t \quad \mathrm{Ax = b} \tag{6}$$

$$\mathrm{x} \geq 0 \tag{7}$$

where (5) is a objective function, (6) is a constraint function, and (7) is a limitation that the decision variable of the problem is of non-negative value (Rao, 2009).

In the standard form of Linear Programming, it can be characterized that the Linear Programming problem fulfill the conditions (Rao, 2009):

1) The objective function is the minimization function..
2) All constraint functions are equation functions with a sign (=).
3) All decision variables are non-negative.

2. 4. *Robust* Optimization

Robust optimization method is a mathematical method for various problems related to finding optimal solutions that are robust from problems involving uncertainty data so that the problems studied are categorized as *optimization under uncertainty* (Chaerani, D., 2016).

If parameters $\mathbf{c}, \mathbf{A}$, and $\mathbf{b}$ in equations (5), (6), and (7) are assumed as uncertain value with $\mathbf{c}, \mathbf{A}, \mathbf{b} \in \mathcal{U}$, where $\mathcal{U}$ is a set of all data uncertainties, then the Linear Programming problem can be obtained the problem of Uncertainty Linear Programming as follows:

$$\min_{x}\{c^{T}x : Ax \leq b, \forall (c, A, b) \in \mathcal{U}\} \tag{8}$$

The main challenge in dealing with uncertainty optimization problems is to answer the question of how and when uncertain problems can be reformulated as computational optimization problems (*computationally tractable*).

The tractability of (8) depends on the selection of the set of $\mathcal{U}$ uncertainty. The following theorem makes it clear that if the specified set of $\mathcal{U}$ can be expressed as one of the linear constraints, quadratic conic constraints, or with semi-definite constraints, then (8) becomes a *computationally tractable* problem.

Theorem 1. Assume that the uncertainty set $\mathcal{U}$ in (8) is given as the affine image of a bounded set $Z = \{\zeta\} \subset R^{N}$, and Z is given either

1) By a system of linear inequalities $P\zeta \leq p$
2) By a system of conic quadratic inequalities $\|P_i\zeta - p_i\|_2 \leq q_i^{T}\zeta - r_i, i = 1, \dots, M$
3) By a linear matrix inequality $P_0 + \sum_{i=1}^{N} \zeta_i P_i \succeq 0$

In the case 2 and 3 we assume that the system of constraints defining $\mathcal{U}$ is strictly feasible. Then the robust counterpart (8) of (5), (6) and (7) is equivalent to

- A linear optimization problem in case 1
- a conic quadratic problem in case 2
- a semidefinite problem in case 3

In all cases, the data of the resulting robust counterpart problem are readily given by m,n and the data specifying the uncertainty set. Moreover, the size of the resulting problem is polynomial in the size of the data specifying the uncertainty set. The proof can be found in (Bental and Nemirovskii, 2002) and (Chaerani and Roos, 2013).

3. RESULT AND DISCUSSION

3. 1. Transformation of *Exponential* Semivariogram Model into Linear Programming Problems

The general form of the *Exponential* Semivariogram model is formulated in equation (4). Using the Taylor Series Expansion, the function $e^{-\frac{h}{a}}$ becomes

$$e^{-\frac{h}{a}} \approx 1 - \frac{h}{a} + \frac{h^2}{2!a^2} - \frac{h^3}{3!a^3} + \cdots \tag{9}$$

Equation (9) becomes

$$\gamma(h) = c_0 + c\left(\frac{h}{a} - \frac{h^2}{2!a^2} + \frac{h^3}{3!a^3} + \cdots\right) \tag{10}$$

Equation (9) can be approximated by

$$\gamma(h) \approx c_0 + c\left(\frac{h}{a} - \frac{h^2}{2!a^2} + \frac{h^3}{3!a^3} + \cdots + \frac{h^{11}}{11!a^{11}}\right) \tag{11}$$

Let $b_0 = c_0, b_1 = \frac{c}{a}, b_2 = -\frac{c}{2!a^2}, \cdots, b_{11} = \frac{c}{11!a^{11}}$, $y = \gamma(h)$, and $x_1 = h,\ x_2 = h^2, \cdots, x_{11} = h^{11}$, equation (1) can be written into a linear equation as follows:

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_{11}x_{11} \tag{12}$$

Let $h_1, h_2, \cdots, h_m$ are observations of m at the distance h, c is *sill*, a is $range$, and $\gamma_m(h)$ is the *Exponential* Semivariogram value at distance h with observations of m following the form of the linear equation below.

$$\begin{cases} x_{11}b_1 + x_{12}b_2 + \cdots + x_{111}b_{11} = y_1 \\ x_{21}b_1 + x_{22}b_2 + \cdots + x_{211}b_{11} = y_2 \\ \vdots \\ x_{m1}b_1 + x_{m2}b_2 + \cdots + x_{m11}b_{11} = y_m \end{cases} \tag{13}$$

with the terms $y_1, \cdots, y_m \neq 0$. Equation (13) is written in the form of a matrix

$$\mathrm{Xb} = \mathrm{y} \tag{14}$$

where $\mathrm{X} = \begin{pmatrix} h_1 & h_1^2 & \dots & h_1^{11} \\ h_2 & h_2^2 & \dots & h_2^{11} \\ \vdots & \vdots & \ddots & \vdots \\ h_m & h_m^2 & \dots & h_m^{11} \end{pmatrix}$, $\mathrm{b} = \begin{pmatrix} \frac{c}{a} \\ -\frac{c}{2!a^2} \\ \vdots \\ \frac{c}{11!a^{11}} \end{pmatrix}$, and $\mathrm{y} = \begin{pmatrix} \gamma_1(h) \\ \gamma_2(h) \\ \vdots \\ \gamma_m(h) \end{pmatrix}$.

If all the desired solutions are non-negative, then choose $\mathrm{b} \geq \theta$ where θ is a vector where each element has zero value. Matrix X can be expressed in the form of row vectors

$\mathrm{X} = \begin{pmatrix} \mathrm{x}_1^{\mathrm{T}} \\ \mathrm{x}_2^{\mathrm{T}} \\ \vdots \\ \mathrm{x}_{\mathrm{m}}^{\mathrm{T}} \end{pmatrix}, \mathrm{x_i} = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{pmatrix}$ where $i = 1, \cdots, m$.

Furthermore, the objective function can be written as follows

$$f(a) = \sum_{i=1}^{m} \left| y_i - \mathrm{x}_\mathrm{i}^\mathrm{T}\mathrm{b} \right| \tag{15}$$

when b ≥ θ, equation (15) can be solved by linear programming as follows

$$\begin{aligned} &min && \begin{pmatrix} \mathrm{u} \\ \theta \end{pmatrix}^\mathrm{T} \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \\ &s.t && \mathrm{D}\begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \geq \begin{pmatrix} \mathrm{y} \\ -\mathrm{y} \end{pmatrix} \\ & && \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \geq \theta \end{aligned} \tag{16}$$

where $\mathrm{D} = \begin{pmatrix} \mathrm{I} & \mathrm{X} \\ \mathrm{I} & -\mathrm{X} \end{pmatrix}$, $\mathrm{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{pmatrix}$, and $\mathrm{u} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$. Here, I is the $m \times n$ matrix, u is the column vector which all elements are equal to 1.

Note to the constraint function in equation (16) that is not yet in the standard form, so subtract the surplus variable

$$\mathrm{D}\begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} - \begin{pmatrix} \mathrm{S}_1 \\ \mathrm{S}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{y} \\ -\mathrm{y} \end{pmatrix} \tag{17}$$

Use the substitution method, so that the equation constraint (16) is obtained as follows

$$\begin{pmatrix} \mathrm{I} & \mathrm{X} & -\mathrm{I} & 0 \\ 0 & 2\mathrm{X} & -\mathrm{I} & \mathrm{I} \end{pmatrix} \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \\ \mathrm{S}_1 \\ \mathrm{S}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{y} \\ 2\mathrm{y} \end{pmatrix} \tag{18}$$

Then, obtained a Linear Programming problem formulation for the *Exponential* semivariogram model as follows.

$$\begin{aligned} &min && \begin{pmatrix} \mathrm{u} \\ \theta \end{pmatrix}^\mathrm{T} \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \\ &s.t && \begin{pmatrix} \mathrm{I} & \mathrm{X} & -\mathrm{I} & 0 \\ 0 & 2\mathrm{X} & -\mathrm{I} & \mathrm{I} \end{pmatrix} \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \\ \mathrm{S}_1 \\ \mathrm{S}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{y} \\ 2\mathrm{y} \end{pmatrix} \\ & && \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \geq \theta. \end{aligned} \tag{19}$$

3. 2. Formulation of the *Exponential* Semivariogram Model Using *Robust* Optimization

To formulate the *Robust* optimization model, the first thing to do is to determine the uncertainty parameters. In the *Exponential* Semivariogram model, assume an uncertain factor, the value of Semivariogram y from (19). That is because the semivariogram value obtained cannot be said to be precise because it is calculated based on the selected basic range that can change. Therefore, it can be assumed y ∈ $\mathcal{U}$. The uncertainty parameter y can be written as

$$y = \bar{y} + P\zeta, \forall\zeta \in Z \tag{20}$$

where $y \in R^n$ is a vector of semivariogram nominal value, $P \in R^{n\times L}$ is a confounding matrix, and $\zeta \in R^L$ is a primitive uncertainty vector.

Consider the constraint equation (19) which contains uncertainty. Because the uncertainty parameter y is on the right-hand side, based on the robust optimization assumption the right-hand of the constraint function must be a certain parameter, so that it can be added $x_{n+1} \leq 1$.. Then substitute (20) into (19), the *Robust Counterpart* model is obtained for the *Exponential* Semivariogram as follows.

$$\begin{aligned} min \quad & \begin{pmatrix} u \\ \theta \end{pmatrix}^T \begin{pmatrix} t \\ b \end{pmatrix} \\ s.t \quad & \begin{pmatrix} I & X & -I & 0 \\ 0 & 2X & -I & I \end{pmatrix} \begin{pmatrix} t \\ b \\ S_1 \\ S_2 \end{pmatrix} - \begin{pmatrix} \bar{y} + P\zeta \\ 2(\bar{y} + P\zeta) \end{pmatrix} (x_{n+1}) = 0 \\ & \begin{pmatrix} t \\ b \end{pmatrix} \geq \theta \end{aligned} \tag{21}$$

3. 3. *Robust Counterpart* for *Exponential* Semivariogram Models with *Box* Uncertainty Set

The set of uncertainty *box* is defined as follows (Gorissen, 2015).

$$Z = \{\zeta : \|\zeta\|_\infty \leq 1\} \tag{22}$$

Robust Counterpart formulation for constraint with uncertain parameters in the *box* uncertainty set is

$$\bar{a}^T x + \|P^T x\|_1 \leq b \tag{23}$$

where $\bar{a} \in R^n$ is a vector of nominal value, $P \in R^{n\times L}$ is a confounding matrix, and $\zeta \in R^L$ is a primitive uncertainty vector.

The *Robust Counterpart* formulation obtained for the *Exponential* Semivariogram model with the *box* uncertainty set in the matrix formulation is as follows.

$$\begin{aligned} min \quad & \tau \\ s.t \quad & \begin{pmatrix} \begin{pmatrix} u \\ \theta \end{pmatrix}^T & 0 & -1 \\ -\begin{pmatrix} I & X & -I & 0 \\ 0 & 2X & -I & I \end{pmatrix} & \begin{pmatrix} \bar{y} + \|P\|_1 \\ 2\bar{y}(1 + \|P\|_1) \end{pmatrix} & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \\ x_{n+1} \\ \tau \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \\ & \begin{pmatrix} t \\ b \end{pmatrix} \geq \theta \end{aligned} \tag{24}$$

3. 4. *Robust Counterpart* for *Exponential* Semivariogram Models with *Ellipsoidal* Uncertainty Set

The set of uncertainty *ellipsoidal* is defined as follows (Gorissen, 2015).

$$Z = \{\zeta : \ \|\zeta\|_2 \leq 1\} \tag{25}$$

Robust Counterpart formulation for constraint with uncertain parameters in the *ellipsoidal* uncertainty set is

$$\bar{a}^T x + \|P^T x\|_2 \leq b \tag{26}$$

where $\bar{a} \in R^n$ is a vector of nominal value, $P \in R^{n\times L}$ is a confounding matrix, and $\zeta \in R^L$ is a primitive uncertainty vector.

The *Robust Counterpart* formulation obtained for the *Exponential* Semivariogram model with the *ellipsoidal* uncertainty set in the matrix formulation is as follows.

$$\begin{array}{ll} min & \tau \\ s.t & \begin{pmatrix} \begin{pmatrix} u \\ \theta \end{pmatrix}^T & 0 & -1 \\ -\begin{pmatrix} I & X & -I & 0 \\ 0 & 2X & -I & I \end{pmatrix} & \begin{pmatrix} \bar{y} + \|P\|_2 \\ 2\bar{y}(1 + \|P\|_2) \end{pmatrix} & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \\ x_{n+1} \\ \tau \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \\ & \begin{pmatrix} t \\ b \end{pmatrix} \geq \theta \end{array} \tag{27}$$

3. 5. Application

The data used as the application of the model in this study are secondary data obtained from Software R. datasets. The data are spatial data of Lead pollutants in the Meuse river flood plain located on the border of the French and Dutch Countries with 164 location points.

Table 1. Lead Concentration Data.

Locations	x(m)	y(m)	Lead (ppm)
1	181072	333611	299
2	181025	333558	277
⋮	⋮	⋮	⋮
3	180627	330190	124

According to Olea (1999), the *Ordinary Kriging* method is a method that assumes the mean is unknown and fulfills the *second order stationary*, this shows that the semivariogram used as a parameter that determines the weight of the *Kriging*, requires normally distributed data. Therefore the Log transformation for Lead data was carried out.

The histogram and the normal Q-Q plot of the transformed Lead data are presented in the following Figure 1 and Figure 2.

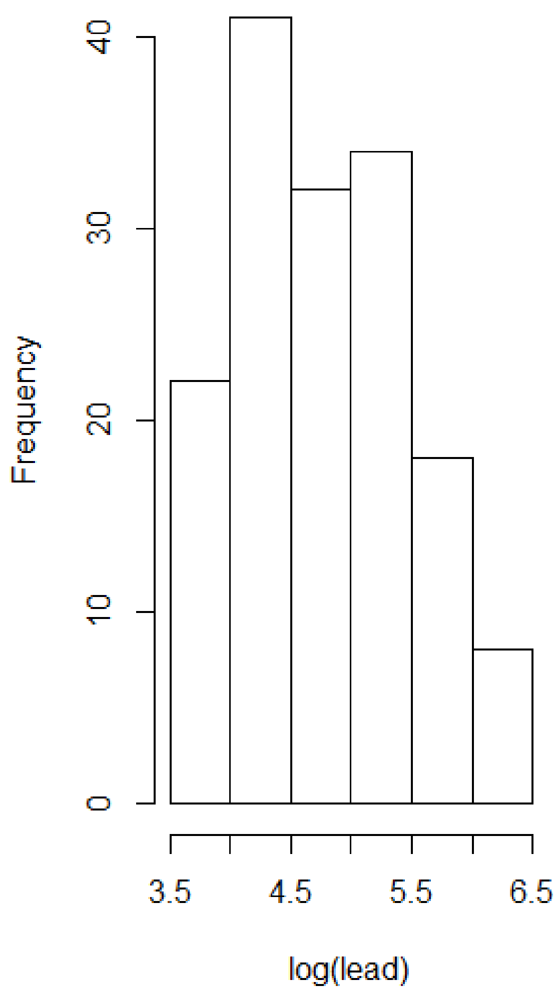

Figure 1. Histogram of Log Transformed Lead Data

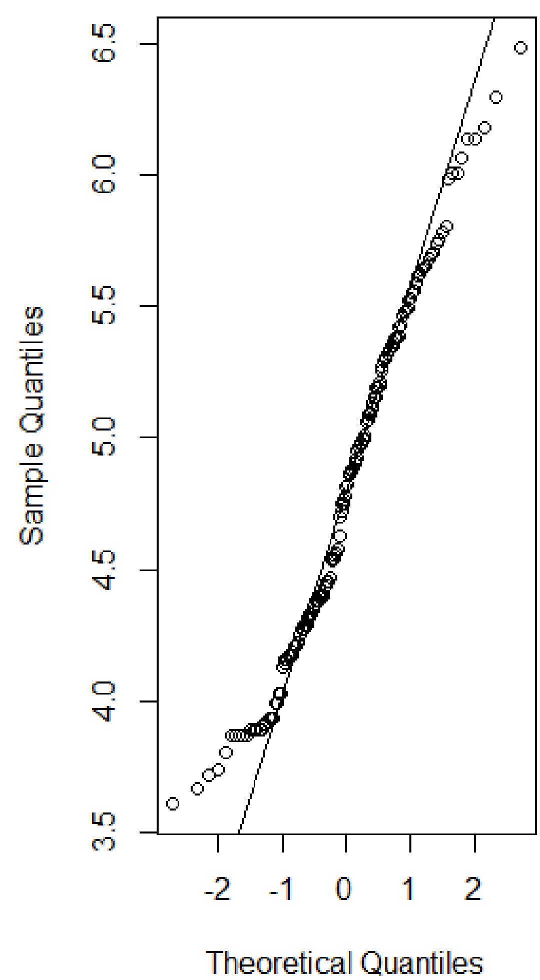

Figure 2. Normal Q-Q Plot of Log Transformed Lead Data

Before fitting the theoretical Semivariogram model, it is necessary to carry out structural analysis which is a process of matching experimental semivariograms with theoretical semivariograms.

Using Software R, the calculation results of the experimental semivariogram plot in Figure 3 andexperimental semivariogram presented in Table 2

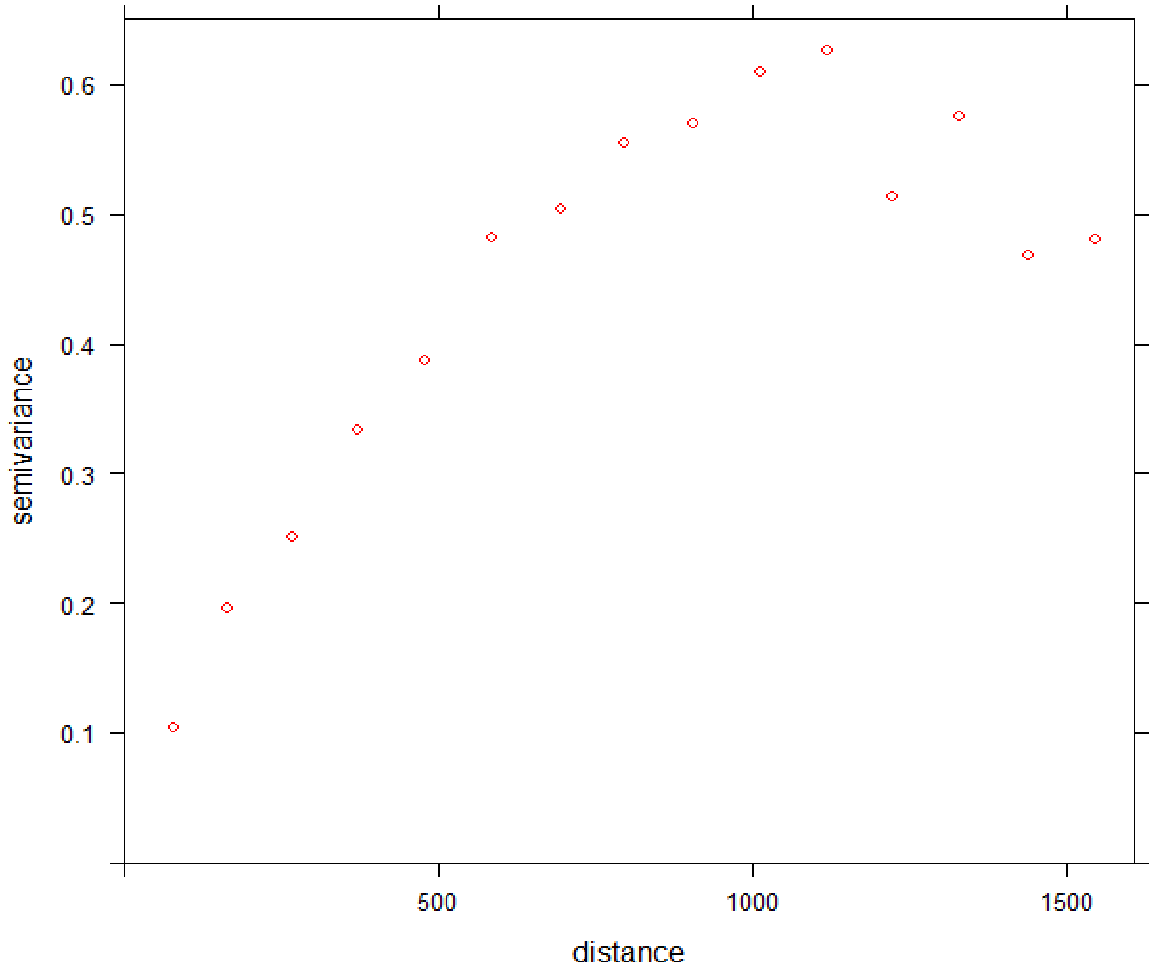

Figure 3. Plot the Experimental Semivariogram of Lead Data

Table 2. Experimental Semivariogram.

No.	Np	Distance	Experimental Semivariogram
1	57	79.29244	0.1046520
2	299	163.97367	0.1965929
3	419	267.36483	0.2507668
4	457	372.73542	0.3330690
5	547	478.47670	0.3875716

6	533	585.34058	0.4817750
7	574	693.14526	0.5031432
8	564	796.18365	0.5545787
9	589	903.14650	0.5693882
10	543	1011.29177	0.6098806
11	500	1117.86235	0.6253271
12	477	1221.32810	0.5126165
13	452	1329.16407	0.5755737
14	457	1437.25620	0.4676728
15	415	1543.20248	0.4804887

By observing the plot, we conclude that *sill* = 0.4441557 and range=800. Based on these three parameter, then we choose the best theoretical semivariogram models by fitting semivariogram models into experimental semivariogram. In fitting semivariogram models, it is used three models that is *Spherical*, *Exponential* and *Gaussian*. The best theoretical semivariogram model with the minimum sum of square error (*SSE*) is *Exponential* model based on table below.

Table 3. *SSE* the Theoretical Semivariogram Model.

***SSE* The Theoretical Semivariogram Model**		
Spherical	***Gaussian***	***Exponential***
3.175989e-05	0.0001764282	2.089753e-05

Based on Table 3, from the three Theoretical Semivariogram models obtained that the *Exponential* Model which has the minimum *SSE* that is 2.089753e-05.

After selecting the *Exponential* semivariogram model as the best model, the Linear Programming model will be applied to the *Exponential* semivariogram of equation (21). Substitute the values of h and $\gamma(h)$ to obtain the **X** and **Y** matrices in equation (21). The following is the Linear Programming model for the *Exponential* semivariogram of the Lead data.

$$
\begin{array}{ll}
min & t \\
s.t & t + 799.9843b_1 + 639974.829b_2 + 511969795.2b_3 + 4.09568 \times 10^{11}b_4 \\
& +3.27648 \times 10^{14}b_5 + 2.62113 \times 10^{17}b_6 + 2.09686 \times 10^{20}b_7 \\
& +1.67746 \times 10^{23}b_8 + 1.34194 \times 10^{26}b_9 + 1.07353 \times 10^{29}b_{10} \\
& +8.58808 \times 10^{31}b_{11} - S_1 = 1278947.241 \\
& 1599.968536b_1 + 1279949.658b_2 + 1023939590b_3 + 8.19136 \times 10^{11}b_4 \\
& +6.55296 \times 10^{14}b_5 + 5.24226 \times 10^{17}b_6 + 4.19373 \times 10^{20}b_7 \\
& +3.35492 \times 10^{23}b_8 + 2.68388 \times 10^{26}b_9 + 2.14706 \times 10^{29}b_{10} \\
& +1.71762 \times 10^{32}b_{11} - S_1 + S_2 = 1881668.909 \\
& t, b_1, b_2, \dots, b_{11} \geq 0.
\end{array} \tag{28}
$$

Based on equation (28), calculation result using Software R obtained an optimal solution for the *Exponential* Semivariogram model using Linear Programming is $Z^* = 3.381128\text{e} - 05$ with a decision variable $b_3 = 1.837676\text{e} - 03$.

Next, an optimal solution will be sought from the uncertain *Exponential* Semivariogram with the *box* uncertainty set with formulations (24).

Substitute the values of h and $\gamma(h)$ to obtain the **X** and **Y** matrices in equation (24). The *Robust* Optimization model with the *box* uncertainty set as follows.

$$
\begin{array}{ll}
min & \tau \\
s.t & t - \tau \leq 0 \\
& t + 799.9843b_1 + 639974.829b_2 + 511969795.2b_3 + 4.09568 \times 10^{11}b_4 \\
& +3.27648 \times 10^{14}b_5 + 2.62113 \times 10^{17}b_6 + 2.09686 \times 10^{20}b_7 \\
& +1.67746 \times 10^{23}b_8 + 1.34194 \times 10^{26}b_9 + 1.07353 \times 10^{29}b_{10} \\
& +8.58808 \times 10^{31}b_{11} - S_1 + P = 1278947.241 \\
& 1599.968536b_1 + 1279949.658b_2 + 1023939590b_3 + 8.19136 \times 10^{11}b_4 \\
& +6.55296 \times 10^{14}b_5 + 5.24226 \times 10^{17}b_6 + 4.19373 \times 10^{20}b_7 \\
& +3.35492 \times 10^{23}b_8 + 2.68388 \times 10^{26}b_9 + 2.14706 \times 10^{29}b_{10} \\
& +1.71762 \times 10^{32}b_{11} - S_1 + S_2 + 2P = 1881668.909 \\
& t, \tau, b_1, b_2, \dots, b_{11} \geq 0
\end{array} \tag{29}
$$

Suppose the parameter P is a random number that is defined as follows.

$$P = \frac{1}{2}(l - u) \tag{30}$$

Based on equation (29), calculation result using Software R obtained an optimal solution for the uncertain *Exponential* Semivariogram model with the *box* uncertainty set is $Z^* = 3.596326\text{e} - 05$ with a decision variable $b_3 = 6.846227\text{e} - 03$.

Lastly, an optimal solution will be sought from the uncertain *Exponential* Semivariogram with the *ellipsoidal* uncertainty set with formulations (27).

Substitute the values of h and $\gamma(h)$ to obtain the **X** and **Y** matrices in equation (27). The *Robust* Optimization model with the *ellipsoidal* uncertainty set as follows.

$$\begin{aligned} &min && \tau \\ &s.t && t - \tau \leq 0 \end{aligned}$$

$$\begin{gathered} t + 799.9843b_1 + 639974.829b_2 + 511969795.2b_3 + 4.09568 \times 10^{11}b_4 \\ +3.27648 \times 10^{14}b_5 + 2.62113 \times 10^{17}b_6 + 2.09686 \times 10^{20}b_7 \\ +1.67746 \times 10^{23}b_8 + 1.34194 \times 10^{26}b_9 + 1.07353 \times 10^{29}b_{10} \\ +8.58808 \times 10^{31}b_{11} - S_1 + \sqrt{P} = 1278947.241 \end{gathered} \tag{31}$$

$$\begin{gathered} 1599.968536b_1 + 1279949.658b_2 + 1023939590b_3 + 8.19136 \times 10^{11}b_4 \\ +6.55296 \times 10^{14}b_5 + 5.24226 \times 10^{17}b_6 + 4.19373 \times 10^{20}b_7 \\ +3.35492 \times 10^{23}b_8 + 2.68388 \times 10^{26}b_9 + 2.14706 \times 10^{29}b_{10} \\ +1.71762 \times 10^{32}b_{11} - S_1 + S_2 + 2\sqrt{P} = 1881668.909 \end{gathered}$$

$$t, \tau, b_1, b_2, \dots, b_{11} \geq 0$$

where the value of the parameter P is a random number that is defined as in equation (30).

Based on equation (29), calculation result using Software R obtained an optimal solution for the uncertain *Exponential* Semivariogram model with the *ellipsoidal* uncertainty set is $Z^* = 4.096326\text{e} - 05$ with a decision variable $b_3 = 7.822847\text{e} - 03$.

4. CONCLUSIONS

An *Exponential* semivariogram model can be built into an uncertain semivariogram model using *Robust* optimization with the uncertainty of parameter is Semivariogram value. The uncertain Semivariogram model using Robust Optimization solved by the *box* and *ellipsoidal* uncertainty set approach. Application of the Experimental Semivariogram Model and Theoretical Semivariogram on the Lead Data at Meuse River gives the results of the *Exponential* Semivariogram Model as the chosen model based on *SSE*. Next, the experimental Semivariogram model using Linear Programming and the Semivariogram model with $\boldsymbol{y}$ uncertainties using the *box* and *ellipsoidal* uncertainty set approach are used. The optimal robust solution with the *box* and *ellipsoidal* uncertainty set gives greater objective function values than the optimal solution for the *Exponential* Semivariogram model using Linear Programming.

Computationally, the results obtained using *Robust* Optimization with the set of uncertainty box and *ellipsoidal* are better, in this case resistant to onterference.

References

[1] Armstrong, M. (1998). *Basic Linear Geostatistics.* New York: Springer

[2] Ben-Tal, A., Nemirovskii, A. Robust Optimization methodology and Applications. *Mathematical Programming Article*, Vol 92(3, Ser. B) pp. 453-480. (2002). doi:10.1007/s101070100286

[3] Chaerani, D., dan Roos, C. Handling Optimization under Uncertainty Problem Using Robust Counterpart Methodology. *Jurnal Teknik Industri*, Vol. 15, No. 2. (2013). doi:10.9744/jti.15.2.111-118.

[4] Chaerani, D., Anggriani N., Firdaniza. On the Robust Optimization to the Uncertain Vaccination Strategy Problem. *Symposium on Biomathematics (Symomath 2013). AIP Conf. Proc.* 1587, 34-37. (2014). doi:10.1063/1.4866528.

[5] Chaerani, D., Ruchjana, B. N., Setiawan, A., Rejito, J., Dewanto, S. P., Dharmawan, I. A., Rosandi, Y. Determining Robust Counterpart of Spatial Optimization Model for Water Supply Allocation. *AIP Conf*. Proc. 17116, 02002. (2016). http://dx.doi.org/10.1063/1.4942985

[6] Chen, Y. dan Jiao X. Semivariogram fitting with linear programming. *Computers & Geosciences Journal*, Volume 27, pp. 71-76. (2001). PII: S0098-3004(00)00022-4.

[7] Eze, PN. dan Kumahor, SK. Gaussian process simulation of soil Zn micronutrient spatial heterogeneity and uncertainty – A performance appraisal of three semivariogram models. *Scientific African 5 (2019)*. Published by Elsevier B.V. on behalf of African Institute of Mathematical Sciences. (2019). https://doi.org/10.1016/j.sciaf.2019.e00110

[8] Falah, A. N., Abdullah, A. S., Parmikanti, K., dan Ruchjana, B. N. Prediction of Cadmium Pollutant With Ordinary Point Kriging Method Using GStatR. *AIP Conference Proceedings: The 2nd International Conference on Applied Statistics (ICAS II 2016)*, 1827, 020019, (2017). https://doi.org/10.1063/1.4979435

[9] Fitriani, R. dan Sumarminingsih, E. (2014). The Dynamic of Spatial Extent of Land Use in the Fringe of Jakarta Metropolitan: A Semivariogram Analysis. *5th International Confrence on Environmental Science and Development - ICESD 2014*, 198-202. doi:10.1016/j.apcbee.2014.10.038

[10] Gabrel, V., Lacroix, M., dan Remli, N. Robust Location Transportation Problems Under Uncertain Demands. *Discrete Applied Mathematics*. Volume 164, Part 1, pp. 100-111. (2014). https://doi.org/10.1016/j.dam.2011.09.015

[11] Gorissen, B. L., Yamkoglu, I., dan Hertog, D. d. A Practical Guide to Robust Optimization. *Omega: The International Journal of Management Science*, Volume 53, pp. 124-137, (2015). https://doi.org/10.1016/j.omega.2014.12.006

[12] Li, S. dan Lu, W. Automatic Fit of the Variogram. *Third International Conference on Information and Computing*, Volume 4, 129–132, (2010). doi:10.1109/ICIC.2010.303

[13] Li, Z., Zhang, X., dan Clarke, KC. An automatic variogram modeling method with high reliability fitness and estimates. *Computers & Geosciences* 120: 48–59, (2018). doi:10.1016/j.cageo.2018.07.011

[14] Olea, R. A. (1999). Geostatistics for engineers and earth scientists. New York: Kluwer Academic Publishers.

[15] Patzold, J. dan Schobel, A. Approximate cutting plane approaches for exact solutions to robust optimization problems. *European Journal of Operational Research* Volume 239, Issue 3, (2019). https://doi.org/10.1016/j.ejor.2019.11.059

[16] Rao, S. S. Engineering Optimization Theory and Practice Fourth Edition. (2009). USA, New Jersey: John Wiley & Sons, Inc.

[17] Solana-Gutiérrez, J. dan Merino-de-Miguel, S. A Variogram Model Comparison for Predicting Forest Changes. *Procedia Environmental Sciences 7: Mapping Global Change,* Volume 7, 383–388, (2011). https://doi.org/10.1016/j.proenv.2011.07.066

[18] Wkm, T.B., Chaerani, D., dan Ruchjana, B.N. Eksplorasi Software R Untuk Fitting Semivariogram Spherical Menggunakan Pemrograman Linear dan Uji Analisis Sensitivitas. *Jurnal Matematika Integratif*. Volume 12 No 2, pp. 75-82, (2016). doi:10.24198/jmi.v12.n2.11918.75-82

[19] Youkuo, C., Yongguo, Y., dan Wangwen, W. Coal Seam Thickness Prediction based on Least Squares Support Vector Machines and Kriging Method. *Electronic Journal of Geotechnical Engineering*, Vol. 20 Bund. 1 167-176. (2015).