Minimax Approach for Semivariogram Fitting in Ordinary Kriging

ANDIE SETIYOKO[1,2], T. BASARUDDIN[1], AND ANIATI MURNI ARYMURTHY[1]

[1]Faculty of Computer Science, Universitas Indonesia, Depok 16424, Indonesia
[2]Remote Sensing Technology and Data Centre-LAPAN, Jakarta 13710, Indonesia

Corresponding authors: Andie Setiyoko (andie.setiyoko@ui.ac.id) and Aniati Murni Arymurthy (aniati@cs.ui.ac.id)

ABSTRACT This research paper aims to analyze the minimax approach used in the semivariogram fitting process that forms one stage of the kriging operation performed for interpolation. The conventional method uses the weighted least squares fit for various theoretical functions such as stable, exponential, spherical. However, several recent approaches have been developed using machine learning regression techniques. This research employs the ordinary kriging technique where the proposed minimax approach is expected to increase the accuracy of the interpolation resulted by reducing the error of the final result. Kriging, which is based on the stochastic method, is widely used for spatial values and has been proven to be a better predicting process than deterministic methods. The novel approach to ordinary kriging discussed here, the minimax approach, is able to increase result accuracy based on the experiments performed. Minimax can predict the weights of the semivariogram values better than the weighted least-squares method and performs faster than machine learning approaches.

INDEX TERMS Minimax techniques, interpolation, approximation methods.

I. INTRODUCTION

A. BACKGROUND

Kriging is an interpolation process based on a stochastic approach invented by Danie G. Krige and developed by Georges Matheron. It is a geostatistical field applied for earth phenomenon prediction [1] initially applied to mineral observation in the field of mining. Interpolation is particularly required when the known or measured points are not well distributed within an area (commonly known as discrete data). This has to be implemented for a discrete data values to produce a continuous data. The principle of kriging states that values of closer samples have a more substantial effect than values at more distant locations [2], known as the regionalized variables theory. There are several types of kriging interpolation, including block kriging, disjunctive kriging, simple kriging, ordinary kriging, universal kriging, and cokriging [3]. The development of kriging has further increased these types to include adaptive kriging and regression kriging [4]. Kriging has advantages in predicting values in comparison to other deterministically based methods such as spline, IDW (Inverse Distance Weighting), and polynomial

methods [5]. Kriging, as the stochastic techniques enables the deficiencies of deterministic methods to be addressed [3]. Several implementations and developments of kriging interpolation have been applied to the digital elevation model (DEM) interpolation [6]. A modification of kriging has been applied to mapping the organic carbon stock of soil [7], and ordinary kriging has been proven as an interpolation method with relatively accepted accuracy for the prediction of land total phosphorus [4]. The kriging process performs value prediction at the unsampled area. In the field of control and approximation, several techniques have been applied for robotics or tracking control such as adaptive neural fuzzy inference system (ANFIS) [8] and fuzzy output tracking controller [9].

The kriging interpolation process comprises of five steps: data representation, data exploration, model fitting, diagnostics, and model comparison [10]. Data exploration is conducted by calculating semivariance of known points and fitting the semivariogram. After the weights are determined from this stage, the kriging interpolation process is conducted. The final step is a result comparison.

A semivariogram represents the differences between two variables, distance and semivariance values. Measured or known points produce semivariance values on a discrete

---

The associate editor coordinating the review of this manuscript and approving it for publication was Qi Zhou.

semivariogram. Semivariogram calculation is conducted by calculating semivariance values at a specific distance between two known points and graphically presenting the results. Semivariogram fitting requires the calculation of a continuous semivariogram model to generate precise weight predictions for a particular distance. The fitting is a process of modeling using a functional approach that represents the shape of the distribution of semivariogram values. Several works of literature state that the semivariogram fitting process plays a vital role in operational kriging. Various semivariogram parameters affect ordinary kriging weights [11].

B. REVIEW THE SEMIVARIOGRAM FITTING DEVELOPMENT

Variogram modeling is an essential part of the kriging interpolation process [12]. The initial approach for this stage is defining a function considering values of, for example, nugget, range, sill, and other variables. These initial approaches type of equations solved by applying weighted least squares [3]. The problem using the weighted least squares is the subjectivity of choosing the available options. The options for this method are stable, spherical, exponential, etc. [12]. Improvement of the semivariogram modeling is still a wide-open topic in the data-regression fields of various disciplines. Since kriging itself requires a variety of practices [11], especially for semivariogram modeling.

Early development of semivariogram modeling is started with the weighted least squares followed by machine learning regression approaches. This approach is applied to the semivariogram fitting process in preference to conventional methods. It is aimed to improve the weight determination aspect of semivariogram modeling. Previously, support vector regression has been applied to variogram modeling [13] and the least squares support vector machine (LS-SVM), invented for optimal control of the SVM [14], has been used to interpolate missing oceanic data [12] and coal seam thickness [13]. SVM and gaussian process regression (GPR) can improve result accuracy compared to conventional methods for DEM interpolation [15]. LS-SVM can be used to improve the cokriging fusion process [16].

In minimizing the uncertainty of the semivariogram modeling, robust optimization has been developed for weighted least square models [17].

C. PROBLEM FORMULATION AND MOTIVATION

There are two problems in the two previous method approaches. The first is the subjective selection of functions at the weighted least square and the second is the length of computational time using machine learning regression. Even though machine learning approach to semivariogram regression increases result accuracy, the computational processing effort is still higher than in weighted least squares techniques. The traditional approach to the semivariogram fitting process using weighted least squares can be replaced with other methods to increase result accuracy with less computational effort. To this end, different approaches to the norm, such as

minimax, can be implemented for this case by expecting the reduction of the average error. Minimax is assumed to be able to minimize the maximum error in the semivariogram fitting process with the approximation approach.

D. NOVELTY AND CONTRIBUTION

Based on the systematic literature study that has been conducted previously, the minimax approach has not been implemented for kriging interpolation. Therefore, the minimax approach is a novel approach for a semivariogram fitting process, particularly for kriging interpolation.

The minimax approach is based on approximation theory. This type of approximation technique is occasionally called ∞-norm, and the 2-norm is sometimes called the least-squares norm [18]. The minimax approach has been applied to various mathematical experiments, and some of the literature is convinced that minimax is better than least squares for such operations [19], [20], [21] and that minimax provides better accuracy than least squares [22], [23]. The minimax means squared error (MSE) estimator can significantly outperform the least-squares estimator [24], and the minimax approach can minimize the maximum error of regression in such functions. In terms of computational effort, minimax is assumed to be faster than machine learning for the regression process. It is because a learning stage is required for the machine learning process. The effort also depends on the size of data as an input.

E. RESEARCH OBJECTIVE AND SCOPE

This research paper aims to implement a novel approach to the semivariogram fitting process of ordinary kriging interpolation by using a minimax approach. The novelty of this research focuses on developing a minimax exchange algorithm to approximate the semivariance distribution which previously using least square method. Minimax has a different type of norm approach compared to the least square. The contribution of this research project is that minimax, as a novel approach, can add alternatives in semivariogram modeling in the kriging process. Secondly, minimax is considered to be a better fit and improving the accuracy of the final result.

The data used in this research is the elevation point in the form of a cloud point that commonly used for developing DEM. The proposed method can also be implemented for another similar type of data.

The hypothesis is that minimax manages to increase result accuracy compared to the traditional weighted least squares approach. Also, it is able to run faster than the machine learning regression approach. The structure of this research paper is as follows: Section 1 introduces the background of the research, problem statement, novel solution approach, and the goal of the study. Section 2 provides a theoretical review of the process and the innovative approach. Section 3 explains the experimental methodology, the specifications of the data being used and the limitations of the research. Section 4 gives the results and brief analysis of the results in comparison to the theoretical review and updated literature. And the last

section discusses the whole research result and opportunities for further research.

II. THEORETICAL REVIEW
This section explains in detail the academic background of ordinary kriging, experimental semivariogram, the fitting process, minimax approximation, basis functions, and validation methods.

A. ORDINARY KRIGING
The most commonly used form of this technique is ordinary kriging [25], which estimates or predicts the value at a point of a region where a variogram is determined. The determination is conducted using measured data surrounding the estimation location. In this study, ordinary kriging is chosen as the applied technique because of its popularity. Also, the accuracy of estimated values is determined by the sample size and distribution of sampling points [4]. It is a proper technique for this data (height value generated from satellite imagery) because it features uniformity and density of sampling points in complicated topographical areas and that the auxiliary variable is not defined.

The formula for ordinary kriging is as follows [4], [25]:

$$Z^* (x_0) = \sum_{\sigma=1}^{n} w_\alpha Z(x_\alpha) \qquad (1)$$

where x_0 is a value to be predicted using the data values from n surrounding the sample points of x_α using weights w_α.

The weights values w_α are calculated from the value of semivariance predictions. The calculation is conducted using a function that fits the distribution of semivariance values $\gamma (h)$. The values are from measured points. The matrix operation for weights determination is as follows [25]:

$$\begin{pmatrix} \gamma (x_1 - x_1) & \cdots & \gamma (x_1 - x_1) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma (x_1 - x_1) & \cdots & \gamma (x_1 - x_1) & 1 \\ 1 & \ldots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1^{OK} \\ \vdots \\ w_n^{OK} \\ \mu OK \end{pmatrix}$$
$$= \begin{pmatrix} \gamma (x_1 - x_0) \\ \vdots \\ \gamma (x_n - x_0) \\ 1 \end{pmatrix} \qquad (2)$$

where h is the distance between points and n is the number of pairs of sample-point measurements of the values at separated distance h.

B. EXPERIMENTAL SEMIVARIOGRAM
The experimental semivariogram is a plot of semivariance values $\gamma (h)$ against distance h. The first step toward a qualitative description of the regionalized variables is the building of an experimental variogram [3]. The semivariance is defined by the following initial formula [26]:

$$\gamma (x_i - x_j) = \frac{1}{2} var(Zx_i - Zx_j) \qquad (3)$$

where *var* is the variance between points *i* and *j*. The semivariogram calculation of measured points in detail is then based on the formula of semivariance [3]

$$\gamma^* (h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} [Z (x_i) - Z(x_i + h)]^2 \qquad (4)$$

where N_h is the number of pairs separated by lag distance, *h*.

The distribution of semivariance values against lag distance is then plotted into an experimental semivariogram.

C. SEMIVARIOGRAM FITTING
The experimental semivariogram illustrates as a distribution of the semivariance values at a specific lag. Semivariogram fitting is conducted to create a continuous semivariogram model. Fitting a semivariogram by using such a theoretical model is a vital process since the variogram model provides useful information for interpolation, sampling optimization, and spatial pattern determination [3]. The fitting process can be defined as a form of regression process analysis. 50 to 100 semivariance values are suggested to construct a stable variogram. Therefore it is required to determine semivariogram fitting functions [3].

Four values determine the characterization of the spatial data used: lag (h), nugget, sill, and range [3]. Nugget is the variance of error from measurement added to that drawn from spatial variation at much shorter distances than the sample spacing. Sill is the value of the lag where there is no spatial dependence between semivariance values at a particular separation distance. The range is a value of h where $\gamma (h)$ rises to the sill. The initial semivariogram model uses these approaches and is fitted using selected functions. The options for purposes include spherical, exponential, stable, etc. The standard procedure usually proceeds using a weighted least squares technique where the weights are calculated from the numbers of pairs [3].

D. MINIMAX APPROXIMATION
An alternative approach to the least squares technique for fitting discrete data is minimax approximation. Both approaches are based on approximation theory and methods [18]. The approximation is the process of estimating unknown values from the best approximation function developed from known values. The differences between approximation and interpolation functions are [27]:
- Interpolating functions pass exactly through the known height points used for calculating the functions.
- Approximating functions do not necessarily pass through the known height points but approach them in a controlled manner.

The approximation functions introduced by Powell [18] are polynomial, minimax, least-squares, and b-splines. These are continuously updated and have been developed recently for many applications, including spatial analysis. The vector norm is based on the mathematical concept behind the minimax and least squares theories. Norms on the real line furnish

a measure of distance. R^n, together with a norm on R^n, defines a metric space. Familiar notions including neighborhood, open sets, convergence and continuity are used when working with vectors and vector-valued functions [28].

The vector norm on R^n is a function of $f: R^n \rightarrow R$ that satisfies the following conditions:

$$f(x) \geq 0 \quad x \in R^n, (f(x) = 0 \text{ if } x = 0$$
$$f(x+y) \leq f(x) + f(y) \quad x, y \in R^n$$
$$f(\propto x) = |\propto| \quad f(x) \propto \in R, x \in R^n \quad (5)$$

In terms of notation, in the function $f(x) = \|x\|$, the double bar is used to distinguish between various norms. Useful classes of vector norms are the *p-norms* defined by:

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{\frac{1}{p}} p \geq 1 \quad (6)$$

Of these the 1, 2, and ∞ norms are the most important:

$$\|x\|_1 = |x_1| + \ldots + |x_n|$$
$$\|x\|_2 = (|x_1|^2 + \ldots + |x_n|^2)^{\frac{1}{2}} = (x^T x)^{\frac{1}{2}}$$
$$\|x\|_\infty = |x_i| \quad (7)$$

A unit vector concerning the norm $\|\cdot\|$ is a vector x that satisfies $\|x\| = 1$.

Based on Powell [18], if f and A are in the space $l[ab]$, this is the set of continuous real-valued functions that are defined on the interval $[a, b]$ of the real line. Turning to discrete problems, where f and A are in R^m this is the set of real m-components vectors. All the spaces are linear and there are choices of norms, namely L_p-norms in the cases when $p = 1, 2, \infty$. Then if we consider finite p, the L_p-norms in $f \in [a, b]$ are defined to have the value

$$\|f\|_p = \left[\int_a^b |f(x)|^p \, dx \right]^{1/p} \quad (8)$$

Then for $p = 1, 2, \infty$

$$p = 1 \rightarrow \|f\|_1 = \int_a^b |f(x)| \, dx$$
$$p = 2 \rightarrow \|f\|_2 = \left[\int_a^b f(x)^2 dx \right]^{1/2}$$
$$p = \infty \rightarrow \|f\|_\infty = \max |f(x)| \quad (9)$$

The ∞ norm provides the foundation of much of approximation theory in specific theorems showing that, if it succeeds in finding an approximation $a \in A$ such that the ∞-norm distance function $d(f, a)$ is small, then the 2-norm and 1-norm distance functions are also small [18].

Minimax calculation requires basis function and initial reference determination based on the data being used. The steps of the exchange algorithm for minimax approximation are shown in Figure 1 [29].

Several iterations are performed until convergence is accomplished. Convergence is defined if one of the following criteria is fulfilled:
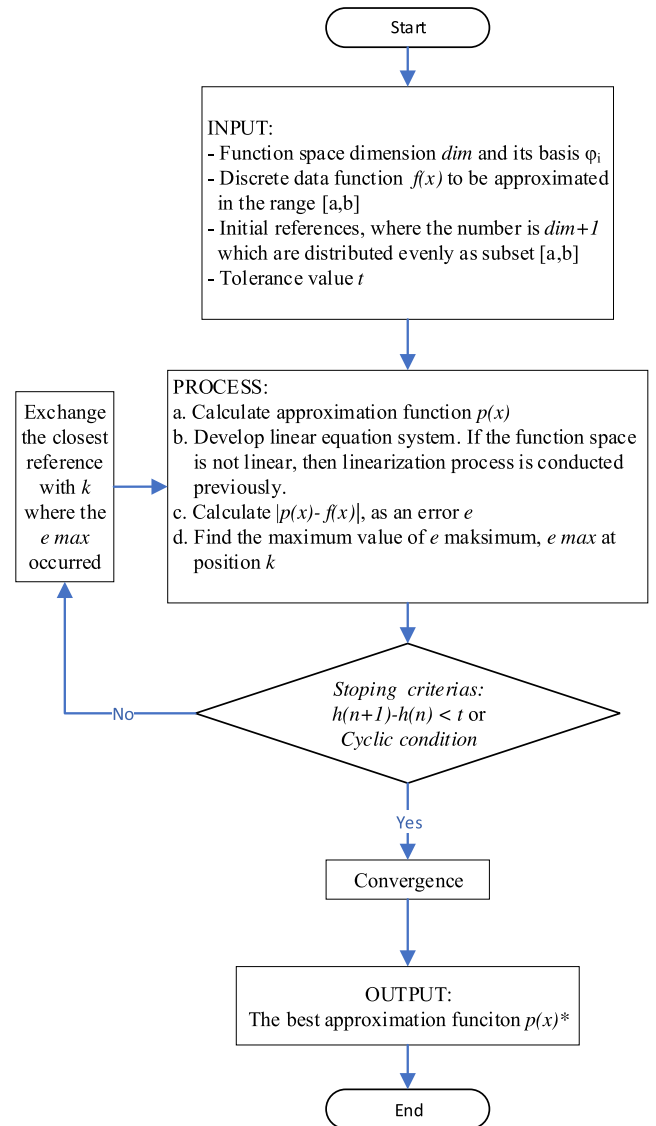

FIGURE 1. Exchange algorithm [29].

1. Conditions when the value of h and the coefficients of the equation are stable. Also, the value of the difference from the iteration n and the previous iteration *n-1*, Δh is lower than the tolerance value, *t*.
2. Cyclic condition or reentry is when the reference point that has been exited in the previous iteration process re-enters the next iteration process.

The output for this process is in the form of function coefficients, $p(x)^*$ after the process reaches convergence. The computational complexity of the exchange algorithm is n^3.

E. BASIS FUNCTIONS
Minimax approximation is applied using basis function as an input. There are several options for functions that might be injected into the process, and exponential and polynomial bases can be used as function inputs. The exponential basis

formula is:

$$f(x) = \alpha + \beta \exp x + \gamma \exp(2x) \qquad (10)$$

$$f(x) = \exp(\alpha + \beta x + \gamma x^2) \qquad (11)$$

while the polynomial option formula is:

$$f(x) = \alpha x^2 + \beta x + \gamma \qquad (12)$$

$$f(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta \qquad (13)$$

The rational reason why these basis function types are chosen is due to the character of semivariogram point distribution. These basis functions are assumed to represent the distribution of data. So, these might require soft analytical or intuition in selecting the right basis function in the expert domain.

F. VALIDATION METHODS

Validation of the output method is conducted by applying mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) to evaluate the differences between height values from the empirical data and the estimated results. The following error formulas have been used in similar research [4], [16]:

$$ME = \frac{1}{n} \sum_{i=1}^{n} \left[z(x_i) - \hat{z}(x_i) \right] \qquad (14)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| z(x_i) - \hat{z}(x_i) \right| \qquad (15)$$

$$RMSE = \left[\frac{\sum_{i=1}^{n} (z(x_i) - \hat{z}(x_i))^2}{n} \right]^{1/2} \qquad (16)$$

where $z(x_i)$ = height value at reference data and $\hat{z}(x_i)$ is predicted height value.

III. METHODOLOGY

The data uses in this experiment is height data derived from satellite acquisition as height point maps. The experimental processing will transform this discrete data into continuous data in the form of a DEM.

The experimental method consists of three major stages: data preparation, the processing experiment, and result validation. Data preparation is a standard image processing method in which the input is raw image data, and the output is a height point map for specific horizontal and vertical data. The processing experiment is shown in Figure 2 in which the input is height data in the form of point clouds and semivariance values are then calculated using formula 4.

Semivariogram fitting or modeling is carried out using three approaches: minimax, machine learning, and weighted least squares. Minimax experimentation uses functions 10, 11, and 12. As the final analysis, the validation stage is conducted using formulas 14, 15, and 16, comparing predicted height value and height value at reference data. The reference data as the most valid value is measured from relative GPS geodetic measurements.
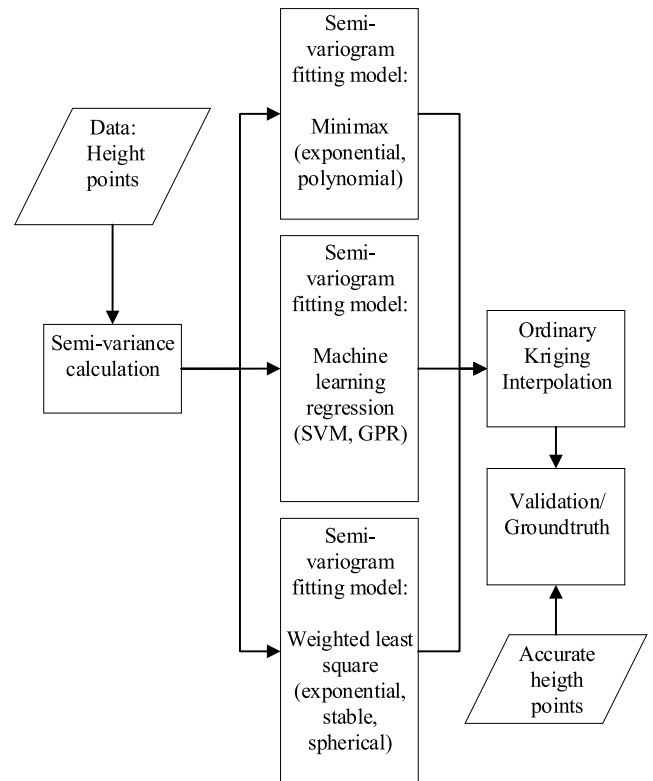

FIGURE 2. Experimental flowchart.

The general algorithm for the research experiment is as follows:
Semivariance calculation:
1   Define the study area.
2   Coordinate transformation into the same horizontal and vertical datum.
3   Calculate semivariance values of selected height points using formula 4.
4   Plot semivariance values into the semivariogram.
Semivariogram fitting model (minimax):
1   Calculate the average of semivariance values for 50–100 values.
2   Define basis function to perform minimax approximation: polynomial (formulas 12 and 13) or exponential (formulas 10 and 11).
3   Define initial references based on the selected basis function.
4   Run the pseudo algorithm as shown in Figure 1 until the iterations reach convergence.
5   Plot the exchange algorithm into the semivariogram fitting model.
Ordinary kriging interpolation:
1   Define the semivariance values based on distances of pair points to solve matrix operation using formula 2.
2   Calculate weight values using formula 2.
3   Define the surrounding measured or known points of the point to be predicted.
4   Calculate the predicted height value of the point using formula 1.

Post-process validation/ground truth:
1 Coordinate transformation into the same horizontal and vertical datum.
2 Calculate the error using formulas 13, 14, and 15.

Weighted least square is implemented using all measured data by applying the following functions: stable, spherical, and exponential. Meanwhile, machine learning regression is implemented on all data by applying several types of SVM and GPR (Gaussian process regression). SVM types being used are SVM linear (SVML), SVM Gaussian (SVMG), SVM quadratic (SVMQ), SVM cubic (SVMC), SVM fine Gaussian (SVMGF), SVM medium Gaussian (SVMGM), SVM coarse Gaussian (SVMGC) and least squares SVM (LS-SVM). While the types of GPR being used are Gaussian process regression, Gaussian process regression rational quadratic (GPRRQ), Gaussian process regression squared exponential (GPRSE), Gaussian process regression matem 5/2 (GPRM), and Gaussian process regression exponential (GPRE). The machine learning technique has been conducted in previous experiments [15], [16].

The research experiment is conducted using several tools, whilst the algorithm is implemented and developed using Matlab software and its apps. Another utilized tool is the ArcGIS Geostatistical Analyst.

IV. DATASET AND STUDY AREA

The novel approach to the semivariogram fitting process is applied to two datasets. Both datasets contain similar data: point clouds generated from optical stereo satellite imagery and validation points drawn from differential Geodetic Global Navigation Satellite System (GNSS) measurements and the national base coordinate system. The detailed specifications of both datasets are shown in Table 1 [30], [31].

Height points in the form of point cloud data are extracted from the stereo imagery using ERDAS Imagine. Typically, the point cloud contains height information (z) of a particular point in the coordinate (x, y). The horizontal and vertical data are set as the same type of datum. The point cloud generated from each stereo satellite imagery is shown in Figure 3.

The study area is selected from data representing both flat and hilly areas. The study area of dataset 1 is located where height (z) varies from 623.23 m to 635.36 m in a relatively flat area covered by the plantations. In Figure 3, the trend projection of dataset 1 is plotted in red and blue lines and shows that the domain is relatively flat. In the study area of dataset 2, the height (z) varies from 700.26 m to 717.86 m, a relatively hilly area mostly covered by plantations. The trend projection of dataset 2 in Figure 4 is plotted in red and blue lines, which also shows that the domain is relatively steep or hilly. The statistical description of the point cloud is shown in Table 2.

V. RESULTS AND ANALYSIS

A. SEMIVARIOGRAM

The experimental methodology is applied to the processing of both datasets. This section describes the results of the

TABLE 1. Dataset specifications.

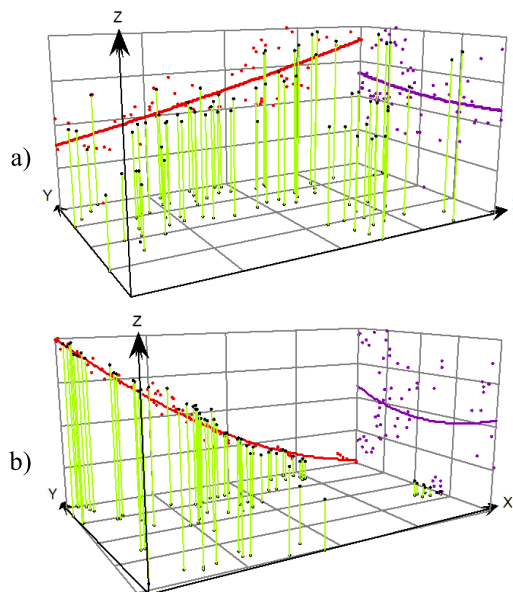| Dataset 1	Dataset 2
Stereo satellite imagery	
Satellite data 1	Satellite data 2
- Imagery: 2-pair stereo panchromatic data - Base to height ratio: 0.62 - Spatial resolution: 2.5 m - Radiometric resolution: 10 bits - Wavelength: 0.5–0.85 μm - Path/row: 0526/0258 - Observation date: October 2, 2005 - Ellipsoid/datum: Everest/Indian 1954 - GCPs: GNSS measurement points	- Imagery: triplet stereo panchromatic data - Base to height ratio: 1.0 (forward view + backward view) and 0.5 (sidelong view + nadir view) - Spatial resolution: 2.5 m (at nadir) - Radiometric resolution: 8 bits - Wavelength: 0.52–0.77 μm - Observation date: August 7, 2006 - Ellipsoid/datum: WGS 84 - Projection system: south universal transverse mercator (SUTM), zone no. 48 - GCPs: GNSS measurement points
Validation data	
Precise GNSS points	National base coordinate system points
- Measurement technique: differential GNSS at a sufficient time of measurement - Tools: geodetic GNSS receiver - Measured time: 2006 - Ellipsoid/datum: Everest/Indian 1954	- High order of national base coordinate system used for national mapping - Source: National map agency database - Horizontal datum WGS 84 - Vertical datum WGS 84

FIGURE 3. Height point cloud of datasets.

processes mainly in terms of the semivariogram fitting model. The initial result is a semivariogram that represents the semivariance values of each point cloud. It is shown in Figure 4.

A semivariogram is a plot of semivariance distribution. This semivariogram will be an input for the semivariogram
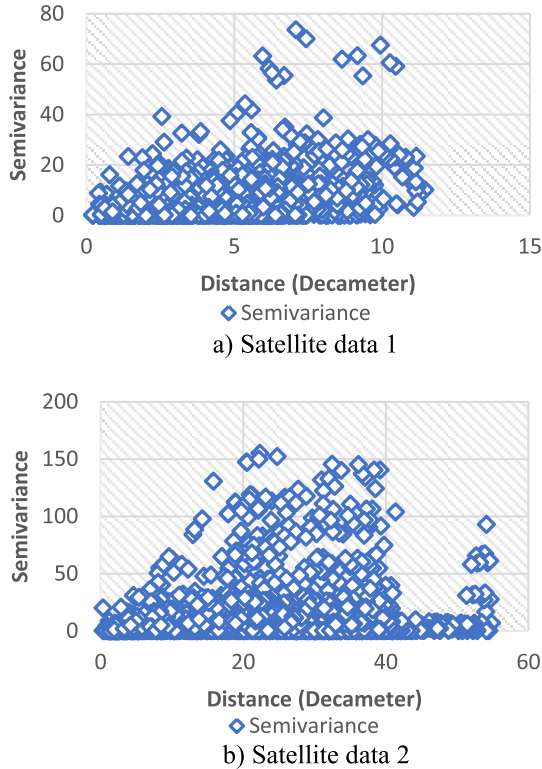
FIGURE 4. Semivariogram of point cloud datasets.

TABLE 2. Point cloud statistical description of the *z* value.

Statistical description	Satellite data 1	Satellite data 2
Min	623.23 m	700.26 m
Max	635.36 m	717.86 m
Mean	630.70 m	704.59 m
Median	630.07 m	703.96 m
Standard deviation	2.3155	3.7121

fitting process. In this experiment, the average value of semivariance for a particular lag is defined. The average value is determined as follows: 20, 30, 40, 50, 60, 70, 80, 90, 100. The values are determined because 50–100 semivariance values are required to build stable semivariogram fitting process [3]. These experiments aim to prove how fewer than 50 semivariance values can define such fitting functions.

B. EXCHANGE ALGORITHM
Minimax approximation is performed to produce an exchange algorithm and each process pass through several iterations until it reaches convergence. The output is in the form of model function coefficients, α, β, γ for each function. The results are shown in Table 3 for dataset 1 and Table 4 for dataset 2.

The minimax process commenced by defining both types of polynomial and exponential functions. The initial reference is based on basis function. In the case of both functions, four initial references are selected randomly over measured points that represents the distribution of the data.

**TABLE 3.** Exchange algorithm for the semivariogram model of dataset 1.

n of $\overline{\gamma^*(h)}$	Polynomial $f(x) = \alpha x^2 + \beta x + \gamma$	Exponential $f(x) = \exp(\alpha + \beta x + \gamma x^2)$
20	$\alpha \ \beta \ \gamma =$ -0.2486 1.6934 -0.0489	$\alpha \ \beta \ \gamma =$ -0.8277 0.1068 0.3948
30	$\alpha \ \beta \ \gamma =$ 0.0463 0.8001 0.8309	$\alpha \ \beta \ \gamma =$ 0.1585 0.3124 0.0230
40	$\alpha \ \beta \ \gamma =$ -0.5122 3.7040 -1.3955	$\alpha \ \beta \ \gamma =$ -0.8339 1.2506 -0.1486
50	$\alpha \ \beta \ \gamma =$ -0.3806 3.2114 -1.2105	$\alpha \ \beta \ \gamma =$ -0.1074 0.8499 -0.0970
60	$\alpha \ \beta \ \gamma =$ 0.1056 0.4434 1.1169	$\alpha \ \beta \ \gamma =$ -0.3629 0.5642 -0.0276
70	$\alpha \ \beta \ \gamma =$ 0.0397 1.0720 -0.4951	$\alpha \ \beta \ \gamma =$ -1.3068 0.9668 -0.0655
80	$\alpha \ \beta \ \gamma =$ 0.1130 0.2524 1.1717	$\alpha \ \beta \ \gamma =$ -0.7546 0.5968 -0.0305
90	$\alpha \ \beta \ \gamma =$ 0.0945 -0.1303 2.9399	$\alpha \ \beta \ \gamma =$ 0.6922 0.1180 0.0074
100	$\alpha \ \beta \ \gamma =$ -0.0096 0.8444 1.0647	$\alpha \ \beta \ \gamma =$ 0.4464 0.2907 -0.0119

**TABLE 4.** Exchange algorithm for the semivariogram model of dataset 2.

n of $\overline{\gamma^*(h)}$	Polynomial $f(x) = \alpha x^2 + \beta x + \gamma$	Exponential $f(x) = \exp(\alpha + \beta x + \gamma x^2)$
20	$\alpha \ \beta \ \gamma =$ 0.0094 0.2474 0.2012	$\alpha \ \beta \ \gamma =$ 1.9474 0.0450 -2.4969e-04
30	$\alpha \ \beta \ \gamma =$ 0.0034 0.5286 3.4387	$\alpha \ \beta \ \gamma =$ 1.1667 0.0914 -8.2254e-04
40	$\alpha \ \beta \ \gamma =$ 0.0184 -0.4280 11.5297	$\alpha \ \beta \ \gamma =$ 0.7196 0.0966 -8.3351e-04
50	$\alpha \ \beta \ \gamma =$ -0.0308 1.7748 -2.8887	$\alpha \ \beta \ \gamma =$ 1.3025 0.0759 -7.3599e-04
60	$\alpha \ \beta \ \gamma =$ 0.0119-0.0414 5 .9430	$\alpha \ \beta \ \gamma =$ 0.1911 0.1247 -0.0012
70	$\alpha \ \beta \ \gamma =$ 0.0191 -0.3140 8.8074	$\alpha \ \beta \ \gamma =$ 0.4727 0.1132 -0.0011
80	$\alpha \ \beta \ \gamma =$ -0.0239 1.5044 -4.7764	$\alpha \ \beta \ \gamma =$ 1.3409 0.0890 -0.0011
90	$\alpha \ \beta \ \gamma =$ -0.0352 1.9870 -5.8391	$\alpha \ \beta \ \gamma =$ 0.5068 0.0960 -3.2916e-04
100	$\alpha \ \beta \ \gamma =$ -0.0450 2.2948 -3.2590	$\alpha \ \beta \ \gamma =$ 0.8992 0.1359 -0.0018

Generally, three to five iterations are needed until convergence is achieved. Several attempts of the iteration produced cyclic conditions, mostly at large numbers of $\overline{\gamma^*(h)}$. This type of circumstance is categorized as an acceptable result of the minimax operation. The samples of both polynomial and exponential semivariogram fitting are shown in Figure 5.

The weight of ordinary kriging is then calculated using a semivariogram model using formula 2. Ordinary kriging interpolation is conducted to produce height prediction based on the weight, using equation 1. The final experiment is the validation process, carried out by measuring residual error using the RMSE technique. This process is conducted by calculating the difference between predicted height values and reference height values. The output results are shown in Table 5 as an error report summary of the values of the minimax experiment and Table 6 as an accuracy report summary of the weighted least square and machine learning experiments.
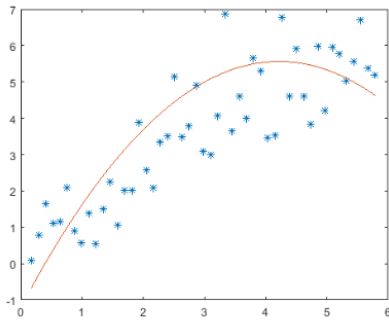
a) Polynomial with n of $\overline{\gamma^*(h)}$=50
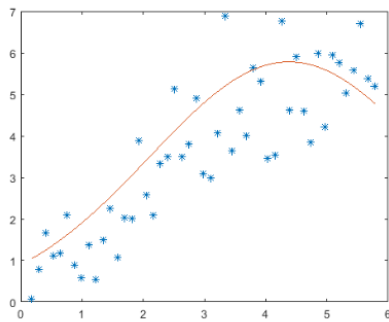
b) Exponential with n of $\overline{\gamma^*(h)}$=50

FIGURE 5. Minimax semivariogram fitting of dataset 1.

TABLE 5. Validation of each minimax semivariogram fitting function of dataset 1.

n of $\overline{\gamma^*(h)}$	Error (m)	
	Polynomial (m)	Exponential (m)
20	0.89	0.87
30	1.06	7.59
40	0.52	0.51
50	**0.45**	0.75
60	1.34	1.27
70	3.68	1.39
80	1.31	1.14
90	1.65	1.49
100	1.02	1.24

C. PERFORMANCE ANALYSIS

The results for dataset 1 show that several attempts of the minimax operation is able to improve accuracy compared to the weighted least-squares approach at a smaller number of $\overline{\gamma^*(h)}$. The least residual error is produced by polynomial minimax at 50 $\overline{\gamma^*(h)}$ of 0.45 m. This accuracy achievement is better than machine learning, which has the best accuracy of 0.46 m. In a machine learning experiment for semivariogram fitting conducted previously [15], machine learning regression processes are set at five-folds.

The overall result of minimax height prediction accuracy is shown in Figure 6 for accuracy trends over different values of $\overline{\gamma^*(h)}$. The figure shows that the polynomial function trend is slightly increasing as the value of $\overline{\gamma^*(h)}$ increases and

TABLE 6. Validation of weighted least square and machine learning experiment of dataset 1.

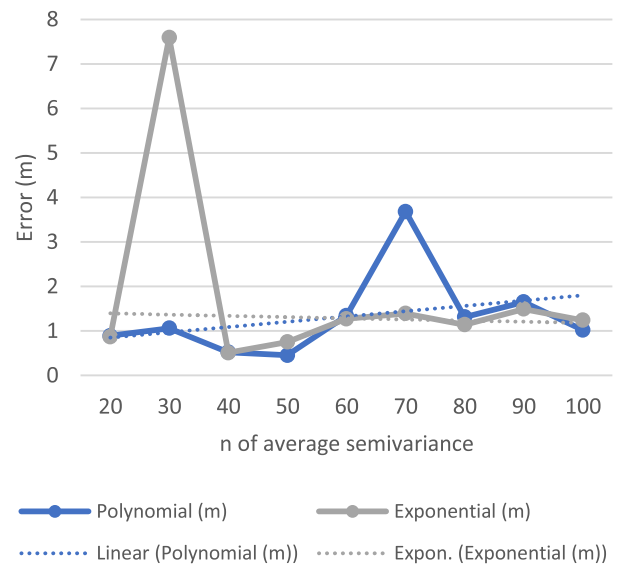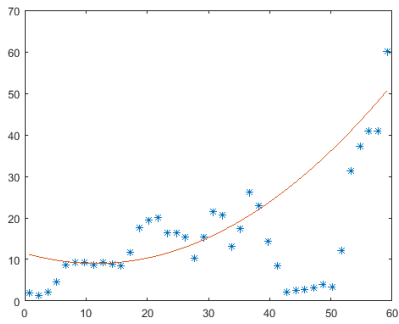| Method	Type	Error (m)
Weighted least square	Stable	1.37
	Exponential	1.34
	Spherical	1.16
Machine learning regression	SVML	0.84
	SVMQ	1.15
	SVMC	7.65
	SVMGM	0.98
	SVMGF	1.37
	SVMGC	0.93
	LS-SVM	0.59
	GPRE	0.46
	GPRM	0.68
	GPRRQ	0.60
	GPRSE	0.79

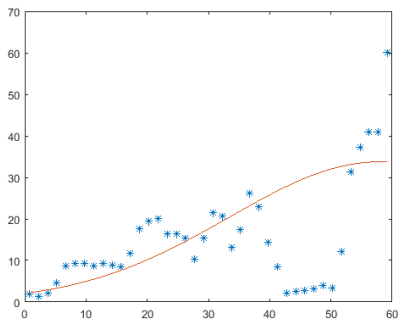**FIGURE 6.** The error of height prediction based on minimax (dataset 1).

the exponential function is relatively flat from overall values of $\overline{\gamma^*(h)}$. There is an extreme error shown at polynomial with n of $\overline{\gamma^*(h)}$ 30 (7.65 m). This phenomenon reflects that minimax is somehow not stable as a possible result of several factors, of which initial reference determination can be one.

Further experimental work is conducted for dataset 2 to gain more analysis and depth of understanding of the minimax operation for semivariogram fitting. The sample results of semivariogram fitting for dataset 2 are shown in Figure 7 and the overall results are shown in Table 7 for the minimax method. Table 8 shows the weighted least squares and machine learning methods.

The results of the minimax operation for semivariogram fitting on dataset 2 are similar to dataset 1. The best result is reached at polynomial with n of $\overline{\gamma^*(h)}$ of 70 (2.06 m). The accuracy is better than most weighted least squares methods but is still less accurate than machine learning regression at 1.62 achieved by the SVMGF method. An unstable

a) Polynomial with n of $\overline{\gamma^*(h)}$=40

b) Exponential with n of $\overline{\gamma^*(h)}$=40
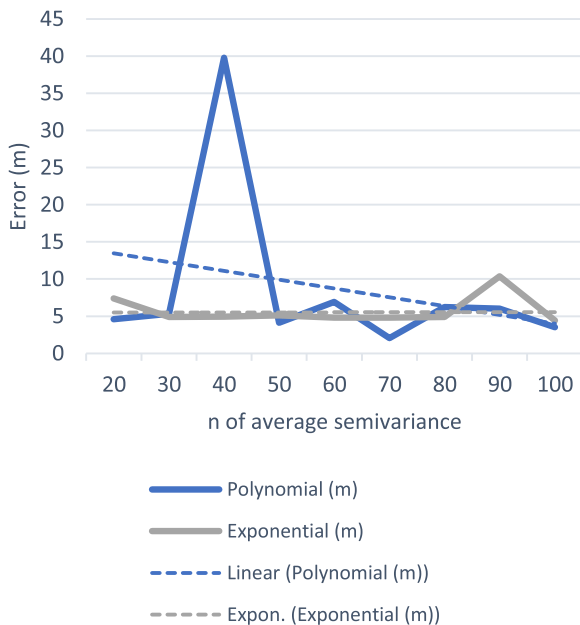
FIGURE 7. Minimax semivariogram fitting of dataset 2.

FIGURE 8. The error of height prediction based on minimax (dataset 2).

TABLE 7. Validation of each minimax semivariogram fitting function of dataset 2.

n of $\overline{\gamma^*(h)}$	Error (m)	
	Polynomial (m)	Exponential (m)
20	4.59	7.38
30	5.31	4.91
40	39.78	4.93
50	4.12	5.12
60	6.91	4.80
70	**2.06**	4.82
80	6.24	4.91
90	6.02	10.36
100	3.51	4.41

TABLE 8. Validation of weighted least square and machine learning experiment of dataset 2.

Method	Type	Error (m)
Weighted least squares	Stable	4.47
	Exponential	4.15
	Spherical	4.40
Machine learning regression	SVML	1.73
	SVMQ	4.55
	SVMC	2.42
	SVMGM	3.56
	SVMGF	1.62
	SVMGC	4.94
	LS-SVM	6.36
	GPRE	4.84
	GPRM	2.41
	GPRRQ	4.48
	GPRSE	27.56

Figure 8 shows that the polynomial function trend is decreasing as the value of $\overline{\gamma^*(h)}$ increases, perhaps caused by the residual error of n of $\overline{\gamma^*(h)}$ of 40. The exponential function is relatively flat on all values of $\overline{\gamma^*(h)}$ and this is similar to the accuracy trend seen for dataset 1.

D. BASIS FUNCTION ANALYSIS
In a broader range of data the accuracy test is conducted with 2nd order polynomial, 3rd order polynomial, and exponential. Error of the three basis functions as shown in Figure 9.

In general, the semivariogram distribution characters that exist in both dataset 1 and dataset 2, exponential function (11) produce the smallest errors then followed by 3rd order polynomial (13) and 2nd order polynomial (12). Therefore it can be said that the exponential function is more precisely applied to the dataset used in this study.

E. COMPUTATIONAL EFFORT
In addition to residual error or accuracy analysis, the measurement of computational effort for minimax and machine learning are also reported. The effort is measured using the processing time needed for the same CPU specification. The computation is performed using the Matlab tool and programming on CPU with specifications as follows: Intel

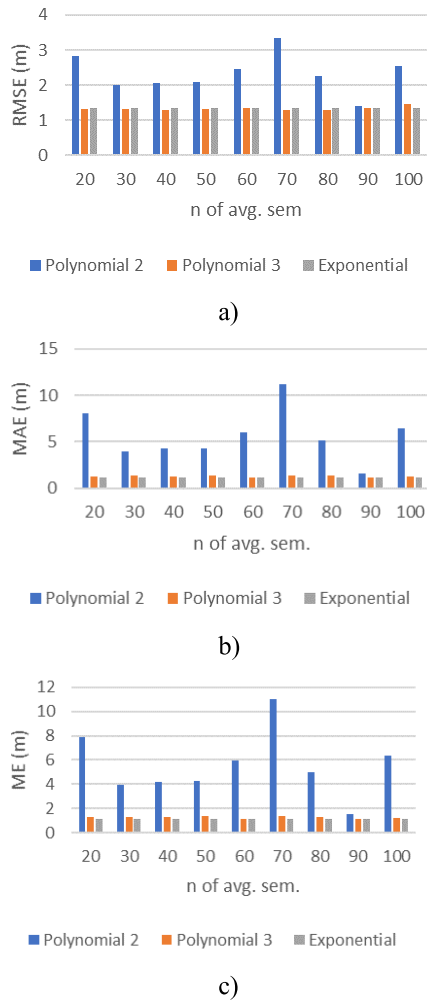phenomenon similar to the previous analysis occurs in the accuracy list where residual increases significantly for the polynomial with n of $\overline{\gamma^*(h)}$ of 40 as shown in Figure 8. The figure presents the overall accuracy of minimax ordinary kriging for dataset 2.

FIGURE 9. Error (m) of each basis function.

TABLE 9. The computational effort of the semivariogram fitting process.

Method	Processing time(s)	
	Dataset 1	Dataset 2
Minimax Poly n of $\overline{\gamma^*(h)}$ =20	6.26	5.26
Minimax Poly n of $\overline{\gamma^*(h)}$ =30	5.31	7.31
Minimax Poly n of $\overline{\gamma^*(h)}$ =40	6.38	3.26
Minimax Poly n of $\overline{\gamma^*(h)}$ =50	6.33	5.29
Minimax Poly n of $\overline{\gamma^*(h)}$ =60	4.29	3.27
Minimax Poly n of $\overline{\gamma^*(h)}$ =70	6.34	4.17
Minimax Poly n of $\overline{\gamma^*(h)}$ =80	6.36	7.27
Minimax Poly n of $\overline{\gamma^*(h)}$ =90	5.34	5.35
Minimax Poly n of $\overline{\gamma^*(h)}$ =100	4.31	4.20
Minimax Exp n of $\gamma^*(h)$ =20	3.28	4.31
Minimax Exp n of $\gamma^*(h)$ =30	4.16	4.29
Minimax Exp n of $\gamma^*(h)$ =40	5.22	3.26
Minimax Exp n of $\gamma^*(h)$ =50	5.31	3.28
Minimax Exp n of $\gamma^*(h)$ =60	4.29	3.26
Minimax Exp n of $\gamma^*(h)$ =70	4.26	4.29
Minimax Exp n of $\gamma^*(h)$ =80	5.32	4.30
Minimax Exp n of $\gamma^*(h)$ =90	5.21	5.33
Minimax Exp n of $\gamma^*(h)$ =100	4.29	4.35
SVML	12.99	13.53
SVMQ	187.07	53.35
SVMC	449.82	459.00
SVMGM	1.82	1.30
SVMGF	2.02	1.60
SVMGC	1.31	1.34
LS-SVM	198.01	177.86
GPRE	121.33	116.39
GPRM	100.47	121.34
GPRRQ	308.03	263.29
GPRSE	140.59	123.88

Core i7-2640M CPU @ 2.80 GHz and 4 RAM. The results are shown in Table 9.

The overall results show that minimax performs faster than machine learning regression for the semivariogram fitting process with the quickest method achieved by SVMGC at 1.31 s for dataset 1 and 1.34 s for dataset 2.

VI. DISCUSSION

Based on the analysis result of the experiment, the minimax approximation can contributes as an alternative to improve kriging interpolation. The benefits or advantages of this technique are accuracy improvement compared to traditional weighted least square and rapid process compared to machine learning approaches. Moreover, minimax approximation guarantees process convergence without a derivative optimization requirement.

Even though this novel approach contributes to decreasing the resulting error, difficulty arise when running the process. Minimax is highly dependent on the selection process of the right functional space. A sufficient understanding of the distribution of the input data is required to select the right functional one. The knowledge lies in the domain of the expert assuming that the expert of the field character knows the real condition.

Both polynomial and exponential basis functions are able to decrease the error. Nevertheless, it has an extreme error at 30-40 numbers of average semivariograms. As the average amount is increasing, the error is declining for dataset 2 but growing for polynomial function for dataset 1. This phenomenon leads to uncertainty in minimax approximation for kriging operation which will be a challenge for further research.

The proposed exchange algorithm can be used for defining weight value for kriging interpolation, as shown in Figure 10. The figure shows that the procedure of the operation is divided into four stages. First stage starts with preprocessing the data where the input is measured or known spatial data points and the output is semivariogram. Then, second stage is the main point of this research where the detail process is explained in Figure 1 and also the production in the form of function coefficients $p(x)^*$. Subsequently, third stage is the kriging process where the input is $p(x)^*$ for weight calculation in kriging operation. And the last stage is validation for determining the error of the final result.

The proposed method might be called as minimax kriging that can be applied for data prediction. It might be
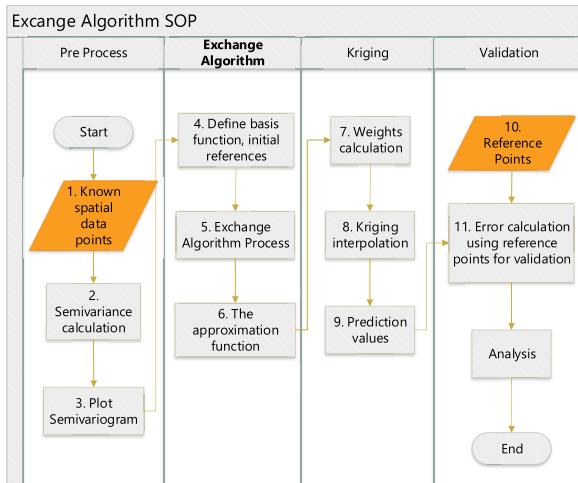
FIGURE 10. Systematic operational procedure of exchange algorithm.

useful for practitioners and researchers in the field of spatial data analysis. The experiment in this research is still developed for generating height points from satellite stereo imagery. The proposed method is deployable for other field of research. Particularly, those who make continuous data from a discrete type of data, for example, height data, temperature data, bathymetry.

VII. CONCLUSION

In summary, overall minimax approximation manages to improve accuracy compared to the weighted least squares method for the semivariogram fitting process. This approach can lead to novelty in kriging interpolation that previously used weighted least squares. Even though minimax is relatively better than the conventional method if compared to the machine learning approach, a similar level of accuracy is achieved. The overall error of both approaches is not significantly different. The advantage of minimax compared to the machine learning approach is less computational effort.

Factors that affects the minimax results are function selection based on the distribution of data and initial reference determination. Both factors can be analyzed in further studies of minimax approximation.

Besides minimizing the uncertainty of the minimax exchange algorithm, the most challenging future work of this research is developing weighted minimax by considering three factors: lag (h), nugget, sill, and range [3] for kriging interpolation.

Further works in the mathematical approach is to build a local basis function by using a specific technique such as spline or wavelet.

CONTRIBUTIONS

Each author contributed to the research experiment based on their expertise. TB is an expert in numerical computing in its contribution to mathematical modeling, AMA is an expert in the field of image processing, and AS has a background and experience in image processing and spatial statistics which contributes to developing programs for implementing novel concepts in spatial areas. This research paper is a multidiscipline collaboration for spatial analysis.

REFERENCES

[1] G. Matheron, *Traité De Géostatistique Appliquée*. Paris, France: Éditions Technip, 1962.

[2] K. T. Chang, *Introduction to Geographic Information Systems*. New York, NY, USA: McGraw-Hill, 2006.

[3] P. A. Burrough and R. A. McDonnell, *Principles of Geographical Information Systems*. Oxford, U.K.: Oxford Univ. Press, 2000.

[4] Q. Shen, Y. Wang, X. Wang, X. Liu, X. Zhang, and S. Zhang, "Comparing interpolation methods to predict soil total phosphorus in the Mollisol area of Northeast China," *CATENA*, vol. 174, pp. 59–72, Mar. 2019.

[5] A. Setiyoko, "Analysis of spatial interpolation techniques for DEM generation using IRS-1C data," in *Proc. 34th Asian Conf. Remote Sens.*, 2013, pp. 1264–1269.

[6] A. Setiyoko and A. M. Arymurthy, "Accuracy analysis of DEM generated from coKriging interpolators," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 88, Oct. 2017, Art. no. 012021.

[7] S. Kumar, R. Lal, and D. Liu, "A geographically weighted regression Kriging approach for mapping soil organic carbon stock," *Geoderma*, vols. 189–190, pp. 627–634, Nov. 2012.

[8] Y. Wang, W. Zhou, J. Luo, H. Yan, H. Pu, and Y. Peng, "Reliable intelligent path following control for a robotic airship against sensor faults," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 6, pp. 2572–2582, Dec. 2019.

[9] Y. Wang, H. R. Karimi, H.-K. Lam, and H. Yan, "Fuzzy output tracking control and filtering for nonlinear discrete-time descriptor systems under unreliable communication links," *IEEE Trans. Cybern.*, early access, Jun. 17, 2019, doi: 10.1109/TCYB.2019.2920709.

[10] K. Johnston, J. M. Ver Hoef, K. Krivoruchko, and N. Lucas, *Using ArcGIS Geostatistical Analyst*, vol. 1. New York, NY, USA: ESRI, 2001.

[11] E. H. Isaaks and R. M. Srivastava, *An Introduction to Applied Geostatistics*. New York, NY, USA: Oxford Univ. Press, 1989.

[12] Z. Huang, H. Wang, and R. Zhang, "An improved Kriging interpolation technique based on SVM and its recovery experiment in oceanic missing data," *Amer. J. Comput. Math.*, vol. 2, no. 1, pp. 56–60, 2012.

[13] C. Youkuo, Y. Yongguo, and W. Wangwen, "Coal seam thickness prediction based on least squares support vector machines and Kriging method," *EJGE*, vol. 20, no. 1, pp. 167–176, 2015.

[14] J. A. K. Suykens, J. Vandewalle, and B. De Moor, "Optimal control by least squares support vector machines," *Neural Netw.*, vol. 14, no. 1, pp. 23–35, Jan. 2001.

[15] A. Setiyoko, A. M. Arymurthy, T. Basaruddin, and R. Arief, "Semivariogram fitting based on SVM and GPR for DEM interpolation," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 311, Aug. 2019, Art. no. 012076.

[16] A. Setiyoko, A. M. Arymurthy, and T. Basaruddin, "DEM fusion concept based on the LS-SVM coKriging method," *Int. J. Image Data Fusion*, vol. 10, no. 4, pp. 244–262, Oct. 2019.

[17] A. Azizah, B. N. Ruchjana, and D. Chaerani, "Uncertain semivariogram model using robust optimization for application of lead pollutant data," *World Sci. News*, vol. 143, no. 2, pp. 155–169, 2020.

[18] M. J. D. Powell, *Approximation Theory and Methods*. Cambridge, U.K.: Cambridge Univ. Press, 1981.

[19] K. Schmidt, "A comparison of minimax and least squares estimators in linear regression with polyhedral prior information," *Acta Applicandae Mathematicae*, vol. 43, no. 1, pp. 127–138, Apr. 1996.

[20] Z. Ben-Haim and Y. C. Eldar, "Minimax estimators dominating the least-squares estimator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2005, vol. 4, no. 7, pp. 53–56.

[21] Z. Ben-Haim and Y. C. Eldar, "Blind minimax estimation," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3145–3157, Sep. 2007.

[22] Z. Ben-Haim and Y. C. Eldar, "Blind minimax estimators: Improving on least-squares estimation," in *Proc. IEEE/SP 13th Workshop Stat. Signal Process.*, Jul. 2005, pp. 545–550.

[23] G. Leibovitz, A. Elron, Z. Ben-Haim, and Y. C. Eldar, "Recursive blind minimax estimation: Improving mse over recursive least squares," in *Proc. IEEE 25th Conv. Electr. Electron. Engineers Isr.*, Mar. 2008, pp. 548–552.

[24] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.

[25] H. Wackernagel, "Ordinary Kriging," in *Multivariate Geostatistics: An Introduction With Applications*. Berlin, Germany: Springer, 1995, pp. 74–81.

[26] K. Johnston, J. M. Ver Hoef, K. Krivoruchko, and N. Lucas, *ArcGIS 9: Using ArcGIS Geostatistical Analyst*, vol. 5, no. 2. Redlands, CA, USA: ESRI Press, 2001, pp. 131–166.

[27] H. Maitre, *Image Processing*. Hoboken, NJ, USA: Wiley, 2008.

[28] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The John Hopkins Univ. Press, 1996.

[29] C. Arévalo and C. Führer, "Numerical approximation the exchange algorithm," Lund Univ., Lund, Sweden, Tech. Rep. 10, 2012.

[30] P. V. Radhadevi, V. Nagasubramanian, A. Mahapatra, S. S. Solanki, K. Sumanth, and G. Varadan, "Potential of high-resolution Indian remote sensing satellite imagery for large scale mapping," in *Proc. ISPRS*, 2009, pp. 1–5.

[31] H. Hasegawa, K. Matsuo, M. Koarai, N. Watanabe, and H. Masaharu, "DEM accuracy and the base to height (B/H) ratio of stereo images," *Int. Arch. Photogramm. Remote Sens.*, vol. 33, no. 4, pp. 356–359, 2000.

ANDIE SETIYOKO received the B.Eng. degree from the Geodetic Engineering, Institute Technology Bandung, and the M.Tech. degree from the Indian Institute of Remote Sensing. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science, Universitas Indonesia. He is currently a Researcher with the Remote Sensing Technology and Data Center, National Institute of Aeronautics and Space (LAPAN). He has experimented in spatial statistics to perform best interpolation for generating DEMs (digital elevation models) from satellite imagery data.

T. BASARUDDIN is currently a Professor of computer science with the Faculty of Computer Science, Universitas Indonesia. He is also an Expert in numerical computation, modeling, and simulation. His skills and expertise are in optimizations, mathematical modeling, mathematical programming, distributed systems, nonlinear optimization, and optimization theory. He has published scientific articles based on his expertise.

ANIATI MURNI ARYMURTHY is currently a Professor of computer science with the Faculty of Computer Science, Universitas Indonesia. She is an Expert in image processing and spatial data. Her skills and expertise are in classifications, feature extraction, algorithms, feature selection, algebra, genetic algorithms, image enhancement, wavelet, and self-organizing maps. She has published a large number of scientific articles in her fields.

• • •