Optimal prediction in isotropic spatial process under spherical type variogram model with application to corn plant data

W Somayasa\textsuperscript{1(\ast)}, D K Sutiari\textsuperscript{2} and W Sutisna\textsuperscript{3}

\textsuperscript{1,3)Department of Mathematics, Halu Oleo University
\textsuperscript{2)Department of Medical Electrical Engineering, Mandala Waluya School of Health Sciences

E-mail: \textsuperscript{1)wayan.somayasa@uho.ac.id (corresponding author), \textsuperscript{2)sutiaridesak@gmail.com, \textsuperscript{3)wilansutisna@gmail.com

Abstract. A central problem in modelling of spatial data is the construction of an optimal prediction map of a numerical quantity or variable under study. In geostatistics this study is commonly called universal kriging. In this paper we study weighted universal kriging method in which a weight function is used for improving the efficiency of the predictor. We consider isotropic spatial process under a variogram function belongs to the parametric family of spherical type. The parameters of the postulated variogram model are estimated by applying ordinary least squares method in that the squared distance between the variogram model and the associated variogram sample is minimized. Numerical approximation for finding the solution of the least squares equations is conducted by using graphical approach. The application of the method to a corn plant data results in the kriging map of the maximum height that can be achieved by the corn plants planted over a rectangular farm land. This investigation result can in advance describe the map of the fertility level of the farm land where the corn have been planted.

1. Introduction
One of important statistical problem in spatial modelling is optimal prediction of future observation and other quantities representing the characteristics of some phenomena. By a valid prediction formula, practitioners can construct a map figuring the behavior of the observations on unobserved places. This is one reason why prediction techniques of spatial data have been extensively studied in various disciplines, such as in agriculture, economics, biology, atmospheric sciences, and geosciences including geology and geostatistics, see [1, 2, 3, 4, 5]. In the field of Geostatistics, optimal linear prediction is usually called kriging which refers to the South African mining engineer Danie G. Krige (1919-2013) who pioneered the work in mining industry.

The present paper aims at the derivation of a type of universal kriging formula for stationary spatial process in which the kriging coefficients are computed based on a weight function represented by a kernel function, so that the coefficients are calculated locally instead of globally. As documented in [1, 6] and recently in [7], the prediction of an unknown observation is calculated based on the whole sample collected over an experimental region. This kind of prediction can lead to inefficiency as well as overestimation. In order to avoid such kind of result, we propose in this work to consider kernel weighted prediction method for isotropic spatial process with
spherical type variogram model. The notion of isotropy refers to a type of intrinsic stationarity in which the covariogram function between any two observations is expressed as the function of the Euclidean length of the lag vector connecting both observations, cf. [1, 2, 3, 4, 5]. Isotropy is sometimes not reasonable in practice, however it is the basic of advance theory in spatial modelling, cf. [5]. Therefore advance relevant theory in this field is still developed by many researchers especially in diagnostic method for checking the existence of isotropy. Classical descriptive method using rose plot as well as asymptotic and nonparametric test methods can be found in the literatures of spatial analysis, such as those investigated in [6, 8, 9].

To understand the problem in detail let $Z(x, y)$ be an observed random variable measured on the point $(x, y) \in D$ admitting the following decomposition:

$$Z(x, y) = h(x, y) + \delta(x, y),$$

(1)

where $h$ is an unknown deterministic function which has bounded variation on the experimental region $D = [a_1, b_1] \times [a_2, b_2]$. The spatial process $\{\delta(x, y) : (x, y) \in D\}$ is a second-order stationary random field, with $E(\delta(x, y)) = 0, \forall (x, y) \in D$ and $Cov(\delta(x, y), \delta(x + a_1, y + a_2)) = Cov(\delta(0, 0), \delta(a_1, a_2)) = C_\delta(a_1, a_2)$, for some function $C_\delta(\cdot)$ on $D$ and for any spatial lag $a = (a_1, a_2) \in D$, cf. [1], p. 53. The function $C_\delta(\cdot)$ is called the covariogram function of $\delta$. Variogram function of a stationary random field $\delta$ is defined by

$$2\gamma_\delta(a) = Var(\delta(x, y) - \delta(x + a_1, y + a_2))$$

$$= E(\delta(x, y) - \delta(x + a_1, y + a_2))^2, \ a = (a_1, a_2) \in D.$$ 

The function $\gamma_\delta(\cdot)$ is called semivariogram function of $\Delta$. In other word, the covariogram and the variogram of a stationary random field depend only on the lag vector connecting the points. They are related each other by the formula

$$2\gamma_\delta(a) = 2 \left( C_\delta(0, 0) - C_\delta(a) \right).$$

(2)

As an immediate consequent of these properties of $\delta$, we get that the observed process $\{Z(x, y) : (x, y) \in D\}$ is not stationary unless the trend $h$ is constant on $D$. However, it can be shown that the covariogram and variogram functions of $Z$ are the same with those of $\delta$, in the sense $C_Z(a) = C_\delta(a)$ and $2\gamma_Z(a) = 2\gamma_\delta(a)$, for any $a \in D$. In universal kriging it is commonly assumed that the observed variable $Z$ follows a linear regression model

$$Z(x, y) = \sum_{j=1}^{p+1} \beta_j f_j(x, y) + \delta(x, y), \ (x, y) \in D,$$

(3)

where the known regression functions $f_0, f_1, \ldots, f_p$ are linearly independent in $L_2(D)$. Let $\{Z(x_i, y_i) : (x_i, y_i) \in D, \ i = 1, \ldots, n\}$ be a set of observation of Model 2. Our goal is to establish optimal linear prediction (universal kriging) of $Z(x_0, y_0)$ in any point $(x_0, y_0) \in D$ taking the functional form given by

$$\hat{Z}(x_0, y_0) := \sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)Z(x_k, y_k),$$

(4)

where $\alpha_1, \ldots, \alpha_n$ are real constants and $\varphi_0(\cdot)$ is a weight function. The mean square error (MSE) of $\hat{Z}(x_0, y_0)$ is defined by

$$MSE(\hat{Z}(x_0, y_0)) = E(\hat{Z}(x_0, y_0) - Z(x_0, y_0))^2.$$
The predictor $\hat{Z}(x_0, y_0)$ is said to be optimal, if and only if it minimizes $MSE(\hat{Z}(x_0, y_0))$, subject to the condition

$$
\sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)f_{j-1}(x_k, y_k) = f_{j-1}(x_0, y_0), \; j = 1, \ldots, p + 1. \tag{5}
$$

When the weight function is a constant function we get the linear predictor defined in [1, 6] which shared the weight uniformly to all of the observations. Such kind of prediction may result in overestimation. One way of overcoming this problem is by attaching a weight function in the prediction formula with the intention to restrict the samples included in the prediction. Practical examples of weight functions which are usually adopted in the spatial modelling have been documented in [10, 11], such as exponential weight defined as

$$
\varphi_0(x, y) := \exp \left\{-\frac{1}{\lambda} \| (x_0, y_0) - (x, y) \|^2 \right\}, \; (x, y) \in D
$$
or a bisquare function, that is

$$
\varphi_0(x, y) := \left(1 - \frac{1}{\theta^2} \| (x_0, y_0) - (x, y) \|^2 \right)^2, \; (x, y) \in D.
$$

In this paper we aim at finding the prediction coefficients $\alpha_1, \ldots, \alpha_n$ defined in (4) that minimize $MSE(\hat{Z}(x_0, y_0))$ subject to (5), so that we can build a map of prediction. We notice that in the statistical decision theory the notion of MSE is studied as a quantity for measuring the efficiency of an estimator, see [12, 13]. An estimator is called efficient if its MSE is well defined and no other estimator has a smaller MSE.

Throughout this paper we assume that the trend $h$ in (1) is a polynomial of two variables with $f_0 \equiv 1$. Standard method using step-wise method based on F-test and T-test documented in [14, 15] or some methods based partial sums process of the residuals proposed in [16, 17, 18] can not be applied for checking the validity of Model 3. This is because of the existence of the covariogram function. By this reason, the problem of establishing model validity check whether or not Model 3 holds true can be very difficult. Therefore an alternative approach based on the asymptotic result due to Somayasa and Wibawa [19, 20] will be investigated by beforehand estimating the unknown covariogram function $C_Z(\cdot)$ as well as the variogram function $2\gamma_Z(\cdot)$ using least squares method, see also [1, 6].

We organize the remainder of the paper as follows. In Section 2 we derive the formula for computing the weighted universal kriging for stationary spatial process in term of variogram as well as covariogram functions. Least square estimation formula for the parameters of spherical varigram function is also discussed in this section. An application of the kriging procedure in real data is presented in Section 3 by considering the rate of growth of corn plants over a 16 x 21 regular lattice. The paper is closed with some important conclusions and assessments for future research in this field of study.

2. Kernel weighted universal kriging

In this section we investigate the solution of the minimization problem mentioned in the preceding section. We show first that under (4) and under the additional assumption $f_0 \equiv 1$, the predictor $\hat{Z}(x_0, y_0)$ is unbiased to $h(x_0, y_0)$, that is $E(\hat{Z}(x_0, y_0)) = h(x_0, y_0)$, provided Model 3 is fulfilled. Furthermore, it holds $\sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k) = 1$. By recalling all these supporting results, we get the following expression for the MSE of $\hat{Z}(x_0, y_0)$:

$$
MSE(\hat{Z}(x_0, y_0)) = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \alpha_k \alpha_\ell \varphi_0(x_k, y_k) \varphi_0(x_\ell, y_\ell) E(\delta(x_k, y_k) - \delta(x_0, y_0)) (\delta(x_\ell, y_\ell) - \delta(x_0, y_0)).
$$
Based on the simple algebraic manipulation \(2ab = a^2 + b^2 - (a-b)^2\), for \(a := (\delta(x_k, y_k) - \delta(x_0, y_0))\) and \(b := \delta(x_\ell, y_\ell) - \delta(x_0, y_0)\) we can get the following equation

\[
\mathbb{E}[(\delta(x_k, y_k) - \delta(x_0, y_0))(\delta(x_\ell, y_\ell) - \delta(x_0, y_0))]
= \gamma \delta(x_k - x_0, y_k - y_0) + \gamma \delta(x_\ell - x_0, y_\ell - y_0) - \gamma \delta(x_k - x_\ell, y_k - y_\ell).
\]

However, for Model 1 it holds \(\gamma_0(\cdot) = \gamma_Z(\cdot)\), so, the last equation can also be written as follows:

\[
\mathbb{E}[(\delta(x_k, y_k) - \delta(x_0, y_0))(\delta(x_\ell, y_\ell) - \delta(x_0, y_0))]
= \gamma_Z(x_k - x_0, y_k - y_0) + \gamma_Z(x_\ell - x_0, y_\ell - y_0) - \gamma_Z(x_k - x_\ell, y_k - y_\ell).
\]

Consequently, the MSE of \(\hat{Z}(x_0, y_0)\) can also be expressed in term of \(\gamma_Z(\cdot)\) as below:

\[
MSE(\hat{Z}(x_0, y_0)) = 2 \sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)\gamma_Z(x_k - x_0, y_k - y_0)
- \sum_{k=1}^{n} \sum_{\ell=1}^{n} \alpha_k \alpha_\ell \varphi_0(x_k, y_k)\varphi_0(x_\ell, y_\ell)\gamma_Z(x_k - x_\ell, y_k - y_\ell),
\]

where

\[
\gamma_Z(x_k - x_\ell, y_k - y_\ell) := \text{Var}(Z(x_k, y_k) - Z(x_\ell, y_\ell))
\gamma_Z(x_k - x_0, y_k - y_0) := \text{Var}(Z(x_k, y_k) - Z(x_0, y_0)).
\]

Theorem 2.1 below gives the formula for computing the values of the constants \(\alpha_1, \ldots, \alpha_n\) that minimize \(MSE(\hat{Z}(x_0, y_0))\) subject to (5). For the case of constant weight the solution has been established in [1], p.153 and [6], p.26 by applying Lagrange method.

**Theorem 2.1** Let \(\Psi := (\varphi_0(x_i, y_j)f_j-1(x_i, y_j))_{i=1}^{n}_{j=1}^{p}\) be the \(n \times (p+1)\) dimensional weighted design matrix of Model 3 whose entry in the \(i\)-th row and \(j\)-th column is given by \(\varphi_0(x_i, y_j)f_j-1(x_i, y_j)\), for \(j = 1, \ldots, p+1\) and \(i = 1, \ldots, n\), with \(\text{rank}(\Psi) = p+1\). Let \(\mathbf{G} = (\varphi_0(x_\ell, y_k)\gamma_Z(x_\ell - x_k, y_\ell - y_k)\varphi_0(x_k, y_k))_{\ell=1}^{n}_{k=1}^{n}\) be the \(n \times n\) dimensional matrix of the weighted semivariogram function of \(\{Z(x_\ell, y_k) : (x_\ell, y_k) \in \mathbb{D}\}\). Furthermore, let \(\gamma_Z(0) := (\varphi_0(x_1, y_1)\gamma_Z(x_1 - x_0, y_1 - y_0), \ldots, \varphi_0(x_n, y_n)\gamma_Z(x_n - x_0, y_n - y_0))^\top\). Then the value of \(\alpha := (\alpha_1, \ldots, \alpha_n)^\top\) that minimize \(MSE(\hat{Z}(x_0, y_0))\), subject to (5) is given by

\[
\alpha^\top = \left\{\gamma_Z(0) + \Psi(\Psi^\top \mathbf{G}^{-1} \Psi)^{-1}(f_0 - \Psi^\top \mathbf{G}^{-1} \gamma_Z(0))\right\}^\top \mathbf{G}^{-1},
\]

where \(f_0 := (f_0(x_0, y_0), f_1(x_0, y_0), \ldots, f_p(x_0, y_0))^\top\).

**Proof:** Let \(\mathcal{G} : \mathbb{R}^{n(p+1)} \mapsto \mathbb{R}\) be a real-valued function defined on \(\mathbb{R}^{n(p+1)}\), given by

\[
\mathcal{G}(\alpha_1, \ldots, \alpha_n, L_0, \ldots, L_p) = 2 \sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)\gamma_Z(x_k - x_0, y_k - y_0)
- \sum_{k=1}^{n} \sum_{\ell=1}^{n} \alpha_k \alpha_\ell \varphi_0(x_k, y_k)\varphi_0(x_\ell, y_\ell)\gamma_Z(x_k - x_\ell, y_k - y_\ell)
- 2 \sum_{j=1}^{p+1} L_{j-1} \left( \sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)f_{j-1}(x_k, y_k) - f_{j-1}(x_0, y_0) \right),
\]
where \( \mathbf{L} := (L_0, \ldots, L_p)^\top \) is the vector of Lagrange multipliers associated with the \( p+1 \) constrains \( \sum_{k=1}^n \alpha_k \varphi_0(x_k, y_k) f_{j-1}(x_k, y_k) = f_{j-1}(x_0, y_0) \), for \( j = 1, \ldots, p + 1 \). Then for \( i = 1, \ldots, n \), we get

\[
\frac{\partial \mathcal{G}(\alpha_1, \ldots, \alpha_n, L_0, \ldots, L_p)}{\partial \alpha_i} = 0 \iff 2\varphi_0(x_i, y_i) \gamma_Z(x_i - x_0, y_i - y_0)
\]

\[
-2 \sum_{k=1}^n \alpha_k \varphi_0(x_i, y_i) \varphi_0(x_k, y_k) \gamma_Z(x_i - x_k, y_i - y_k) - 2 \sum_{j=1}^{p+1} L_{j-1} \varphi_0(x_i, y_i) f_{j-1}(x_i, y_i) = 0
\]

\[
\iff \sum_{k=1}^n \alpha_k \varphi_0(x_i, y_i) \varphi_0(x_k, y_k) \gamma_Z(x_i - x_k, y_i - y_k) + \sum_{j=1}^{p+1} L_{j-1} \varphi_0(x_i, y_i) f_{j-1}(x_i, y_i) = \varphi_0(x_i, y_i) \gamma_Z(x_i - x_0, y_i - y_0), \tag{7}
\]

for \( i = 1, \ldots, n \). Next, for \( j = 1, \ldots, p + 1 \), it holds

\[
\frac{\partial \mathcal{G}(\alpha_1, \ldots, \alpha_n, L_0, \ldots, L_p)}{\partial L_{j-1}} = 0 \iff -2 \left( \sum_{k=1}^n \alpha_k \varphi_0(x_k, y_k) f_{j-1}(x_k, y_k) - f_{j-1}(x_0, y_0) \right) = 0
\]

\[
\iff \sum_{k=1}^n \alpha_k \varphi_0(x_k, y_k) f_{j-1}(x_k, y_k) = f_{j-1}(x_0, y_0). \tag{8}
\]

Both system of linear equations (7) and (8) can be written in the following block matrix and bloc vector forms:

\[
\begin{pmatrix}
\mathbf{G} & \mathbf{\Psi} \\
\mathbf{\Psi}^\top & \mathbf{O}
\end{pmatrix}
\begin{pmatrix}
\alpha \\
\mathbf{L}
\end{pmatrix}
= \begin{pmatrix}
\gamma_Z(0) \\
\mathbf{f}_0
\end{pmatrix}, \tag{9}
\]

where \( \mathbf{O} \) is \((p+1) \times (p+1)\)-dimensional zero matrix. The solution of (9) is given by

\[
\begin{pmatrix}
\alpha \\
\mathbf{L}
\end{pmatrix}
= \begin{pmatrix}
\mathbf{G} & \mathbf{\Psi} \\
\mathbf{\Psi}^\top & \mathbf{O}
\end{pmatrix}^{-1}
\begin{pmatrix}
\gamma_Z(0) \\
\mathbf{f}_0
\end{pmatrix}.
\]

Furthermore, by applying the well-known inverse formula of block matrix, see e.g. [21, 22], we finally get the following results

\[
\begin{pmatrix}
\alpha \\
\mathbf{L}
\end{pmatrix}
= \begin{pmatrix}
\mathbf{G}^{-1} - \mathbf{G}^{-1} \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{\Psi}^\top \mathbf{G}^{-1} & \mathbf{G}^{-1} \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \\
(\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{\Psi}^\top \mathbf{G}^{-1} & -(\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1}
\end{pmatrix}
\begin{pmatrix}
\gamma_Z(0) \\
\mathbf{f}_0
\end{pmatrix}
= \begin{pmatrix}
\mathbf{G}^{-1} (\gamma_Z(0) + \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0)) \\
-(\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0)
\end{pmatrix}
= \begin{pmatrix}
\mathbf{G}^{-1} (\gamma_Z(0) + \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0)) \\
-(\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0)
\end{pmatrix}.
\]

Hence, we obtain

\[
\alpha = \mathbf{G}^{-1} (\gamma_Z(0) + \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi} (\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0))
\]

\[
\mathbf{L} = -(\mathbf{\Psi}^\top \mathbf{G}^{-1} \mathbf{\Psi})^{-1} \mathbf{f}_0 - \mathbf{\Psi}^\top \mathbf{G}^{-1} \gamma_Z(0),
\]

finishing the proof of the theorem.
It is worth mentioning that the kriging equation can also be expressed in terms of the covariogram function of $Z$ by recalling identity (2), so we get

$$MSE\left(\hat{Z}(x_0, y_0)\right) = -2\sum_{k=1}^{n} \alpha_k \varphi_0(x_k, y_k)C_Z(x_k - x_0, y_k - y_0)$$

$$\quad + \sum_{k=1}^{n} \sum_{\ell=1}^{n} \alpha_k \alpha_\ell \varphi_0(x_k, y_k) \varphi_0(x_\ell, y_\ell)C_Z(x_k - x_\ell, y_k - y_\ell).$$

By applying differential method as in the case of kriging using variogram function, we get the solution of the minimization problem as follows:

$$\alpha^\top = \left\{C_Z(0) + \Psi(\Psi^\top \Sigma^{-1} \Psi)^{-1}(f_0 - \Psi^\top \Sigma^{-1} C_Z(0))\right\}^\top \Sigma^{-1} \Psi^\top \Sigma^{-1}$$

$$L^\top = \left(f_0 - \Psi^\top \Sigma^{-1} C_Z(0)\right)^\top \left(\Psi^\top \Sigma^{-1} \Psi\right)^{-1},$$

where

$$\Sigma := (\varphi_0(x_k, y_k)C_Z(x_k - x_\ell, y_k - y_\ell)\varphi_0(x_\ell, y_\ell))_{k=1, \ell=1}^{n, n}$$

$$C_Z(0) := (\varphi_0(x_1, y_1)C_Z(x_1 - x_0, y_1 - y_0), \ldots, \varphi_0(x_n, y_n)C_Z(x_n - x_0, y_n - y_0))^\top,$$

see also [1, 6, 7] for constant weight.

3. Spherical model

The variogram as well as covariogram functions embedded in the kriging formulas must belong to a valid model. Bochner theorem gives guarantee that a variogram function of a stationary spatial process is said to be valid if and only if it is conditionally negative definite, see [1], p. 84 and [2]. More concretely, a variogram function $\gamma_Z(\cdot)$ is conditionally negative definite, if only if for every points $s_1, \ldots, s_n$ on $D$ and constants $\kappa_1, \ldots, \kappa_n$ with $\sum_{i=1}^{n} \kappa_i = 0$, it implies $\sum_{i=1}^{n} \sum_{j=1}^{n} \kappa_i \kappa_j \gamma_Z(s_i - s_j) < 0$.

For the stationary spatial process $\{Z(x, y) \mid (x, y) \in D\}$, a variogram function $\gamma_Z(\cdot)$ is said to be spherical type if for every points $(x_k, y_k)$ and $(x_\ell, y_\ell)$, with $(x_k - x_\ell, y_k - y_\ell) = h$, it holds

$$\gamma_Z(h) := \begin{cases} 
0 & \|h\| = 0 \\
c_0 + c_s \left(\frac{3}{2}(\|h\|/a_s) - \frac{1}{2}(\|h\|/a_s)^3\right) & 0 < \|h\| \leq a_s \\
c_0 + c_s & a_s < h
\end{cases}$$

with $0 \leq c_0, 0 \leq c_s$ and $0 \leq a_s$. Spherical model belongs to parametric isotropic variogram model since it depends only on the Euclidean distance of any two considered points. To show that it is negative definite, let $s_1 = (x_1, y_1), \ldots, s_n = (x_n, y_n)$ be any $n$ pints on $D$ and $\kappa_1, \ldots, \kappa_n$ be any real constants, such that $\sum_{i=1}^{n} \kappa_i = 0$. Suppose that $\lambda_{ij} := \|s_i - s_j\|$, for $i, j = 1, \ldots, n$. We assume that $\max_{1 \leq i, j \leq n} \|s_i - s_j\| = b > 0$, with $0 < b < a_s$. Then we have by recalling $\gamma_Z(s_i - s_j) = 0$, for $i = j$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \kappa_i \kappa_j \gamma_Z(s_i - s_j) = c_0 \left(\sum_{i=1}^{n} \alpha_i\right)^2 + 2 \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} \alpha_i \alpha_j c_s \left\{\frac{3}{2}(\lambda_{ij}/a_s) - \frac{1}{2}(\lambda_{ij}/a_s)^3\right\}$$

$$\leq 2 \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} \alpha_i \alpha_j c_s \left\{\frac{3}{2}(b/a_s) - \frac{1}{2}(\lambda_{ij}/a_s)^3\right\}$$

$$\leq \frac{3c_s}{2}(b/a_s) \left(\sum_{i=1}^{n} \alpha_i^2 - \frac{\sum_{i=1}^{n} \alpha_i^2}{n}\right) = -\frac{3c_s}{2}(b/a_s) \sum_{i=1}^{n} \alpha_i^2 < 0.$$
Thus, by Bochner theorem (cf. [1], p. 84 and [2]) we can conclude that spherical model is a valid variogram model.

In the application, the parameters of the model $c_0, c_s$ and $a_s$ are usually unknown. The estimation of the unknown parameters are obtained by fitting the variogram model to the sample variograms which are defined by

$$2\hat{\gamma}_Z(h) := \frac{1}{|N(h)|} \sum_{(x_i, y_i) \in N(h)} (Z(x_i, y_i) - Z(x_j, y_j))^2,$$

where $N(h) := \{(x_i, y_i), (x_j, y_j) : (x_i - x_j, y_i - y_j) = h\}$, that is the number of pairs having the lag $h$, see also [1, 6]. Let $\{h_1, h_2, \ldots, h_m\}$ be a set of $m$ lag vectors on $D$. Let $\{2\hat{\gamma}_Z(h_1), 2\hat{\gamma}_Z(h_2), \ldots, 2\hat{\gamma}_Z(h_m)\}$ and $\{2\gamma_Z(h_1; c), 2\gamma_Z(h_2; c), \ldots, 2\gamma_Z(h_m; c)\}$ be the corresponding sets of sample and theoretical parametric variogram models of spherical type, respectively, with the parameter vector $c := (c_0, c_s, a_s)^T \in \mathbb{R}^3$. Least squares estimator $\hat{c} = (\hat{c}_0, \hat{c}_s, \hat{a}_s)^T$ of $c := (c_0, c_s, a_s)^T$ is the vector of constants that satisfies

$$L(\hat{c}_0, \hat{c}_s, \hat{a}_s) := \min_{(c_0, c_s, a_s) \in \mathbb{R}^3} \sum_{i=1}^{m} (2\hat{\gamma}_Z(h_i) - 2\gamma_Z(h_i; c))^2.$$

By applying differential method, the least square estimators of $c_0, c_s$ and $a_s$ is obtained by solving the following differential equations system

$$\frac{\partial L(c_0, c_s, a_s)}{\partial c_0} = 0, \quad \frac{\partial L(c_0, c_s, a_s)}{\partial c_s} = 0, \quad \text{and} \quad \frac{\partial L(c_0, c_s, a_s)}{\partial a_s} = 0.$$ 

By simple algebraic derivation, from the first equation we get

$$c_0 = \frac{1}{m} \sum_{j=1}^{m} \hat{\gamma}_Z(h_j) - c_s \frac{1}{m} \sum_{j=1}^{m} \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right) \quad (10)$$

From the second equation, we find

$$\frac{\partial L(c_0, c_s, a_s)}{\partial c_s} = 0 \iff \sum_{j=1}^{m} \hat{\gamma}_Z(h_j) \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right) - c_s \sum_{j=1}^{m} \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)^2 = 0$$

$$\iff c_s = \frac{\sum_{j=1}^{m} \hat{\gamma}_Z(h_j) \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)}{\sum_{j=1}^{m} \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)^2} \quad (11)$$

Since we have $c_0$ in (10), then by the substitution of $c_0$ in (11) with that in (10), it leads to

$$c_s = \frac{\sum_{j=1}^{m} \hat{\gamma}_Z(h_j) \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)}{\sum_{j=1}^{m} \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)^2 - \left( \sum_{j=1}^{m} \frac{3}{2} \|h_j\|/a_s - \frac{1}{2}(\|h_j\|/a_s)^3 \right)^2} \quad (12)$$

Simultaneous substitution of $c_s$ to (10) results in the following closed form of $c_0$ as the function of $a_s$ and $\hat{\gamma}_Z(\cdot)$:

$$c_0 = \frac{1}{m} \sum_{j=1}^{m} \hat{\gamma}_Z(h_j)$$
The rate of growth of two weeks old corn plants have been measured in cm/day over a period of two weeks. The data can be regarded as a realization of the spherical variogram model in direction west to east, where the covariance function is given by

$$\gamma(h) = \frac{1}{2} \|h\|/a_s - \frac{1}{2} \left( \frac{3}{2} \left( \frac{3}{2} \left[ \frac{3}{2} \left( \frac{3}{2} \|h\|/a_s - \frac{1}{2} (\|h\|/a_s)^3 \right) + \frac{1}{2} (\|h\|/a_s)^2 \right] \right) \right).$$

(13)

For the last derivative, we get

$$\frac{\partial L(c_0, c_s, a_s)}{\partial a_s} = 0 \iff 2 \sum_{j=1}^{m} \left( \tilde{\gamma}_Z(h_j) - c_0 - c_s \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2} (\|h_j\|/a_s)^3 \right) \right) \times \left( -c_s \frac{3}{2} \|h_j\|/a_s^2 \left( (\|h_j\|/a_s)^2 - 1 \right) \right) = 0$$

$$\iff \sum_{j=1}^{m} \left( \tilde{\gamma}_Z(h_j)c_s^2 \left( \frac{3}{2} \|h_j\|/a_s^2 \left( (\|h_j\|/a_s)^2 - 1 \right) \right) - c_0 c_s^2 \left( \frac{3}{2} \|h_j\|/a_s^2 \left( (\|h_j\|/a_s)^2 - 1 \right) \right) - c_s^2 \left( \frac{3}{2} \|h_j\|/a_s - \frac{1}{2} (\|h_j\|/a_s)^3 \right) \left( \frac{3}{2} \|h_j\|/a_s^2 \left( (\|h_j\|/a_s)^2 - 1 \right) \right) \right) = 0.$$

Unfortunately, the last equation is nonlinear in $a_s$, so that it can not be solved explicitly for $a_s$. In contrast to $a_s$, Equation (12) and (13) show that the constants $c_0$ and $c_s$ are expressible as the function of $a_s$ and $\tilde{\gamma}_Z(h_j)$. By this reason we solve the equations using numerical approximation based on graphical method. In the first stage of finding the approximated solutions, we specify several lags $h_j$ and calculate the corresponding value of $\tilde{\gamma}_Z(h_j)$. Second, we determine the values of $a_s$ in the interval $(0, \infty)$. For every chosen value of $a_s$ we compute simultaneously the associated values of $c_0$, $c_s$ and $L(c_0, c_s, a_s)$ and immediately draw the graph of $(c_s, L(c_0, c_s, a_s))$ for several chosen values of $a_s$. The values of $a_s$ that corresponds to the smallest value of $L(c_0, c_s, a_s)$ will be chosen as the estimate $\hat{a}_s$. The corresponding values of $c_0$ and $c_s$ will give the estimates $\hat{c}_0$ and $\hat{c}_s$.

4. Application

In this section we study the weighted universal kriging of corn plant data presented in [7, 20] in which the rate of growth of two weeks old corn plants have been measured in cm/day over a rectangular region of size 12m x 15.75m. The observations consist of 16 x 21 measurements ranging from the west to the east and from the south to the north with a regular distance between two neighboring points is 0.75m. Thus the data can be regarded as a realization of the sample $\{Z(x_i, y_j) : 1 \leq i \leq 16, 1 \leq j \leq 21\}$, where $x_i - x_{i-1} = 0.75m$ and $y_j - y_{j-1} = 0.75m$, see [7] for more information regarding the data. Here $Z$ represents the rate of growth of corn plant measured in cm/day unit.

In attempt of estimating the parameters of the spherical variogram model using least square method we consider a set of lag vectors $\{h_1, \ldots, h_{20}\}$ in west-east direction, where $h_j = 0.75j\mathbf{e}_1$, with $\mathbf{e}_1 = (1, 0)$, for $j = 1, \ldots, 20$. For all this specified lags we compute the empirical variogram functions $\tilde{\gamma}_Z(h_j)$. Next, by using graphical method as discussed at the end of Section 2, we obtain the estimation to the spherical variogram model in direction west to east, given by

$$\tilde{\gamma}_Z(h) = 1.95613 + 2.12125 \left( \frac{3}{2} \left( \frac{3}{2} \left( \frac{3}{2} \left( \frac{3}{2} \|h\|/12.50 - \frac{1}{2} (\|h\|/12.50)^3 \right) \right) \right), 0 < \|h\| \leq 12.50.$$
Figure 1. The graphs of $a_s$ versus $L(c_0, c_s, a_s)$, for $a_s$ varies in $(0, 25)$. The minimum value of $L(c_0, c_s, a_s)$ is 0.85597 which is attained at $a_s = 12.5$, $c_0 = 1.95613$ and $c_s = 2.12125$.

Figure 2. The graphs of the estimated spherical model $\hat{\gamma}_Z(h)$ (smooth line) and empirical variogram function (dotted line).

nugget parameter $\hat{c}_0 = 1.95613$, the sill parameter $\hat{c}_s = 2.12125$ and the range $\hat{a}_s = 12.50$ are obtained by graphical method as discussed in Section 2, in which the values of $c_0(a_s)$ and
The distance between two neighboring points is 0.75 m.

\( c_s(a_s) \) are computed using Equation (12) and (13) for several chosen values of \( a_s \in (0, 25) \).

Next we plot the graph of \( L(c_0(a_s), c_s(a_s), a_s) \) by drawing a smooth line connecting the pairs \((a_s, L(c_0(a_s), c_s(a_s), a_s))\), as presented in Figure 1. The graph shows that the minimum value of \( L(c_0(a_s), c_s(a_s), a_s) \) is attained when \( a_s = 12.50 \). For this value of \( a_s \) we obtain \( \hat{c}_0 = 1.95613 \) and \( \hat{c}_s = 2.12125 \).

Before we proceed to the weighted universal kriging for the rate of growth of corn plants, a step-wise test regarding the plausibility of the regression model must be conducted. By applying the partial sums method proposed in [19, 20], reasonable model for describing the behavior of the rate of growth over the farmland is a first-order polynomial model with two variables. That is \( Z(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \delta(x, y) \), for \((x, y) \in D\), where \( \beta_0, \beta_1, \beta_2 \) are unknown constants, see the hypothesis test performed in [19, 20]. By this result, the design matrix involved in the kriging formula is constructed based on the known regression functions \( f_0(x, y) = 1 \), \( f_1(x, y) = x \), \( f_2(x, y) = y \), together with the weight function which is given in this case by the exponential weight \( \varphi_0(x, y) = \exp\{-\|(x_0, y_0) - (x, y)\|/4.5\} \), for \((x, y) \in D\) and fixed \((x_0, y_0) \in D\). Given a point \((x_0, y_0) \in D\), the vector of prediction coefficients \( \alpha = (\alpha_1, \ldots, \alpha_n)^T \) in \( \hat{Z}(x_0, y_0) \) can be computed by using the formula given in Theorem 2.1. In this application we write a computer program using R for computing the prediction values of \( \hat{Z}(x_0, y_0) \) for 16 x 20 predetermined points \((x_0, y_0) \in D\). For the first case we develop a prediction map based on the collection of predicted points over a regular lattice of size 16 x 21 with neighboring distance 0.75 m, whereas in the second one we build the prediction map over the same region with 16 x 21 points with the distance between the two neighboring points is 1 m. Both prediction maps are presented in Figure 3 and Figure 4, respectively.

The prediction map in Figure 3 shows almost similar results as the contour plot of the original data as it should be. Because in the first case we predict the rate of growth in the same points
Figure 4. The contour plot of the weighted kriging for the rate of growth under the estimated spherical model \( \hat{\gamma}(h) = 1.95613 + 2.12125 \left\{ \frac{3}{2} (\|h\|/12.50) - \frac{1}{2} (\|h\|/12.50)^3 \right\} \), with exponential weight \( \varphi_0(x, y) = \exp\{-\|(x_0, y_0) - (x, y)\|/4.5\} \). The distance between two neighboring points is 1m.

where the data have been collected, see [20]. As it can be seen in Figure 3, the corn plants with greater rate of growth lay in the northwest region, whereas those with smaller rate of growth are positioned in the southwest region.

5. Concluding remark
In the present paper we have successfully establish an optimal linear prediction method by incorporating a weight function in the formula of the predictor. Base on the estimated predictor, we can build a prediction map for the rate of growth of corn plant observed over a regular lattice on a rectangular region. Under the assumption that a first-order polynomial model is plausible we derive the formula for computing the kriging coefficients of isotropic spatial process are computed under a first-order polynomial function and with the parametric variogram function of spherical type. In this work it is assumed that the variogram function is isotropic. However many spatial process encountered in the application seem to be anisotropic. By this reason we need a formal inference as well as optimal prediction procedure under anisotropic condition, even for non stationary spatial process.

Acknowledgments
The authors are grateful to the Indonesian Ministry of Research and Higher Education for awarding them a research grant. The authors thank the Department of Mathematics of Halu Oleo University for hospitality.

References
[1] Cressie N A C 1991 Statistics for Spatial Data (New York: John Wiley & Sons)
[2] Gneiting T 2002 J Amer Stat Assoc 97(458) 590
[3] Wackernagel H 2003 *Multivariate Geostatistics* (Berlin: Springer-Verlag)
[4] Schabenberger O and Gotway C A 2005 *Statistical Methods for Spatial Data Analysis* (London: Chapam-Hall)
[5] Gneiting T, Kleiber W and Schlather M 2010 *J Amer Stat Assoc* 105(491) 1167
[6] Sherman M 2011 *Spatial Statistics and Spatio-Temporal Data* (New York: John Wiley & Sons)
[7] Somayasa W, Wibawa G A, Ruslan and Sutiai D K 2019 *J. Phys.: Conf. Ser.* 1341 062029
[8] Guan Y, Sherman M and Calvin J A 2004 *J. Amer. Stat. Assoc.* 99 (467) 810–821
[9] Maity A and Sherman M 2012 *J Stat Plan Inference* 142(5) 1081
[10] Brunsdon C, Fotheringham S and Charlton M 1998 *Journal of the Royal Statistical Society: Series D (The Statistician)* 47 (3) 431–443

[11] Leung Y, Mei C L and Zhang W X 2000 *Environment and Planning A: Economy and Space* 32 (1) 9-32
[12] Casella G and Berger R L 1990 *Statistical Inference* (California: Wadsworth, Inc.)
[13] Longford N T 2013 *Statistical Decision Theory* (New York: Springer-Verlag)
[14] Arnold S F 1981 *The Theory of Linear Models and Multivariate Analysis* (New York: John Wiley & Sons)
[15] Seber G A F and Lee A J 2003 *Linear Regression Analysis* (2nd ed.) (New York: John Wiley & Sons)
[16] Bischoff W and Somayasa W 2009 *J. Multivariate Analysis* 100 2167–77
[17] Somayasa W, Ruslan, Cahyono E and Engkoimani L O 2015 *Fareast J. Math. Sci.* 96(8) 933
[18] Somayasa W 2016 *IOP Conf. Proceedings* 1707 080018
[19] Somayasa W and Wibawa G A 2016 *J. Phys.: Conf. Ser.* 725 012002
[20] Somayasa W 2019 *WSEAS Trans. Math.* 18 62
[21] Strang G 2016 *Introduction to Linear Algebra, 5th edition* (New York: Wellesley-Cambridge Press)
[22] Magnus J N and Neudecker H 2007 *Matrix Differential Calculus with Applications in Statistics and Econometrics* (3rd edition) (New York: John Wiley & Sons)