Universal kriging of corn plant data under isotropic power type variogram model

W Somayasa¹, G A Wibawa², Ruslan³ and D K Sutiari¹
¹,²,³Department of Mathematics, Halu Oleo University, Jl. Mokodompit, Kendari 93118, Indonesia
⁴Department of Medical Electrical Engineering, Mandala Waluya School of Health Sciences, Kendari 93232, Indonesia
E-mail: ¹wayan.somayasa@uvo.ac.id(Corresponding author),²gnawibawa@gmail.com, ³ruslan@uvo.ac.id, ⁴sutiaridesak@gmail.com

Abstract. Kriging is a well known notion in geostatistics which aim at the construction of optimal prediction for unobserved quantities of interest. In this paper we derive optimal prediction of a second-order stationary spatial process by minimizing the quadratic risk or mean square error of the predictor subject to some additional conditions. The coefficients of the linear predictor are calculated by a formula which is obtained by applying Lagrange method. The spatial process under study is assumed to be isotropic with a variogram model belongs to quadratic family. Least squares estimation to the parameters of the variogram model is calculated by graphical method. The validity of the mean function is investigated by utilizing asymptotic test based on the partial sums process of weighted least squares residuals. The application of the method to spatial process of rate of growth of corn plants results in the kriging map of the process. The maps are generated under two different variogram models: power and linear (Tent) models.

1. Introduction
Optimal prediction of future observations in spatial analysis was studied in various areas of studies and disciplines, such as in geosciences, agriculture, economics, biology, atmospheric sciences, environmental monitoring and other areas, see Gneiting [1] and Gneiting and et al. [2]. The purpose of prediction is to obtain accurate information regarding the value of a spatial process in any unobserved sites or positions. In the context of geostatistics, the terminology kriging refers to the optimal prediction of spatial variable based on spatial data sampled over some region. Rigorous discussion of kriging and other spatial techniques for univariate spatial data can be found in Cressie [3], Schabenberger and Gotway [4] and Sherman[5] and Maity and Sherman [6], whereas those for multivariate case have been studied in Wackernagel [7].

In the present paper we aim to demonstrate the application of universal kriging technique to obtain optimum prediction of the rate of growth of corn plants observed over a rectangular experimental region, see also the data presented in Somayasa [8]. The inherent existence of spatial dependence structure among the observations suggests us to handle the analysis using geostatistics approach by regarding the sample as a certain values of a random phenomena in the space which is commonly called spatial process (random field) with fixed index set, cf. [3].

To fix the idea let \( Y(x, y) \) be a random variable represents the rate of growth of corn plant,
where the index \((x, y) \in D\) denotes the coordinate of the position on which the variable was measured and \(D = [a_1, b_1] \times [a_2, b_2]\) is the farm land region where the corn plants were grown up. By following the common approach, the set of observations \(\{Y(x_i, y_j) : 1 \leq i \leq 16, 1 \leq j \leq 21\}\) is considered as a realization of a random field \(\{Y(x, y) : (x, y) \in D\}\) admitting the following decomposition:

\[
Y(x, y) = h(x, y) + \Delta(x, y), \ (x, y) \in D
\]

where \(h\) is an unknown deterministic function which has bounded variation on \(D\) and \(\{\Delta(x, y) : (x, y) \in D\}\) is a second-order stationary random field, with

\[
\begin{align*}
E(\Delta(x, y)) &= 0, \ \forall \ (x, y) \in D, \\
\text{Cov}(\Delta(x, y), \Delta(x + a_1, y + a_2)) &= \text{Cov}(\Delta(0, 0), \Delta(a_1, a_2)) = C_\Delta(a_1, a_2),
\end{align*}
\]

for some function \(C_\Delta(\cdot)\) on \(D\) and for any spatial lag \(a = (a_1, a_2) \in D\). 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

\[
2\gamma_\Delta(a) = 2(C_\Delta(0, 0) - C_\Delta(a)).
\]

We refer the reader to [3], p. 53 for these notions. In universal kriging it is assumed that the regression function \(h\) is linear and expressible as a linear combination of \(p + 1\) known linearly independent regression functions \(\{f_0, f_1, \ldots, f_p\} \subset L_2(D)\). That is

\[
Y(x, y) = \sum_{j=1}^{p+1} \beta_{j-1} f_{j-1}(x, y) + \Delta(x, y), \ (x, y) \in D.
\]

This means that the spatial observations \(\{Y(x_i, y_j) : 1 \leq i \leq 16, 1 \leq j \leq 21\}\) follow a linear spatial regression model with some unknown covariogram and variogram functions \(C_Y(\cdot)\) and \(2\gamma_Y(\cdot)\), see [3, 5]. One well-known example is polynomial model order 2 with two variables:

\[
h(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 x^2 + \beta_4 xy + \beta_5 y^2.
\]

In this case we set \(f_0(x, y) = 1, f_1(x, y) = x, f_2(x, y) = y, f_3(x, y) = x^2, f_4(x, y) = xy\) and \(f_5(x, y) = y^2\), for \((x, y) \in D\). Throughout this paper we assume that \(h\) is a polynomial of two variables with \(f_0 \equiv 1\). In contrast to classical linear regression model studied in Arnold [9] and Seber and Lee [10], where the observations are assumed to be stochastically independent and normally distributed, Model (3) consists not only correlated observations with unknown correlation structure, but also the probability distribution model underlying the spatial process is not clearly specified. By this reason, the problem of establishing model validity check whether or not Model 3 holds true can be very difficult. Standard method using F-test documented in [9, 10] or some methods based partial sums process of the residuals proposed in Bischoff and Somayasa [11], Somayasa and et al. [12] and Somayasa [13] can not be applied. Therefore an alternative approach based on the asymptotic result due to Somayasa and Wibawa [14] will be
investigated by beforehand estimating the unknown covariogram function $C_Y(\cdot)$ as well as the variogram function $2\gamma_Y(\cdot)$ using least squares method, see also [3, 5]. The main objective of this work is to build prediction map for the unobserved rate of growth of corn plants based on universal kriging method under isotropic variogram function in which the variogram depends only on the length of the lag vector but not on the direction. Descriptive method using rose plot as well as asymptotic method for testing isotropy have been proposed by [5, 6]. Although isotropy are sometimes not reasonable, however it is the basic of advance theory in statistical spatial analysis ranging from anisotropy to nonstationarity.

The remainder of the present paper is organized as follows. In Section 2 we derive the kriging equation in term of variogram as well as covariogram function by using Lagarange method. The discussion includes the derivation of the least squares estimator of the power type variogram model in which the solution of the minimization problem is obtained by combining differential and graphical method. We also give a review regarding asymptotic test procedure based on the partial sums process of the weighted least squares residuals for checking the validity the mean function. In Section 3 we demonstrate the application of the kriging procedure to the rate of growth of corn plants. We build prediction map under power type variogram model and first-order polynomial mean function. The paper is closed in Section 4 by some conclusions and remark for future research.

2. Universal kriging

Suppose for the moment that Model 3 holds true with the variogram function $2\gamma_\Delta(\cdot)$. Let the set \{\(Y(x_1, y_1), \ldots, Y(x_n, y_n)\)\} be a sample of size \(n\) taken from Model 3. Linear predictor of \(Y(x_0, y_0)\) in the point \((x_0, y_0)\) is denoted by \(\hat{Y}(x_0, y_0)\), defined by \(\hat{Y}(x_0, y_0) := \sum_{k=1}^{n} \kappa_k Y(x_k, y_k)\), such that under the condition

\[
\sum_{k=1}^{n} \kappa_k f_j(x_k, y_k) = f_j(x_0, y_0), \quad j = 0, 1, \ldots, p, \tag{4}
\]

the mean square error (MSE) of \(\hat{Y}(x_0, y_0)\), defined by

\[
MSE(\hat{Y}(x_0, y_0)) := E (\hat{Y}(x_0, y_0) - Y(x_0, y_0))^2 = E \left( \sum_{k=1}^{n} \kappa_k Y(x_k, y_k) - Y(x_0, y_0) \right)^2 \tag{5}
\]

attains minimum value with respect to the unknown constants \(\kappa_1, \ldots, \kappa_n\). For this reason \(\hat{Y}(x_0, y_0)\) is called optimum predictor. In the statistical decision context, \(MSE(\hat{Y}(x_0, y_0))\) is also known as the quadratic risk. Hence \(\hat{Y}(x_0, y_0)\) is an optimal predictor by the reason it minimizes the quadratic risk, see Casella and Berger [15]. Moreover, when (4) is true, it can be shown that \(\hat{Y}(x_0, y_0)\) is unbiased to \(h(x_0, y_0)\), since we have

\[
E(\hat{Y}(x_0, y_0)) = E \left( \sum_{k=1}^{n} \kappa_k h(x_k, y_k) \right) \quad \text{and} \quad E \left( \sum_{k=1}^{n} \kappa_k Y(x_k, y_k) \right) = \sum_{k=1}^{n} \kappa_k h(x_k, y_k)
\]

\[
= \sum_{k=1}^{n} \kappa_k \sum_{j=1}^{p+1} \beta_{j-1} f_{j-1}(x_k, y_k) = \sum_{j=1}^{p+1} \beta_{j-1} \sum_{k=1}^{n} \kappa_k f_{j-1}(x_k, y_k)
\]

\[
= \sum_{j=1}^{p+1} \beta_{j-1} f_{j-1}(x_0, y_0) = h(x_0, y_0). \tag{6}
\]

By recalling (4) and (5), the \(MSE(\hat{Y}(x_0, y_0))\) can be further expressed as follows

\[
MSE \left( \hat{Y}(x_0, y_0) \right) = Var \left( \hat{Y}(x_0, y_0) - Y(x_0, y_0) \right) + \left[ E \left( \hat{Y}(x_0, y_0) - Y(x_0, y_0) \right) \right]^2
\]
\[ f(x, y) = \text{var} \left( \sum_{k=1}^{n} \kappa_k h(x_k, y_k) + \sum_{k=1}^{n} \kappa_k \Delta(x_k, y_k) - h(x_0, y_0) - \Delta(x_0, y_0) \right) \]

\[ = \text{var} \left( \sum_{k=1}^{n} \kappa_k \Delta(x_k, y_k) - \Delta(x_0, y_0) + \sum_{k=1}^{n} \kappa_k h(x_k, y_k) - h(x_0, y_0) \right) \]

\[ = \text{var} \left( \sum_{k=1}^{n} \kappa_k \Delta(x_k, y_k) - \Delta(x_0, y_0) \right) \]

\[ = \text{var} \left( \sum_{k=1}^{n} \kappa_k \Delta(x_k, y_k) - \sum_{k=1}^{n} \kappa_k \Delta(x_0, y_0) \right) \]

\[ = \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell \text{E} \left[ (\Delta(x_k, y_k) - \Delta(x_0, y_0))(\Delta(x_\ell, y_\ell) - \Delta(x_0, y_0)) \right]. \quad (7) \]

The last two equations follow by incorporating Condition (4). Indeed, when \( f_0 \equiv 1 \), then under condition \( \sum_{i=1}^{n} \kappa_i f_0(x_i, y_i) = f_0(x_0, y_0) \), it holds \( \sum_{i=1}^{n} \kappa_i = 1 \). This means that when \( p = 1 \) and \( f_0 \equiv 1 \), universal kriging reduces to ordinary kriging, cf. [3], p.152. Next, by applying the following relation

\[ 2\gamma_\Delta (x_k - x_\ell, y_k - y_\ell) = 2\gamma_\Delta (x_k - x_0, y_k - y_0) + 2\gamma_\Delta (x_\ell - x_0, y_\ell - y_0) - 2\text{E} (\Delta(x_k, y_k) - \Delta(x_0, y_0))(\Delta(x_\ell, y_\ell) - \Delta(x_0, y_0)) \]

and by recalling the fact that \( \sum_{k=1}^{n} \kappa_k = 1 \), Equation 7 reduces to

\[ \text{MSE} (\hat{Y}(x_0, y_0)) = 2 \sum_{k=1}^{n} \kappa_k \gamma_\Delta (x_k - x_0, y_k - y_0) - \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell \gamma_\Delta (x_k - x_\ell, y_k - y_\ell). \]

Furthermore, since, \( \gamma_Y = \gamma_\Delta \), the last equation can also be written by

\[ \text{MSE} (\hat{Y}(x_0, y_0)) = 2 \sum_{k=1}^{n} \kappa_k \gamma_Y (x_k - x_0, y_k - y_0) - \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell \gamma_Y (x_k - x_\ell, y_k - y_\ell), \quad (8) \]

where \( \gamma_Y (x_k - x_0, y_k - y_0) = \text{var}(Y(x_k, y_k) - Y(x_0, y_0)) \).

The following theorem presents the formula of computing the values of the constants \( \kappa_1, \ldots, \kappa_n \) that minimize \( \text{MSE} (\hat{Y}(x_0, y_0)) \) subject to Condition (4). The solution has been established by applying Lagrange method, see also [3], p.153 and [5], p.26. In this paper we give the proof of the theorem in more rigorous form. Concise extension of the result to multivariate spatial process has been considered in [2, 7].

**Theorem 2.1** ([3], p.153 and [5], p.26) Let \( X \) be the design matrix of Model 3 with dimension \( n \times (p + 1) \) and \( \text{rank}(X) = p + 1 \), whose entry in the \( i \)-th row and \( j \)-th column is given by \( f_{j-1}(x_i, y_i) \), for \( j = 1, \ldots, p + 1 \) and \( i = 1, \ldots, n \). Let \( \Gamma = (\gamma_Y(\ell, k))_{i=1}^{n, i=1} \) be the \( n \times n \) dimensional matrix of the semivariogram function of \( \{Y(x, y) : (x, y) \in \mathbb{D}\} \), where \( \gamma_Y(\ell, k) := \gamma_Y(Y(x_\ell, y_\ell) - Y(x_k, y_k)) \), for \( k, \ell = 1, \ldots, n \). Let \( \gamma_Y(0) := (\gamma_Y(1, 0), \gamma_Y(2, 0), \ldots, \gamma_Y(n, 0))^{\top} \), where \( \gamma_Y(0) := \gamma_Y(x_0 - x_k, y_0 - y_k) \), \( k = 1, \ldots, n \). Then the values of \( \kappa := (\kappa_1, \ldots, \kappa_n)^{\top} \) that minimize (5) subject to Condition 4 is given by

\[ \kappa^{\top} = \left\{ \gamma_Y(0) + X(X^{\top} \Gamma^{-1} X)^{-1}(x_0 - X^{\top} \Gamma^{-1} \gamma_Y(0)) \right\} \Gamma^{-1}, \quad (9) \]
where 
\[ x_0 := (f_0(x_0, y_0), f_1(x_0, y_0), \ldots, f_p(x_0, y_0))^\top. \]

Proof: Let \( L := (L_0, \ldots, L_p)^\top \) be the vector of Lagrange multipliers associated with the \( p + 1 \) constrains \( \sum_{k=1}^{n} \kappa_k f_{j-1}(x_k, y_k) = f_{j-1}(x_0, y_0) \), for \( j = 1, \ldots, p + 1 \). Let \( G : \mathbb{R}^{n+p+1} \to \mathbb{R} \), defined by

\[
G(\kappa_1, \ldots, \kappa_n, L_0, \ldots, L_p) = 2 \sum_{k=1}^{n} \kappa_k \gamma_Y(k, 0) - \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell \gamma_Y(k, \ell)
- 2 \sum_{j=1}^{p+1} L_{j-1} \left( \sum_{k=1}^{n} \kappa_k f_{j-1}(x_k, y_k) - f_{j-1}(x_0, y_0) \right).
\]

Then for \( i = 1, \ldots, n \), we get

\[
\frac{\partial G(\kappa_1, \ldots, \kappa_n, L_0, L_1, \ldots, L_p)}{\partial \kappa_i} = 0 \iff 2 \gamma_Y(i, 0) - 2 \sum_{k=1}^{n} \kappa_k \gamma_Y(i, k) - \sum_{j=1}^{p+1} L_{j-1} f_{j-1}(x_i, y_i) = 0
\]

and for \( j = 1, \ldots, p + 1 \), it holds

\[
\frac{\partial G(\kappa_1, \ldots, \kappa_n, L_0, L_1, \ldots, L_p)}{\partial L_{j-1}} = 0 \iff -2 \left( \sum_{k=1}^{n} \kappa_k f_{j-1}(x_k, y_k) - f_{j-1}(x_0, y_0) \right) = 0
\]

Next, let \( M \) be an \((n + p + 1) \times (n + p + 1)\)-dimensional square matrix, defined by

\[
M := \begin{bmatrix}
\gamma_Y(1, 1) & \gamma_Y(1, 2) & \cdots & \gamma_Y(1, n) & f_0(x_1, y_1) & f_1(x_1, y_1) & \cdots & f_p(x_1, y_1) \\
\gamma_Y(2, 1) & \gamma_Y(2, 2) & \cdots & \gamma_Y(2, n) & f_0(x_2, y_2) & f_1(x_2, y_2) & \cdots & f_p(x_2, y_2) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\gamma_Y(n, 1) & \gamma_Y(n, 2) & \cdots & \gamma_Y(n, n) & f_0(x_n, y_n) & f_1(x_n, y_n) & \cdots & f_p(x_n, y_n) \\
f_0(x_1, y_1) & f_0(x_2, y_2) & \cdots & f_0(x_n, y_n) & 0 & 0 & \cdots & 0 \\
f_1(x_1, y_1) & f_1(x_2, y_2) & \cdots & f_1(x_n, y_n) & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
f_p(x_1, y_1) & f_p(x_2, y_2) & \cdots & f_p(x_n, y_n) & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

and let \( \kappa_L \) and \( \gamma_{Y,f} \) be \( n + p + 1 \) dimensional vector given respectively by

\[
\kappa_L := (\kappa_1, \kappa_2, \cdots, \kappa_n, L_0, L_1, \cdots, L_p)^\top
\]

\[
\gamma_{Y,f} := (\gamma_Y(1, 0), \gamma_Y(2, 0), \cdots, \gamma_Y(n, 0), f_0(x_0, y_0), f_1(x_0, y_0), \cdots, f_p(x_0, y_0))^\top.
\]

The system of linear equation consisting of \( n + p + 1 \) equations in both (10) and (11) can be simultaneously written in matrix and vector form as \( M \kappa_L = \gamma_{Y,f} \) having the solution \( \kappa_L = M^{-1} \gamma_{Y,f} \). Furthermore, since \( M, \kappa_L \) and \( \gamma_{Y,f} \) can be partitioned respectively as

\[
M = \begin{bmatrix}
I & X \\
X^\top & 0
\end{bmatrix}, \quad \kappa_L = \begin{bmatrix} \kappa \\ L \end{bmatrix} \quad \text{and} \quad \gamma_{Y,f} = \begin{bmatrix} \gamma_Y(0) \\ x_0 \end{bmatrix},
\]

where

\[
x_0 := (f_0(x_0, y_0), f_1(x_0, y_0), \ldots, f_p(x_0, y_0))^\top.
\]
then by applying the inverse of block matrix, studied in Harville [16] and Magnus and Neudecker [17], we get the following equations

\[ \kappa^\top = \left\{ \gamma_Y(0) + X(X^\top \Gamma^{-1} X)^{-1}(x_0 - X^\top \Gamma^{-1} \gamma_Y(0)) \right\}^\top \Gamma^{-1} \]

\[ L^\top = - \left( x_0 - X^\top \Gamma^{-1} \gamma_Y(0) \right)^\top \left( X^\top \Gamma^{-1} X \right)^{-1}, \]

establishing the proof.

**Remark 2.2** It is noticed that the kriging equation can also be expressed by using covariogram function. By utilizing equation \( 2\gamma_Y(x_k - x_\ell, y_k - y_\ell) = 2(C_Y(0,0) - C_Y(x_k - x_\ell, y_k - y_\ell)) \), cf. Equation 2, we can write (10) as follows

\[
MSE \left( \hat{Y}(x_0, y_0) \right) = 2 \sum_{k=1}^{n} \kappa_k \left( C_Y(0,0) - C_Y(k,0) \right) - \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell \left( C_Y(0,0) - C_Y(k, \ell) \right)
\]

\[ = -2 \sum_{k=1}^{n} \kappa_k C_Y(k,0) + \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell C_Y(k, \ell), \]

where \( C_Y(k, \ell) \) refers to \( C_Y(x_k - x_\ell, y_k - y_\ell) \), for \( k, \ell = 1, \ldots, n \). So the kriging coefficients \( \kappa_1, \kappa_2, \ldots, \kappa_n \) can be obtained by differentiating

\[ F(\kappa_1, \kappa_2, \ldots, \kappa_n, L_0, L_1, \ldots, L_p) = -2 \sum_{k=1}^{n} \kappa_k C_Y(k,0) + \sum_{k=1}^{n} \sum_{\ell=1}^{n} \kappa_k \kappa_\ell C_Y(k, \ell), \]

\[ -2 \sum_{j=1}^{p+1} L_{j-1} \left( \sum_{k=1}^{n} \kappa_k f_{j-1}(x_k, y_k) - f_{j-1}(x_0, y_0) \right) \]

with respect to \( \kappa_i \) and \( L_{j-1} \), for \( i = 1, \ldots, n \) and \( j = 1, \ldots, p + 1 \). By applying the similar manner as that of kriging using variogram, it holds

\[ \kappa^\top = \left\{ C_Y(0) + X(X^\top \Sigma^{-1} X)^{-1}(x_0 - X^\top \Sigma^{-1} C_Y(0)) \right\}^\top \Sigma^{-1} \]

\[ L^\top = \left( x_0 - X^\top \Sigma^{-1} C_Y(0) \right)^\top \left( X^\top \Sigma^{-1} X \right)^{-1}, \]

(12)

where \( \Sigma := (C_Y(k, \ell))_{k=1, \ell=1}^{n,n} \) and \( C_Y(0) := (C_Y(1,0), C_Y(2,0), \ldots, C_Y(n,0))^\top \), see also [3, 5].

The variogram and the covariogram functions involved respectively in the kriging formulas (9) and (12) are usually unknown. Therefore in the practice \( \gamma_Y(\cdot) \) and \( C_Y(\cdot) \) must be estimated before they are used in the computation of the kriging coefficients. As documented in the literatures of geostatistics, the sample variogram and covariogram functions are defined respectively by

\[ 2\hat{\gamma}_Y(a) := \frac{1}{|N(a)|} \sum_{(x_i, y_i), (x_j, y_j), (x_i - x_j, y_i - y_j) = a} \left( Y(x_i, y_i) - Y(x_j, y_j) \right)^2, \quad a = (a_1, a_2) \in D, \]

\[ \hat{C}_Y(a) := \frac{1}{|N(a)|} \sum_{(x_i, y_i), (x_j, y_j), (x_i - x_j, y_i - y_j) = a} \left( Y(x_i, y_i) - \bar{Y} \right) \left( Y(x_j, y_j) - \bar{Y} \right), \quad a = (a_1, a_2) \in D, \]

where \( N(a) := \{(x_i, y_i), (x_j, y_j) : (x_i - x_j, y_i - y_j) = a\} \) and \( |N(a)| \) stands for the number of distinct pairs in \( N(a) \). However both \( 2\hat{\gamma}_Y(\cdot) \) and \( 2\hat{C}_Y(\cdot) \) are not valid variogram and covariogram
functions in the sense of [1, 2, 3] by the fact they do not satisfy the well-known theorem of Bochner, cf. [1] and [3], p.84.

There are several valid isotropic parametric variogram models which are commonly used in the practice: linear (Tent model), spherical, exponential, rational, wave, power and Matérn model. All of these models need to be estimated before they are used in kriging by either least squares, weighted least squares or maximum likelihood method, see [1, 2, 3, 5]. Let \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\} be a set of \(m\) lag vectors in \(\mathbf{D}\). Let \(\{2\gamma_Y(\mathbf{a}_1), 2\gamma_Y(\mathbf{a}_2), \ldots, 2\gamma_Y(\mathbf{a}_m)\}\) and \(\{2\gamma_Y(\mathbf{a}_1; \mathbf{c}), 2\gamma_Y(\mathbf{a}_2; \mathbf{c}), \ldots, 2\gamma_Y(\mathbf{a}_m; \mathbf{c})\}\) be the corresponding sets of estimated and theoretical parametric variogram models, respectively with the parameter vector \(\mathbf{c} := (c_1, \ldots, c_p)^\top \in \mathbb{R}^p\).

Least squares estimate \(\hat{\mathbf{c}} = (\hat{c}_1, \ldots, \hat{c}_p)^\top\) of \(\mathbf{c} := (c_1, \ldots, c_p)^\top\) is the vector of constants that satisfies

\[
L(\hat{c}_1, \ldots, \hat{c}_p) := \min_{\mathbf{c} \in \mathbb{R}^p} \sum_{i=1}^{m} (\hat{\gamma}_Y(\mathbf{a}_i) - \gamma_Y(\mathbf{a}_i; \mathbf{c}))^2.
\]

In this paper we propose to consider power model as the valid variogram function for the rate of growth of corn plants which is defined by

\[
2\gamma_Y(\mathbf{a}; c_0, b_p, \lambda) := \begin{cases} 
0, & \mathbf{a} = \mathbf{0} \\
c_0 + b_p \|\mathbf{a}\|^\lambda, & \mathbf{a} \neq \mathbf{0}
\end{cases}
\]

where \(c_0 \geq 0, b_p \geq 0\) and \(0 \leq \lambda < 2\). When \(\lambda = 1\), power model reduces to linear model, cf. [3, 5]. The least squares estimator of \(c_0\), \(b_p\) and \(\lambda\) are the simultaneous solution of the following equations:

\[
\frac{\partial L(c_0, b_p, \lambda)}{\partial c_0} = 0, \quad \frac{\partial L(c_0, b_p, \lambda)}{\partial b_p} = 0 \quad \text{and} \quad \frac{\partial L(c_0, b_p, \lambda)}{\partial \lambda} = 0,
\]

where

\[
L(c_0, b_p, \lambda) = \sum_{i=1}^{m} \left(\hat{\gamma}_Y(\mathbf{a}_i) - c_0 - b_p \|\mathbf{a}_i\|^\lambda\right)^2.
\]

By some basic algebraic computations, we get

\[
c_0(\lambda) = \frac{1}{m} \left(\sum_{i=1}^{m} \hat{\gamma}_Y(\mathbf{a}_i) - b_p \sum_{i=1}^{m} \|\mathbf{a}_i\|^\lambda\right)
\]

\[
b_p(\lambda) = \left(\frac{1}{m} \sum_{i=1}^{m} \hat{\gamma}_Y(\mathbf{a}_i) \|\mathbf{a}_i\|^\lambda \right) \left(\sum_{i=1}^{m} \|\mathbf{a}_i\|^\lambda \right) - \left(\sum_{i=1}^{m} \hat{\gamma}_Y(\mathbf{a}_i) \|\mathbf{a}_i\|^\lambda \right) \frac{1}{2} \left(\sum_{i=1}^{m} \|\mathbf{a}_i\|^\lambda \right) - \left(\sum_{i=1}^{m} \|\mathbf{a}_i\|^2 \right) \frac{1}{2} \left(\sum_{i=1}^{m} \|\mathbf{a}_i\|^\lambda \right),
\]

\[
\sum_{i=1}^{m} \hat{\gamma}_Y(\mathbf{a}_i) \|\mathbf{a}_i\|^\lambda \ln \|\mathbf{a}_i\| = \sum_{i=1}^{m} \left( c_0 \|\mathbf{a}_i\|^\lambda \ln \|\mathbf{a}_i\| + b_p \|\mathbf{a}_i\|^{2\lambda} \ln \|\mathbf{a}_i\| \right)
\]

It is seen that analytic solution for \(\lambda\) cannot be calculated explicitly. Since \(c_0\) and \(b_p\) depend on \(\lambda\), consequently, both are also analytically uncomputable. Approximate solution using graphical method can be applied by scatting the graph of the pairs \((\lambda, L(c_0(\lambda), b_p(\lambda), \lambda))\), for several chosen values of \(\lambda \in [0, 2]\). The smallest value of \(L(c_0(\lambda), b_p(\lambda), \lambda)\) corresponds to the approximate values of \(\hat{\lambda}, \hat{c}_0\) and \(\hat{b}_p\).

Basic procedure of kriging a spatial data is firstly estimating either the variogram or covariogram function. This depends on whether to use (9) or (12). Second, checking the assumed model for the unknown mean function \(h\). Third, computing the kriging coefficients. Fourth, building contour map of the predictors. Finally, giving interpretation of the kriging result. Among those steps the most crucial is the second one, since beforehand, the proposed model \(h(x, y) = \sum_{j=1}^{p+1} \beta_j f_{j-1}(x, y)\) must be shown to be a valid model for representing the mean of the observations. The problem becomes more difficult by the existence of the covariogram matrix. We note that when we are lucky in that the observations are normally distributed, we
can apply $F$ test by beforehand transforming the vector of the observations so that the new vector consists of probabilistically independent components, see [9, 10]. When the underlying probability distribution is not known, such classical $F$ test is not eligible anymore, unless the transformed observations are assumed probabilistically independent in the sense of Chow and Teicher [18]. For this case asymptotic test procedures proposed in Arnold [9] can be adopted for conducting overall check of the adequateness of the model.

In the present work we investigate the application of partial sums (cumulative sums) method for second-order stationary random field. To see the idea in more detail, let Model 1 be observed over an equally spaced design with $n_1 \times n_2$ design points over $D$, given by $\Omega := \{(x_i, y_j) \in D : 1 \leq i \leq n_1, 1 \leq j \leq n_2\}$, where $D$ is a closed rectangle. Let $Y_{n_1n_2} := (Y(x_1, y_1), \ldots, Y(x_{n_1}, y_1), \ldots, Y(x_{n_1}, y_{n_2}))^\top$ be the $n_1n_2$-dimensional vector of random observations, $H_{n_1n_2} := (h(x_1, y_1), \ldots, h(x_{n_1}, y_1), \ldots, h(x_{n_1}, y_{n_2}))^\top$ be the $n_1n_2$-dimensional vector of means and $\Delta_{n_1n_2} := (\Delta(x_1, y_1), \ldots, \Delta(x_{n_1}, y_1), \ldots, \Delta(x_{n_1}, y_{n_2}))^\top$ be the $n_1n_2$-dimensional vector of random errors with $n_1n_1 \times n_1n_2$ dimensional matrix of covariogram function $\Sigma(\theta)$, where the $(i,j)$-th entry of $\Sigma(\theta)$ is defined by $\sigma_{ij} = C_Y(x_j - x_i, y_j - y_i)$, for $1 \leq i \leq n_1$ and $1 \leq j \leq n_2$. Clearly $\Sigma(\theta)$ is non-singular provided $C_Y(\cdot)$ is a valid covariogram model, cf. [1, 2, 3]. Then the transformed model, defined by

$$
\Sigma(\theta)^{-1/2} Y_{n_1n_2} = \Sigma(\theta)^{-1/2} H_{n_1n_2} + \Sigma(\theta)^{-1/2} \Delta_{n_1n_2}
$$

has the centered error term $\Sigma(\theta)^{-1/2} \Delta_{n_1n_2}$, with

$$
Cov \left( \Sigma(\theta)^{-1/2} \Delta_{n_1n_2} \right) = \Sigma(\theta)^{-1/2} \Sigma(\theta) \Sigma(\theta)^{-1/2} = I_{n_1n_2}.
$$

For brevity, we write the model throughout the paper as

$$
Y_{n_1n_2}^{tr} = H_{n_1n_2}^{tr} + \Delta_{n_1n_2}^{tr}. \tag{14}
$$

So, when the underlying distribution of the observation is normal the new model behaves like that in the classical regression model in which the components of the transformed error term
are probabilistically independent, cf. Johnson and Wichern [19]. Otherwise, it consists only of uncorrelated components. Let \( V_{p_1p_2} \) be a linear subspace of \( \mathcal{R}^{p_1p_2} \) generated by \( \{ f_0, f_1, \ldots, f_p \} \), where \( f_{j-1} := (f_{j-1}(x_1), x_2), \ldots, f_{j-1}(x_{n_1}, y_1), \ldots, f_{j-1}(x_{n_1}, y_{n_2}) \) \( \in \mathcal{R}^{p_1p_2} \), for \( j = 1, \ldots, p + 1 \). When the assumed model is true, then by following [9], the vector of least squares residuals \( \mathbf{R}_{n_1n_2} \) is given by the orthogonal projection of the error vector to the complement of the transformed subspace

\[
\mathbf{V}_{n_1n_2}^{tr} := \Sigma(\theta)^{-1/2} \mathbf{V}_{n_1n_2} := \left[ \Sigma(\theta)^{-1/2} f_0, \Sigma(\theta)^{-1/2} f_1, \ldots, \Sigma(\theta)^{-1/2} f_p \right],
\]

given by

\[
\mathbf{R}_{n_1n_2}^{tr} = pr(\mathbf{V}_{n_1n_2}^{tr}) \mathbf{Y}_{n_1n_2}^{tr} = \left( \mathbf{I}_{n_1n_2} - \mathbf{X}_{tr} (\mathbf{X}_{tr}^{\top} \mathbf{X}_{tr})^{-1} \mathbf{X}_{tr}^{\top} \right) \mathbf{Y}_{n_1n_2}^{tr},
\]

where \( \mathbf{X}_{tr} := \Sigma(\theta)^{-1/2} \mathbf{X} \). Furthermore, when the model is true, we get

\[
\mathbf{R}_{n_1n_2}^{tr} = \left( \mathbf{I}_{n_1n_2} - \mathbf{X}_{tr} (\mathbf{X}_{tr}^{\top} \mathbf{X}_{tr})^{-1} \mathbf{X}_{tr}^{\top} \right) \mathbf{Y}_{n_1n_2}^{tr} = pr(\mathbf{V}_{n_1n_2}^{tr}) \mathbf{Y}_{n_1n_2}^{tr}.
\]

Next, by following [14], we transform the vector of residuals to a random field by using a linear operator, defined by

\[
S(\mathbf{R}_{n_1n_2}^{tr})(t, s) := \sum_{i=1}^{[n_1t]} \sum_{j=1}^{[n_2s]} r_{ij} + (n_1t - [n_1t]) \sum_{j=1}^{[n_2s]} r_{[n_1t] + 1j} + (n_2s - [n_2s]) \sum_{i=1}^{[n_1t]} r_{i, [n_2s] + 1} + (n_1t - [n_1t]) (n_2s - [n_2s]) r_{[n_1t] + 1, [n_2s] + 1}, \quad (t, s) \in \mathbf{I} := [0, 1] \times [0, 1].
\]

Under the assumption that \( \{ Y(x_i, y_j) : (x_i, y_j) \in \mathbf{D} \} \) are normally distributed, it was shown in [14] by the geometric technique of Bischoff and Somayasa [11], see also Somayasa and et al. [20] that,

\[
S(pr(\mathbf{V}_{n_1n_2}^{tr}), \mathbf{Y}_{n_1n_2}^{tr})(\cdot) \Rightarrow W(\cdot) - \sum_{j=1}^{p+1} \int_{\mathbf{I}} f_{j-1}(x, y) dW(x, y) g_{f_{j-1}}(\cdot),
\]

where \( W \) is the Brownian sheet with sample path in the space of continuous functions on \( \mathbf{I} \) and \( g_{f}(x, y) := \int_{[0,x] \times [0,y]} f_{j-1}(t, s) dtds \). The Kolmogorov-Smirnov and Cramér-von Mises type statistics defined respectively by

\[
KS_{n_1n_2} := \max_{1 \leq i \leq n_1; 1 \leq j \leq n_2} \left| \frac{1}{\sqrt{n_1n_2}} S(pr(\mathbf{V}_{n_1n_2}^{tr}), \mathbf{Y}_{n_1n_2}^{tr}) \left( \frac{i}{n_1}, \frac{j}{n_2} \right) \right|
\]

\[
CM_{n_1n_2} := \frac{1}{n_1n_2 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left( \frac{1}{\sqrt{n_1n_2}} S(pr(\mathbf{V}_{n_1n_2}^{tr}), \mathbf{Y}_{n_1n_2}^{tr}) \left( \frac{i}{n_1}, \frac{j}{n_2} \right) \right)^2}
\]

have been used in checking the validity of the proposed model whose limits can be obtained by applying the well-known continuous mapping theorem, cf. Theorem 5.1 in Billingsley [21]. Thus in the model check stage, it must be verified beforehand whether or not the transformed vector of the observations \( \mathbf{Y}_{n_1n_2}^{tr} \) normally distributed is.

3. Prediction for corn plant data

We now illustrate the application of the universal kriging method on corn plant data. The data consists of the rate of growth of 16 \times 21 corn plants measured in cm/day over a rectangular region of size 12m \times 15.75 m ranging from the west to the east and from the south to the north.
The positions of the corn plants was set according to a regular lattice where the size of each sub rectangle is 0.75m x 0.75m. Thus the data can be represented as the realization of the sample \( \{Y(x_i,y_j) : 1 \leq i \leq 16, 1 \leq j \leq 21\} \), where \( x_i - x_{i-1} = 0.75 \) and \( y_j - y_{j-1} = 0.75 \). The drop line scatter plot of the 336 measurements together with their associated positions is presented in Figure 1. The height of each line corresponds to the amount of the rate of decay where it is observed. The figure shows a tendency that the rate of decay get larger as the positions move away from the origin which is positioned in the southwest corner. The points correspond with the largest rate are in the northeast subregion. This description can also be seen in Figure 2 which represents the plot of the contour constants of the data. The positions of points which have the same rate of growth are laid in the same curve, where the amount of the corresponding rate of growth are indicated by the attached figures. Similarly, Figure 2 shows that the points having poor rate mainly belong to the southwest region, whereas those with rich rates belong to the northeast region.

By observing Figure 1 and Figure 2, it is clear that the behavior of the process in the west-east direction is different from that in the north-south direction. By this reason, variogram stimators are calculated in these two directions. Figure 3 exhibits the scatter plot of the variogram estimates in the south-north directions \( 2\hat{\gamma}_Y(a) \), for \( a = \{a(1)e_1, a(2)e_1, \ldots, a(20)e_1\} \), where \( e_1 \) is the vector with \( \|e_1\| = 1 \) and direction north-south (\( \pi/2 \) radian); \( a(j) = 0.75j \), for \( j = 1, 2, \ldots, 20 \). Figure 4 presents the similar plot of \( 2\hat{\gamma}_Y(a) \) in the direction west-east, for \( a = \{a(1)e_2, a(2)e_2, \ldots, a(15)e_2\} \), where \( e_2 \) is the vector with \( \|e_2\| = 1 \) and direction west-east (0 radian); \( a(j) = 0.75j \), for \( j = 1, 2, \ldots, 15 \).

The least squares estimate of the variogram model of power type in the direction south-north is given by \( 2\hat{\gamma}_Y(a) = 4.62079 + 0.04131\|a\|^{1.375} \). The graph of this fitted model is superimposed in Figure 3 represented by the smooth line. The estimate of the intercept
The 3rd International Conference On Science
Journal of Physics: Conference Series 1341 (2019) 062029
IOP Publishing
doi:10.1088/1742-6596/1341/6/062029

Figure 4. Estimated variogram in the west-east direction. The smooth line is the least squares estimate $2\hat{\gamma}_Y(a) = 4.06737 + 0.52245\|a\|^{0.875}$.

$c_0 = 4.62079$ and the slope $\hat{b}_p = 0.04131$ are obtained by graphical method as discussed in Section 2, in which the values of $c_0(\lambda_i)$ and $b_p(\lambda_i)$ are computed using Equation 13 for several chosen values of $\lambda_i \in [0, 2)$. Next we plot the graph of $L(c_0(\lambda), b_p(\lambda), \lambda)$ by drawing smooth line connecting the pairs $(\lambda_i, L(c_0(\lambda_i), b_p(\lambda_i), \lambda_i))$, as presented in Figure 5. The graph shows that a minimum value of $L(c_0(\lambda), b_p(\lambda), \lambda)$ is attained when $\lambda = 1.375$. For this value of $\lambda$ we obtain $\hat{c}_0 = c_0(1.375) = 4.62079$ and $\hat{b}_p = b_p(1.375) = 0.04131$. Similarly, by using graphical method we get the least squares estimate of variogram model of power type in the west-east direction as given by $2\hat{\gamma}_Y(a) = 4.06737 + 0.52245\|a\|^{0.875}$ whose graph is superimposed as smooth line in Figure 4. The graph of $L(c_0(\lambda), b_p(\lambda), \lambda)$ for this case is presented in Figure 4 showing the minimum value for $\lambda$ approximately equal to 0.875. Hence, $\hat{c}_0 = c_0(0.875) = 4.06737$ and $\hat{b}_p = b_p(0.875) = 0.52245$. Thus the variogram estimates of the rate of growth in the south-north and west-east directions are different. By this reason it can be concluded that the spatial process $\{Y(x, y) : (x, y) \in D\}$ is not isotropic. However we assume for convenience that the process is isotropic with the variogram function $2\hat{\gamma}_Y = 4.06737 + 0.52245\|a\|^{0.875}$ for any lag vector $a \in D$.

For testing model validity we apply the asymptotic F test and the partial sums method explained at the end of Section 2 by firstly conducting a goodness of fit test for the normality of the process based on the transformed vector $\hat{\Sigma}^{-1/2}Y_{336}$, where $\hat{\Sigma}$ is the $336 \times 336$ dimensional variogram matrix, whose $(i, j)$-th entry is given by

$$\hat{\sigma}_{ij} = 2\hat{\gamma}(i, j) = 4.06737 + 0.52245\|(x_i - x_j, y_i - y_j)\|^{0.875}, \quad i, j = 1, \ldots, 336.$$ 

For our case, we have $x_i = 0.75i$, for $i = 1, 2, \ldots, 16$ and $y_j = 0.75j$, for $j = 1, 2, \ldots, 21$. By using the command ks.test in R for the Kolmogorov-Smirnov goodness of fit test, we get the result $ks = 0.0439$ and the corresponding $p-value = 0.5$. This leads us to the conclusion that the rate of growth of corn plant is normally distributed, see also [8]. Next, we firstly propose to test that a constant model is adequate, that is we test $H_0 : Y(x, y) = \beta_0 + \Delta(x, y)$, for
some unknown parameter $\beta_0$. Both F test defined in [9, 10] and the asymptotic partial sums method using $KS_{16 \times 21}$ and $CM_{16 \times 21}$ defined in the end of Section 2 result in the rejection of $H_0$, meaning that constant model is not adequate. For this reason we further propose to test $H_0$: $Y(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \Delta(x, y)$, that is first order model holds true. By the similar way, we get for the alternative $H_1$: second order model holds true,

$$F = \frac{\|pr(W^\tau)Y^\tau_{336}\|^2}{\|pr(W^\tau)Y^\tau_{336}\|^2} = 1.08012$$

and the .95-th quantile of the Snedecor’s F distribution with 3 and 330 degrees of freedoms $f_{3,330}(0.95) = 2.631975$. Hence, $H_0$ is not rejected for $\alpha = 5\%$ by the fact $F < f_{3,330}(0.95)$, see also [9, 10] for the rejection criterion of the F test. Under the partial sums method, we get $KS_{16 \times 21} = 1.5012$ and $CM_{16 \times 21} = 0.2502$ with the corresponding approximated $p$-values given by 0.8921 and 0.9999, respectively. By these large value of the $p$-values, both statistics lead to the acceptance of $H_0$. Thus we are convinced that first-order polynomial model is plausible for the rate of growth of corn plant.

The next step is to build prediction map which gives the prediction values of the rate of growth of corn plant in unobserved or unobservable positions under first-order polynomial mean function. By using the variogram estimate $2\hat{\gamma}_{Y}(a) = 4.06737 + 0.52245\|a\|^{0.875}$ and the kriging equation (11), we get the prediction map as presented in Figure 7. The map which is build using R exhibits the contour constant of $\{\hat{Y}(x_0, y_0) : (x_0, y_0) \in D\}$. The prediction points are given by a regular lattice of size 16 x 20 with the distance between neighboring points are 1 m. Figure 8 represents the prediction map for the same prediction points and mean function, but the variogram estimate is a linear function $2\hat{\gamma}_{Y}(a) = 4.30006 + 0.37116\|a\|$. It can be seen in both figures that the prediction results are slightly the same with samples. The rate of growth in the north-east region is in average larger than those in other sub region. The contour map of the prediction results reflect the fertility level of the land farm.

4. Concluding remarks
In this work we successfully build a prediction map for the rate of growth of corn plant observed over a regular lattice on a rectangular region. The kriging coefficients are computed under a
first-order polynomial function and the variogram of power type. For checking the plausibility of the mean function, asymptotic test method based on the partial sums process of the weighted least squares residuals as well as the well known F test are proposed. In this work it is assumed that the variogram function is isotropic. However least squares estimation of variogram function in two different directions shows that the process is anisotropic. We need a formal inference procedure for detecting anisotropy. In a forthcoming paper we consider kriging procedure for multivariate spatial data under multivariate Matér covariogram model.

**Acknowledgments**

The authors would like to thank the Indonesian Ministry of Research and Higher Education for providing financial support. The authors is grateful to all staffs of the Department of Mathematics of Halu Oleo University for hospitality.

**References**

[1] Gneiting T 2002 *J Amer Stat Assoc* **97**(458) 590  
[2] Gneiting T, Kleiber W and Schlather M 2010 *J Amer Stat Assoc* **105**(491) 1167  
[3] Cressie N A C 1991 *Statistics for Spatial Data* (New York: John Wiley & Sons)  
[4] Schabenberger O and Gotway C A 2005 *Statistical Methods for Spatial Data Analysis* (London: Chapam-Hall)  
[5] Sherman M 2011 *Spatial Statistics and Spatio-Temporal Data* (New York: John Wiley & Sons)  
[6] Maity A and Sherman M 2012 *J Stat Plan Inference* **142**(5) 1081  
[7] Wackernagel H 2003 *Multivariate Geostatistics* (Berlin: Springer-Verlag)  
[8] Somayasa W 2019 *WSEAS Trans. Math.* **18** 62  
[9] Arnold S F 1981 *The Theory of Linear Models and Multivariate Analysis* (New York: John Wiley & Sons)  
[10] Seber G A F and Lee A J 2003 *Linear Regression Analysis* (2nd ed.) (New York: John Wiley & Sons)  
[11] Bischoff W and Somayasa W 2009 *J Multivariate Analysis* **100** 2167–77  
[12] Somayasa W, Ruslan, Cahyono E and Engkoimani L O 2015 *Far East J. Math. Sci.* **96**(8) 933  
[13] Somayasa W 2016 *IOP Conf. Proceedings* **1707** 080018  
[14] Somayasa W and Wibawa G A 2016 *J. Phys.: Conf. Ser.* **725** 012002  
[15] Casella G and Berger R L 1990 *Statistical Inference* (California: Wadsworth, Inc.)  
[16] Harville D 1997 *Matrix Algebra from a Statistician’s Perspective* (Berlin: Springer Verlag)  
[17] Magnus J N and Neudecker H 2007, *Matrix Differential Calculus with Applications in Statistics and Econometrics* (3rd edition) (New York: John Wiley & Sons)
[18] Chow Y S and Teicher H 2003  Probbility Theory: Independence, Interchangeablity, Martingales (3rd edition) (Berlin: Springer)

[19] Johnson R A and Wichern D W 2007  Applied Multivariate Statistical Analysis (New York: Prentice-Hall)

[20] Somayasa W, Wibawa G A and Pasolon Y B 2015  Int. J. Math. Model and Method Appl. Sci. 9 700

[21] Billingsley P 1999  Convergence of Probability Measures (2nd ed.) (New York: John Wiley & Sons)