Parameter Estimation of Spatial Error Model – Multivariate Adaptive Generalized Poisson Regression Spline

Septia Devi Prihastuti Yasmirullah, Bambang Widjanarko Otok*, Jerry Dwi Trijoyo Purnomo, and Dedy Dwi Prastyo

***Abstract*— The non-parametric regression method becomes an alternative that prioritizes flexibility. Therefore, it is possible to obtain a regression curve model when its shape is not yet known. Multivariate adaptive regression spline (MARS) is one of the non-parametric approaches. In 1991, MARS was introduced by Friedman. The MARS approach, which uses nonparametric regression, can consider additive and interactive effects between predictor variables. MARS modeling has typically been used to model continuous or categorical data. However, researchers in the health sector not only encounter data with continuous or categorical responses but also count data. The original MARS method did not support count data with varying variances and means. Therefore, this study aims to develop the Spatial Error Model—Multivariate Adaptive Generalized Poisson Regression Spline (SEM-MAGPRS), which combines the MARS method with the generalized Poisson regression method with spatial effects.**

Keywords: count data; generalized Poisson regression; MARS; SEM; spatial regression.

I. Introduction

One of the popular methods in mathematical modeling is regression analysis, which is applied in various fields. Regression analysis encompasses three distinct approaches, namely parametric, semi-parametric, and non-parametric. The utilization of the parametric regression is employed in instances where the regression curve is already established. These conditions make parametric regression inflexible in data modeling, especially when data with nonlinear patterns and high dimensionality exist. Non-parametric regression is considered a viable alternative that emphasizes flexibility, i.e., it is possible to obtain a regression curve model when its shape is not yet known. The development of non-parametric regression has been examined by several researchers, i.e., nonparametric identification [1], development of local linear kernel estimators for nonparametric regression [2], multi-response nonparametric regression [3], nonparametric regression for employing best input [4], nonparametric multiple imputation [5], and a nonparametric approach for forecasting [6].

MARS is one of the non-parametric regression techniques. Friedman first introduced MARS in 1991. This approach, which uses nonparametric regression, can consider additive and interactive effects between predictor variables [7]. The unknown pattern of correlations between the response and predictors can be used by MARS to obtain good predictions for the regression curve's form [8]. The functional relation form of the predictor variables and response is not also assumed by MARS. It has a form that is adaptable and practical. MARS has a knot that is capable of processing data that exhibits altered behavior patterns at specific sub-intervals.

The MARS approach combines the truncated spline and recursive partitioning regression (RPR). When using several predictors, the truncated spline approach has issues estimating the position and the number of knots. There will be many combinations for the predictor count, knot location, and knot count. Meanwhile, the process of determining knots in MARS involves an adaptive approach, as opposed to seeking them individually through combination. In this instance, the limitation of a truncated spline can be addressed through this approach [7], [8].

The adaptive process in MARS has been executed through a stepwise algorithm that incorporates both forward and backward procedures [9]. The forward stepwise method is a technique used to construct a model with the minimum number of basis functions, where truncated spline basis functions (including knots and interactions) are incrementally added to the model. Subsequently, the forward stepwise basis function is selected during the backward stepwise process to derive a parsimonious model that relies on the minimum value of the generalized cross-validation (GCV) as it exerts the most notable impact on the estimated response [10].

Generally, MARS modeling has been applied to modeling continuous or categorical data. In addition, the researchers in the health sector encounter not only continuous or categorical responses but also count data. Therefore, the development of the count data analysis needs particular

Manuscript received December 23, 2022; revised June 20, 2023.

S. D. P. Yasmirullah is a Doctoral student of Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, and an Assistant Lecturer of Data Science Technology Study Program, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya, Indonesia. (e-mail: septiadevipy@gmail.com; septia.devi@ftmm.unair.ac.id).

B. W. Otok is a Professor of Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. (corresponding author: +62 812-3110-6646; e-mail: dr.otok.bw@gmail.com; bambang_wo@statistika.its.ac.id).

J. D. T. Purnomo is a Lecturer of Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. (e-mail: jerry@statistika.its.ac.id).

D. D. Prastyo is a Lecturer of Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. (e-mail: dedy-dp@statistika.its.ac.id).

concern. However, the MARS method has not supported the count data type with different variances and means. Therefore, this study aims to combine the MARS method with the generalized Poisson regression method to develop the Multivariate Adaptive Generalized Poisson Regression Spline (MAGPRS).

In many cases, data have spatial effects. When observational data has a relationship between regions, called spatial data, the method of data analysis usually has limitations in making assumptions, namely those related to correlated error issues and/or heterogeneity issues. Since observations at one location have a strong relationship or dependency with observations at another location nearby, they are called spatial effects [11]. Spatial effects can be divided into spatial autocorrelation and spatial heterogeneity. The existence of dependence (spatial error correlation) in the cross-sectional data causes spatial autocorrelation, while spatial heterogeneity is due to the random effect of the region, namely the differences in features from one region to another. Therefore, we need a method that can accommodate dependency factors and spatial heterogeneity. Spatial regression with the area approach can be an alternative to resolving the effects of dependence and spatial heterogeneity [11]. One method that can be used in spatial area regression is the Spatial Error Model (SEM).

The kind of discrete or count data that exhibits spatial correlation is what drives us to build a new Spatial Error Model – Multivariate Adaptive Generalized Poisson Regression Spline (SEM-MAGPRS). The objective of this suggested model is to use a nonparametric regression strategy to account for types of count data that have spatial effects. The estimator can be used with other count distributions even though it was created under the generalized Poisson distribution assumption. This study proposes an application of the Ordinary Least Square (OLS) method to estimate the SEM-MAGPRS model within an information-theoretic framework.

II. LITERATURE REVIEW

A. MARS

Friedman came up with a method for nonparametric multivariate regression called MARS. It has a flexible functional structure, and it is assumed that the functional relationship between response and predictor variables is unknown [7]. With this method, the problem with Recursive Partition Regression (RPR), which is that the model it creates is not continuous on knots, is fixed. Knots on MARS are not identified individually from these combinations but rather through an adaptive process. A backward and forward stepwise algorithm is used in the adaptive process of MARS. Therefore, MARS is a technique used to solve classification and regression problems to predict response variables based on several predictors [8], [12]–[14].

The general MARS model can be written as the following equation:

$$y_i = f\left(\mathbf{x}_i\right) + \varepsilon_i, i = 1, \ldots, n \tag{1}$$

where,

$$\begin{aligned} f(\mathbf{x}_i) &= a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m} \left[s_{km} \left(x_{v(k,m)\mathrm{i}} - t_{km} \right) \right]_+ \\ &= a_0 + \sum_{m=1}^{M} a_m B_{mi}\left(\mathbf{x}_i\right) \\ &= \mathbf{Ba} \end{aligned}$$

Therefore,

$$\mathbf{y} = \mathbf{Ba} + \boldsymbol{\varepsilon} \tag{2}$$

where,

$\mathbf{y}_{n\times 1} = \left(y_1, \ldots, y_n\right)^T$,

$\mathbf{a}_{(M+1)\times 1} = \left(a_0, \ldots, a_M\right)^T$,

$\boldsymbol{\varepsilon}_{n\times 1} = \left(\varepsilon_1, \ldots, \varepsilon_n\right)^T$,

$$\mathbf{B}_{n\times(M+1)} = \begin{pmatrix} 1 & \prod_{k=1}^{K_1}\left[s_{k1}\left(x_{v(k,1)1} - t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_M}\left[s_{kM}\left(x_{v(k,M)1} - t_{kM}\right)\right] \\ 1 & \prod_{k=1}^{K_1}\left[s_{k1}\left(x_{v(k,1)2} - t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_M}\left[s_{kM}\left(x_{v(k,M)2} - t_{kM}\right)\right] \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1}\left[s_{k1}\left(x_{v(k,1)n} - t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_M}\left[s_{kM}\left(x_{v(k,M)n} - t_{kM}\right)\right] \end{pmatrix}$$

Then,

$\mathbf{a}$	:	parameter of basis function
M	:	number of maximum basis function
K_m	:	maximum interaction of m-th basis function
s_{km}	:	sign of basis function in the k-th interaction and m-th basis function, where s_{km} is +1 or -1
$x_{v(k,m)i}$	:	v-th predictor variable, where v is an index of predictor variables related to k-th interaction and m-th basis function in MARS function
t_{km}	:	value of knot in k-th interaction and m-th basis function

The Ordinary Least Squares (OLS) method for estimating the parameters of the MARS model will minimize the following functions:

$$\begin{aligned} \boldsymbol{\varepsilon}^{\mathbf{T}}\boldsymbol{\varepsilon} &= \left(\mathbf{y} - \mathbf{Ba}\right)^T\left(\mathbf{y} - \mathbf{Ba}\right) \\ &= \mathbf{y}^T\mathbf{y} - \mathbf{y}^T\left(\mathbf{Ba}\right) - \left(\mathbf{a}^T\mathbf{B}^T\right)\mathbf{y} + \left(\mathbf{a}^T\mathbf{B}^T\right)\left(\mathbf{Ba}\right) \\ &= \mathbf{y}^T\mathbf{y} - \left(\mathbf{a}^T\mathbf{B}^T\right)\mathbf{y} - \left(\mathbf{a}^T\mathbf{B}^T\right)\mathbf{y} + \left(\mathbf{a}^T\mathbf{B}^T\right)\left(\mathbf{Ba}\right) \\ &= \mathbf{y}^T\mathbf{y} - 2\left(\mathbf{a}^T\mathbf{B}^T\right)\mathbf{y} + \left(\mathbf{a}^T\mathbf{B}^T\right)\left(\mathbf{Ba}\right) \end{aligned} \tag{3}$$

The next step is to get the first derivative of the function $\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}$ with respect to $\mathbf{a}$, then the result is equal to zero.

$$\frac{\partial\left(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}\right)}{\partial\left(\mathbf{a}\right)} = \frac{\partial\left[\mathbf{y}^T\mathbf{y} - 2\left(\mathbf{a}^T\mathbf{B}^T\right)\mathbf{y} + \left(\mathbf{a}^T\mathbf{B}^T\right)\left(\mathbf{Ba}\right)\right]}{\partial\left(\mathbf{a}\right)}$$

$$0 = -2\mathbf{B}^T\mathbf{y} + 2\mathbf{B}^T\left(\mathbf{Ba}\right)$$

$$2\mathbf{B}^T\mathbf{Ba} = 2\mathbf{B}^T\mathbf{y}$$

$$\hat{\mathbf{a}} = \left(\mathbf{B}^T\mathbf{B}\right)^{-1}\left(\mathbf{B}^T\mathbf{y}\right)$$

Then, the MARS model estimator is:

$$\hat{\mathbf{a}}_{OLS} = \left(\mathbf{B}^T\mathbf{B}\right)^{-1}\left(\mathbf{B}^T\mathbf{y}\right) \tag{4}$$

B. Generalized Poisson Distribution

Generalized Poisson distribution has two parameters, i.e., μ and ϕ as dispersion parameter. If $\phi = 0$, then it calls equidispersion cases. If $\phi > 0$, then it calls overdispersion cases. If $\phi < 0$, then it calls underdispersion cases [15]. Suppose $Y_i \sim GP(\mu_i, \phi)$, where $i = 1, 2, ..., n$, is a random sample, then the probability density function of the generalized Poisson distribution is [16]:

$$f(y_i \mid \mu_i, \phi) = \left[\frac{\mu_i}{1+\phi\mu_i}\right]^{y_i} \frac{(1+\phi y_i)^{y_i-1}}{y_i!} \exp\left(-\frac{\mu_i(1+\phi y_i)}{1+\phi\mu_i}\right) \tag{5}$$

The mean and variance of *Generalized Poisson* distribution are $E(Y_i) = \mu_i$ dan $V(Y_i) = \mu_i(1+\phi\mu_i)^2$.

C. Generalized Poisson Regression

Generalized Poisson Regression (GPR) is a Generalized Linear Model (GLM) and therefore requires a link function component. The link function of the Generalized Poisson regression has been obtained in the following ways [17]:

a. Do the logarithms of the two sides of the function

$$\log\left(f(y_i \mid \mu_i, \phi)\right) = \log\left(\begin{array}{c}\left[\frac{\mu_i}{1+\phi\mu_i}\right]^{y_i} \frac{(1+\phi y_i)^{y_i-1}}{y_i!} \\ \exp\left(-\frac{\mu_i(1+\phi y_i)}{1+\phi\mu_i}\right)\end{array}\right)$$

b. Do the exponential of the two sides of the equation obtained in the first step

$$\exp\left\{\log\left(f(y_i \mid \mu_i, \phi)\right)\right\} = \frac{1}{y_i!}\frac{(1+\phi y_i)^{y_i-1}}{(1+\phi\mu)^{y_i}} \exp\left\{y_i \log\mu_i - \frac{\mu_i(1+\phi y_i)}{1+\phi\mu_i}\right\}$$

c. Do the mathematical manipulations

$$f(y_i \mid \mu_i, \phi) = \frac{1}{y_i!}\frac{(1+\phi y_i)^{y_i-1}}{(1+\phi\mu)^{y_i}} \exp\left\{y_i \log\mu_i - \frac{\mu_i(1+\phi y_i)}{1+\phi\mu_i}\right\}$$

Therefore, the link function of the Generalized Poisson Regression is $\log(\mu)$. The general model of the GPR is:

$$\begin{aligned} \log(\mu_i) &= \beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi} \\ \mu_i &= \exp\left(\beta_0 + \beta_1 x_{1i} + ... + \beta_p x_{pi}\right); i = 1, 2, ..., n \\ \log(\boldsymbol{\mu}) &= \mathbf{X}^T\boldsymbol{\beta} \\ \boldsymbol{\mu} &= \exp\left(\mathbf{X}^T\boldsymbol{\beta}\right) \end{aligned} \tag{6}$$

Where,

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{p1} \\ 1 & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

D. MAGPRS

The Multivariate Adaptive Generalized Poisson Regression Spline (MAGPRS) is a mix of the MARS and the Generalized Poisson Regression methods. This equation can be used to express the general MAGPRS model [18], [19]:

$$\begin{aligned} Y_i &\sim GP(\mu, \phi) \\ \ln\mu_i &= f(\mathbf{x}_i) = a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m}\left[s_{km}\left(x_{v(k,m)\mathrm{i}} - t_{km}\right)\right]_+ \\ \ln\mu_i &= a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m}\left[s_{km}\left(x_{v(k,m)\mathrm{i}} - t_{km}\right)\right]_+ \\ \mu_i &= \exp\left(a_0 + \sum_{m=1}^{M} a_m \prod_{k=1}^{K_m}\left[s_{km}\left(x_{v(k,m)\mathrm{i}} - t_{km}\right)\right]_+\right) \\ &= \exp\left(a_0 + \sum_{m=1}^{M} a_m B_{mi}(\mathbf{x}_i)\right) \\ &= \exp(\mathbf{Ba}) \end{aligned} \tag{7}$$

E. Spatial Error Model (SEM)

Spatial are things that have to do with a place or region, and spatial data are observations that have to do with a place or region. Locations close to each other often affect links between areas, so spatial analysis is needed to figure out how locations and responses affect each other.

The spatial regression model is the development of a simple regression model. In 1998, a general spatial model was developed using cross-sectional spatial data [11]. One of the spatial regression models is the spatial error model.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \delta\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \tag{8}$$

where,

$$\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^T$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdots & w_{1n} \\ w_{21} & w_{22} & w_{23} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & w_{n3} & \cdots & w_{nn} \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^T$$

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix}^T$$

III. METHODS

This study develops MAGPRS with a spatial model. This model is the development of MARS and Generalized Poisson Regression considering spatial influences in the process modeling, which is SEM-MAGPRS. Estimation of the SEM-MAGPRS parameters can be performed using the Ordinary Least Squares (OLS) method by following these steps:

1. Define SEM model for count data.

$$\boldsymbol{\mu} = \exp\left(\mathbf{X}\boldsymbol{\beta} + \delta\mathbf{W}\mathbf{u}\right) \tag{9}$$

2. Define the SEM-MAGPRS model.

$$\begin{aligned} \ln \mu_i^* &= f(\mathbf{x}_i) \\ &= a_0 + \sum_{m=1}^{M} a_m B_{mi}(\mathbf{x}_i) + \delta \mathbf{W} u_i \\ \mu_i^* &= \exp\left(a_0 + \sum_{m=1}^{M} a_m B_{mi}(\mathbf{x}_i) + \delta \mathbf{W} u_i \right) \\ \boldsymbol{\mu}^* &= \exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right) \end{aligned} \tag{10}$$

3. Construct the least squares function.

$$\begin{aligned} S(a_0, a_1, \ldots, a_m) &= \sum_{i=1}^{n} \varepsilon_i^2 \\ &= \sum_{i=1}^{n} \left(y_i - \mu_i^* \right)^2 \\ &= \left(\mathbf{y} - \boldsymbol{\mu}^* \right)^2 \end{aligned} \tag{11}$$

4. Find the first derivation of the least squares function.

IV. RESULTS AND DISCUSSION

A. Least Square Function

According to equation (10) and (11), $\boldsymbol{\varepsilon}$ for the SEM-MAGPRS model is:

$$\boldsymbol{\varepsilon} = \mathbf{y} - \exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right), \tag{12}$$

then the least squares function of the SEM-MAGPRS is:

$$\begin{aligned} & S(a_0, a_1, \ldots, a_m) \\ &= \left(\mathbf{y} - \exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right) \right)^2 \\ &= \left(\mathbf{y} - \exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right) \right)' \left(\mathbf{y} - \exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right) \right) \\ &= \begin{bmatrix} \mathbf{y}'\mathbf{y} - \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} - \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ + \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{y}'\mathbf{y} - 2\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ + \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \end{bmatrix} \end{aligned} \tag{13}$$

Finding the estimated model parameters is a little bit simpler once the least squares function has been obtained.

B. Estimation of the basis function coefficient

Lemma 1. When the response variable is considered to have a Generalized Poisson distribution and the MARS model can accommodate the count data type and integrate spatial dependence, the model is SEM-MAGPRS. Additionally, the following is the estimated parameter of the basis function for SEM-MAGPRS:

$$\begin{aligned} \hat{\mathbf{a}} = & \left[\left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right) \right]^{-1} \\ & \left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \end{aligned} \tag{14}$$

If Lemma 1 is satisfied, the estimated parameter of the basis function for SEM-MAGPRS can be found by the least squares function in equation (13).

$$\frac{\partial\left(S(a_0, a_1, \ldots, a_m)\right)}{\partial(\mathbf{a})} = \frac{\partial \begin{bmatrix} \mathbf{y}'\mathbf{y} - 2\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ + \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \end{bmatrix}}{\partial(\mathbf{a})}$$

$$\begin{aligned} 0 = & -2\mathbf{B}\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ & +2\mathbf{B}\mathbf{a}\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \\ \hat{\mathbf{a}} = & \left[\left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right) \right]^{-1} \\ & \left(\exp\left(\mathbf{B}\hat{\mathbf{a}} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \end{aligned}$$

In Appendix 2, there is a more detailed explanation of how to estimate the basis function parameters for SEM-MAGPRS by using OLS.

C. Estimation of the spatial lag coefficient on the error

Lemma 2. When the response variable is considered to have a Generalized Poisson distribution, and the MARS model can accommodate the count data type and integrate spatial dependence, the model is SEM-MAGPRS. Additionally, the following is the estimated parameter of the spatial lag on the error for SEM-MAGPRS:

$$\begin{aligned} \hat{\delta} = & \left[\left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right) \right]^{-1} \\ & \left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \end{aligned} \tag{15}$$

If Lemma 2 is satisfied, the estimated parameter of the spatial lag on the error for SEM-MAGPRS can be find through least square function in equation (13).

$$\frac{\partial\left(S(a_0, a_1, \ldots, a_m)\right)}{\partial(\delta)} = \frac{\partial \begin{pmatrix} \mathbf{y}'\mathbf{y} - 2\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ + \left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \end{pmatrix}}{\partial(\delta)}$$

$$\begin{aligned} 0 = & -2\mathbf{W}\mathbf{u}\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \\ & +2\delta\mathbf{W}\mathbf{u}\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \delta\mathbf{W}\mathbf{u}\right)\right) \\ \hat{\delta} = & \left[\left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right)'\left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right) \right]^{-1} \\ & \left(\exp\left(\mathbf{B}\mathbf{a} + \hat{\delta}\mathbf{W}\mathbf{u}\right)\right)'\mathbf{y} \end{aligned}$$

In Appendix 3, there is a more detailed explanation of how to estimate the spatial lag parameters for SEM-MAGPRS by using OLS.

D.Simulation Studies

In these simulation studies, we generated the data using this scenario:

Table I Simulation Data

Variables	Scenario
X_1	Uniform (0.1, 5)
X_2	Uniform (0.1, 5)
X_3	Uniform (0.1, 5)
X_4	Uniform (0.1, 5)
X_5	Uniform (0.1, 5)
Y	Poisson $(\mu = 4)$

The data generated is given in Appendix 1. Moreover, the data visualization has been shown in this figure.

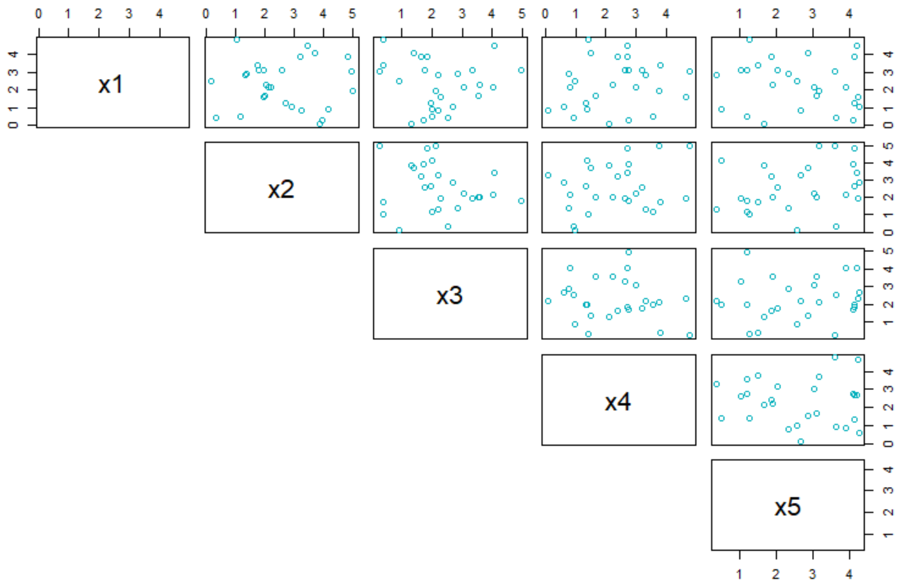

Figure 1 Data Visualization

Figure 1 demonstrates that the predictor variables are independent and lack a discernible pattern. Then, Figure 2 shows the relationship between predictor variables and response variable.

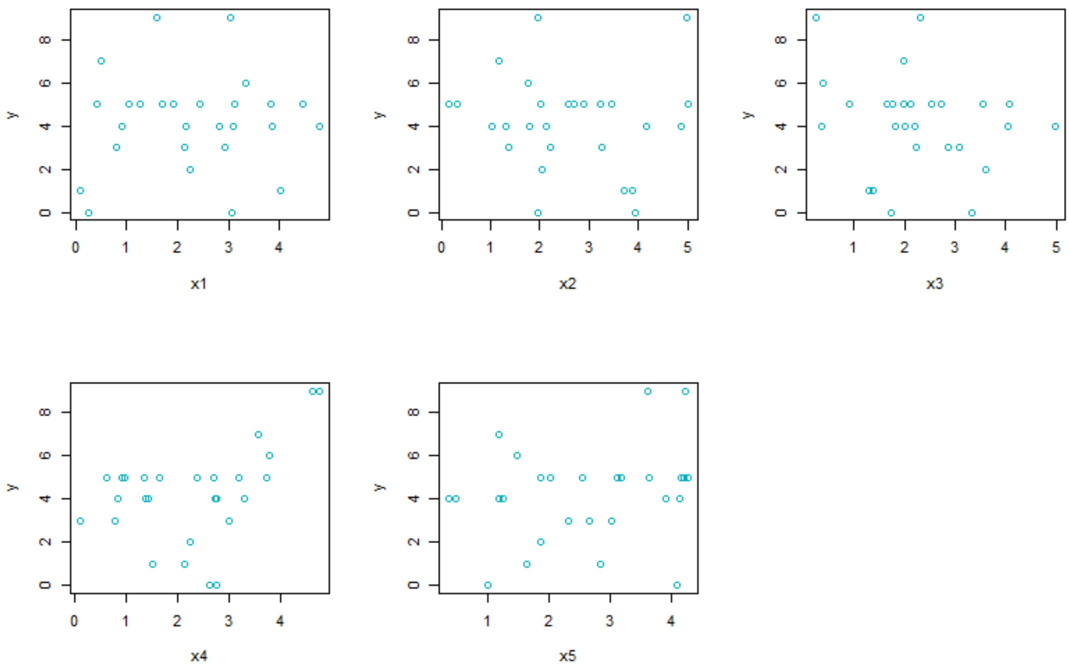

Figure 2 Response and Predictor Variables

Visually, the relationship between response and predictor variables tends to appear unknown. Therefore, a non-parametric approach is appropriate instead of a parametric approach. But in this study, we will compare parametric and non-parametric approaches.

We examined the simulation data using some methods: SEM-MAGPRS, MAGPRS, and a generalized linear model. The generalized linear model is a parametric approach for non-normal data, such as generalized Poisson data. The results of the generalized linear model have been shown in Table II.

Table II Modeling Using a Generalized Linear Model

Variables	Estimate
(Intercept)	1.04754
x1	0.01796
x2	-0.14253
x3	-0.09097
x4	0.19598
x5	0.1618

Next, when modeling using MAGPRS and SEM-MAGPRS, it is classified as a non-parametric approach. The results of the MAGPRS methods will be presented in this table.

Table III Modeling Using MAGPRS

Variables	Estimate
(Intercept)	1.234257
h(x4-2.75099)	0.505807

Furthermore, the following table presents the results of modeling using SEM-MAGPRS methods.

Table IV Modeling Using SEM-MAGPRS

Variables	Estimate
(Intercept)	3.37662
h(x4-2.75099)	2.79305

A comparison of these methods using AIC criteria is given in this table:

Table V Comparison of methods

Methods	**AIC**
Generalized Linear Model	121.37
MAGPRS	111.40
SEM-MAGPRS	**106.18**

According to the AIC criterion, the best model is the one with the lowest AIC value. The model with the lowest AIC value is the SEM-MAGPRS model, with a value of 106.18. So, the non-parametric model is better than the parametric model. This result is consistent with previous research, which shows that the non-parametric model is better than the parametric model [18].

V. Conclusion

The Spatial Error Model - Multivariate Adaptive Generalized Poisson Regression Spline (SEM-MAGPRS) is a spatial and non-parametric regression model that we proposed in this study. The purpose of this proposed model is to accommodate the types of count data that have spatial effects.

We have demonstrated how to estimate the parameter of SEM-MAGPRS using ordinary least squares (OLS) approaches. It is easy to see how the same process used to create the SEM-MAGPRS model may also be used to construct other MARS models based on various distributions.

Based on a comparison of the simulation studies, it can be shown that the SEM-MAGPRS model is the best model with the lowest AIC value.

VI. Abbreviations

The following abbreviations are used in this article:

MARS	Multivariate Adaptive Regression Spline
MAGPRS	Multivariate Adaptive Generalized Poisson Regression Spline
SEM	Spatial Error Model
SEM-MAGPRS	Spatial Error Model – Multivariate Adaptive Generalized Poisson Regression Spline
GCV	Generalized cross-validation
OLS	Ordinary Least Squares

Appendix 1 Simulation data

Y	**X1**	**X2**	**X3**	**X4**	**X5**
3	2.136	2.226	3.068	2.994	3.035
0	3.074	1.954	3.327	2.627	1.004
5	1.059	2.900	2.711	0.609	4.272
5	1.699	2.008	3.552	1.650	3.111
5	4.468	3.455	4.069	2.698	4.204
4	2.820	1.313	2.192	3.305	0.364
2	2.252	2.049	3.603	2.241	1.878
5	0.428	0.310	2.538	0.921	3.632
7	0.510	1.180	1.985	3.555	1.190
4	0.912	4.158	1.993	1.374	0.489
1	4.029	3.708	1.373	1.512	2.854
0	0.268	3.947	1.726	2.751	4.099
3	2.928	1.374	2.861	0.781	2.319
9	3.037	4.971	0.244	4.765	3.619
3	0.814	3.267	2.213	0.104	2.660
5	3.128	2.590	1.768	3.174	2.040
4	3.848	4.852	1.821	2.713	4.133
5	3.832	3.224	1.654	2.384	1.866
6	3.350	1.769	0.388	3.782	1.477
1	0.102	3.886	1.302	2.118	1.650
4	3.089	1.792	4.981	2.756	1.196
5	2.453	0.143	0.892	0.961	2.561
5	1.932	4.996	2.122	3.740	3.178
5	1.254	2.689	1.975	1.350	4.156
4	4.797	1.027	0.361	1.429	1.265
4	2.161	2.141	4.042	0.830	3.911
9	1.586	1.966	2.303	4.631	4.244

Appendix 2 Estimation of Basis Function Parameters for SEM-MAGPRS

$$\frac{\partial\left(S\left(a_0,a_1,\ldots,a_m\right)\right)}{\partial(\mathbf{a})}=\frac{\partial\begin{bmatrix}\mathbf{y}'\mathbf{y}-2\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}\\+\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\end{bmatrix}}{\partial(\mathbf{a})}$$

$$0=-2\mathbf{B}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}$$
$$+2\mathbf{Ba}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$2\mathbf{B}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=2\mathbf{Ba}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\mathbf{B}^{-1}\mathbf{B}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\mathbf{B}^{-1}\mathbf{Ba}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]^{-1}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\mathbf{a}\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]^{-1}$$
$$\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]$$

$$\hat{\mathbf{a}}=\left[\left(\exp\left(\mathbf{B}\hat{\mathbf{a}}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{B}\hat{\mathbf{a}}+\delta\mathbf{Wu}\right)\right)\right]^{-1}\left(\exp\left(\mathbf{B}\hat{\mathbf{a}}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}$$

Appendix 3 Estimation of Spatial Lag Parameter for SEM-MAGPRS

$$\frac{\partial\left(S\left(a_0,a_1,\ldots,a_m\right)\right)}{\partial(\delta)}=\frac{\partial\begin{pmatrix}\mathbf{y}'\mathbf{y}-2\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}\\+\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\end{pmatrix}}{\partial(\delta)}$$

$$0=-2\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}$$
$$+2\delta\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$2\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=2\delta\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\delta\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\left(\mathbf{Wu}\right)^{-1}\left(\mathbf{Wu}\right)\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\delta\left(\mathbf{Wu}\right)^{-1}\mathbf{Wu}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\delta\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)$$

$$\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]^{-1}\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\mathbf{y}=\delta\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]^{-1}$$
$$\left[\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\delta\mathbf{Wu}\right)\right)\right]$$

$$\hat{\delta}=\left[\left(\exp\left(\mathbf{Ba}+\hat{\delta}\mathbf{Wu}\right)\right)'\left(\exp\left(\mathbf{Ba}+\hat{\delta}\mathbf{Wu}\right)\right)\right]^{-1}\left(\exp\left(\mathbf{Ba}+\hat{\delta}\mathbf{Wu}\right)\right)'\mathbf{y}$$

REFERENCES

[1] G. Koshkin and A. Kitayeva, “Nonparametric Identification of Static and Dynamic Production Functions,” IAENG International Journal of Applied Mathematics, vol. 41, no. 3, pp228-234, 2011.

[2] L. S. Dharmasena, P. Zeephongsekul, and B. M. de Silva, “Sample Size Determination for Kernel Regression Estimation Using Sequential Fixed-Width Confidence Bands,” IAENG International Journal of Applied Mathematics, vol. 38, no. 3, pp129-135, 2008.

[3] D. P. Rahmawati, I. N. Budiantara, D. D. Prastyo, and M. A. D. Octavanny, “Mixture Spline Smoothing and Kernel Estimator in Multi-Response Nonparametric Regression,” IAENG International Journal of Applied Mathematics, vol. 51, no. 3, pp695-706, 2021.

[4] R. E. Caraka et al., “Employing Best Input SVR Robust Lost Function with Nature-Inspired Metaheuristics in Wind Speed Energy Forecasting,” IAENG International Journal of Computer Science, vol. 47, no. 3, pp572-584, 2020.

[5] I. A. Gheyas and L. S. Smith, “A Novel Nonparametric Multiple Imputation Algorithm for Estimating Missing Data,” Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2009, 1-3 July, 2009, London, U.K., pp1281-1286.

[6] P. Mishra, “Forecasting Natural Gas Price - Time Series and Nonparametric Approach,” Lecture Notes in Engineering and Computer Science: Proceedings of the World Congress on Engineering 2012, 4-6 July, 2012, London, U.K., pp490-497.

[7] J. H. Friedman, “Multivariate Adaptive Regression Splines (MARS),” The Annals of Statistics, vol. 19, no. 1, pp. 1–67, Mar. 1991.

[8] A. P. Ampulembang, B. W. Otok, A. T. Rumiati, and Budiasih, “Bi-responses nonparametric regression model using MARS and its properties,” Applied Mathematical Sciences, vol. 9, no. 29–32, pp. 1417–1427, 2015, doi: 10.12988/ams.2015.5127.

[9] J. Stoklosa and D. I. Warton, “A Generalized Estimating Equation Approach to Multivariate Adaptive Regression Splines,” Journal of Computational and Graphical Statistics, vol. 27, no. 1, pp. 245–253, Jan. 2018, doi: 10.1080/10618600.2017.1360780.

[10] C. Kooperberg, “Multivariate Adaptive Regression Splines.” Wiley StatsRef: Statistics Reference Online, Seattle, WA, USA, 2014.

[11] L. Anselin, Spatial Econometrics: Methods and Models, no. 411. Santa Barbara: Springer, 1988.

[12] B. W. Otok, R. Y. Putra, and S. D. P Yasmirullah, “BOOTSTRAP AGGREGATING MULTIVARIATE ADAPTIVE REGRESSION SPLINE FOR OBSERVATIONAL STUDIES IN DIABETES CASES,” 2020.

[13] S. D. P. Yasmirullah, B. W. Otok, J. D. T. Purnomo, and D. D. Prastyo, “Parameter Estimation of Multivariate Adaptive Regression Spline (MARS) with Stepwise Approach to Multi

Drug-Resistant Tuberculosis (MDR-TB) Modeling in Lamongan Regency," in Journal of Physics: Conference Series, Feb. 2021, vol. 1752, no. 1. doi: 10.1088/1742-6596/1752/1/012017.

[14] S. D. P. Yasmirullah, B. W. Otok, J. D. T. Purnlmo, and D. D. Prastyo, "Multivariate adaptive regression spline (MARS) methods with application to multi drug-resistant tuberculosis (mdr-tb) prevalence," in AIP Conference Proceedings, Feb. 2021, vol. 2329. doi: 10.1063/5.0042145.

[15] F. Famoye, J. T. Wulu, and K. P. Singh, "On the Generalized Poisson Regression Model with an Application to Accident Data," 2004.

[16] F. Famoye, "Restricted generalized poisson regression model," Commun Stat Theory Methods, vol. 22, no. 5, pp. 1335–1354, Jan. 1993, doi: 10.1080/03610929308831089.

[17] P. McCullagh and J. A. Nelder, Generalized Linear Models, 2nd ed. London: Chapman and Hall, 1989.

[18] S. D. P. Yasmirullah, B. W. Otok, J. D. Trijoyo Purnomo, and D. D. Prastyo, "Modification of Multivariate Adaptive Regression Spline (MARS)," in Journal of Physics: Conference Series, Apr. 2021, vol. 1863, no. 1. doi: 10.1088/1742-6596/1863/1/012078.

[19] S. Hidayati, B. W. Otok, and Purhadi, "Parameter Estimation and Statistical Test in Multivariate Adaptive Generalized Poisson Regression Splines," in IOP Conference Series: Materials Science and Engineering, Jul. 2019, vol. 546, no. 5. doi: 10.1088/1757-899X/546/5/052051.

Septia Devi Prihastuti Yasmirullah is an Assistant Lecturer at the Faculty of Advanced and Multidisciplinary Technology, Universitas Airlangga. She obtained Bachelor's and Master's degrees in Statistics from the Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember (ITS), Indonesia. She is currently preparing for a doctorate in statistics at Institut Teknologi Sepuluh Nopember Surabaya. Her research interests include nonparametric regression, multivariate adaptive regression spline (MARS), and spatial modeling.

Bambang Widjanarko Otok received Dr., M.Si. S.Si. degree in Statistics from Institut Teknologi Sepuluh Nopember in 1994, Statistics from Institut Pertanian Bogor in 2000, and Mathematics from Universitas Gadjah Mada in 2008, respectively. Currently, He is a Professor at Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya. His research interest includes MARS, SEM, Classification Method and Machine Learning.

Jerry Dwi Trijoyo Purnomo received the B.Sc. and M.Sc. degrees in statistics from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, respectively, and the Ph.D. degree in statistics from the National Chiao Tung University, Taiwan. He is currently a Lecturer at the Department of Statistics, ITS. His research interests include biostatistics, clinical trial and epidemiology.

Dedy Dwi Prastyo received the B.Sc. and M.Sc. degrees in statistics from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, in 2006 and 2008, respectively, and the Ph.D. degree in statistics from the Ladislaus von Bortkiewicz Chair of Statistics, Institute of Statistics and Econometrics, Humboldt-Universität zu Berlin (HUB), Germany, in 2015. He was a Research Assistant at HUB, during his doctoral study, where he was responsible for research, teaching, and other managerial tasks. He is currently a Lecturer at the Department of Statistics, ITS. He has an additional position as the Head of Postgraduate Program in statistics, ITS, whereas previously, he was the Deputy Head of the Program. He has contributed to more than 60 research papers that are published in reputable journals or conferences. He published numerous articles that can be found in his Scopus account. His research interests include (but not limited to) computational statistics and finance, financial econometrics, machine learning for classification and forecasting, and statistical process control.