# MODELING CASE FATALITY RATE OF COVID-19 IN INDONESIA USING TIME SERIES SEMIPARAMETRIC REGRESSION BASED ON LOCAL POLYNOMIAL ESTIMATOR

VITA FIBRIYANI[1,2], NUR CHAMIDAH[3,4,*], TOHA SAIFUDIN[3,4]

[1]Doctoral Study Program of Mathematics and Natural Sciences, Faculty of Science and Technology, Airlangga

University, Surabaya 60115, Indonesia

[2]Department of Management, Faculty of Economic, Universitas Merdeka Pasuruan, Pasuruan 67129, Indonesia

[3]Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia

[4]Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Airlangga University,

Surabaya 60115, Indonesia

*Corresponding author: nur-c@fst.unair.ac.id

**Abstract:** At the beginning of 2020, the world was shocked by the COVID-19 which causes acute and contagious respiratory problems. Several countries in the world have declared the COVID-19 outbreak a pandemic, including Indonesia. The number of patients who have infected by the COVID-19 and died due to this virus continues to increase every day. In this study, we model the Case Fatality Rate (CFR) of COVID-19 in one city in Indonesia by using semiparametric regression model approach based on time series local polynomial estimator where the local polynomial estimator is used to accommodate fluctuations of daily COVID-19 death rate data. The daily data on CFR of COVID-

19 during a certain period in one city were used as the basis for the analysis. In this approach, we combine parametric regression and nonparametric regression components to identify factors that influence on the CFR. The parametric component represents factors that influence on the CFR with a known regression function, while the nonparametric component represents factors that influence on the COVID-19 death rate with an unknown regression function. In this study, it is known that the parametric component is the COVID-19 death rate one day earlier, while the nonparametric component is a time series. Result shows that the time series semiparametric regression approach based on local polynomial estimator can model the CFR of COVID-19 well, and in the future the obtained model can be used to predict the CFR of COVID-19 in Indonesia, especially in Pasuruan city, for supporting one of the (Sustainable Development Goals) SDGs namely control the pandemic. Also, the obtained model can be used to predict the death rate caused by COVID-19 or similar disease outbreaks.

**Keywords:** fatality rate; COVID-19, time series; semiparametric regression; local polynomial estimator.

**2020 AMS Subject Classification:** 62G05, 62G08, 62P10.

## 1. INTRODUCTION

In December 2019, Coronavirus disease 2019 (COVID-19) was first discovered and reported in Wuhan, China, and then it was later reported in other countries. The COVID-19 is a novel viral disease caused by severe acute respiratory syndrome Coronavirus-2 (SARS-Cov-2) which led to a global pandemic [1]. The COVID-19 spread rapidly to other countries in the following weeks and was eventually classified as a global pandemic on March 11, 2020 [2]. It has rapidly spread around the world, posing enormous health, economic, environmental and social challenges to the entire human population [3].

In Indonesia, COVID-19 case was first announced to public on March 2, 2020. Based on data at that time, it is known that the number of confirmed cases of COVID-19 was 1528 cases with the number of deaths was 136 [4]. The Case Fatality Rate (CFR) has especially become an important thing that was discussed by many people when the pandemic occurred. Sipahutar and Eryando [5] explained that on social media, people have been debating the national CFR of COVID-19 and even compared it with CFR of other countries such as China and Vietnam. This

comparison was followed by comparing the performance of Indonesian government in handling the COVID-19 pandemic. Therefore, a proper analysis is needed to model the CFR of COVID-19 by using statistical modeling techniques. One of the statistical techniques used to describe functional relationship between response variables and predictor variables is regression model. There are three common types of regression models namely parametric regression model, i.e. parametric, nonparametric regression model, and semiparametric regression model. The parametric regression model has rigid assumptions such as has known certain form of regression function and other rigid assumptions. While the nonparametric regression model assumes that the regression function is unknown and contained in the Sobolev space [6,7]. The nonparametric regression model approach has high flexibility because the regression function is not specified in a particular form but it is assumed to be smooth, so that it can be estimated by using certain smoothing methods based on data patterns [8]. Furthermore, the semiparametric regression model is a combination regression model between parametric regression model and nonparametric regression model [9–11]. One of the parametric approaches used for predicting purpose is the linear regression analysis which has been done by Suleiman et al. [12]. Additionally, another parametric approach is negative binomial model which has been done by Pan et al. [13] for identifying factors that may explain the variation in CFR across countries. Some previous studies have discussed nonparametric regression models for some cases and applications. For example, Nidhomuddin et al.[14] used local linear estimator for modeling case increase and case fatality rate COVID-19 in Indonesia, Siregar et al. [15] used penalized spline estimator to model world crude oil price on the effects of COVID-19 pandemic. Next, in the nonparametric regression and semiparametric regression models, several smoothing techniques which are often used are kernel [16,17], local linear [18–21], local polynomial [22–27], and Spline [28–37]. According to Chamidah et al. [23], the local polynomial estimator is a popular approach and has exceptional cases that are if the degree of local polynomial equals to zero then it is called as kernel, and if the polynomial degree equals to one, then it is called as local linear. Furthermore, we use the local linear estimator when the data

patterns are generally monotonous [38]. According to Loader [39], we can approach locally every differentiable function by a straight line.

Time series data is a set of observational data taken at different times and collected periodically at a certain time interval [40]. Time series data includes one research object or individual, such as the confirmed cases, fatality rate, mortality rate in several periods, such as daily, weekly, monthly, or yearly. Time series data modeling usually uses classical models such as Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA). These models are linear models in the time series, which are very commonly used in public health and biomedical data [41]. There are previous researchers that have modeled case fatality rate of COVID-19 for example Meimela et al. [42] and Somyanonthanakul [43], used ARIMA, ARIMAX and association rule mining to forecasting COVID-19 cases.

The nonparametric regression approach for time series data has been developed by several researchers, for example, Gao and Gijbel [44], Chen et al. [45], Gao and King [46] carried out modeling based on the kernel estimator. Also, some researchers such as Fernandez and Cao [47], Wang and Phillips [48], and Li et al. [49] developed nonparametric regression approach to time series data based on local linear estimators. Dong and Gao [50] have used the truncated least square estimator to model time series data. In addition, several researchers have developed semiparametric regression approach to time series data, including Gao and Hawthorne [51]; Perez and Vieu [52], and Gao [53] modeled time series data by using kernel estimator.

Local polynomial estimator is a popular new approach to smoothing technique which has a special form for polynomial degree ($d$) equal to zero which is called kernel estimator or constant local estimator. Liang and Chen [38] it is explained that if the data has a monotonous pattern, a degree polynomial of one ($d=1$) can be used. But if the patterned data is not monotonous with a local maximum or minimum, then the degree of polynomial one cannot overcome the sharpness of the curve in the data so that a local polynomial estimator with a higher degree of polynomial is needed to get a suitable estimate.

This research aims to develop a model of local polynomial semiparametric regression for time series data and it's applied to modeling a case fatality rate of COVID-19 in Kota Pasuruan, East Java, Indonesia. This research theoretically discusses how we estimate the model using polynomial local estimator for semiparametric regression approaches. The result of this research is expected to be applied in modeling a CFR of COVID-19 in Indonesia.

## 2. PRELIMINARIES

Given a nonparametric regression model as follows [54]:

$$(1) \qquad y_t = g(v_t) + \varepsilon_t, \quad t = 1, 2, \ldots, T$$

where $y$ is response variable, $g(.)$ is unknown smooth function and $\varepsilon_t$ is zero mean and $\sigma^2$ variance random error. To estimate function $g(v_t)$ in Equation (1), we use local polynomial estimator. Function $g(v_t)$ can be approximated by Taylor expansion as follows:

$$(2) \qquad g(v_t) \approx \sum_{d=0}^{r} \frac{m^{(r)}(v_0)}{d!}(v_t - v_0)^d = \sum_{d=0}^{r} \beta_d(v_0)(v_t - v_0)^d$$

where $g^{(d)}(v_0)$ is the $d$-derivative of $g(v_0)$ at $v_0$, $v_t \in (v_0 - h, v_0 + h)$, and $\dfrac{g^{(d)}(v_0)}{d!} = \beta_d(v_0)$.

If Equation (2) is written in matrix notation, we obtain

$$(3) \qquad g(v_t) = \mathbf{V}_{v_0} \boldsymbol{\beta}(v_0)$$

where $\qquad \mathbf{V}_{v_0} = \begin{bmatrix} 1 & (v_t - v_0) & \ldots & (v_t - v_0)^r \end{bmatrix}, v_t \in (v_0 - h, v_0 + h)$ $\qquad$ and

$\boldsymbol{\beta}(v_0) = \begin{bmatrix} \beta_0(v_0) & \beta_1(v_0) & \ldots & \beta_r(v_0) \end{bmatrix}^T$.

Based on equation (3), we can write equation (1) as:

$$(4) \qquad y_t = \mathbf{V}_{v_0} \boldsymbol{\beta}(v_0) + \varepsilon_t$$

By giving $T$ samples paired data $\{v_t, y_t\}_{t=1}^{T}$, equation (4) can be written as follows:

$$
\left.\begin{aligned}
y_1 &= \beta_0\left(v_0\right)+\beta_1\left(v_0\right)\left(v_1-v_0\right)+\beta_2\left(v_0\right)\left(v_1-v_0\right)^2+\ldots+\beta_r\left(v_0\right)\left(v_1-v_0\right)^r+\varepsilon_1 \\
y_2 &= \beta_0\left(v_0\right)+\beta_1\left(v_0\right)\left(v_2-v_0\right)+\beta_2\left(v_0\right)\left(v_2-v_0\right)^2+\ldots+\beta_r\left(v_0\right)\left(v_2-v_0\right)^r+\varepsilon_2 \\
&\qquad\vdots \\
y_T &= \beta_0\left(v_0\right)+\beta_1\left(v_0\right)\left(v_T-v_0\right)+\beta_2\left(v_0\right)\left(v_T-v_0\right)^2+\ldots+\beta_r\left(v_0\right)\left(v_T-v_0\right)^r+\varepsilon_T
\end{aligned}\right\}
$$
(5)

Hence, equation (5) can be expressed in matrix notation:

(6)
$$
\underset{\sim}{\mathbf{y}} = \mathbf{V}_{v_0}\underset{\sim}{\boldsymbol{\beta}}\left(v_0\right)+\underset{\sim}{\boldsymbol{\varepsilon}}
$$

where

$$
\mathbf{V}_{v_0} = \begin{bmatrix} 1 & \left(v_1-v_0\right) & \cdots & \left(v_1-v_0\right)^r \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \left(v_T-v_0\right) & \cdots & \left(v_T-v_0\right)^r \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix}
$$

To estimate $\underset{\sim}{\boldsymbol{\beta}}\left(v_0\right)$, we apply weighted least square (WLS) optimization as follows:

(7)
$$
Q\left(v_0\right) = Min\left(\underset{\sim}{\mathbf{y}}-\mathbf{V}_{v_0}\underset{\sim}{\boldsymbol{\beta}}\left(v_0\right)\right)^T \mathbf{K}_h\left(v_0\right)\left(\underset{\sim}{\mathbf{y}}-\mathbf{V}_{t_0}\underset{\sim}{\boldsymbol{\beta}}\left(v_0\right)\right)
$$

where

$$
\mathbf{K}_h\left(v_0\right) = \begin{bmatrix} K_h\left(v_1-v_0\right) & 0 & \cdots & 0 \\ 0 & K_h\left(v_2-v_0\right) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_h\left(v_T-v_0\right) \end{bmatrix}
$$

$K_h\left(.\right)$ is kernel function with bandwidth $h$ defined by [55] as follows:

$$
K_h\left(v\right) = \frac{1}{h}K\left(\frac{v}{h}\right); -\infty < v < \infty \text{ and } h > 0
$$

In this study, we use Gaussian Kernel defined as follows [56]:

$$
K\left(v\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2}{2}\right)
$$

Next, we take derivative of equation (7) with respect to $\beta\left(v_0\right)$ and then equalizing to zero, such that we get:

$$\frac{\partial Q(v_0)}{\partial \beta(v_0)} = -2\mathbf{V}_{v_0}{}^T\mathbf{K}_h(v_0)\mathbf{y} + 2\mathbf{V}_{v_0}{}^T\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\beta(v_0) = 0$$

$$\mathbf{V}_{v_0}^T K_h(v_0)\mathbf{V}_{v_0}\beta(v_0) = \mathbf{V}_{v_0}^T K_h(v_0)\mathbf{V}_{v_0}\beta(v_0)\mathbf{y}, \text{ so}$$

(8)
$$\widehat{\beta}(v_0) = \left(\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)\mathbf{y}$$

Based on equations (3) and (8), the local polynomial estimator for $\widehat{g}(v_t)$ is

(9)
$$\widehat{g}(v_t) = v_{v_0}\left(\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)\mathbf{y}$$

To get the best estimation, one of the most important things is to choose an optimal bandwidth with associated Kernel function. This can be done using Generalized Cross-Validation criterion with formula [56]:

(10)
$$GCV_{(h)} = \frac{1}{n}\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\left[1 - n^{-1}trace\mathbf{A}_{(h)}\right]^2}$$

dengan, $\mathbf{A}_{(h)} = v_{v_0}\left(\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T\mathbf{K}_h(v_0)$.

The error rate measurement to compare the best estimator is based the value of:

(11)
$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{|y_t - \hat{y}_t|}{y_t}\times 100\%$$

where $T$ is the size of the sample, $\hat{y}_t$ is the value predicted by the model for time point $t$ and $y$ is the value observed at time point $t$. The criteria for MAPE values are shown in the Table 1[57].

**Table 1**. MAPE Value Criteria

| MAPE | Definition |
| --- | --- |
| < 10 | Highly Accurate |
| 10 – 20 | Accurate |
| 20 – 50 | Reasonable |
| >50 | Inaccurate |

## 3. MAIN RESULTS

In this section, we provide the theoretical results on estimating local polynomial semiparametric regression model for time series data including estimating model for estimating

local polynomial semiparametric regression model and implementation to modeling a case fatality rate (CFR) of COVID-19.

## 3.1 Estimating Model

Suppose given a paired data set $(y_{t-1}, v_{t-1}, y_t)$, where $t = 1, 2, 3, \ldots, T$ is the number of observation, $y_{t-1}$ is the autoregressive from response variable, $v_{t-1}$ is the autoregressive from predictor variable, and the data meet the local polynomial semiparametric regression model as follows:

$$(12) \qquad y_t = \theta_1 y_{t-1} + g(v_{t-1}) + \varepsilon_t$$

where $y_t$ is response variable, $\theta_1 y_{t-1}$ is component of parametric, $g(v_{t-1})$ is component of nonparametric, $\varepsilon_t$ is the measurement error with mean is $\mathbf{0}$ and variance is $\Sigma_t$.

First, we estimate the nonparametric regression model where the component of parametric assumed known such that the following equation is obtained:

$$(13) \qquad y_t^* = g(v_{t-1}) + \varepsilon_t$$

where $y_t^* = y_t - \theta_1 y_{t-1}$. Next, we use a local polynomial estimator to estimate the nonparametric function in estimating semiparametric regression model. Equation (13) is smooth function assumed to have an unknown form and is estimated using a nonparametric approach based on local polynomial estimator. The nonparametric function $g(v_{t-1})$ is a smooth function with continous and differentiable properties. The differentiable function can be approximated by Taylor series expansion. The Taylor series for $g(v_{t-1})$ around $v_0$ can be expressed as follows:

$$(14) \qquad g(v_{t-1}) \approx \sum_{d=0}^{r} \frac{g^{(d)}(v_0)}{d!} (v_{t-1} - v_0); d = 0, 1, 2, \ldots, d$$

$$(15) \qquad g(v_{t-1}) = \sum_{d=0}^{r} \beta_d(v_0)(v_{t-1} - v_0)$$

where $g^{(d)}(v_0)$ is $d-$th derivative of $g(v_0)$ to $v_0$, for $v_{t-1} \in (v_0 - h, v_0 + h)$. The equation (13)

can be written in matrix form as follows:

(16) $$g(v_{t-1}) = \mathbf{v}_{(t)v_0}\boldsymbol{\beta}(v_0)$$

where $\mathbf{v}_{(t)v_0} = \begin{bmatrix} 1 & (v_{t-1}-v_0) & (v_{t-1}-v_0)^2 & \cdots & (v_{t-1}-v_0)^r \end{bmatrix}, v_{t-1} \in (v_0-h, v_0+h)$ and

$$\boldsymbol{\beta}(v_0) = \begin{bmatrix} \beta_0(v_0) & \beta_1(v_0) & \beta_2(v_0) & \cdots & \beta_r(v_0) \end{bmatrix}^T$$

Based on equation (16), the model on equation (13) can be written as follows:

(17) $$\mathbf{y}_t^* = \mathbf{v}_{(t)v_0}\boldsymbol{\beta}(v_0) + \boldsymbol{\varepsilon}_t$$

Estimation of $\underset{\sim}{\beta}(v_0)$ based on local polynomial estimator on equation (17) obtained by taking a

paired sample $\{v_{t-1}, y_t^*\}_{t=2}^T$ so that based on equation (17) can be written as follows:

(18)
$$\begin{cases} y_2^* = \beta_0(v_0) + \beta_1(v_0)(v_1-v_0) + \beta_2(v_0)(v_1-v_0)^2 + \ldots + \beta_r(v_0)(v_1-v_0)^r + \varepsilon_2 \\ y_3^* = \beta_0(v_0) + \beta_1(v_0)(v_2-v_0) + \beta_2(v_0)(v_2-v_0)^2 + \ldots + \beta_r(v_0)(v_2-v_0)^r + \varepsilon_3 \\ y_4^* = \beta_0(v_0) + \beta_1(v_0)(v_3-v_0) + \beta_2(v_0)(v_3-v_0)^2 + \ldots + \beta_r(v_0)(v_3-v_0)^r + \varepsilon_4 \\ \vdots \\ y_T^* = \beta_0(v_0) + \beta_1(v_0)(v_{T-1}-v_0) + \beta_2(v_0)(v_{T-1}-v_0)^2 + \ldots + \beta_r(v_0)(v_{T-1}-v_0)^r + \varepsilon_T \end{cases}$$

Equation (18) can be expressed in matrix form as:

(19) $$\mathbf{y}^* = \mathbf{V}_{v_0}\boldsymbol{\beta}(v_0) + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y}^* = \begin{bmatrix} y_2^* \\ y_3^* \\ y_4^* \\ \vdots \\ y_T^* \end{bmatrix}; \quad \mathbf{V}_{v_0} = \begin{bmatrix} 1 & (v_1-v_0) & (v_1-v_0)^2 & \cdots & (v_1-v_0)^r \\ 1 & (v_2-v_0) & (v_2-v_0)^2 & \cdots & (v_2-v_0)^r \\ 1 & (v_3-v_0) & (v_3-v_0)^2 & \cdots & (v_3-v_0)^r \\ \vdots & & & \ddots & \vdots \\ 1 & (v_{T-1}-v_0) & (v_{T-1}-v_0)^2 & \cdots & (v_{T-1}-v_0)^r \end{bmatrix}; \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

$$\boldsymbol{\beta}(v_0) = \begin{bmatrix} \beta_0(v_0) & \beta_1(v_0) & \beta_2(v_0) & \cdots & \beta_r(v_0) \end{bmatrix}^T;$$

The estimator of $\boldsymbol{\beta}(v_0)$ is obtained based on local polynomial estimator using the Kernel function

$K_h(v_{t-1}-v_0)$ as weighting. The weight shape is determined by the kernel function, while the

weight size is determined by the value of the parameter $h$ called bandwidth. Estimation of $\hat{\boldsymbol{\beta}}(v_0)$

in equation (19) using the weighted least square (WLS) can be obtained by minimizing the function

(20)
$$Q(v_0) = \left(\mathbf{y}^* - \mathbf{V}_{v_0}\boldsymbol{\beta}(v_0)\right)^T \mathbf{K}_h(v_0)\left(\mathbf{y}^* - \mathbf{V}_{v_0}\boldsymbol{\beta}(v_0)\right)$$

where

$$K_h(v_0) = \begin{bmatrix} K_h(v_1 - v_0) & 0 & 0 & \dots & 0 \\ 0 & K_h(v_2 - v_0) & 0 & \dots & 0 \\ 0 & 0 & K_h(v_3 - v_0) & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & 0 \\ 0 & 0 & 0 & \dots & K_h(v_{T-1} - v_0) \end{bmatrix}$$

is the matrix containing the weight function, and $K_h(.)$ is kernel function.

The estimated value for $\boldsymbol{\beta}(v_0)$ is $\hat{\boldsymbol{\beta}}(v_0)$, if substituted in Equation (20) will minimize $Q(v_0)$.

This estimated value can be obtained by differentiating equation (20) with respect to $\boldsymbol{\beta}(v_0)$. The

minimum value of $Q(v_0)$ is reached when $\dfrac{\partial Q(v_0)}{\partial \boldsymbol{\beta}(v_0)} = \mathbf{0}$. Hence, we have:

(21)
$$\hat{\boldsymbol{\beta}}(v_0) = \left(\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{y}^*$$

Based on equation (16) and (21), local polynomial estimator for $\hat{g}(V_{t-1})$ can be written as follows:

(22)
$$\hat{g}(v_{t-1}) = \mathbf{v}_{(t)v_0}\left(\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{y}^*$$

Equation (22) can be written as follows:

(23)
$$\hat{g}(v_{t-1}) = \mathbf{a}_{(t)v_0}\mathbf{y}^*$$

where $\mathbf{a}_{(t)v_0} = \mathbf{v}_{v_0}\left(\mathbf{V}^T_{v_0}K_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)$.

Based on the estimation results on equation (23), semiparametric regression model based on local

polynomial estimator can be written as follows:

$$y_t^* = \mathbf{a}_{(t)v_0}y_t^* + \varepsilon_t$$

where $y_t^* = y_t - \theta y_{t-1}$, so that we have:

$$y_t - \theta y_{t-1} = \mathbf{a}_{(t)v_0}\left(y_t - \theta y_{t-1}\right) + \varepsilon_t$$

(24)
$$y_t = \theta y_{t-1} + \mathbf{a}_{(t)v_0}\left(y_t - \theta y_{t-1}\right) + \varepsilon_t .$$

Semiparametric regression model on equation (24) still contains a parameter $\theta$, so we need to estimate it. The equation (24) can be written as follows:

(25)
$$\mathbf{y}_t = \theta \mathbf{y}_{t-1} + \mathbf{A}_{v_0}\left(\mathbf{y}_t - \theta \mathbf{y}_{t-1}\right) + \varepsilon_t$$

where $\mathbf{y}_t = \begin{bmatrix} y_2 \\ y_3 \\ \vdots \\ y_T \end{bmatrix}$; $\mathbf{y}_{t-1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{T-1} \end{bmatrix}$; $\mathbf{A}_{v_0} = \begin{bmatrix} \mathbf{a}_{(2)v_0}, \mathbf{a}_{(3)v_0}, \ldots, \mathbf{a}_{(T)v_0} \end{bmatrix}^T$; $\varepsilon_t = \begin{bmatrix} \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_T \end{bmatrix}$

To get the estimator of $\theta$ can be done by minimizing a sum of square $S$ of semiparametric regression model given in equation (25) as follows:

$$S = \varepsilon^T \varepsilon$$

(26)
$$S = \left(\mathbf{y}_t - \theta\mathbf{y}_{t-1} - \mathbf{A}_{v_0}\left(\mathbf{y}_t - \theta\mathbf{y}_{t-1}\right)\right)^T \left(\mathbf{y}_t - \theta\mathbf{y}_{t-1} - \mathbf{A}_{v_0}\left(\mathbf{y}_t - \theta\mathbf{y}_{t-1}\right)\right)$$

Hence, the equation (26) can be written as follows:

(27)
$$S = \mathbf{y}_t{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t - 2\theta^T\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t + \theta^T\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\theta\mathbf{y}_{t-1}$$

The estimated value of $\theta$ is $\hat{\theta}$. This estimated value can be obtained by differentiating equation (27) with respect to $\theta$. The minimum value of $S$ is reached when $\dfrac{\partial S}{\partial \theta} = \mathbf{0}$. The estimated value of $\theta$ can be described as follows:

$$\frac{\partial S}{\partial \theta} = \mathbf{0}$$

$$\frac{\partial\left(\mathbf{y}_t{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t - 2\theta^T\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t + \theta^T\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\theta\mathbf{y}_{t-1}\right)}{\partial \theta} = \mathbf{0}$$

$$0 - 2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t + 2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\hat{\theta}\mathbf{y}_{t-1} = \mathbf{0}$$

$$-2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t + 2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\hat{\theta}\mathbf{y}_{t-1} = \mathbf{0}$$

$$2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\hat{\theta}\mathbf{y}_{t-1} = 2\mathbf{y}_{t-1}{}^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^T\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_t$$

$$\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\hat{\theta}\mathbf{y}_{t-1}=\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t}$$

$$\hat{\theta}=\left(\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t-1}\right)^{-1}\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t}$$

Hence,

$$(28)\qquad \hat{\theta}=\left(\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t-1}\right)^{-1}\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t}\ .$$

The estimation results of the semiparametric regression model based on the local polynomial estimator are obtained by substituting equation (28) into equation (24).

$$\hat{\mathbf{y}}_{t}=\hat{\theta}\mathbf{y}_{t-1}+\mathbf{A}_{v_0}\left(\mathbf{y}_{t}-\hat{\theta}\mathbf{y}_{t-1}\right)+\boldsymbol{\varepsilon}_{t}$$

$$(29)\qquad
\begin{aligned}
\hat{\mathbf{y}}_{t}&=\mathbf{y}_{t-1}\left(\left(\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t-1}\right)^{-1}\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t}\right)\\
&\quad+\mathbf{A}_{v_0}\left(\mathbf{y}_{t}-\mathbf{y}_{t-1}\left(\left(\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t-1}\right)^{-1}\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t}\right)\right)+\boldsymbol{\varepsilon}_{t}
\end{aligned}$$

Hence, the equation (23) can be written as follows:

$$(30)\qquad \hat{\mathbf{y}}_{t}=\mathbf{C}\mathbf{y}_{t}+\mathbf{A}_{v_0}\left(\mathbf{I}-\mathbf{C}\right)\mathbf{y}_{t}$$

where $\mathbf{C}=\mathbf{y}_{t-1}\left(\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)\mathbf{y}_{t-1}\right)^{-1}\mathbf{y}_{t-1}{}^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)^{T}\left(\mathbf{I}-\mathbf{A}_{v_0}\right)$,

$\mathbf{A}_{v_0}=\left[\mathbf{a}_{(2)v_0},\mathbf{a}_{(3)v_0},\ldots,\mathbf{a}_{(T)v_0}\right]^{T}$, and $\mathbf{a}_{(t)v_0}=\mathbf{v}_{(t)v_0}\left(\mathbf{V}^{T}{}_{v_0}K_{h}\left(v_0\right)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^{T}{}_{v_0}K_{h}\left(v_0\right)$.

### 3.2 Implementation on Case Fatality Rate of COVID-19 Data

In 2020, East Java was in the second highest for confirmed cases of COVID-19 in Indonesia. One area that has contributed is Kota Pasuruan. Below, we will model a CFR COVID-19 using semiparametric regression model based on local polynomial estimator. The data used is daily data on case fatality rate (CFR) and case growth rate (CGR). The CFR as response variable and as predictor are CFR and CGR at one period previous. This research uses secondary data obtained from Dinas Kesehatan Kota Pasuruan from 2 October 2020 until 12 August 2021, 315 observations. The data is divided into two parts, 80% data (or 252 observations) to form a model and 20% data (or 63 observations) for testing. The following are the results of descriptive analysis of variables:

**Table 2**. Variable Descriptive Statistics

| Variable | Statistic | | | |
| --- | --- | --- | --- | --- |
| | **Mean** | **Variance** | **Min.** | **Max.** |
| Case Fatality Rate (%) | 10.440 | 1.813 | 6.407 | 11.924 |
| Case Growth Rate (%) | 0.549 | 0.447 | 0.000 | 3.812 |

Since the first two cases of COVID-19 were announced on March 2, 2020 in Indonesia. This announcement seemed to be a sign of major changes and adjustments. measures to prevent virus transmission and social restrictions were put in place, something that had never existed before in the country. COVID-19 continues to spread in the country until it has infected all provinces in Indonesia (34 provinces) and 510 districts/cities as of the end of 2020. Difference in the conditions of the COVID-19 pandemic in each province/district/city have resulted in variations in policy for handling it. At the end of 2020, several regions in Indonesia began to loosen a PSBB (Pembatasan Sosial Skala Besar) or begin a transition period [58].

Based on Table 2, it is known that October 2020 to June 2021 is a transition period for living together with the COVID-19 virus, the average of CFR has been 10.440%, with a comparatively large diversity of 1.813%. The spread of COVID-19 in Kota Pasuruan during the transition period can be seen from CGR value with the mean as 0.549. In August 4, 2021, Kota Pasuruan had the lowest CFR as 6.407% but the highest CFR occurred on October 2, 2020.

The functional link between CFR and CGR at one period previous is depicted in Figure 1. We can observe that it deviates from the traditional definition of a functional relationship. We can see that the pattern of a relationship between CFR and CGR in the previous day is unknown, so that it can be modeling using nonparametric approach. Figure 2 explains that the functional relationship between the CFR and the previous period's CFR is linear, so it can be modeling using a parametric approach. If we combine these in one case, we can use semiparametric approach.
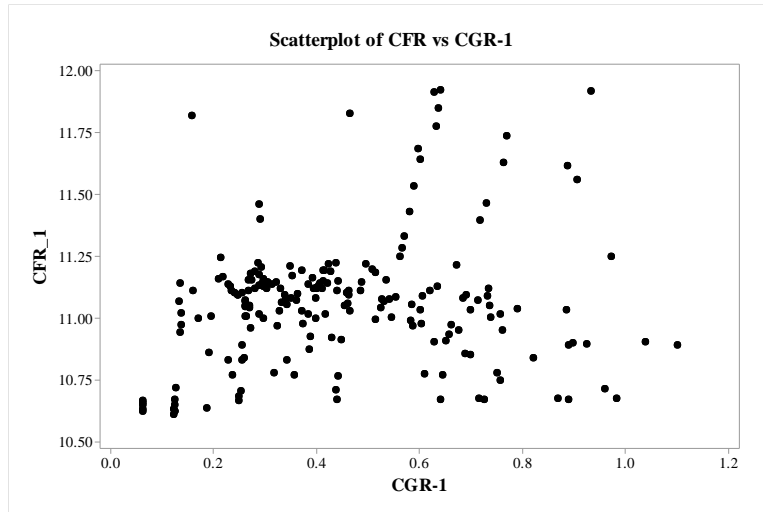
**Figure 1.** Scatter Plot of CFR vs CGR in the previous day Kota Pasuruan
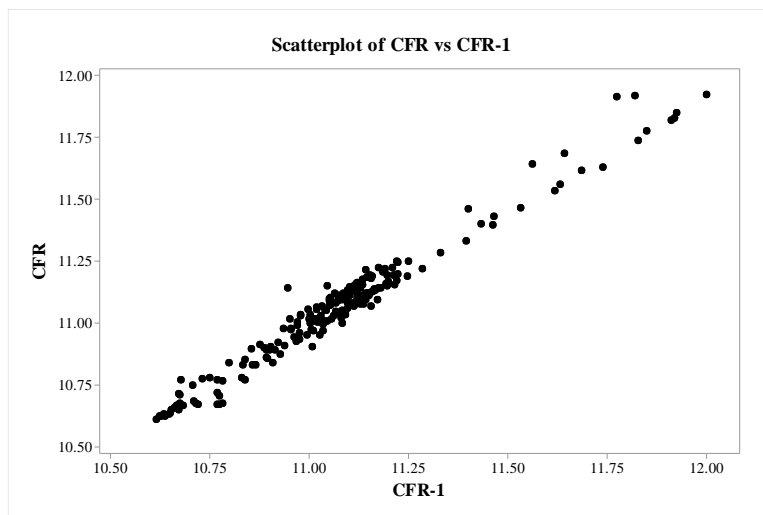


**Figure 2.** Scatter Plot of CFR vs CFR in the previous day Kota Pasuruan

From the information given above, the semiparametric model is written as follows:

(31) $$CFR_t = \theta CFR_{t-1} + \beta_0 + \beta_1\left(CGR_{t-1} - CGR_0\right), CGR_{t-1} \in \left(CGR_0 - h, CGR_0 + h\right)$$

Case fatality rate is being predicted by researchers using semiparametric regression approach based on a local polynomial estimator. A semiparametric regression model based on a local polynomial estimator with polynomial $(p = 1)$ as the degree is created in this study, we can say that

it has a local linear model. The optimal bandwidth for a researcher's local linear model is 0.20 with minimum GCV as 0.0025. The researcher then uses the MAPE criterion to determine the prediction accuracy.

**Table 3** Semiparametric regression approach using MAPE

|  | **MAPE** | **MSE** | $R^2$ |
|---|---|---|---|
| Training | 0.3407 | 0.0024 | 95.8017 |
| Testing | 0.2385 | 0.0037 | 97.5321 |
| Overall | 0.2896 | 0.0031 | 96.6669 |

Table 3 shows that MAPE value for training is 0.3407% and testing is 0.2385%. According to Moreno et al. [57], if the MAPE of a semiparametric regression technique is less than 10, the model provides a highly accuracy for modeling.
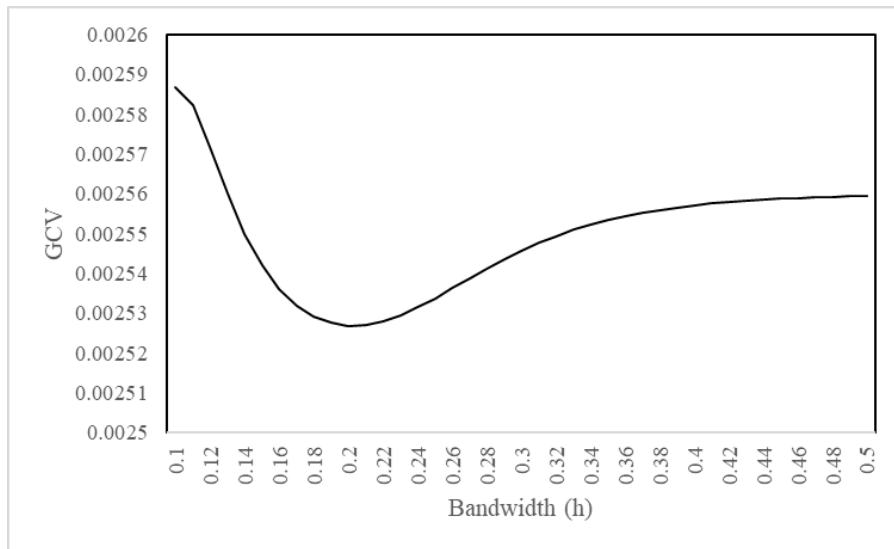


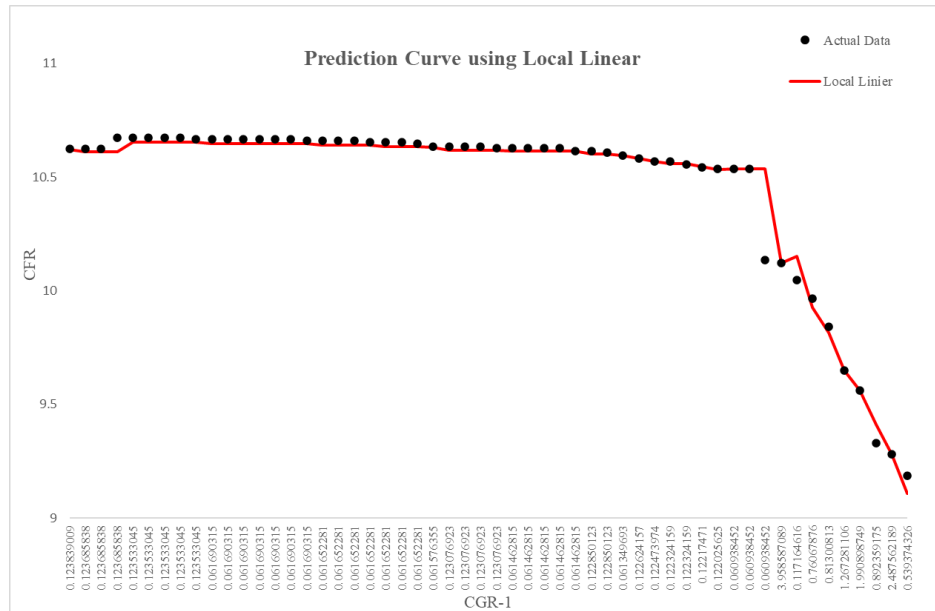**Figure 3.** Plot of bandwidth versus GCV for the Gaussian function type (*d*=1)

**Figure 4**. Prediction Curve using Local Linear Model

According to Table 6, a semiparametric regression model based on local polynomial estimator has a MAPE value less than 10%, so the prediction using a local linear model is highly accurate. Figure 4 show that the plot between actual data of daily CFR COVID-19 and its predicted value. We can see that the predicted value using a semiparametric regression model which is shown the red line (using local linear model) appears to coincide with the actual data. It shows that the prediction error has obtained by a semiparametric regression approach too small.

Based on the result of analysis, we know that prediction of CFR COVID-19 using semiparametric regression model has a small MAPE value, less than 10%. It means that predicted value from this model is highly accurate. It can be used as reference for government to modeling or predicting a CFR of a disease or outbreak. So that, the government can determine that policies will be made to reduce a CFR of disease or extraordinary event.

Estimating a semiparametric regression model is equivalent to estimating a regression function that describes the functional relationship between the response variable and the predictor variables. The regression function of the semiparametric regression model consists of parametric components and nonparametric components. Theoretically, in estimating a semiparametric

regression model for time series data using a local polynomial estimator, we first estimate the parametric component parameter by taking the solution to the weighted least square (WLS) optimization function. In other word we determine the values parameters in the parametric component which minimize the WLS function, so that we obtained the estimator for these parameters as presented in equation (19). Next, based on equation (19), we estimate the nonparametric component regression function of the model such that we obtained the estimated local polynomial semiparametric regression model for time series data as presented in (22).

## 4. CONCLUSIONS

Theoretically, the estimated semiparametric regression model based on local polynomial estimator can be used to modeling a CFR of COVID-19. In the future, the estimated model can be used to predict a CFR of COVID-19 which is influenced by the parametric component such as CFR in the previous period and the nonparametric component such as CGR in the previous period. Furthermore, the obtained model estimate can be used to predict CFR in Indonesia for supporting one of the SDGs, control the pandemic.

### CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

## REFERENCES

[1] K.O. Kwok, V.W.Y. Wong, W.I. Wei, et al. Epidemiological characteristics of the first 53 laboratory-confirmed cases of COVID-19 epidemic in Hong Kong, 13 February 2020, Eurosurveillance. 25 (2020), 2000155. https://doi.org/10.2807/1560-7917.es.2020.25.16.2000155.

[2] WHO, Novel coronavirus (2019-nCoV) situation report-12 (2020), World Health Organization, Geneva, 2020.

[3] I. Chakraborty, P. Maity, COVID-19 outbreak: Migration, effects on society, global environment and prevention, Sci. Total Environ. 728 (2020), 138882. https://doi.org/10.1016/j.scitotenv.2020.138882.

[4] A. Susilo, C.M. Rumende, C.W. Pitoyo, et al. Coronavirus disease 2019: Tinjauan literatur terkini, J. Ilmu Penyakit Dalam. 7 (2020) 45–76. https://doi.org/10.7454/jpdi.v7i1.415.

[5] T. Sipahutar, T. Eryando, COVID-19 case fatality rate and detection ability in Indonesia, Nat. Public Health J. Spec. Iss. 1 (2020), 14–17. https://doi.org/10.21109/kesmas.v15i2.3936.

[6] B. Lestari, Fatmawati, I. N. Budiantara, Spline Estimator and Its Asymptotic Properties in Multiresponse Nonparametric Regression Model, Songklanakarin J. Sci. Technol. 42(3) (2020) 533–548.

[7] B. Lestari, N. Chamidah, D. Aydin, et al. Reproducing kernel Hilbert space approach to multiresponse smoothing spline regression function, Symmetry. 14 (2022), 2227. https://doi.org/10.3390/sym14112227.

[8] R.L. Eubank, Nonparametric regression and spline smoothing, 2nd ed., CRC Press, Boca Raton, 1999.

[9] H.F.F. Mahmoud, Parametric versus semi and nonparametric regression models, Int. J. Stat. Prob. 10 (2021), 90-109. https://doi.org/10.5539/ijsp.v10n2p90.

[10] N. Chamidah, B. Lestari, I.N. Budiantara, et al. Consistency and asymptotic normality of estimator for parameters in multiresponse multipredictor semiparametric regression model, Symmetry. 14 (2022), 336. https://doi.org/10.3390/sym14020336.

[11] B. Lestari, N. Chamidah, I.N. Budiantara, et al. Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator, J. King Saud Univ. - Sci. 35 (2023), 102664. https://doi.org/10.1016/j.jksus.2023.102664.

[12] A.A. Suleiman, A. Suleiman, U.A. Abdullahi, et al. Estimation of the case fatality rate of COVID-19 epidemiological data in Nigeria using statistical regression analysis, Biosafe. Health. 3 (2021), 4–7. https://doi.org/10.1016/j.bsheal.2020.09.003.

[13] J. Pan, J.M. St. Pierre, T.A. Pickering, et al. Coronavirus disease 2019 (COVID-19): a modeling study of factors driving variation in case fatality rate by country, Int. J. Environ. Res. Public Health. 17 (2020), 8189. https://doi.org/10.3390/ijerph17218189.

[14] Nidhomuddin, N. Chamidah, A. Kurniawan, Confidence interval of the parameter on multipredictor biresponse longitudinal data analysis using local linear estimator for modeling of case increase and case fatality rates COVID-19 in Indonesia: A theoretical discussion, Commun. Math. Biol. Neurosci. 2022 (2022), 23. https://doi.org/10.28919/cmbn/6900.

[15] N.R.A.A. Siregar, Farida, S. Falasifah, et al. Pemodelan harga minyak mentah dunia berdsasarkan efek pandemi COVID-19 dengan estimator penalized spline, MUST: J. Math. Edu., Sci. Technol. 7 (2022), 152–166.

[16] T.W. Utami, A. Lahdji, Modeling of local polynomial kernel nonparametric regression for COVID daily cases in Semarang City, Indonesia, Media Stat. 14 (2022), 206–215. https://doi.org/10.14710/medstat.14.2.206-215.

[17] S. Saidi, N. Herawati, K. Nisa, et al. Nonparametric modeling using kernel method for estimation of the COVID-19 data in Indonesia during 2020, Int. J. Math. Trends Technol. 67 (2021), 136–144. http://repository.lppm.unila.ac.id/id/eprint/34008.

[18] N. Chamidah, B. Zaman, L. Muniroh, et al. Designing local standard growth charts of children in East Java Province using a local linear estimator, Int. J. Innov., Creat. Change. 13 (2020), 45–67.

[19] N. Chamidah, Y.S. Yonani, E. Ana, et al. Identification the number of mycobacterium tuberculosis based on sputum image using local linear estimator, Bull. Electric. Eng. Inform. 9 (2020), 2109–2116. https://doi.org/10.11591/eei.v9i5.2021.

[20] N. Chamidah, E. Tjahjono, A.R. Fadilah, et al. Standard growth charts for weight of children in East Java using local linear estimator, J. Phys.: Conf. Ser. 1097 (2018), 012092. https://doi.org/10.1088/1742-6596/1097/1/012092.

[21] E. Ana, N. Chamidah, P. Andriani, et al. Modeling of hypertension risk factors using local linear of additive nonparametric logistic regression, J. Phys.: Conf. Ser. 1397 (2019), 012067. https://doi.org/10.1088/1742-6596/1397/1/012067.

[22] N. Chamidah, K.H. Gusti, E. Tjahjono, et al. Improving of classification accuracy of cyst and tumor using local polynomial estimator, TELKOMNIKA. 17 (2019) 1492-1500. https://doi.org/10.12928/telkomnika.v17i3.12240.

[23] N. Chamidah, B. Lestari, Estimation of covariance matrix using multi-response local polynomial estimator for designing children growth charts: A theoretically discussion, J. Phys.: Conf. Ser. 1397 (2019), 012072. https://doi.org/10.1088/1742-6596/1397/1/012072.

[24] A. Delaigle, J. Fan, R.J. Carroll, A design-adaptive local polynomial estimator for the errors-in-variables problem, J. Amer. Stat. Assoc. 104 (2009), 348–359. https://doi.org/10.1198/jasa.2009.0114.

[25] M. Francisco-Fernández, J.M. Vilar-Fernández, Local polynomial regression estimation with correlated errors, Commun. Stat. - Theory Methods. 30 (2001), 1271–1293. https://doi.org/10.1081/sta-100104745.

[26] K. Benhenni, D. Degras, Local polynomial estimation of the mean function and its derivatives based on functional data and regular designs, ESAIM Prob. Stat. 18 (2014), 881–809. https://doi.org/10.1051/ps/2014009.

[27] C.B. Kikechi, On local polynomial regression estimators in finite populations, Int. J. Stats. Appl. Math. 5 (2020), 58–63.

[28] B. Lestari, Fatmawati, I.N. Budiantara, et al. Estimation of regression function in multi-response nonparametric regression model using smoothing spline and kernel estimators, J. Phys.: Conf. Ser. 1097 (2018), 012091. https://doi.org/10.1088/1742-6596/1097/1/012091.

[29] B. Lestari, N. Chamidah, T. Saifudin, Estimasi fungsi regresi dalam model regresi nonparametrik birespon menggunakan estimator smoothing spline dan estimator kernel, J. Mat. Stat. Komput. 15 (2019), 20–24. https://doi.org/10.20956/jmsk.v15i2.5710.

[30] N. Chamidah, B. Lestari, T. Saifudin, Modeling of blood pressures based on stress score using least square spline estimator in bi-response non-parametric regression, Int. J. Innov., Creat. Change. 5 (2019), 1200–1216.

[31] Fatmawati, I.N. Budiantara, B. Lestari, Comparison of smoothing and truncated spline estimators in estimating blood pressure models, Int. J. Innov., Creat. Change. 5 (2019), 1177–1199.

[32] N. Chamidah, B.Zaman, L. Muniroh, et al. Estimation of median growth charts for height of children in East Java Province of Indonesia using penalized spline estimator, in: Proceeding of Global Conference on Engineering and Applied Science Sapporo, Hokkaido, Japan, (GCEAS, 2019), 68-78.

[33] W. Ramadan, N. Chamidah, B. Zaman, et al. Standard growth chart of weight for height to determine wasting nutritional status in East Java based on semiparametric least square spline estimator, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052063. https://doi.org/10.1088/1757-899x/546/5/052063.

[34] B. Lestari, Fatmawati, I.N. Budiantara, et al. Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches, J. Phys.: Conf. Ser. 1397 (2019), 012064. https://doi.org/10.1088/1742-6596/1397/1/012064.

[35] N. Chamidah, B. Lestari, A. Massaid, et al. Estimating mean arterial pressure affected by stress scores using spline nonparametric regression model approach, Commun. Math. Biol. Neurosci. 2020 (2020), 72. https://doi.org/10.28919/cmbn/4963.

[36] N. Chamidah, B. Lestari, A.Y. Wulandari, et al. Z-score standard growth chart design of toddler weight using least square spline semiparametric regression, AIP Conf. Proc. 2329 (2021), 060031. https://doi.org/10.1063/5.0042285.

[37] D. Aydın, E. Yılmaz, N. Chamidah, et al. Right-censored partially linear regression model with error in variables: application with carotid endarterectomy dataset, Int. J. Biostat. (2023). https://doi.org/10.1515/ijb-2022-0044.

[38] H. Liang, J.D.Z. Chen, Assessment of the esophageal pressure in gastroesophageal reflux disease by the local regression, Ann. Biomed. Eng. 33 (2005), 847–853. https://doi.org/10.1007/s10439-005-2866-8.

[39] C. Loader, Local regression and likelihood, Springer, New York, (1999).

[40] D. Gujarati, Basic econometrics, Fourth Ed., McGraw-Hill, New York, (2003).

[41] S.L. Zeger, R. Irizarry, R.D. Peng, On time series analysis of public health and biomedical data, Annu. Rev. Public Health. 27 (2006), 57–79. https://doi.org/10.1146/annurev.publhealth.26.021304.144517.

[42] A. Meimela, S.S.S. Lestari, I.F. Mahdy, et al. Modeling of covid-19 in Indonesia using vector autoregressive integrated moving average, J. Phys.: Conf. Ser. 1722 (2021), 012079. https://doi.org/10.1088/1742-6596/1722/1/012079.

[43] R. Somyanonthanakul, K. Warin, W. Amasiri, et al. Forecasting COVID-19 cases using time series modeling and association rule mining, BMC. Med. Res. Methodol. 22 (2022), 281. https://doi.org/10.1186/s12874-022-01755-x.

[44] J. Gao, I. Gijbels, Bandwidth selection in nonparametric kernel testing, J. Amer. Stat. Assoc. 103 (2008), 1584–1594. https://doi.org/10.1198/016214508000000968.

[45] S.X. Chen, J. Gao, C.Y. Tang, A test for model specification of diffusion processes, Ann. Stat. 36 (2008), 167–198. https://doi.org/10.1214/009053607000000659.

[46] J. Gao, M. King, Adaptive testing in continuous-time diffusion models, Econ. Theory. 20 (2004), 844–882. https://doi.org/10.1017/s0266466604205023.

[47] J. Vilar-Fernández, R. Cao, Nonparametric forecasting in time series - a comparative study, Commun. Stat. - Simul. Comput. 36 (2007), 311–334. https://doi.org/10.1080/03610910601158377.

[48] Q. Wang, P.C.B. Phillips, Asymptotic theory for zero energy functionals with nonparametric regression applications, Econom. Theory. 27 (2010), 235–259. https://doi.org/10.1017/s0266466610000277.

[49] D. Li, J. Chen, J. Gao, Non-parametric time-varying coefficient panel data models with fixed effects, Econom. J. 14 (2011), 387-408. https://doi.org/10.1111/j.1368-423x.2011.00350.x.

[50] C. Dong, J. Gao, Specification testing driven by orthogonal series for nonlinear cointegration with endogeneity, Econom. Theory. 34 (2017), 754–789. https://doi.org/10.1017/s0266466617000238.

[51] J. Gao, K. Hawthorne, Semiparametric estimation and testing of the trend of temperature series, Econom. J. 9 (2006), 332–355. https://doi.org/10.1111/j.1368-423x.2006.00188.x.

[52] G. Aneiros-Pérez, P. Vieu, Nonparametric time series prediction: A semi-functional partial linear modeling, J. Multivar. Anal. 99 (2008), 834–857. https://doi.org/10.1016/j.jmva.2007.04.010.

[53] J. Gao, Identification, estimation and specification in a class of semi-linear time series models, SSRN. (2012). https://doi.org/10.2139/ssrn.2027271.

[54] P.J. Green, B.W. Silverman, Nonparametric regression and generalized linear models, Chapman &Hall/CRC, Boca Raton, (1994).

[55] V. Fibriyani, N. Chamidah, Prediction of inflation in indonesia using nonparametric regression approach based on local polynomial estimator, in: Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019), Indonesia, 2020. https://doi.org/10.2991/assehr.k.201010.013.

[56] V. Fibriyani, N. Chamidah, Prediction of inflation rate in indonesia using local polynomial estimator for time series data, J. Phys.: Conf. Ser. 1776 (2021), 012065. https://doi.org/10.1088/1742-6596/1776/1/012065.

[57] J.J. Montaño Moreno, A. Palmer Pol, A. Sesé Abad, Using the R-MAPE index as a resistant measure of forecast accuracy, Psicothema. (2013), 500–506. https://doi.org/10.7334/psicothema2013.23.

[58] Kementerian Keuangan Indonesia, Pengaruh COVID-19 atas kondisi sosial ekonomi di Indonesia. https://pen.kemenkeu.go.id/in/page/pengaruhcovid.