HOSTED BY

ELSEVIER

Full Length Article

Estimating time series semiparametric regression model using local polynomial estimator for predicting inflation rate in Indonesia

Vita Fibriyani [a,b], Nur Chamidah [c,d,*], Toha Saifudin [c,d]

[a] Doctoral Study Program of Mathematics and Natural Sciences, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia
[b] Department of Management, Faculty of Economics, Universitas Merdeka Pasuruan, Pasuruan 67129, Indonesia
[c] Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia
[d] Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia

ABSTRACT

A model built from a parametric regression model and a nonparametric regression model is called a semi-parametric regression (SR) model. The main problem in the SR model is the estimation of the regression function. In this study, we develop the SR model for time series data that is called Time Series Semiparametric Regression (TSSR) model, and discuss estimation of the TSSR model by using local polynomial. Also, we apply it to data of inflation rate (IR) in Indonesia where IR is as response variable, and both IR and money supply in the previous periods are as predictor variables. Next, we compare the results of estimating the IR using the TSSR with the classical method, namely the ARIMA. Also, the TSSR has high accurate criterion for predicting the IR in Indonesia. The results of this study are useful for analyzing Indonesia's economic growth rate, which is one of the Sustainable Development Goals (SDGs).

1. Introduction

Economic growth (EG) is an increase about national income or production in a country from year to year. To measure the EG of a country, we can look at the level of gross domestic product (GDP) of the country. In macroeconomic, one of references used to measure a country's economy is inflation that is an economic condition where prices have tendency to rise continuously, or a decrease of money supply in the overall value. (Simanungkalit, 2020). In fundamental economics concept, inflation depends on the circulating quantity of excess money in an economy (Esumanba et al., 2019). In general, inflation is a depreciation of currency or the increasing price of a basic need for instances services and goods. The continuously increasing IR can become an EG obstacle. If a high inflation happened, it would decrease prosperity level and reduce EG, like in Indonesia in 1998. It was caused by the uncontrolled printing of money. Hence, in circulation, the amount of money became huge. Therefore, it is important to keep inflation is to be stable for sustaining EG. Stable condition of inflation is a mandatory prerequisite for sustaining EG, thereby increasing society welfare (Atmaja, 1999; Yu et al., 2023). The IR has a high variance, so there is a fluctuating movement. Hence, an appropriate modeling is required. As one of monetary agencies, Bank of Indonesia controls IR in Indonesia by

evaluating whether the target still sets projections of inflation in the future. Saluza (2015) predicted IR in Indonesia based on some information in the past. IR prediction is required for the government and business to make strategic policies in economic field.

As one of statistical models, regression model, can be used for predicting IR. It has two types, namely, parametric regression (PR) model (Yolanda, 2017; Esumanba et al., 2019), and nonparametric regression (NR) model (Aneiros-Perez & Vieu, 2008; Li et al., 2011; Lestari et al., 2018, 2020, 2022; Clark et al., 2022; Chamidah et al., 2020a, 2022, 2024a, 2024b; Aydin et al., 2024). Yolanda (2017) used Time Series Parametric Regression (TSPR) model for analyzing factors that affect inflation and their impact on development index of human and poverty in Indonesia. Next, Nyoni (2019) used TSPR model, namely, ARIMA for predicting inflation in Burundi. Further, Suparti et al. (2018) used Time Series Nonparametric Regression (TSNR) model to model inflation rate in Indonesia. Prahutama et al. (2018) used TSNR for modelling foods sector inflation in Indonesia. According to Chamidah et al. (2012, 2019a); Chamidah & Lestari (2019); and Fibriyani & Chamidah (2021), local polynomial is a very popular approach that can be used to explore localness properties of data. Local polynomial estimator has two special cases. If it has degree of polynomial, $d = 0$, then it will result a kernel estimator or constant local estimator, and if $d = 1$, then it will result a

local linear estimator (Chamidah et al., 2020b, 2020c). Furthermore, if we combine these two regression types, we will have a Semiparametric Regression (SR) model (Gao & Hawthorne, 2006; Ramadan et al., 2019; Chamidah et al., 2019b, 2022; Lestari et al., 2023; Aydin et al., 2023; Fibriyani et al., 2024). The SR model can be applied to a time series dataset which is in this study we call it a TSSR model. Time series data is a set of observations taken at different times and periodically collected at a certain time interval. Time series data include one object of research or individual research, for instances, stock prices, currency rates, or inflation rates in several periods such as daily, weekly, monthly, or yearly. Since inflation rate data have high fluctuation pattern, then the use of local polynomial estimator for analyzing this data is very appropriate.

Furthermore, the SR model has more complicated calculation than PR model or NR model. There is no researcher presented above that concerned on estimating TSSR model by using local polynomial estimator. Therefore, this study intends to discuss a new estimation method of the TSSR model by using local polynomial estimator. Next, we apply it to data of IR in Indonesia where IR is as response variable, and both IR and money supply in the previous periods are as predictor variables. Then, we compare the results of estimating the IR using the TSSR with the classical method, namely, ARIMA.

2. Materials and methods

In the following we briefly provide materials and methods used in this study such as TSSR model, Generalized Cross Validation, and Mean Absolute Percentage Error.

2.1. TSSR model

The TSSR model is a SR model applied to time series data. It can be expressed as follows:

$y_t = \theta_1 y_{t-1} + g(v_{t-1}) + \varepsilon_t, \ t = 1, 2, \cdots, T$ (1) where y_t is a t-th response variable, θy_{t-1} is a parametric component, θ is an unknown parameter of parametric component, $g(v_{t-1})$ is a nonparametric component that is a smooth function contained in Sobolev space, and ε_t is a t-th measurement error with mean of zero and variance of σ^2.

To estimate $g(v_{t-1})$ by using local polynomial estimator, we employ a Taylor series expansion as follows:

$$g(v_{t-1}) \approx \sum_{d=0}^{r} \frac{g^{(d)}(v_0)}{d!}(v_{t-1} - v_0)^d = \sum_{d=0}^{r} \beta_d(v_0)(v_{t-1} - v_0)^d \quad (2)$$

where $g^{(d)}(\boldsymbol{v}_0)$ is the d-th derivative of $\boldsymbol{g}(\boldsymbol{v}_0)$ at $\boldsymbol{v}_0$, $v_{t-1} \in (v_0 - h, v_0 + h)$, and $\frac{g^{(d)}(v_0)}{d!} = \boldsymbol{\beta_d}(\boldsymbol{v}_0)$. Hence, we can express equation (2) as follows:

$$\boldsymbol{g}(v_{t-1}) = \mathbf{V}_{v_0}\beta(v_0) \quad (3)$$

where $\mathbf{V}_{\boldsymbol{v}_0} = \begin{bmatrix} 1 & (\boldsymbol{v}_{t-1} - \boldsymbol{v}_0) & \cdots & (\boldsymbol{v}_{t-1} - \boldsymbol{v}_0)^r \end{bmatrix}$, $\boldsymbol{v}_{t-1} \in (v_0 - h, v_0 + h)$, and $\beta(\boldsymbol{v}_0) = [\boldsymbol{\beta_0}(\boldsymbol{v}_0) \quad \boldsymbol{\beta_1}(\boldsymbol{v}_0) \quad \cdots \quad \boldsymbol{\beta_r}(\boldsymbol{v}_0)]^T$. So, the TSSR model can be presented as follows:

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

By giving paired dataset $\{\boldsymbol{v}_t, \boldsymbol{y}_t\}_{t=1}^T$, we can elaborate model (4) as follows:

$$\left.\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{\beta_0}(\boldsymbol{v}_0) + \boldsymbol{\beta_1}(\boldsymbol{v}_0)(\boldsymbol{v}_1 - \boldsymbol{v}_0) + \boldsymbol{\beta_2}(\boldsymbol{v}_0)(\boldsymbol{v}_1 - \boldsymbol{v}_0)^2 + \cdots + \boldsymbol{\beta_r}(\boldsymbol{v}_0)(\boldsymbol{v}_1 - \boldsymbol{v}_0)^r + \boldsymbol{\varepsilon}_1 \\ \boldsymbol{y}_2 &= \boldsymbol{\beta_0}(\boldsymbol{v}_0) + \boldsymbol{\beta_1}(\boldsymbol{v}_0)(\boldsymbol{v}_2 - \boldsymbol{v}_0) + \boldsymbol{\beta_2}(\boldsymbol{v}_0)(\boldsymbol{v}_2 - \boldsymbol{v}_0)^2 + \cdots + \boldsymbol{\beta_r}(\boldsymbol{v}_0)(\boldsymbol{v}_2 - \boldsymbol{v}_0)^r + \boldsymbol{\varepsilon}_2 \\ &\vdots \\ \boldsymbol{y}_T &= \boldsymbol{\beta_0}(\boldsymbol{v}_0) + \boldsymbol{\beta_1}(\boldsymbol{v}_0)(\boldsymbol{v}_T - \boldsymbol{v}_0) + \boldsymbol{\beta_2}(\boldsymbol{v}_0)(\boldsymbol{v}_T - \boldsymbol{v}_0)^2 + \cdots + \boldsymbol{\beta_r}(\boldsymbol{v}_0)(\boldsymbol{v}_T - \boldsymbol{v}_0)^r + \boldsymbol{\varepsilon}_T \end{aligned}\right\} \quad (5)$$

Hence, it gives the following TSSR model:

$$\mathbf{y} = \mathbf{V}_{v_0}\beta(v_0) + \varepsilon \quad (6)$$

where $\mathbf{V}_{\boldsymbol{v}_0} = \begin{bmatrix} 1 & (\boldsymbol{v}_1 - \boldsymbol{v}_0) & \cdots & (\boldsymbol{v}_1 - \boldsymbol{v}_0)^r \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (\boldsymbol{v}_T - \boldsymbol{v}_0) & \cdots & (\boldsymbol{v}_T - \boldsymbol{v}_0)^r \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_T \end{bmatrix}$, and $\varepsilon = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_T \end{bmatrix}$.

To estimate $\beta(\boldsymbol{v}_0)$ in (6), we carry out a WLS (weighted least square) optimization:

$$\boldsymbol{MinQ}(\boldsymbol{v}_0) = Min\left(\left(\mathbf{y} - \mathbf{V}_{v_0}\beta(v_0) \right)^{\mathrm{T}} \mathbf{K}_h(v_0) \left(\mathbf{y} - \mathbf{V}_{v_0}\beta(v_0) \right) \right) \quad (7)$$

where $\mathbf{K}_{\boldsymbol{h}}(.)$ represents a Kernel function with bandwidth $\boldsymbol{h}$ given as follows:

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

Hence, $\mathbf{K}_h(v_0) = diag(\mathrm{K}_h(v_1 - v_0) \quad \mathrm{K}_h(v_2 - v_0) \quad \cdots \quad \mathrm{K}_h(v_T - v_0))$.

Here, we use kernel function of Gaussian given as follows:

$$K(v) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2}{2}\right) \quad (8)$$

We take partially derivative of equation (7) with respect to $\beta(\boldsymbol{v}_0)$ and then set it to zero:

$$\frac{\partial \boldsymbol{Q}(\boldsymbol{v}_0)}{\partial \beta(v_0)} = -2\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{y} + 2\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{V}_{v_0}\beta(v_0) = 0$$

Hence, we obtain $\widehat{\beta}(v_0) = \left(\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{y}$. (9)

Finally, we substitute equation (9) into equation (3) and we obtain:

$$\widehat{\boldsymbol{g}}(v_{t-1}) = \mathbf{v}_{v_0}\left(\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{y} \quad (10)$$

2.2. Generalized Cross Validation (GCV)

In estimating TSSR model, we use a criterion of GCV by selecting a minimum value of GCV to obtain an optimal bandwidth (h) value that controls smoothness and goodness of fit to be balance which will affect bias-variance trade-off. Therefore, we obtain:

$$\widehat{\boldsymbol{g}}(v_{t-1}) = \mathbf{A}_{(h_j)}\mathbf{y}$$

where $\mathbf{A}_{(h_j)} = \mathbf{v}_{v_0}\left(\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0)$. The optimal value of h_j can be obtained using GCV criterion as follows (Chamidah et al., 2012):

$$GCV(h) = \frac{T^{-1}\sum_{t=1}^{T}\left(y_t^{(j)} - \widehat{y}_t^{(j)}\right)^2}{\left(T^{-1}tr\left[\mathbf{I} - \mathbf{A}_{(h_j)}\right]\right)^2} \quad (11)$$

2.3. Mean Absolute Percentage error (MAPE)

We often use a MAPE value to determine the forecasting accuracy measurement of the estimated model. MAPE obeys the following mathematical expression (Moreno et al., 2013):

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}\frac{|\widehat{y}_t - y_t|}{y_t} \times 100 \quad (12)$$

where T represents sample size, $\hat{y}_t$ represents the estimated model value at time point t, and y_t represents the observed value of y at time point t. The MAPE values criteria can be found in Table 1.

3. Results

The results of this study consist of theoretical results on estimating TSSR model, and application of TSSR model.

3.1. Estimating TSSR model

Suppose given a paired dataset (y_{t-1}, v_{t-1}, y_t), $t = 1, 2, 3, \cdots, T$ where y_{t-1} represents autoregressive of response variable, v_{t-1} represents autoregressive of predictor variable. The dataset follow the TSSR model:

$$y_t = \theta_1 y_{t-1} + g(v_{t-1}) + \varepsilon_t \tag{13}$$

where y_t is response variable, $\theta_1 y_{t-1}$ is parametric component, $g(v_{t-1})$ is nonparametric component, ε_t is measurement error which has mean of zero and variance of σ_t^2.

First, we estimate the NR model by assuming θ_1 is known. So, we have:

$$y_t^* = g(v_{t-1}) + \varepsilon_t \tag{14}$$

where $y_t^* = y_t - \theta_1 y_{t-1}$. Next, based on local polynomial estimator, we estimate $g(v_{t-1})$ which is a smooth function in the sense of continuous and differentiable function. We can approximate it by using Taylor series expansion at v_{t-1} around v_0 as follows:

$$g(v_{t-1}) \approx \sum_{d=0}^{r} \frac{g^{(d)}(v_0)}{d!}(v_{t-1} - v_0)^d \quad d = 0, 1, 2, \cdots, r \tag{15}$$

$$g(v_{t-1}) = \sum_{d=0}^{r} \beta_d(v_0)(v_{t-1} - v_0) \tag{16}$$

where $g^{(d)}(v_0)$ is the d-th derivative of $g(v_0)$ and $v_{t-1} \in (v_0 - h, v_0 + h)$. We may express model (14) as follows:

$$g(v_{t-1}) = \mathbf{V}_{v_0}\beta(v_0) \tag{17}$$

where $\mathbf{V}_{v_0} = \begin{bmatrix} 1 & (v_{t-1} - v_0) & (v_{t-1} - v_0)^2 & \cdots & (v_{t-1} - v_0)^r \end{bmatrix}$, $v_{t-1} \in (v_0 - h, v_0 + h)$ and $\beta(v_0) = \begin{bmatrix} \beta_0(v_0) & \beta_1(v_0) & \cdots & \beta_r(v_0) \end{bmatrix}^T$.

Based on equation (17), we obtain:

$$y_t^* = \mathbf{V}_{v_0}\beta(v_0) + \varepsilon_t \tag{18}$$

For a paired sample $\{v_{t-1}, y_t\}_{t=2}^T$, model (18) gives:

Table 1
Criteria of MAPE Values.

MAPE	Criteria
< 10	Highly Accurate
$10 - 20$	Accurate
$20 - 50$	Reasonable
> 50	Inaccurate

where $\mathbf{y}^* = \begin{bmatrix} y_2^* \\ y_3^* \\ y_4^* \\ \vdots \\ y_T^* \end{bmatrix}$;

$$\mathbf{V}_{v_0} = \begin{bmatrix} 1 & (v_1 - v_0) & (v_1 - v_0)^2 & \cdots & (v_1 - v_0)^r \\ 1 & (v_2 - v_0) & (v_2 - v_0)^2 & \cdots & (v_2 - v_0)^r \\ 1 & (v_3 - v_0) & (v_3 - v_0)^2 & \cdots & (v_3 - v_0)^r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (v_{T-1} - v_0) & (v_{T-1} - v_0)^2 & \cdots & (v_{T-1} - v_0)^r \end{bmatrix};$$

$\beta(v_0) = \begin{bmatrix} \beta_0(v_0) & \beta_1(v_0) & \cdots & \beta_r(v_0) \end{bmatrix}^T$; and $\varepsilon = \begin{bmatrix} \varepsilon_2 & \varepsilon_3 & \varepsilon_4 & \cdots \varepsilon_T \end{bmatrix}^T$.

The local polynomial estimator of $\beta(v_0)$ in (20) can be obtained by taking value which minimizes the following weighted least square function:

$$\mathbf{Q}(v_0) = (\mathbf{y}^* - \mathbf{V}_{v_0}\beta(v_0))^T \mathbf{K}_h(v_0)(\mathbf{y}^* - \mathbf{V}_{v_0}\beta(v_0)) \tag{21}$$

where

$$\mathbf{K}_h(v_0) = \begin{bmatrix} K_h(v_1 - v_0) & 0 & 0 & \cdots & 0 \\ 0 & K_h(v_2 - v_0) & 0 & \cdots & 0 \\ 0 & 0 & K_h(v_3 - v_0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & K_h(v_{T-1} - v_0) \end{bmatrix},$$

and $K_h(.)$ is kernel function. The local polynomial estimator of $\beta(v_0)$ is $\hat{\beta}(v_0)$ which will minimize the $\mathbf{Q}(v_0)$ in (21). The minimum value of $\mathbf{Q}(v_0)$ is reached when $\frac{\partial \mathbf{Q}(v_0)}{\partial \beta(v_0)} = 0$. Hence, we have:

$$\frac{\partial \mathbf{Q}(v_0)}{\partial \beta(v_0)} = 0 \Leftrightarrow \frac{\partial \left((\mathbf{y}^* - \mathbf{V}_{v_0}\beta(v_0))^T \mathbf{K}_h(v_0)(\mathbf{y}^* - \mathbf{V}_{v_0}\beta(v_0))\right)}{\partial \beta(v_0)}\Big|_{\beta(v_0) = \hat{\beta}(v_0)} = 0$$

$$\Leftrightarrow \hat{\beta}(v_0) = \left(\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0) \mathbf{V}_{v_0}\right)^{-1} \mathbf{V}_{v_0}^T \mathbf{K}_h(v_0) \mathbf{y}^*. \tag{22}$$

Based on equations (17) and (22), we obtain:

$$\begin{rcases} y_2^* = \beta_0(v_0) + \beta_1(v_0)(v_1 - v_0) + \beta_2(v_0)(v_1 - v_0)^2 + \cdots + \beta_r(v_0)(v_1 - v_0)^r + \varepsilon_2 \\ y_3^* = \beta_0(v_0) + \beta_1(v_0)(v_2 - v_0) + \beta_2(v_0)(v_2 - v_0)^2 + \cdots + \beta_r(v_0)(v_2 - v_0)^r + \varepsilon_3 \\ y_4^* = \beta_0(v_0) + \beta_1(v_0)(v_3 - v_0) + \beta_2(v_0)(v_3 - v_0)^2 + \cdots + \beta_r(v_0)(v_3 - v_0)^r + \varepsilon_4 \\ \vdots \\ y_T^* = \beta_0(v_0) + \beta_1(v_0)(v_{T-1} - v_0) + \beta_2(v_0)(v_{T-1} - v_0)^2 + \cdots + \beta_r(v_0)(v_{T-1} - v_0)^r + \varepsilon_T \end{rcases} \tag{19}$$

Hence, the models (19) gives expressed model:

$$\mathbf{y}^* = \mathbf{V}_{v_0}\beta(v_0) + \varepsilon \tag{20}$$

$$\hat{g}(v_{t-1}) = \mathbf{V}_{v_0}\left(\mathbf{V}_{v_0}^T \mathbf{K}_h(v_0) \mathbf{V}_{v_0}\right)^{-1} \mathbf{V}_{v_0}^T \mathbf{K}_h(v_0) \mathbf{y}^* \tag{23}$$

Table 2

Analysis Results Using ARIMA.

Model	Parameter	p-Value	White-Noise	Normal	MAPE		
					In-Samples	Out-Samples	All Samples
SARIMA $(1,1,0)(1,1,1)^{12}$	AR 1	0.042	Yes	No			
	SAR 12	0.000					
	SMA 12	0.000					
SARIMA $(0,1,1)(1,1,1)^{12}$	SAR 12	0.000	Yes	No			
	MA 1	0.004					
	SMA 12	0.000					
SARIMA $(0,1,1)(0,1,1)^{12}$	MA 1	0.000	Yes	No	8.616	10.693	9.655
	SMA 12	0.000					

Model	**Residual Plot**
SARIMA $(1,1,0)(1,1,1)^{12}$	
SARIMA $(0,1,1)(1,1,1)^{12}$	
SARIMA $(0,1,1)(0,1,1)^{12}$	
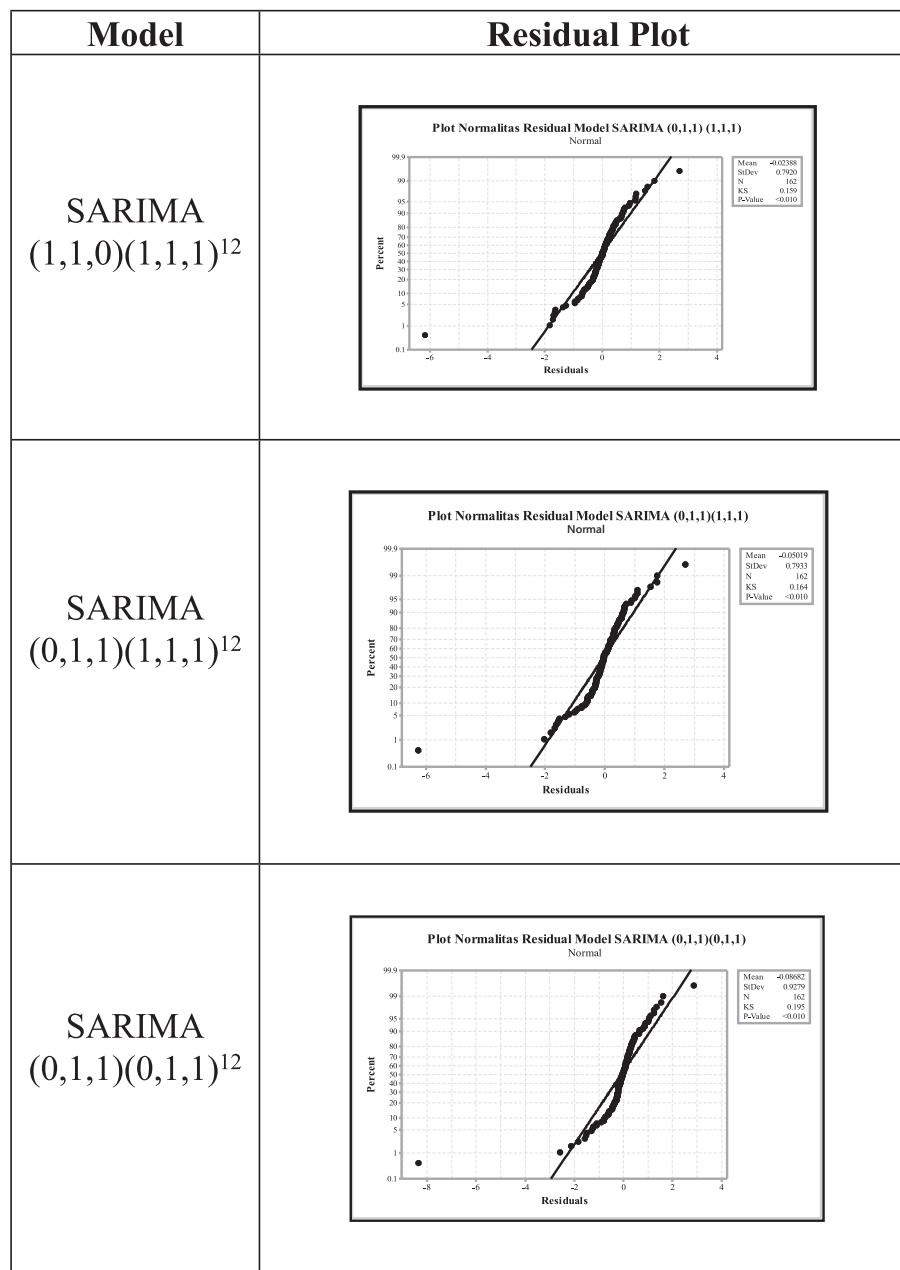
Fig. 1. Results of Kolmogorov-Smirnov Testing.

Table 3
Statistical description.

Variable	Statistic			
	Mean	Variance	Min.	Max.
Inflation Rate	5.636	12.304	1.320	18.380
Money Supply	3.785	4.212	0.995	7.963
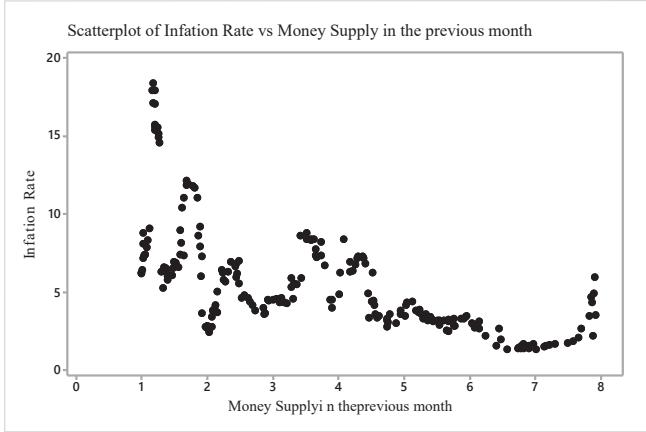
**Fig. 2.** Scatter Plot of IR versus MS in Previous Months.
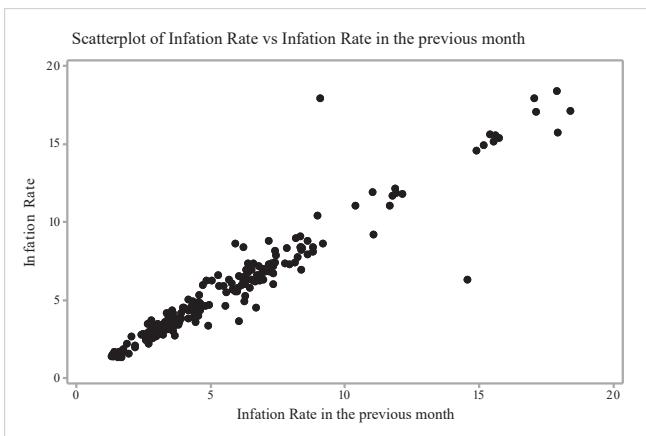
Fig. 3. Scatter Plot of IR versus IR in Previous Months.

By substituting equation (23) into equation (20), we obtain:

$$\widehat{\mathbf{y}}^* = \mathbf{V}_{v_0}\left(\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{y}^* = \mathbf{A}_{(v_0)}\mathbf{y}^* \tag{24}$$

$$\mathbf{A}_{(v_0)} = \mathbf{V}_{v_0}\left(\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)$$

Next, to get the estimator of θ we minimize the following sum of square of TSSR model:

Table 5
Value of Some Criteria for TSSR Model.

Criteria	Orde 1 – Local Linear Model Data		
	Insample	*Outsample*	*overall*
MAPE	0.309	0.474	0.392
R^2	96.796	96.759	96.778
AIC	22.362	−37.672	
Criteria	**Orde 2 – Local Quadratic Model Data**		
	Insample	*Outsample*	*overall*
MAPE	0.306	0.494	0.400
R^2	96.741	96.616	96.778
AIC	23.534	−36.399	
Criteria	**Orde 3 – Local Qubic Model Data**		
	Insample	*Outsample*	*overall*
MAPE	0.311	0.529	0.420
R^2	96.715	96.313	96.514
AIC	23.689	−39.955	

$$S = \varepsilon^T\varepsilon = \left(\mathbf{y}_t - \mathbf{y}_{t-i}\theta - \mathbf{A}_{(v_0)}(\mathbf{y}_t - \mathbf{y}_{t-i}\theta)\right)^T\left(\mathbf{y}_t - \mathbf{y}_{t-i}\theta - \mathbf{A}_{(v_0)}(\mathbf{y}_t - \mathbf{y}_{t-i}\theta)\right) \tag{25}$$

The estimated value of θ is $\widehat{\theta}$ that can be obtained by determining $\frac{\partial S}{\partial \theta} = 0$ which gives:

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

Finally, based on equations (14), (25), and (26), we obtain:

$$\widehat{\mathbf{y}}_t = \mathbf{y}_{t-i}\widehat{\theta} + \mathbf{A}_{(v_0)}(\mathbf{y}_t - \mathbf{y}_{t-i}\widehat{\theta}) = \mathbf{C}\mathbf{y}_t + \mathbf{A}_{(v_0)}(\mathbf{y}_t - \mathbf{C}\mathbf{y}_t)$$
$$= \left(\mathbf{C} + \mathbf{A}_{(v_0)}(\mathbf{I} - \mathbf{C})\right)\mathbf{y}_t \tag{27}$$

where

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

$$\mathbf{A}_{(v_0)} = \mathbf{V}_{v_0}\left(\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)\mathbf{V}_{v_0}\right)^{-1}\mathbf{V}^T_{v_0}\mathbf{K}_h(v_0)$$

3.2. Application

To investigate that the TSSR model is better than the classical time series model, namely, ARIMA, we provide the comparison results of IR estimates by using ARIMA and TSSR approaches.

3.2.1. ARIMA model approach

Dataset of IR in Indonesia from 2004 until 2022 used in this study, shows trend and seasonal patterns. So, the dataset is not stationer in mean and variance. It becomes stationer after doing differencing and Box-Cox transformation. Next, we apply Ljung-Box testing and Kolmogorov-Smirnov testing with significance level of $\alpha = 5\%$ to investigate whether residual satisfies white-noise and Normality assumptions, and the results are given in Table 2 and Fig. 1.

Table 4
Calculation Results Using TSSR Models Approach.

| Optimal Bandwidth	GCV Minimum Value	Degree	Model	MAPE		
				Training	Testing	Overall
0.25	1.583	1	Local Linear	9.019	8.771	8.949
0.31	1.368	2	Local Quadratic	10.273	8.299	9.286
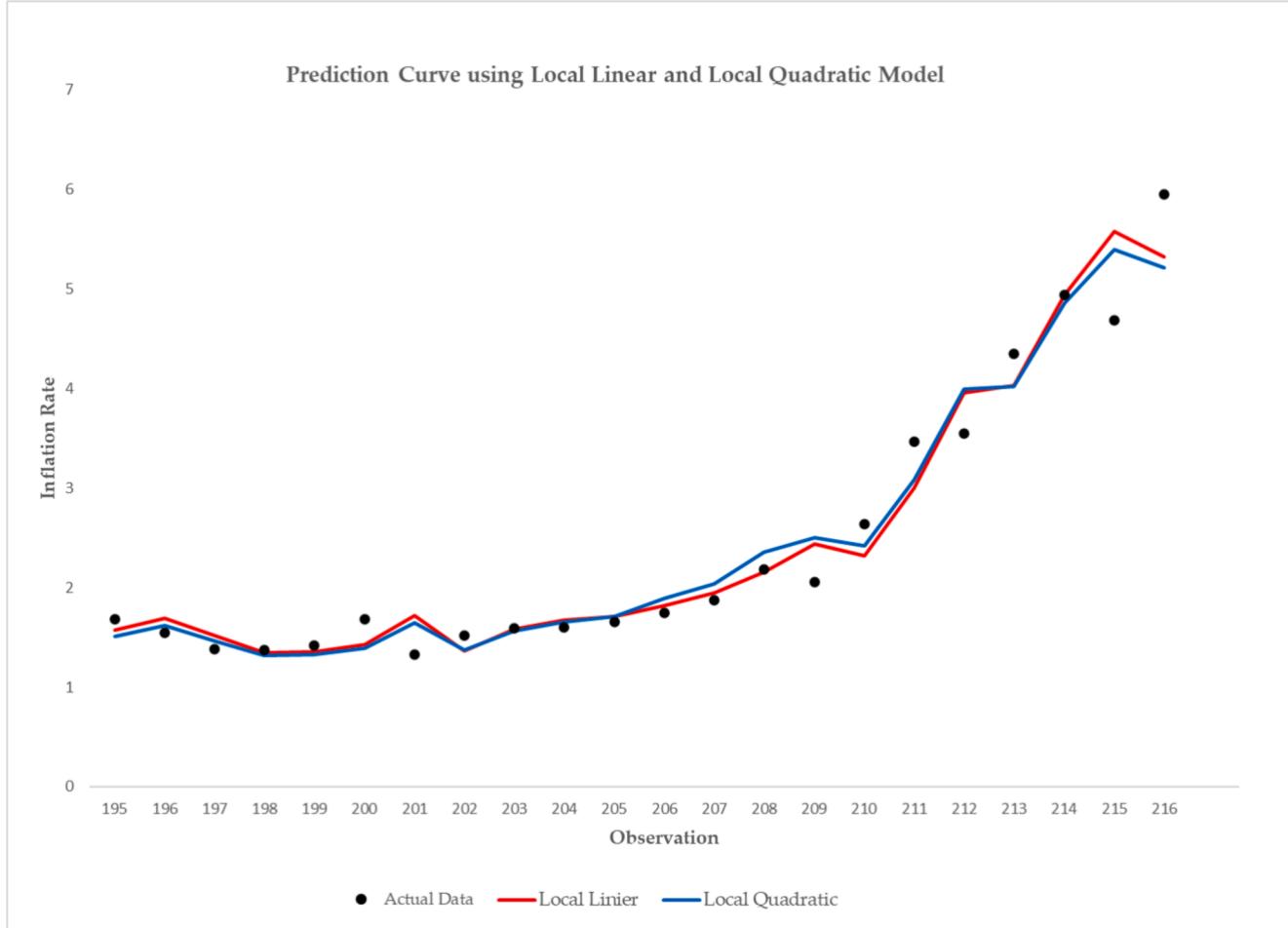
Fig. 4. Prediction Curves of Local Linear (red line) and Local Quadratic (blue line).

3.2.2. TSSR model approach

Statistical description of inflation dataset from 2004 until 2022 is given in Table 5.

It can be observed from Table 3 that from the start of the reformation era to the present, the average IR has been 5.636 %, with a comparatively large diversity of 12.304 %. In November 2005, when global oil prices and interest rates were rising and local monetary stability was under pressure, Indonesia saw its highest IR of 18.380 %. Domestic fuel prices and the value of the rupiah decreased as a result of this situation. In August 2022, Indonesia had the lowest IR as 1.320 %. The functional link between Indonesia's IR and money supply (MS) is depicted in Fig. 2. We can observe that it deviates from the traditional definition of a regression. We can observe that the relationship pattern between IR and MS in the previous month is unknown. It can be modeled by using NR approach.

Fig. 3 shows that the functional relationship between the IR and the previous period's IR is linear, so it can be modeled by using PR approach. If we combine these models in one case, then we can use SR model approach. Hence, the SR model for this case can be expressed as follows:

$$Inf_t = \theta Inf_{t-1} + \beta_0(MS_0) + \beta_1(MS_{t-1} - MS_0) + \varepsilon_t, MS_{t-1}$$
$$\in (MS_0 - h; MS_0 + h) \tag{28}$$

where *Inf* is inflation and *MS* is money supply.

In this case, these 194 observations serve as the model's input data. We create degrees of polynomial of $d = 1$ (local linear), and $d = 2$ (local quadratic). We obtained optimal bandwidths of 0.25 and 0.31 for local

linear and for local quadratic, respectively. The calculation results using TSSR models (local linear and local quadratic) are given in Table 4 and Table 5.

Table 4 and Table 5 show that overall values of MAPE are 8.949 % and 9.286 % for local linear model and for local quadratic model, respectively. This means that both local linear model and local quadratic model are the best models approach. These models have high accurate criterion because these models have MAPE values which are less than 10 % (see Table 1). Thus, the estimated TSSR model presented in (28) is very suitable for predicting IR in Indonesia. Plots of prediction values using TSSR models (local linear and local quadratic) and actual values are given in Fig. 4.

Fig. 4 shows that the plot between actual data of IR in Indonesia and its predicted values where the red line represents prediction curve of local linear approach and the blue line represents prediction curve of local quadratic approach. We can see that the predicted values of IR in Indonesia by using TSSR models based on local polynomial appear to coincide with the actual data. This means that the prediction error obtained by TSSR model approach is too small.

4. Discussion

Theoretically, the estimation result of the TSSR model by using local polynomial estimator presented in equation (27) is a combination between estimation result of parametric component and estimation result of nonparametric component.

In application, we compare the classical time series model approach, namely, ARIMA, with the TSSR model approach. Modeling of monthly

IR in Indonesia by using ARIMA results significant parameters, and it has satisfied white-noise assumption, but three best ARIMA models obtained does not satisfy assumption of normality residual (see Table 2 and Fig. 1). This means that the ARIMA model cannot accommodate factors that affect IR in Indonesia. On the other hand, modeling of monthly IR in Indonesia by using TSSR model based on local polynomial results a small value of MAPE which is less than 10 %. It means that predicted value using TSSR model is highly accurate. Thus, the TSSR model approach is better than ARIMA for prediction of IR in Indonesia.

5. Conclusion

The estimation result of the TSSR model by using estimator of local polynomial is to be combination between estimation result of parametric component and estimation result of nonparametric component of the TSSR model. The classical time series model, namely, ARIMA, cannot be used for prediction IR in Indonesia because its residual does not meet normality assumption. In case of there is no intervention or outlier, the TSSR model approach is suitable for modeling inflation rate in Indonesia. The TSSR can be used to predict inflation rate in Indonesia that is affected by inflation rate and money supply in the previous periods where in this case, the inflation rate in the previous periods is as parametric component, and money supply in the previous periods is as nonparametric component. Also, the predicted values using TSSR model based on local polynomial is highly accurate. It was shown by the MAPE value which is less than 10 %. In addition, it can be used as reference for government in determining policy. Thus, the estimated TSSR model obtained can be used to predict inflation rate in Indonesia for supporting one of the SDGs, namely, economic growth rate.

CRediT authorship contribution statement

Vita Fibriyani: Data curation, Software, Visualization. **Nur Chamidah:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Toha Saifudin:** Data curation, Investigation, Resources, Software, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jksus.2024.103549.

References

Aneiros-Perez, G., Vieu, P., 2008. Nonparametric time series prediction: a semi-functional partial linear modeling. J. Multivar. Anal. 99 (5), 834–857.

Atmaja, A.S., 1999. Inflasi di Indonesia: sumber-sumber penyebab dan pengendaliannya. J. Akuntansi Dan Keuangan 1 (1), 54–67.

Aydin, D., Yilmaz, E., Chamidah, N., Lestari, B., 2023. Right-censored partially linear regression model with error in variables: application with carotid endarterectomy dataset. Int. J. Biostatistics 2023, 1–34. https://doi.org/10.1515/ijb-2022-0044.

Aydin, D., Yilmaz, E., Chamidah, N., Lestari, B., Budiantara, I.N., 2024. Right-censored nonparametric regression with measurement error. Metrika 87 (3). https://doi.org/10.1007/s00184-024-00953-5.

Chamidah, N., Budiantara, I.N., Sunaryo, S., Zain, I., 2012. Designing of child growth chart based on multi-response local polynomial modeling. J. Math. Stats. 8 (3), 342–347.

Chamidah, N., Lestari, B., 2019. Estimating of covariance matrix using multi-response local polynomial estimator for designing children growth charts: A theoretically discussion. J. Phys.: Conf. Ser. 1397 (1), 012072.

Chamidah, N., Lestari, B., Wulandari, A.Y., Muniroh, L., 2019b. Z-Score standard growth chart design of toddler weight using least square spline semiparametric regression. AIP Conference Proceedings 2329, 060031.

Chamidah, N., Lestari, B., Massaid, A., Saifudin, T., 2020a. Estimating mean arterial pressure affected by stress scores using spline nonparametric regression model approach. Commun. Math. Biol. Neurosci. 2020, 72. https://doi.org/10.28919/cmbn/4963.

Chamidah, N., Yonani, Y.S., Ana, E., Lestari, B., 2020b. Identification the number of mycobacterium tuberculosis based on sputum image using local linear estimator. Bullet. Elect. Eng. Inform. 9 (5), 2109–2116.

Chamidah, N., Lestari, B., Budiantara, I.N., Saifudin, T., Rulaningtyas, R., Aryati, A., Wardani, P., Aydin, D., 2022. Consistency and asymptotic normality of estimator for parameters in multiresponse multipredictor semiparametric regression model. Symmetry 14 (2), 336.

Chamidah, N., Lestari, B., Budiantara, I.N., Aydin, D., 2024a. Estimation of multiresponse multipredictor nonparametric regression model using mixed estimator. Symmetry 16 (4), 386.

Chamidah, N., Lestari, B., Saifudin, T., Rulaningtyas, R., Wardhani, P., Budiantara, I.N., 2024b. Estimating the number of malaria parasites on blood smears microscopic images using penalized spline nonparametric Poisson regression. Commun. Math. Biol. Neurosci. 2024, 60.

Clark, Todd, E., Huber, F., Koop, G., Marcellino, M., 2022. Forecasting US Inflation Using Bayesian Nonparametric Models. Federal Reserve Bank of Cleveland, Working Paper No. 22-05. DOI: 10.26509/frbc-wp-202205.

Esumanba, S.V., Danaa, N., Konadu, N.E.A., 2019. The impact of money supply on inflation rate in Ghana. Res. J. Finance and Accounting 10 (16), 156–170.

Fibriyani, V., Chamidah, N., 2021. Prediction of inflation rate in Indonesia using local polynomial estimator for time series data. J. Phys.: Conf. Ser. 1776 (1), 012065.

Fibriyani, V., Chamidah, N., Saifudin, T., 2024. Modeling case fatality rate of Covid-19 in Indonesia using time series semiparametric regression based on local polynomial estimator. Commun. Math. Biol. Neurosci. 2024, 9.

Gao, J., Hawthorne, K., 2006. Semiparametric estimation and testing of the trend of temperature series. The Econometrics J. 9 (2), 332–355.

Lestari, B., Fatmawati, B., I.n., 2020. Spline estimator and its asymptotic properties in multiresponse nonparametric regression model. Songklanakarin J. Sci. Technol. 42 (3), 533–548.

Lestari, B., Fatmawati, B., Chamidah, I.N., 2018. Estimation of regression function in Multi-response nonparametric regression model using smoothing spline and kernel estimators. J. Phys.: Conf. Ser. 1097 (1), 012091.

Lestari, B., Chamidah, N., Aydin, D., Yilmaz, E., 2022. Reproducing kernel Hilbert space approach to multiresponse smoothing spline regression function. Symmetry 14 (11), 2227.

Lestari, B., Chamidah, N., Budiantara, I.N., Aydin, D., 2023. Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator. J. King Saud Univ.–sci. 35 (5), 102664.

Li, D., Chen, J., Gao, J., 2011. Nonparametric time-varying coefficient panel data models with fixed effects. The Econometrics J. 14 (3), 387–408.

Moreno, J.J.M., Pol, A.P., Abad, A.J., Blasco, B.C., 2013. Using the R-MAPE index as a resistant measure of forecast accuracy. Psicothema 25 (4), 500–506.

Nyoni, T., 2019. Modelling and forecasting inflation in Burundi using ARIMA models. Munich Personal RePEc Archive. Retrieved from MPRA Paper No. 92444. https://mpra.ub.uni-muenchen.de/92444.

Prahutama, A., Suparti, U., T. w.,, 2018. Modelling fourier regression for time series data – a case study: modelling inflation in foods sector in Indonesia. J. Phys.: Conf. Ser. 974, 012067.

Ramadan, W., Chamidah, N., Zaman, B., Muniroh, L., Lestari, B., 2019. 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. 456 (5), 052063.

Saluza, I., 2015. Aplikasi metode jaringan syaraf tiruan backpropagation dalam meramal tingkat inflasi di Indonesia. J. Gradien 11 (1), 1075–1078.

Simanungkalit, E.F.B., 2020. Pengaruh inflasi terhadap pertumbuhan ekonomi di Indonesia. J. Management (SME's) 13 (3), 327–340.

Suparti, P., Santosa, A.R., 2018. Modeling inflation in transportation, communication, and financial services using B-Spline time series model. J. Phys.: Conf. Ser. 1025 (1), 012104.

Yolanda, Y., 2017. Analysis of factors affecting inflation and its impact on human development index and poverty in Indonesia. European Res. Stud. J. XX(4B), 38–56.

Yu, D., Chen, L., Li, L., 2023. Nonparametric modeling for the time-varying persistence of inflation. Econ. Lett. 225, 111040.