STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, December 2025, pp 3776–3803.
Published online in International Academic Press (www.IAPress.org)

Least Squares Spline Estimation Method in Semiparametric Time Series Regression for Predicting Indonesia Composite Index

Any Tsalasatul Fitriyah [1,2], Nur Chamidah [3,4,*], Toha Saifudin [3,4], Budi Lestari [4,5], Dursun Aydin [6,7]

[1] *Doctoral Program of Mathematics and Natural Science, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia*
[2] *Department of Mathematics Education, Faculty of Teaching, Training, and Education, Universitas Islam Negeri Mataram, Mataram, 83125, Indonesia*
[3] *Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia*
[4] *Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia*
[5] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, The University of Jember, Jember 68121, Indonesia*
[6] *Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla 48000, Turkey*
[7] *Research Scholar at Department of Mathematics, University of Wisconsin, Oshkosh Algoma Blvd, Oshkosh, WI 54901, USA*

Abstract The Least Squares Spline (LS-Spline) method offers a flexible approach for modeling fluctuating time series data by adaptively positioning knots at points of structural change. This study develops an LS-Spline estimation method for the Semiparametric Time Series Regression (STSR) model, combining an autoregressive structure as the parametric component and multiple nonparametric functions to capture nonlinear effects. The model is applied to predict the Indonesia Composite Index (ICI), a key indicator of sustainable economic growth. In this framework, the ICI at lag-1 is modeled parametrically, while the BI Rate and Inflation are modeled nonparametrically. Four data splitting schemes 6, 12, 18, and 24 months of testing data are used to evaluate forecasting performance over short, medium, and long term horizons. Results show that the LS-Spline STSR model consistently achieves high predictive accuracy, with MAPE and sMAPE below 10% and MASE below 1. Residual diagnostics using ACF and PACF confirm that the model satisfies the white noise assumption. These findings emphasize the potential of the LS-Spline STSR model as an economic forecasting tool that can support policies related to one of poin Sustainable Development Goals (SDGs), namely sustainable economic growth.

Keywords Semiparametric Regression, Least Square Spline, Time Series; Indonesia Composite Index, Inflation, Sustainable Economic Growth.

AMS 2010 subject classifications 62G05, 62G08, 62P20

DOI: 10.19139/soic-2310-5070-2704

1. Introduction

A statistical analysis method that can be used to analyze the relationship between predictor variables and response variables is regression analysis. In the regression analysis, we can use parametric regression, nonparametric regression, and semiparametric regression models approach [1]. Regression with a parametric approach is used if the functional relationship between the predictor variables and response variables is known and has information related to the form of the function. In the parametric approach, assumptions are required to be met. Unlike the nonparametric approach, the nonparametric approach in its use does not require assumptions to be met [2, 3]. In addition, the resulting curve is more flexible, following the data analyzed contained in a particular function

*Correspondence to: Nur Chamidah (Email: nur-c@fst.unair.ac.id). Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya 60115, Indonesia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

[4]. Regression with a combined parametric and nonparametric approach is called semiparametric regression [5]. Smoothing techniques are needed in estimating nonparametric and semiparametric regression models. There are several smoothing techniques that are often used by researchers, including least squares spline [6, 7, 8, 9, 10, 11], truncated spline [12, 13, 14, 15], smoothing spline [16, 17, 18, 19], penalized spline [20, 21, 22], multivariate adaptive regression spline [23], kernel [5, 24, 25, 26, 27, 28], local linear [29, 30, 31, 32, 33, 34, 35], local polynomial [36, 37, 38, 39, 40, 41, 42, 43], local polynomial kernel [44], Fourier series [45, 46, 47, 48], and mixed smoothing of estimator [49, 50].

Regression is not only used to analyze the relationship of variables in cross-section data, but is also widely applied to time series data. Time series data is data generated based on objects observed based on certain time intervals, such as daily, weekly, and even monthly [40]. Time series is used to analyze and forecast trends, cycles, and fluctuations of variables in one period [51]. Commonly used time series data modeling is Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrative Moving Average (ARIMA). Ospina et al. [52] used ARIMA to model Covid-19 in Brazil. Adu et al. [53] also used ARIMA and ARIMAX to predict unemployed in Ghana. Researches related to time series have also been developed in regression with a nonparametric approach. For examples, Gao and King [54] , and Chen et al. [55] used kernel estimators to analyze time series data, and Dong and Gao [56] used spline estimator to determine specification testing driven. Spline is one of the smoothing techniques that has the advantage of overcoming relationship patterns that tend to change by placing knot points and generalizing them into a complex and complicated function [3, 14, 17, 18]. Determination of knot points in the STSR modeling using the LS-Spline is carried out based on the Generalized Cross Validation (GCV) value [16]. The optimal model is produced based on the smallest GCV value [3, 16, 17]. For the semiparametric approach Ahmed et al. [42], and Fibriyani et al.[43] used local polynomial estimators to analyze time series data.

The stock market plays an important role in the economic development of a country [57]. In Indonesia, the condition of the stock market is reflected in the Indonesia Composite Index (ICI), which is a stock index listed on the Stock Exchange [58]. The ICI plays a role in supporting the Sustainable Development Goals (SDGs), namely the point of decent work and economic growth. As an economic barometer, the positive movement of the ICI reflects the rate of economic growth in Indonesia. The ICI data fluctuate and tend to be influenced by the ICI in the previous period [59, 60]. The ICI is also influenced by economic conditions such as the BI Rate and Inflation. The BI rate is a monetary policy in the form of a reference interest rate set by Bank Indonesia to maintain economic stability [61]. Changes in the BI rate have an impact on borrowing costs and affect company profitability and consumer purchasing power. This has an impact on stock price movements. The BI Rate is one of the references for investors to invest in the stock market [62]. In addition to the BI rate, inflation is also a consideration for investors. Inflation is a condition where the price of goods or services increases in general over a certain period [63]. Furthermore, the ICI movements are influenced by many factors, including inflation [64].

In a previous study, Fitriyah [65] compared the ARIMA model and the nonparametric spline model in predicting the Indonesia Composite Index (ICI). The results showed that the spline approach performed well in capturing nonlinear patterns in time series data. Based on these findings, the author was motivated to extend the research by developing a semiparametric spline time series model, which combines the advantages of both parametric and nonparametric approaches. Therefore, this study develops the Least Squares Spline (LS-Spline) estimation method to estimate the Semiparametric Time Series Regression (STSR) model and applies it to real data. The discussion presents the estimation of the STSR model with more than one nonparametric component and an autoregressive component as the parametric part. In applying LS-Spline to real data, this study constructs an STSR model for the Indonesia Composite Index (ICI), which is influenced by the ICI in the previous period (lag-1), the Bank Indonesia (BI) rate, and inflation, and estimates the ICI model using the proposed estimation method, namely LS-Spline.

2. Method details

This section discusses details of the proposed strategy, such as the semiparametric time series regression model, data source, data analysis procedures, and semiparametric time series regression model estimation.

2.1. Semiparametric Time Series Regression Model

In this study, we consider a STSR model as follows [42]:

$$y_t = \mathbf{X}_t\boldsymbol{\beta} + f(z_t) + \varepsilon_t \tag{1}$$

where y_t is a value of the t-th response variable, $\mathbf{X}_t = (X_{1t}, X_{2t}, \ldots, X_{pt})$ is an $(n \times p)$-dimensional matrix of the t-th predictor variable, $\boldsymbol{\beta}$ is a $(p \times 1)$-dimensional vector of parameter of the parametric component, $f(z_t)$ is the t-th unknown function of the nonparametric component, and ε_t is the t-th random error.

Next, if the nonparametric function, $f(z_t)$, is estimated using the LS-Spline with order of p and a knot point of τ_k then the nonparametric function, $f(z_t)$, in the STSR model (1), can be expressed as follows [78]:

$$f(z_t) = \sum_{j=0}^{p} \theta_j z_t^j + \sum_{k=1}^{K} \theta_{p+k}(z_t - \tau_k)_+^p \tag{2}$$

where

$$(z_t - \tau_k)_+^p = \begin{cases} (z_t - \tau_k)^p & \text{for } z_t > \tau_k \\ 0 & \text{for } z_t \leq \tau_k \end{cases}.$$

In the following section we provide a method for estimating the STSR model by using the LS-Spline method.

2.2. Estimating the STSR Model

In this study, we develop an STSR model with predictor variables consisting of an autoregressive component as the parametric component and two unknown functions as the nonparametric component. Next, for estimating the STSR model we use a LS-Spline method. Here, to obtain the estimate, we first consider an STSR model as follows:

$$y_t = \beta y_{t-1} + f_1(z_{1t}) + f_2(z_{2t}) + \varepsilon_t, \quad t = 2, 3, \ldots, T \tag{3}$$

where y_t is the observation value at time t, β is the parameter of the parametric component, y_{t-1} is the observation value at time $(t-1)$ or lag-1, $f_1(z_{1t})$ and $f_2(z_{2t})$ are unknown functions of the nonparametric component, and ε_t is a random error at time t.

In this case, we assume that the estimation of the parametric component parameter is known. Let $\hat{\beta}$ be an estimator of the parametric component parameter β such that the equation (3) can be expressed as follows:

$$y_t - \hat{\beta} y_{t-1} = f_1(z_{1t}) + f_2(z_{2t}) + \varepsilon_t, \quad t = 2, 3, \ldots, T. \tag{4}$$

Hence, the equation (4) can be written as follows:

$$y_t^* = f_1(z_{1t}) + f_2(z_{2t}) + \varepsilon_t, \quad t = 2, 3, \ldots, T \tag{5}$$

where $y_t^* = y_t - \hat{\beta} y_{t-1}$. Note that equation (5) represents a nonparametric regression model, so to estimate the model (5) we can use estimation methods in nonparametric regression. In this study we use the LS-Spline method. Furthermore, we may express the equation (5) into a matrix notation as follows:

$$\mathbf{y}^* = \mathbf{f}_1(\mathbf{z}_1) + \mathbf{f}_2(\mathbf{z}_2) + \boldsymbol{\varepsilon} \tag{6}$$

where

$$\mathbf{y}^* = \mathbf{y} - \hat{\beta}\mathbf{y}_{t-1};\ \mathbf{y} = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_T \end{pmatrix};\quad \mathbf{y}_{t-1} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{T-1} \end{pmatrix};\ \mathbf{f}_1(\mathbf{z}_1) = \begin{pmatrix} f_1(z_{12}) \\ f_1(z_{13}) \\ \vdots \\ f_1(z_{1T}) \end{pmatrix};\quad \mathbf{f}_2(\mathbf{z}_2) = \begin{pmatrix} f_2(z_{22}) \\ f_2(z_{23}) \\ \vdots \\ f_2(z_{2T}) \end{pmatrix};$$

and $\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_T \end{pmatrix}$.

Next, by considering equation (2), we estimate the nonparametric functions, $f_1(z_{1t})$ and $f_2(z_{2t})$, in equation (4) by using the LS-Spline method with order of p_1 and knot points of $\tau_{11}, \tau_{12}, \ldots, \tau_{1K}$ for variable z_1, and with order of p_2 and knot points of $\tau_{21}, \tau_{22}, \ldots, \tau_{2K}$ for variable z_2. In this step, we can express the function $f_1(z_{1t})$ as follows:

$$f_1(z_{1t}) = \sum_{j=0}^{p_1} \theta_{1j} z_{1t}^j + \sum_{k=1}^{K} \theta_{1,p_1+k}(z_{1t} - \tau_{1k})_+^{p_1}, \quad t = 2, 3, \ldots, T. \tag{7}$$

We may express equation (7) into a matrix notation as follows:

$$\mathbf{f_{1\lambda}(z_1)} = \mathbf{Z_{1\lambda}}\, \boldsymbol{\theta_{1\lambda}} \tag{8}$$

where

$$\mathbf{f_{1\lambda}(z_1)} = \begin{pmatrix} f_1(z_{12}) \\ f_1(z_{13}) \\ \vdots \\ f_1(z_{1T}) \end{pmatrix};\ \mathbf{Z_{1\lambda}} = \begin{pmatrix} 1 & z_{12}^1 & z_{12}^2 & \cdots & z_{12}^{p_1} & (z_{12} - \tau_{11})_+^{p_1} & \cdots & (z_{12} - \tau_{1K})_+^{p_1} \\ 1 & z_{13}^1 & z_{13}^2 & \cdots & z_{13}^{p_1} & (z_{13} - \tau_{11})_+^{p_1} & \cdots & (z_{13} - \tau_{1K})_+^{p_1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_{1T}^1 & z_{1T}^2 & \cdots & z_{1T}^{p_1} & (z_{1T} - \tau_{11})_+^{p_1} & \cdots & (z_{1T} - \tau_{1K})_+^{p_1} \end{pmatrix};$$

$$\boldsymbol{\theta_{1\lambda}} = \begin{pmatrix} \theta_{10} \\ \theta_{11} \\ \vdots \\ \theta_{1,p_1+1K} \end{pmatrix};$$

and λ is a smoothing parameter.

Similarly, we can express the function $f_2(z_{2t})$ as follows:

$$f_2(z_{2t}) = \sum_{j=0}^{p_2} \theta_{2j} z_{2t}^j + \sum_{k=1}^{K} \theta_{2,p_2+k}(z_{2t} - \tau_{2k})_+^{p_2}, \quad t = 2, 3, \ldots, T. \tag{9}$$

Hence, we can express equation (9) into a matrix notation as follows:

$$\mathbf{f_{2\lambda}(z_2)} = \boldsymbol{Z_{2\lambda}\theta_{2\lambda}} \tag{10}$$

where

$$\mathbf{f_{2\lambda}(z_1)} = \begin{pmatrix} f_2(z_{12}) \\ f_2(z_{13}) \\ \vdots \\ f_2(z_{1T}) \end{pmatrix};\ \mathbf{Z_{2\lambda}} = \begin{pmatrix} 1 & z_{22}^1 & z_{22}^2 & \cdots & z_{22}^{p_2} & (z_{22} - \tau_{21})_+^{p_2} & \cdots & (z_{22} - \tau_{2K})_+^{p_2} \\ 1 & z_{23}^1 & z_{23}^2 & \cdots & z_{23}^{p_2} & (z_{23} - \tau_{21})_+^{p_2} & \cdots & (z_{23} - \tau_{2K})_+^{p_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_{2T}^1 & z_{2T}^2 & \cdots & z_{2T}^{p_2} & (z_{2T} - \tau_{21})_+^{p_2} & \cdots & (z_{2T} - \tau_{2K})_+^{p_2} \end{pmatrix};$$

$$\boldsymbol{\theta_{1\lambda}} = \begin{pmatrix} \theta_{20} \\ \theta_{21} \\ \vdots \\ \theta_{2,p_2+2K} \end{pmatrix};$$

and λ is a smoothing parameter.

By substituting equations (8) and (10) into equation (6), we obtain:

$$\begin{aligned} \mathbf{y}^* &= \mathbf{f_1(z_1)} + \mathbf{f_2(z_2)} + \varepsilon \\ &= \mathbf{Z_{1\lambda}}\boldsymbol{\theta_{1\lambda}} + \mathbf{Z_{2\lambda}}\boldsymbol{\theta_{2\lambda}} + \varepsilon \\ &= \begin{pmatrix} \mathbf{Z_{1\lambda}} & \mathbf{Z_{2\lambda}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta_{1\lambda}} \\ \boldsymbol{\theta_{2\lambda}} \end{pmatrix} + \varepsilon \\ \mathbf{y}^* &= \mathbf{Z_\lambda}\boldsymbol{\theta_\lambda} \ + \varepsilon \end{aligned} \tag{11}$$

where $\mathbf{Z_\lambda} = \begin{pmatrix}\mathbf{Z_{1\lambda}} & \mathbf{Z_{2\lambda}}\end{pmatrix}$ and $\boldsymbol{\theta_\lambda} = \begin{pmatrix}\boldsymbol{\theta_{1\lambda}} \\ \boldsymbol{\theta_{2\lambda}}\end{pmatrix}$. The next step, we determine a Sum of Squares of Error (SSE) of model presented in (11) to obtain estimation of the model (11). Let S represents the SSE of the model (11), then we have the following equation:

$$S = \boldsymbol{\varepsilon^T}\boldsymbol{\varepsilon} = (\mathbf{y}^* - \mathbf{Z_\lambda}\boldsymbol{\theta_\lambda})^{\mathbf{T}}(\mathbf{y}^* - \mathbf{Z_\lambda}\boldsymbol{\theta_\lambda}) \tag{12}$$

By taking partially derivative of equation (12) with respect to θ_λ and set it equal to zero, we obtain:

$$\begin{aligned}\frac{\partial S}{\partial \boldsymbol{\theta_\lambda}} = \mathbf{0} \Leftrightarrow & \frac{\partial(\mathbf{y}^* - \mathbf{Z_\lambda}\boldsymbol{\theta_\lambda})^{\mathbf{T}}(\mathbf{y}^* - \mathbf{Z_\lambda}\boldsymbol{\theta_\lambda})}{\partial \boldsymbol{\theta_\lambda}} = \mathbf{0} \\ \Leftrightarrow & \hat{\boldsymbol{\theta}}_{\boldsymbol{\lambda}} = (\mathbf{Z}_\lambda^{\mathbf{T}}\mathbf{Z}_\lambda)^{-\mathbf{1}}\mathbf{Z}_\lambda^{\mathbf{T}}\mathbf{y}^*\end{aligned} \tag{13}$$

From equation (11), we have relationship $\mathbf{f_1(z_1)} + \mathbf{f_2(z_2)} = \mathbf{Z}_\lambda\boldsymbol{\theta}_\lambda$. Next, if $\mathbf{f}_\lambda(\mathbf{z}) = \mathbf{f_1(z_1)} + \mathbf{f_2(z_2)}$ then by considering equation (13) we obtain the LS-Spline estimator for $\mathbf{f_\lambda(z)}$, namely $\hat{\mathbf{f}}_{\boldsymbol{\lambda}}(\mathbf{z})$, as follows:

$$\hat{\mathbf{f}}_{\boldsymbol{\lambda}}(\mathbf{z}) = \mathbf{Z_\lambda}\hat{\boldsymbol{\theta}}_{\boldsymbol{\lambda}} = \mathbf{Z_\lambda}(\mathbf{Z_\lambda^T}\mathbf{Z_\lambda})^{-\mathbf{1}}\mathbf{Z_\lambda^T}\mathbf{y}^* \tag{14}$$

We can express equation (14) simply as follows:

$$\hat{\mathbf{f}}_{\boldsymbol{\lambda}}(\mathbf{z}) = \mathbf{H}\mathbf{y}^* \tag{15}$$

where $\mathbf{H} = \mathbf{Z_\lambda}(\mathbf{Z_\lambda^T}\mathbf{Z_\lambda})^{-\mathbf{1}}\mathbf{Z_\lambda^T}$ is called the hat matrix. Furthermore, the LS-Spline estimator of $\mathbf{f_\lambda(z)}$ given in equation (15) is substituted into equation (6) such that we have the following relationship:

$$\begin{aligned}\mathbf{y}^* &= \mathbf{H}\mathbf{y}^* + \boldsymbol{\varepsilon} \\ &\Leftrightarrow \mathbf{y} - \hat{\beta}\mathbf{y_{t-1}} = \mathbf{H}(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}}) + \boldsymbol{\varepsilon} \\ &\Leftrightarrow \mathbf{y} = \hat{\beta}\mathbf{y_{t-1}} + \mathbf{H}(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}}) + \boldsymbol{\varepsilon} \\ &\Leftrightarrow \boldsymbol{\varepsilon} = \mathbf{y} - \hat{\beta}\mathbf{y_{t-1}} - \mathbf{H}(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}}) \\ &\Leftrightarrow \boldsymbol{\varepsilon} = (\mathbf{I} - \mathbf{H})(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}})\end{aligned} \tag{16}$$

Let $S = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}$ be the sum of squares of error. Since $S = [(\mathbf{I} - \mathbf{H})(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}})]^{\mathbf{T}}[(\mathbf{I} - \mathbf{H})(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}})]$, then we have:

$$\begin{aligned}\frac{\partial S}{\partial \hat{\beta}} &= \frac{\partial}{\partial \hat{\beta}}[(\mathbf{I} - \mathbf{H})(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}})]^{\mathbf{T}}[(\mathbf{I} - \mathbf{H})(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}})] = \mathbf{0} \\ &\Leftrightarrow \hat{\beta} = [\mathbf{y_{t-1}^T}(\mathbf{I} - \mathbf{H})^{\mathbf{T}}(\mathbf{I} - \mathbf{H})\mathbf{y_{t-1}}]^{-\mathbf{1}}\mathbf{y_{t-1}^T}(\mathbf{I} - \mathbf{H})^{\mathbf{T}}\mathbf{y}\end{aligned} \tag{17}$$

Hence, we obtain an estimation result of the STSR model using LS-Spline method as follows:

$$\begin{aligned}\hat{\mathbf{y}} &= \hat{\beta}\mathbf{y_{t-1}} + \mathbf{H}(\mathbf{y} - \hat{\beta}\mathbf{y_{t-1}}) \\ &= \mathbf{K}\mathbf{y} + \mathbf{H}(\mathbf{I} - \mathbf{K})\mathbf{y} \\ &= [\mathbf{K} + \mathbf{H}(\mathbf{I} - \mathbf{K})]\mathbf{y} \\ &= \mathbf{M}\mathbf{y}\end{aligned} \tag{18}$$

where

$$\begin{aligned}&\mathbf{M} = \mathbf{K} + \mathbf{H}(\mathbf{I} - \mathbf{K}); \\ &\mathbf{K} = \mathbf{y_{t-1}}[\mathbf{y_{t-1}^T}(\mathbf{I} - \mathbf{H})^{\mathbf{T}}(\mathbf{I} - \mathbf{H})\mathbf{y_{t-1}}]^{-\mathbf{1}}\mathbf{y_{t-1}^T}(\mathbf{I} - \mathbf{H})^{\mathbf{T}}; \text{ and} \\ &\mathbf{H} = \mathbf{Z}_\lambda(\mathbf{Z}_\lambda^{\mathbf{T}}\mathbf{Z}_\lambda)^{-\mathbf{1}}\mathbf{Z}_\lambda^{\mathbf{T}}.\end{aligned}$$

2.3. Data Source

In this study, for an example of the application of the proposed method to real data, we use secondary data obtained from websites, namely www.investing.com and www.bps.go.id. These data are monthly data of the Indonesia Composite Index (ICI), BI (Bank of Indonesia) Rate, and Inflation, recorded from October 2015 to September 2024 that contain 108 observations. In this study, the response variable is the ICI, while predictor variables consist of the parametric component, namely ICI in the previous period lag-1, and the nonparametric components, namely BI rate and inflation. Next, the data will be analysed according to the proposed method with the steps described in the following section.

2.4. Steps for Analyzing Data

In this section, we provide the steps taken in data analysis in an example of the application of the proposed method, namely LS-Spline, to real data. The steps are as follows:

1. Create a scatterplot between response variable and each predictor variable, to visually see the relationship pattern between the response variable and each predictor variable which will be used as a basis to support the suitability of the proposed method, and for determining parametric components and nonparametric components.
2. Calculate the correlation between predictor variables and response variable, to see if there is a significant correlation between the response variables and the predictor variables so that a regression model approach is suitable.
3. Determine parametric components and nonparametric components based on step (1).
4. Determine the number and location of knot points using the quantile method, to see changes in the curve behavior pattern.
5. Select the optimal knot point based on the minimum value of GCV (Generalized Cross-Validation). Together with the results in step (4), the results in this step will be used to build a spline model.
6. Create a STSR model based on the results in step (5).
7. Calculate MAPE, sMAPE, and MASE value to determine the criteria for the goodness of the estimated STSR model obtained by LS-Spline method.

3. Method Validation

This section explains the procedures used to validate the proposed method, namely is the LS-Spline method, applied to real data for prediction purposes. In this study, to validate the proposed method, we used two statistical tools, namely GCV and MAPE.

To build a spline model, we need to determine the number and location of knot points. Knot points are points where changes in the curve pattern occur. The selection of the optimal knot points is carried out to get the best model. The optimal knot points are selected based on the smallest value of the GCV. The GCV formula is given as follows [16, 17, 18, 26, 27, 79]:

$$\text{GCV}(\lambda) = \frac{\text{MSE}(\lambda)}{\left[\frac{1}{n}\text{trace}(\mathbf{I} - \mathbf{A}(\lambda))\right]^2} \tag{19}$$

where $\lambda = \{p, \tau_1, \tau_2, \ldots, \tau_k\}$ is the smoothing parameter, p is the order of spline, τ_k is the knot point, and $\mathbf{A}(\lambda) = \mathbf{Z}_\lambda(\mathbf{Z}_\lambda^\mathbf{T}\mathbf{Z}_\lambda)^{-\mathbf{1}}\mathbf{Z}_\lambda^\mathbf{T}$.

Next, by using the obtained optimal number and location of knot points, we build a STSR model based on the LS-Spline. In addition, to validate the proposed method, namely LS-Spline, on the estimated STSR model obtained, we use MAPE value. The MAPE formula is given as follows [23, 44, 47, 48, 79]:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{y_t}\right| \tag{20}$$

Based on the MAPE value, we can determine the accuracy category of the proposed method (i.e., LS-Spline method). The following table (see Table 1) presents the accuracy categories based on the MAPE values [23, 44, 47, 48, 71].

Table 1. Category of the MAPE Value

MAPE Value	**Category**
MAPE $> 50\%$	Inaccurate
$20\% <$ MAPE $\geq 50\%$	Reasonable
$10\% <$ MAPE $\leq 20\%$	Accurate
MAPE $\leq 10\%$	Highly Accurate

In addition to MAPE, the Symmetric Mean Absolute Percentage Error (sMAPE) was also used to validate the STSR model based on LS-splines produced in this study. is used to measure the accuracy of model predictions, particularly in time series and statistical forecasting. The formula is as follows [66]:

$$\text{sMAPE} = \frac{100\%}{N} \sum_{t=1}^{N} \frac{2|y_t - \hat{y}_t|}{|y_t| + |\hat{y}_t|} \tag{21}$$

The model accuracy categories based on are explained in Table 2 [67, 68]:

Table 2. Category of the sMAPE Value

Value	**Category**
sMAPE $> 50\%$	Inaccurate
$20\% <$ sMAPE $\geq 50\%$	Reasonable
$10\% <$ sMAPE $\leq 20\%$	Accurate
sMAPE $\leq 10\%$	Highly Accurate

As a complement to previous error measures, this study also uses the Mean Absolute Scaled Error (MASE), which is considered more stable and scale-free in assessing forecast model accuracy. Mathematically, the MASE formula is as follows [69]:

$$\text{MASE} = \frac{\frac{1}{N}\sum_{t=1}^{N} |y_t - \hat{y}_t|}{\frac{1}{N-1}\sum_{t=2}^{N} |y_t - y_{t-1}|} \tag{22}$$

MASE is used to assess the accuracy of a model by comparing the average forecast error to the error generated by a simple baseline model. A MASE value < 1 indicates that the proposed model performs better than the baseline model, a MASE value of 1 indicates that the model has an equivalent level of precision, while a MASE value > 1 indicates a larger model error.

4. Application Example of the Proposed Method on the Real Data

In this study, we apply the proposed method (i.e., the LS-spline) to the real data to predict the ICI affected by the BI rate and inflation. The ICI is a stock index that describes changes or movements in stock prices on the Indonesian stock exchange. The increase in ICI is a sign that most things in the stock market are increasing and vice versa. The ICI is also considered an indicator of Indonesia's economic conditions. Investors and analysts, when making investment decisions, look at ICI as a reference to view stock market conditions in Indonesia. The ICI trend is fluctuating and in recent years has tended to increase. In several months in 2020, the ICI experienced a decline, this was due to pandemic conditions. The trend of ICI in the last nine years is shown in Figure 1.

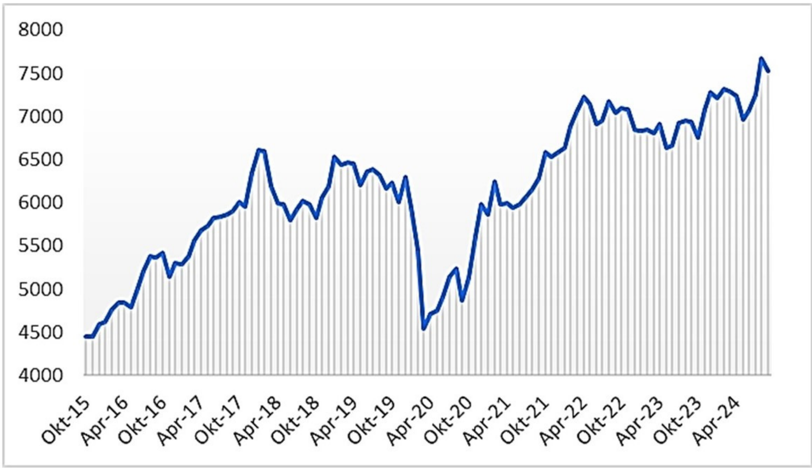


Figure 1. Trend of Indonesia Composite Index (ICI)

The ICI movements are influenced by various factors, including the BI rate and inflation. The BI rate is an economic policy in the form of a reference interest rate set by Bank Indonesia. The role of the BI rate is very important for Indonesian economy. Movements in the BI rate can affect other financial interest rates such as bank loans and deposits. The BI rate is also used to control inflation and maintain the stability of the value of the Rupiah. Another factor discussed in this research is inflation. Inflation is a general and continuous increase in the prices of goods and services over a certain period of time. In the economy, inflationary movements influence many things, such as decreasing consumer purchasing power and increasing interest rates. The BI rate and inflation values are announced every month. Every month Bank Indonesia conducts an evaluation regarding economic conditions as a reason for changes in the BI rate. Meanwhile, inflation is measured and reported monthly. Even though the prices of goods and services change every day, inflation is used as a measure of the average increase in goods and services. Inflation is measured based on the Consumer Price Index. The BI Rate and Inflation trends are depicted in Figure 2.

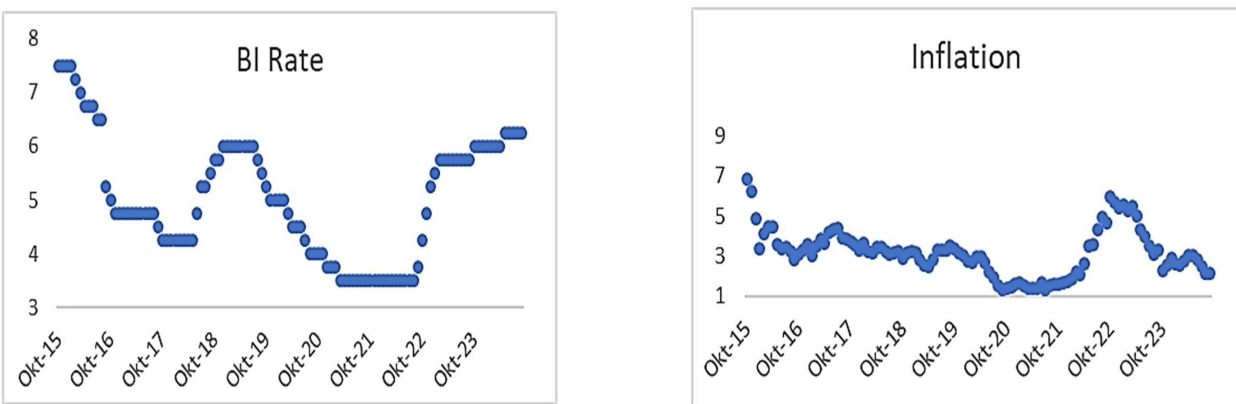


Figure 2. Trends of BI Rate (left) and Inflation (right)

In this study, the variables used are ICI, ICI in the previous period lag-1, BI rate, and inflation. The ICI movements do not occur separately from one period to another so ICI in the previous period has a close relationship to ICI in the future. The BI Rate is a monetary policy determined by Bank Indonesia that can influence investment attractiveness and liquidity in the stock market. Meanwhile, inflation has a direct impact on people's purchasing power as well as market expectations of company performance which influences stock market movements. Descriptive statistics of variables in this study are shown in Table 3.

Table 3. Descriptive Statistics of Variables

Variable	Min	Max	Mean	Variance
ICI	4446.46	7670.73	6124.789	647354.2
BI Rate	3.5	7.5	5.055556	1.267913
Inflation	1.32	6.83	3.17537	1.346672

The ICI movements tend to follow patterns influenced by past share price momentum. Figure 3 shows a scatter plot of the relationship between ICI versus ICI in the previous period lag-1.

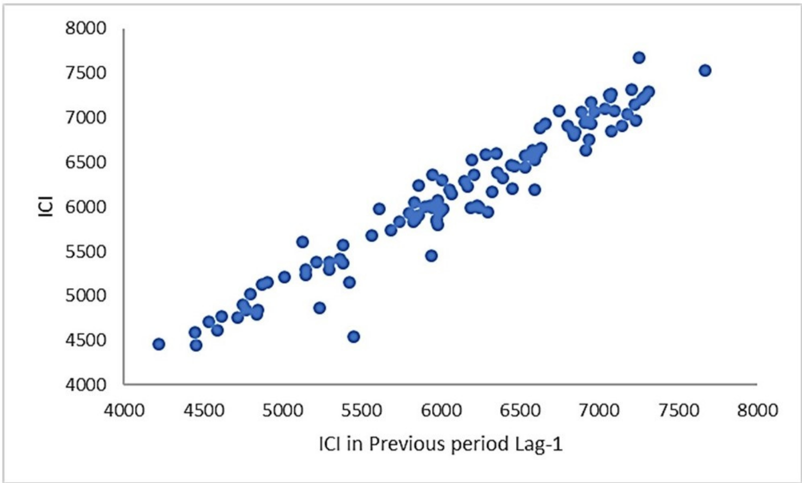
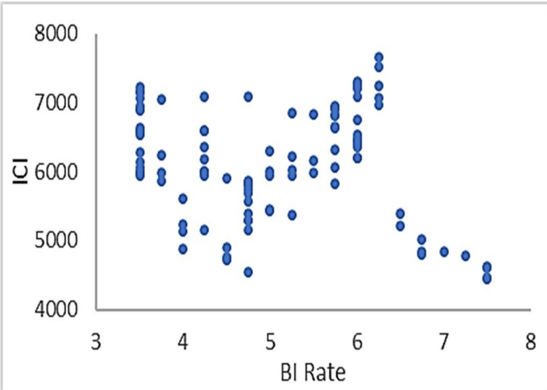


Figure 3. Scatterplot of ICI versus ICI in Previous Period Lag-1

In Figure 3, it can be seen that the relationship between the ICI variable in the previous period lag-1 and ICI has a linear pattern. The correlation coefficient between ICI and the ICI in previous period lag-1 is 0.968 with a significance value of zero which is less than 0.05. This means that the variables ICI and ICI in previous period lag-1 are significantly correlated. Therefore, the ICI variable in the previous period is lag-1 as a parametric component. This is different from the pattern of the relationship between the BI Rate and Inflation variables towards ICI. The pattern of the relationship between these two variables towards ICI has a non-linear pattern. Thus, the BI Rate and Inflation variables are nonparametric components. The relationship pattern between the BI rate and inflation against ICI is shown in Figure 4.



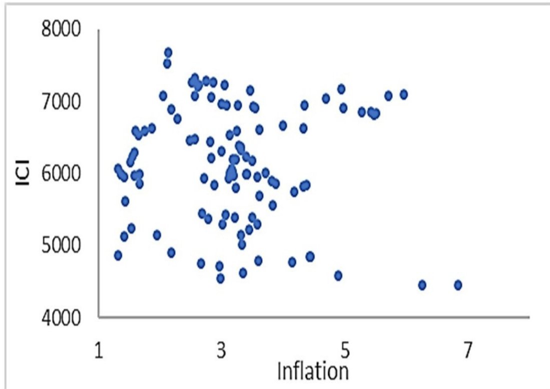


Figure 4. Scatter plots of ICI versus BI rate (left), and ICI versus Inflation (right).

The correlation between the BI rate and ICI has a significance value of 0.104, while the correlation between inflation and ICI has a significance value of 0.661. Both variables have a significant correlation value of more than 0.05, so it can be said that the BI Rate and inflation variables are not correlated with ICI. Therefore, in this study, the variables BI Rate and inflation were chosen as nonparametric components.

In this research, the data is divided into training data and testing data. Training data is used to build the model, while testing data is used to test the model that has been built. In this study, the model evaluation was conducted under several schemes of training and testing data. These schemes included using 102 months of data for training with 6 months for testing, 96 months for training with 12 months for testing, 90 months for training with 18 months for testing, and 84 months for training with 24 months for testing. This partitioning strategy was designed to evaluate the model's performance across different forecast horizons, ranging from short-term to medium-to-long-term, thus demonstrating its robustness in capturing market dynamics and varying economic conditions. To build the model in this study, the optimal knot point was determined by selecting the smallest GCV value. Furthermore, a linear spline (order = 1) was used based on the principle of parsimony. With a relatively limited number of observations (108 monthly data), the use of a higher-order spline (quadratic or cubic) potentially increases the risk of overfitting and produces unrealistic curvature [70]. Linear splines are considered sufficient to maintain model interpretability and stability in economic applications. The use of linear splines is also common in macroeconomic and financial studies, as it provides a balance between flexibility and simplicity.

4.1. Scheme 1

The first scheme is a scheme built based on 102 trading data and 6 testing data. This scheme is designed to evaluate the model's performance in a short-term forecast horizon of six months. The six-month horizon is particularly relevant in the context of the Indonesian stock market, as it reflects short-term market reactions to economic and financial dynamics, including monetary policy adjustments, commodity price fluctuations, and global shocks that often exert significant impacts within a relatively short period. Therefore, this scheme highlights the model's ability to provide timely and adaptive forecasts under rapidly changing market conditions. Descriptive statistics on this scheme are shown in Table 4

To build the STSR model using LS-Spline, knot points are required as the location of the curve change point. In this study, the quantile method is used to determine the knot points. The quantile method is considered more effective compared to other methods such as positioning the knot points according to changes in the curve shape [71, 72]. The quantile method is also said to be able to avoid overfitting [73]. The knot points in this scheme are determined as 3 knot points, namely 4.0625, 4.75, and 5.75 for the BI Rate variable. The knot points for the inflation variable are 2.75; 3.205; and 3.61. Next, exploration of all knot point combinations is carried out and evaluated by calculating the GCV value, so that the optimal knot point combination can be obtained. The results of the STSR Model estimation using LS-Spline with 102 training and 6 testing data are shown in Table 5.

Table 4. Statistic Descriptive of Scheme 1

Data	Variable	Min	Max	Mean	Variance
Training	ICI	4446.46	7316.11	6056.415	597271.9
	BI Rate	3.5	7.5	4.973039	1.206568
	Inflation	1.32	6.25	3.171667	1.266768
Testing	ICI	6970.74	7670.73	7287.157	71799.3
	BI Rate	6	6.25	6.208333	0.010417
	Inflation	1.84	3	2.406667	0.206067

Table 5. Estimation results of STSR Model Using LS-Spline for Scheme 1

Knot point of BI Rate Variable			GCV	Knot point of Inflation Variable			GCV
4.0625			0.069	2.57			0.069
4.75			0.067	3.205			0.070
5.75			0.070	3.61			0.070
4.0625	4.75		0.068	2.57	3.205		0.071
4.0625		5.75	0.070	2.57		3.61	0.071
	4.75	5.75	0.066		3.205	3.61	0.071
4.0625	4.75	5.75	0.067	2.57	3.61	3.75	0.072

Based on Table 5, the optimal knot points for the Bi rate variable are 4.75 and 5.75 with a GCV value of 0.066. For the inflation variable, the optimal knot point was found to be 2.57 with a GCV value of 0.069. So the model obtained is:

$$y_t = 1606.72 + 0.872y_{t-1} - 232.9z_1 + 56.25(z_1 - 4.75)_+ + 452.17(z_1 - 5.75)_+ - 360.46z_2 - 33.28(z_2 - 2.57)_+ \tag{23}$$

Based on equation (23), the relationship between the ICI variable in the previous period lag-1 and the ICI variable has a positive relationship. This is written as follows:

$$y_t = 1606.72 + 0.872y_{t-1} + f_1(z_1) + f_2(z_2) \tag{24}$$

Equation (24) it is stated that the current ICI relationship is influenced by the ICI in the previous period by 0.872, which means that if the ICI in the previous period increases by one unit, then the current ICI will also increase by 0.872. Next, the influence of the BI rate is explained, which is written in equation (25).

$$\begin{aligned} f_1(z_1) &= -232.9z_1 + 56.25(z_1 - 4.75)_+ + 452.17(z_1 - 5.75)_+ \\ &= \begin{cases} -232.9z_1 & \text{for } z_1 \leq 4.75 \\ -267.169 - 176.658z_1 & \text{for } 4.75 < z_1 \leq 5.75 \\ -2867.14 + 275.51z_1 & \text{for } z_1 > 5.75 \end{cases} \end{aligned} \tag{25}$$

The analysis of equation (25) shows that the BI Rate tends to have a negative effect on the ICI. When the BI Rate is below 4.75, a one-unit increase in the interest rate, assuming other variables remain constant, will reduce the ICI by 232.9 points. When the BI Rate is in the range of 4.75 to 5.75, a one-unit increase in the interest rate reduces the ICI by 176.65 points. However, if the BI Rate exceeds 5.75, each one-unit increase actually increases the JCI by 275.51 points. The effect of inflation on the ICI based on the estimation results of the STSR Model Using LS-Spline is explained in equation (26) as follows:

$$f_2(z_2) = -360.46z_2 - 33.28(z_2 - 2.57)_+$$

$$= \begin{cases} -360.46 z_2 & \text{for } z_2 \leq 2.57 \\ 83.208 - 393.745 z_2 & \text{for } z_2 > 2.57 \end{cases} \tag{26}$$

Based on equation (26) Inflation is proven to have a negative effect on ICI. A one unit increase in inflation, whether the value is less than or equal to 2.57 or more than 2.57, causes a decrease in the ICI of 360.46 and 393.745 points, respectively. In addition to producing a model for predicting ICI, a figure is also presented which is a plot of the results of the STSR Model estimation using LS-Spline on both training and testing data (see Figure 5).

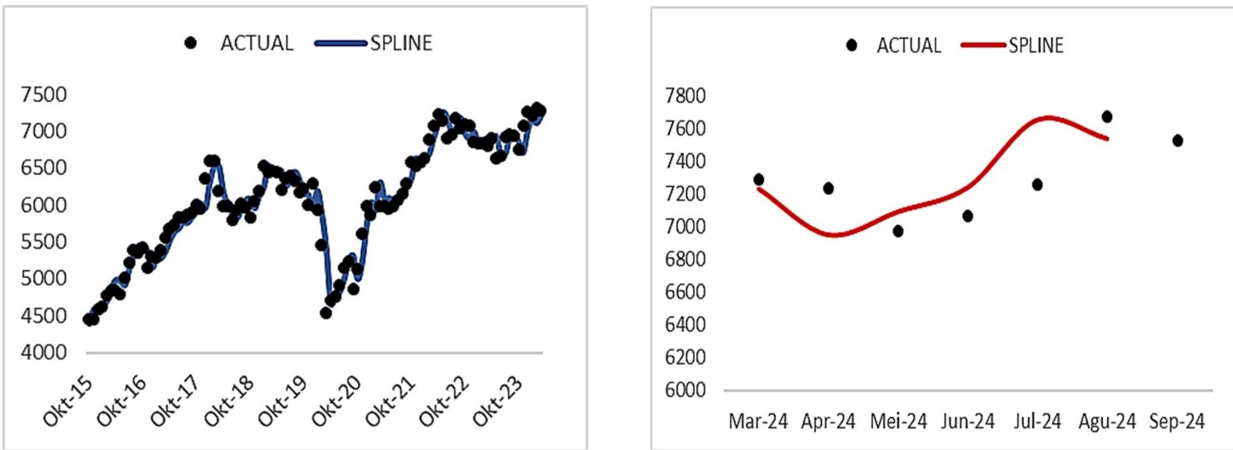


Figure 5. Plot of estimation results of STSR model using LS-Spline for training data of Scheme 1 (left) and testing data of scheme 1 (right)

This study was conducted to predict the ICI. Therefore, for goodness of fit we used the MAPE, sMAPE, and MASE. The calculation results of the MAPE, sMAPE, and MASE values for training, testing, and overall data in scheme 1 are presented in Table 6.

Table 6. MAPE, sMAPE, and MASE value of Scheme 1

Data	**MAPE**	**sMAPE**	**MASE**
Training	2.38%	2.364%	0.942
Testing	0.206%	0.205%	0.067
Overall	1.295%	1.285%	0.504

Based on the MAPE and sMAPE values shown in Table 6, both the MAPE and sMAPE values for the training, testing, and overall data show values of less than 10 %, thus it can be concluded that the resulting model is a highly accurate prediction. In addition to the MAPE and sMAPE values, the MASE values obtained for both the training, testing, and overall data are all less than 1. Because the MASE value is less than 1, it means that the resulting STSR Model estimation using LS-Spline performs better than the baseline model. In addition, as part of the model diagnostic test, the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the residuals are displayed. The purpose of this analysis is to check whether there are still remaining autocorrelation patterns in the error after the model is estimated. If the residual ACF and PACF plots do not show significant autocorrelation at various lags, it can be concluded that the model is adequate in capturing the temporal dependency in the ICI data.

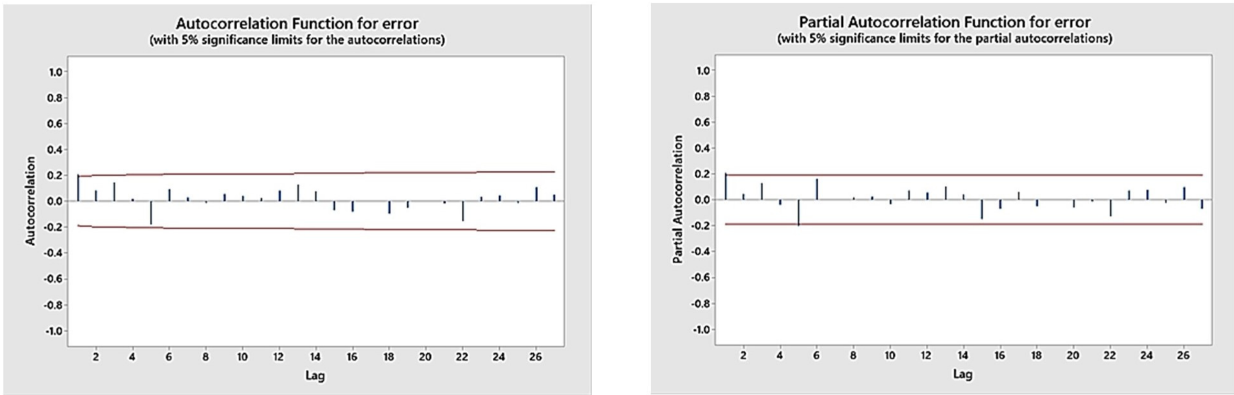


Figure 6. Plot of ACF residual of Scheme 1 (left) and PACF residual of scheme 1 (right)

Based on the ACF and PACF residual plots for this model, depicted in Figure 6, it can be seen that no bars fall outside the confidence limits, thus concluding that there is no significant autocorrelation. The same is true for the PACF residual plot, which means it meets the white noise assumption.

4.2. Scheme 2

The second scheme employs 96 observations for the training data and the last 12 observations for the testing data. The one-year forecast horizon aligns with annual financial reporting cycles and fiscal policies, which are typically planned on a yearly basis. Moreover, the yearly horizon is relevant for capturing seasonal dynamics of the Indonesian economy, such as consumption patterns during Ramadan and Eid, as well as foreign capital flows at the end of the year. This scheme thus enables an assessment of the model's ability to provide forecasts consistent with annual economic cycles. As for the descriptive statistics of the data used in this scheme, both training data and testing data are presented in the table 7

Table 7. Statistic Descriptive of Scheme 2

Data	Variable	Min	Max	Mean	Variance
Training	ICI	4446.46	7228.91	5987.872	551882.5
	BI Rate	3.5	7.5	4.908854	1.212
	Inflation	1.32	6.25	3.199063	1.331897
Testing	ICI	6752.21	7670.73	7220.129	58289.31
	BI Rate	6	6.25	6.104167	0.016572
	Inflation	1.84	3.05	2.57	0.139909

The quantile calculation results for the BI Rate variable in Scheme 2 show three knot points for each of the BI Rate and Inflation variables. For the BI Rate variable, these are 4, 4.75, and 5.75, and for the inflation variable, the knot points obtained are 2.62, 3.24, and 3.75. The results of the training data estimation using Scheme 2 are presented in Table 8.

Based on Table 8, the optimal knot points for the Bi rate variable are 4.75 and 5.75 with a GCV value of 0.072. For the inflation variable, the optimal knot point was found to be 2.62 with a GCV value of 0.075. So the STSR Model estimation using LS-Spline obtained for scheme 2 is:

$$y_t = 2196.58 + 0.81y_{t-1} - 304.74z_1 + 81.59(z_1 - 4.75)_+ + 558.46(z_1 - 5.75)_+ - 486.43z_2 - 29.15(z_2 - 2.63)_+ \quad (27)$$

Based on equation (27), the relationship between the ICI variable in the previous period lag-1 and the ICI variable has a positive relationship. This is written as follows:

Table 8. Estimation results of STSR Model Using LS-Spline for Scheme 2

Knot point of BI Rate Variable			GCV	Knot point of Inflation Variable			GCV
4			0.075	2.62			0.075
4.75			0.073	3.24			0.076
5.75			0.076	3.75			0.077
4	4.75		0.074	2.62	3.24		0.076
4		5.75	0.077	2.62		3.75	0.076
	4.75	5.75	0.072		3.24	3.75	0.077
4	4.75	5.75	0.073	2.62	3.24	3.75	0.078

$$y_t = 2196.58 + 0.81 y_{t-1} + f_1(z_1) + f_2(z_2) \tag{28}$$

In equation (28) it can be explained that the relationship between ICI in the previous period lag-1 and the ICI variable was 0.81. This means that if other variables are considered constant and the ICI variable in the previous period lag-1 increases by one unit, then the ICI variable will increase by 0.81. Apart from being influenced by ICI in the previous period, the ICI is also influenced by the function of the BI rate and inflation variables. The function that explains the BI rate variable is written as follows:

$$f_1(z_1) = -304.74 z_1 + 81.59(z_1 - 4.75)_+ + 558.46(z_1 - 5.75)_+$$

$$= \begin{cases} -304.74 z_1 & \text{for } z_1 \leq 4.75 \\ -387.55 - 223.15 z_1 & \text{for } 4.75 < z_1 \leq 5.75 \\ -3598.7 + 335.32 z_1 & \text{for } z_1 > 5.75 \end{cases} \tag{29}$$

The BI rate variable tends to have a negative influence on the ICI. If the BI rate is less than 4.75 the BI rate increases by one unit and other variables are considered constant, then the ICI will decrease by 304.74. If the BI rate is more than 4.75 and less than 5.75, increases by one unit, and other variables are considered constant, then the ICI will decrease by 223.15. If other variables are constant, the BI rate increases by one unit and is more than 5.75, then the ICI will increase by 335.32. Apart from the BI Rate, inflation also has a negative influence on the ICI. This is explained in the following equation (30):

$$f_2(z_2) = -486.43 z_2 - 29.15(z_2 - 2.62)_+$$

$$= \begin{cases} -360.46 z_2 & \text{for } z_2 \leq 2.62 \\ 76.373 - 515.58 z_2 & \text{for } z_2 > 2.62 \end{cases} \tag{30}$$

If other variables are considered constant and inflation is less than and equal to 2.62, then if inflation increases by one unit, the ICI will decrease by 486.43. If the inflation value is more than 2.62 and increases by one unit while other variables are considered constant, then the ICI will decrease by 515.58. The estimation results of STSR model using LS-Spline are also depicted in the form of a plot diagram. Figure 7 shows plot of estimation results of STSR model using LS-Spline for training and testing data.

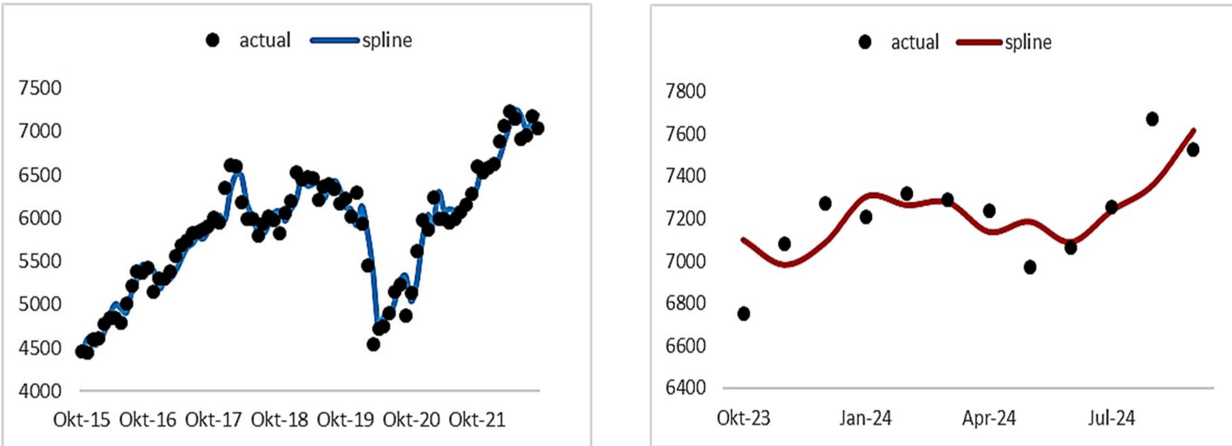


Figure 7. Plot of estimation results of STSR model using LS-Spline for training data of Scheme 2 (left) and testing data of scheme 2 (right)

Figure 7 shows that the plot produced by STSR model based on LS-Spline (described by the blue line) is close to the actual value (described by the black point). The same result is also explained in the plot of the STSR model based on LS-Spline for testing data depicted in Figure 7 (left). The values obtained from the estimation results of the testing data are shown with a red line. The results of calculating the MAPE, sMAPE, and MASE values in scheme 2 are shown in Table 9

Table 9. MAPE, sMAPE, and MASE value of Scheme 2

Data	MAPE	sMAPE	MASE
Training	2.37 %	2.35 %	0.929
Testing	1.79 %	1.79 %	0.753
Overall	2.08 %	2.074 %	0.842

Based on the calculation of MAPE and sMAPE values on the training, testing, and overall data in scheme 2, the obtained values are less than 10%, which means the resulting model has a highly accurate prediction. Similarly, the MASE value is less than 1, meaning the STSR model using LS-Spline has better performance compared to the baseline model.

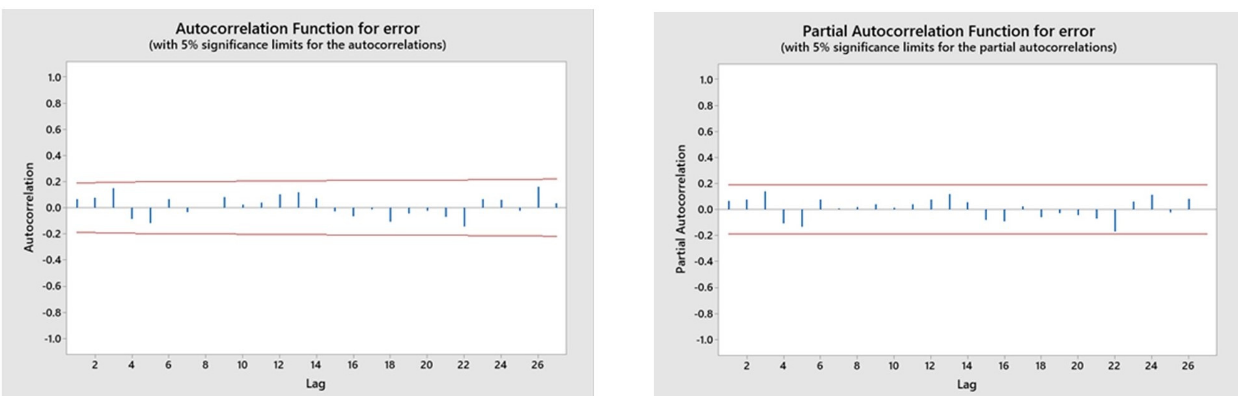


Figure 8. Plot of ACF residual of Scheme 2 (left) and PACF residual of scheme 2 (right)

Figure 8 shows the ACF and PACF plots for the residual data in scheme 2. Based on Figure 8, it can be concluded that the STSR model using LS-Spline produced is free from autocorrelation and meets the white noise assumption.

4.3. Scheme 3

The third model consists of 90 observations for training data and the last 18 observations for testing data. The 18-month out-of-sample model was chosen to represent the intermediate period, which is important for capturing the lagged effects of monetary and fiscal policy. In practice, changes in Bank Indonesia's benchmark interest rate or inflationary pressures are not always immediately reflected in the ICI but are only visible after 12 to 24 months. Therefore, this model provides an overview of the model's ability to predict market dynamics over the intermediate period, reflecting the transition from an annual cycle to a longer policy cycle. The table 10 presents descriptive statistics for the training and testing data for this model.

Table 10. Statistic Descriptive of Scheme 3

Data	Variable	Min	Max	Mean	Variance
Training	ICI	4446.46	7228.91	5931.115	535715.9
	BI Rate	3.5	7.5	4.85	1.237079
	Inflation	1.32	6.25	3.184778	1.389039
Testing	ICI	6633.26	7670.73	7093.162	78404.4
	BI Rate	5.75	6.25	6	0.036765
	Inflation	1.84	4.33	2.851111	0.0392

An exploration of knot points was conducted to determine the optimal knots in estimating the STSR model using spline. Based on the results obtained from Scheme 3, the optimal number of knots for the BI Rate variable was found to be two, located at 4.75 and 5.75, with a GCV value of 0.0760. Meanwhile, for the inflation variable, the optimal knot point was one, located at 2.5, with a GCV value of 0.0790. The results of the optimal knot selection are presented in Table 11

Table 11. Estimation results of STSR Model Using LS-Spline for Scheme 3

Knot point of BI Rate Variable			GCV	Knot point of Inflation Variable			GCV
4			0.0786	2.5			0.079
4.75			0.0761	3.22			0.080
5.75			0.0796	3.61			0.080
4	4.75		0.0774	2.5	3.22		0.081
4		5.75	0.0804	2.5		3.61	0.081
	4.75	5.75	0.0760		3.22	3.61	0.081
4	4.75	5.75	0.0772	2.5	3.22	3.61	0.082

Based on the optimal knot points obtained in Table 11, the STSR Model using Spline Scheme 3 is explained in equation (31):

$$y_t = 2251.38 + 0.803y_{t-1} - 316.856z_1 + 99.209(z_1 - 4.75)_+ + 562.67(z_1 - 5.75)_+ - 478.65z_2 - 44.65(z_2 - 2.5)_+ \tag{31}$$

Based on equation (31), the relationship between the ICI variable and the ICI variable in the previous period can be explained in equation (32) as follows:

$$y_t = 2251.38 + 0.803y_{t-1} + f_1(z_1) + f_2(z_2) \tag{32}$$

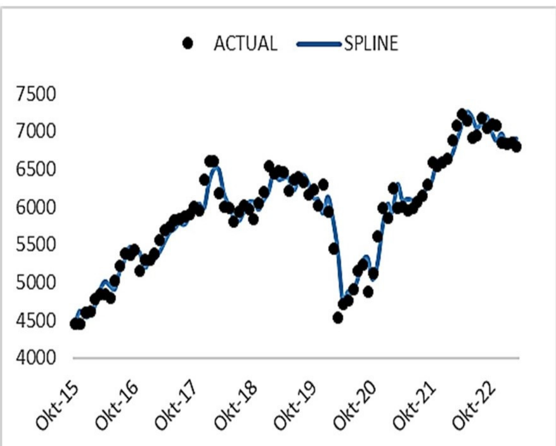
Equation (32) explains the positive relationship between the ICI variable in the previous period and the ICI variable of 0.803. This means that if other variables are considered constant and the ICI in the previous period increases by one unit, then the ICI will increase by 0.803. Equation (33) explains the relationship between the BI rate variable and the ICI.

$$\begin{aligned} f_1(z_1) &= -316.856z_1 + 99.209(z_1 - 4.75)_+ + 562.67(z_1 - 5.75)_+ \\ &= \begin{cases} -316.856z_1 & \text{for } z_1 \leq 4.75 \\ -471.247 - 217.647z_1 & \text{for } 4.75 < z_1 \leq 5.75 \\ -3706.64 + 345.03z_1 & \text{for } z_1 > 5.75 \end{cases} \end{aligned} \tag{33}$$

The estimation results in equation (33) show that the BI Rate has a different effect on the ICI in scheme 3, which is no different from the results in the previous scheme. At interest rate levels below 5.75, an increase in the BI Rate tends to reduce the ICI, whereas when the BI Rate exceeds 5.75, an increase in the interest rate actually encourages an increase in the ICI. The effect of inflation on the ICI is explained in equation (34).

$$\begin{aligned} f_2(z_2) &= -478.65z_2 - 44.65(z_2 - 2.5)_+ \\ &= \begin{cases} -478.65z_2 & \text{for } z_2 \leq 2.5 \\ 111.632 - 523.308z_2 & \text{for } z_2 > 2.5 \end{cases} \end{aligned} \tag{34}$$

The estimation results indicate that inflation has a negative effect on the ICI. When the inflation rate is below 2.5, the negative effect is 478.65, implying that an increase of one unit in inflation leads to a decrease in the ICI by 478.65. Similarly, when the inflation rate is above 2.5, the negative effect is 523.308, meaning that a one-unit increase in inflation results in a decrease in the ICI by 523.308. Therefore, it can be concluded that rising inflation consistently leads to a decline in the ICI under both low and high inflation conditions. The plot of the estimation results of the STSR model using the LS-Spline method for training and testing data in Scheme 3 is presented in Figure 9.



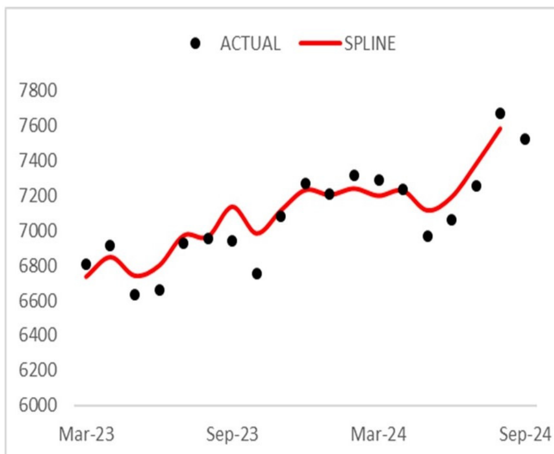


Figure 9. Plot of estimation results of STSR model using LS-Spline for training data of Scheme 3 (left) and testing data of scheme 3 (right)

The MAPE and sMAPE values obtained for the training, testing, and overall data demonstrate the model's ability to produce predictions with excellent accuracy. Furthermore, the MASE value, which is less than one, strengthens the evidence that the STSR model using LS-Spline approach has better predictive performance than the baseline model. The results of the MAPE, sMAPE, and MASE calculations for scheme 3 are shown in Table 12.

The ACF and PACF residual plots of the STSR model using LS-Spline for scheme 3 are shown in Figure 10.

Table 12. MAPE, sMAPE, and MASE value of Scheme 3

Data	**MAPE**	**sMAPE**	**MASE**
Training	2.4%	2.38%	0.921
Testing	1.76%	1.75%	0.784
Overall	2.08%	2.06%	0.852

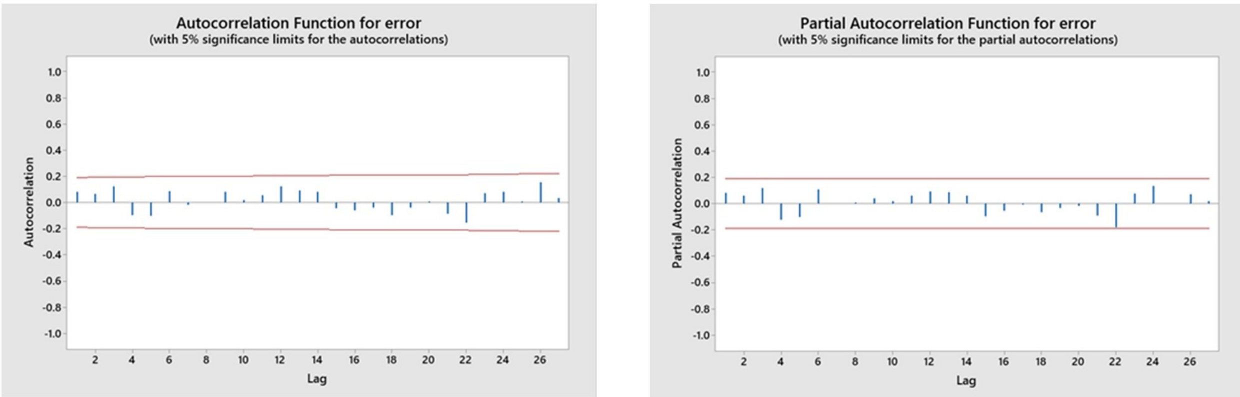


Figure 10. Plot of ACF residual of Scheme 3 (left) and PACF residual of scheme 3 (right)

Test results using ACF and PACF residual plots show no significant lags exceeding the confidence interval. This indicates that the model residuals are random and do not contain any specific patterns, thus the estimated model adequately explains data variation without leaving any autocorrelation.

4.4. Scheme 4

The fourth scheme uses 84 observations as training data and the last 24 observations as testing data. The two-year testing period was chosen because it reflects the medium to long term, which is important for assessing the model's robustness in the face of structural economic dynamics. Factors such as political cycles (elections), changes in global commodity prices, and medium-term government economic policies typically require longer periods to test the stability of predictions. With a 24-month time horizon, the model is tested not only under short-term market conditions but also in the context of more fundamental economic changes. Descriptive statistics for the training and testing data in this scheme are presented in Table 13.

Table 13. Statistic Descriptive of Scheme 4

Data	Variable	Min	Max	Mean	Variance
Training	ICI	4446.46	7228.91	5860.496	497670.5
	BI Rate	3.5	7.5	4.806548	1.28818
	Inflation	1.32	6.25	3.027024	1.107893
Testing	ICI	6633.26	7670.73	7049.817	67675.9
	BI Rate	4.75	6.25	5.864583	0.119452
	Inflation	1.84	5.71	3.486667	1.582232

To build the STSR Model Using LS-Spline, optimal knot points are needed to obtain a good model. The estimated results of the STSR Model using LS-Spline in this scheme are presented in Table 14.

Table 14. Estimation results of STSR Model Using LS-Spline for Scheme 4

Knot point of BI Rate Variable			GCV	Knot point of Inflation Variable			GCV
3.75			0.0907	2.1875			0.0920
4.75			0.0867	3.18			0.0910
5.75			0.0910	3.5575			0.0919
3.75	4.75		0.0883	2.1875	3.18		0.0931
3.75		5.75	0.0928	2.1875		3.5575	0.0941
	4.75	5.75	0.0857		3.18	3.5575	0.0923
3.75	4.75	5.75	0.0873	2.1875	3.18	3.5575	0.0941

Based on table 14, the optimal knot points for the BI rate variable are 2 knot points, namely 4.75 and 5.75 with a GCV value of 0.0857. For the inflation variable, the optimal knot point is 1 knot point, namely 3.18 with a GCV value of 0.091. Thus, the STSR Model Using LS-Spline for scheme 4 is obtained as follows:

$$y_t = 2586.82 + 0.776y_{t-1} - 367.85z_1 + 105.52(z_1 - 4.75)_+ + 669.607(z_1 - 5.75)_+ - 575.34z_2 - 63.287(z_2 - 3.18)_+ \tag{35}$$

The relationship between the ICI variable and the ICI variable in the previous period has a positive relationship of 0.776. In other words, if the ICI in the previous period increases by one unit and other variables are considered constant, then the ICI will increase by 0.776. This relationship is written in the following equation (36):

$$y_t = 2586.82 + 0.776y_{t-1} + f_1(z_1) + f_2(z_2) \tag{36}$$

The relationship between the BI rate and the ICI in scheme 4 has consistent results with the results in schemes 1, 2 and 3. The relationship between the BI rate and the ICI has a negative tendency. This is written in equation (37):

$$\begin{aligned} f_1(z_1) &= -367.856z_1 + 105.52(z_1 - 4.75)_+ + 669.607(z_1 - 5.75)_+ \\ &= \begin{cases} -367.856z_1 & \text{for } z_1 \leq 4.75 \\ -501.205 - 262.334z_1 & \text{for } 4.75 < z_1 \leq 5.75 \\ -4351.45 + 407.274z_1 & \text{for } z_1 > 5.75 \end{cases} \end{aligned} \tag{37}$$

Based on equation (37) the BI rate has a negative influence of 367,856 on the ICI when the BI rate is less than or equal to 4.75. Likewise, if the BI rate is in the interval of 4.75 to 5.75, it has a negative influence of 262,334. However, it is different when the BI rate is more than 5.75, it has a positive influence of 407,274. The relationship between inflation and the ICI in scheme 4 is explained in equation (38):

$$\begin{aligned} f_2(z_2) &= -575.34z_2 - 63.287(z_2 - 3.18)_+ \\ &= \begin{cases} -575.34z_2 & \text{for } z_2 \leq 3.18 \\ 201.239 - 638.6258z_2 & \text{for } z_2 > 3.18 \end{cases} \end{aligned} \tag{38}$$

The relationship explained from the STSR model using LS-Spline in scheme 4 shows that inflation has a negative effect on ICI of 575.34 for inflation less than or equal to 3.18, meaning that if inflation increases by one unit, then ICI results in a decrease in ICI of 575.34. A negative effect of 638.625 also occurs when inflation is more than 3.18, meaning that if inflation increases by one unit, then ICI will decrease by 638.625. The plot of estimation results of the STSR model using LS-Spline for training and testing for scheme 4 is depicted in Figure 11.

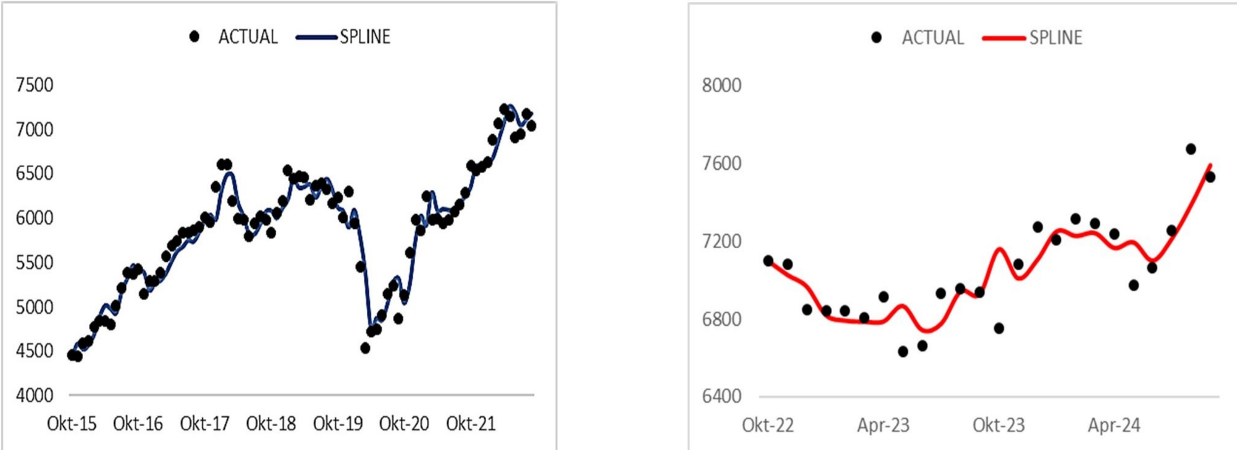


Figure 11. Plot of estimation results of STSR model using LS-Spline for training data of Scheme 4 (left) and testing data of scheme 4 (right)

The accuracy testing of the model in Scheme 4 is explained through the results of the MAPE, sMAPE, and MASE value calculations. The MAPE and sMAPE values for both the training, testing, and overall data show values of less than 10 %, indicating a highly accurate model prediction. A MASE value of less than 1 indicates good performance of the STSR model using LS-Spline. The results of the MAPE, sMAPE, and MASE calculations for the training, testing, and overall data in Scheme 4 are presented in Table 15.

Table 15. MAPE, sMAPE, and MASE value of Scheme 4

Data	MAPE	sMAPE	MASE
Training	2.48%	2.46%	0.906
Testing	1.43%	1.42%	0.746
Overall	1.96%	1.94%	0.826

The residual model test using the ACF and PACF plots of the STSR model using LS-Spline for scheme 4 shown in Figure 12 shows that the resulting model has met the assumptions for time series, namely no autocorrelation and the fulfillment of the white noise assumption. Therefore, it can be concluded that the obtained model is good for use in predicting ICI.

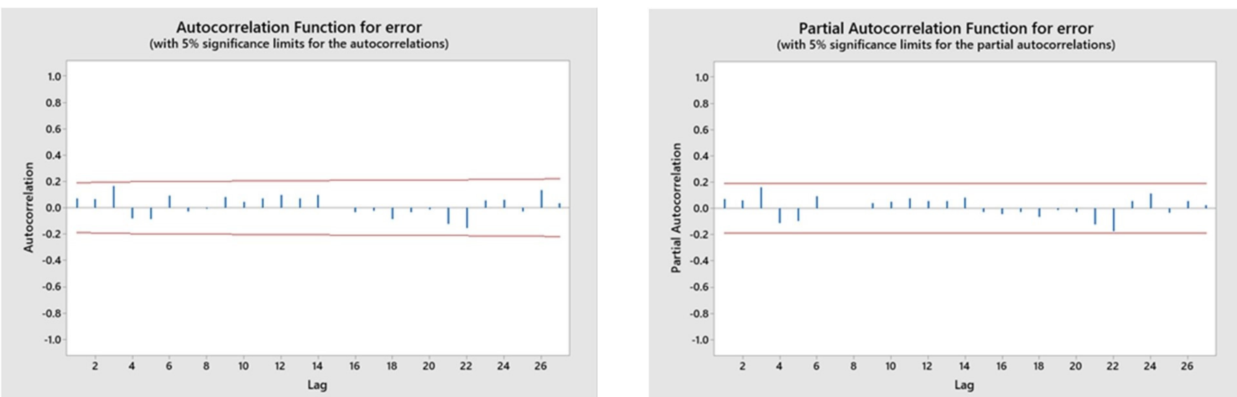


Figure 12. Plot of ACF residual of Scheme 4 (left) and PACF residual of scheme 4 (right)

The results of the model evaluation on the four data sharing schemes, namely with testing data of 6, 12, 18, and 24 months, obtained MAPE and sMAPE values for training, testing, and overall data that are below 10 %, and MASE values are less than 1. These results indicate that the STSR model using LS-Spline has excellent and consistent predictive capabilities across various time horizons, both short, medium, and long term. This consistency indicates that the model is able to adapt to changes in market dynamics in different time periods, without losing predictive accuracy. In addition, the results of the residual diagnostic test through the ACF and PACF plots show that there are no bars that go outside the confidence interval limits, which indicates that the residuals are random (white noise) and free from autocorrelation. This condition strengthens the validity of the model because it shows that the temporal dependency structure in the JCI data has been successfully captured completely by the STSR model using LS-Spline. Thus, the STSR model using LS-Spline is proven to be able to provide stable and reliable prediction results, both for short-term reactive and long-term strategic forecasting in the context of Indonesian financial market analysis.

4.5. ARIMA and ARIMAX for Predicting Indonesia Composite Index

In this study, to provide a more comprehensive evaluation of the proposed STSR Model Using LS-Spline, a comparison of time series models frequently used for forecasting the ICI is included. These models are ARIMA and ARIMAX. ARIMA is the most popular time series method used for forecasting, especially financial and economic data. ARIMA combines three components: Autoregressive, Integrated, and Moving Average. Autoregressive explains the relationship between the response variable and the variable itself in the previous period. Furthermore, integrated shows the number of differencing processes required to make the data stationary, and Moving Average models the relationship between the current value and the error in the next period [74, 75].

The initial step in ARIMA is to identify data stationarity, both in variance and mean. This is done using the Box-Cox test, and mean stationarity is identified through the ACF plot. After second-order transformation and differencing, the data in this study exhibited stationarity. Figure 13 shows a time series plot of ICI and output differencing. Figure 14 shows the ACF and PACF plots after differencing twice.

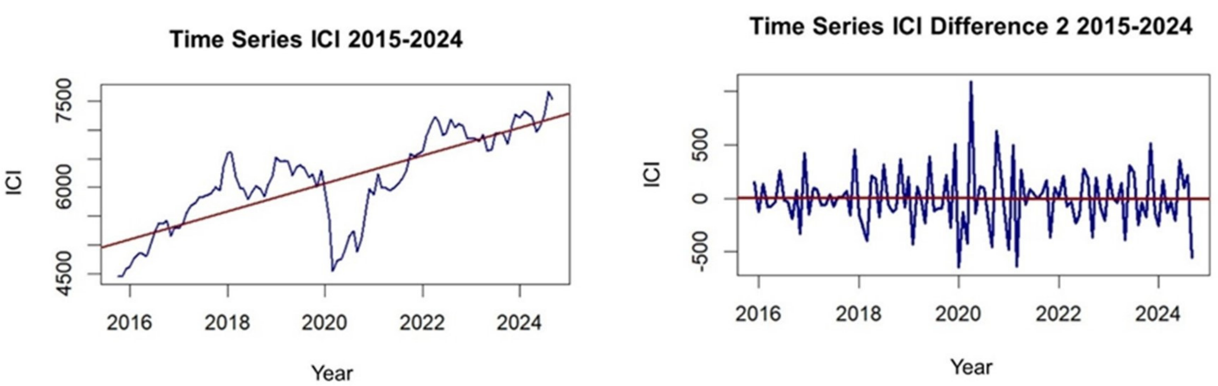
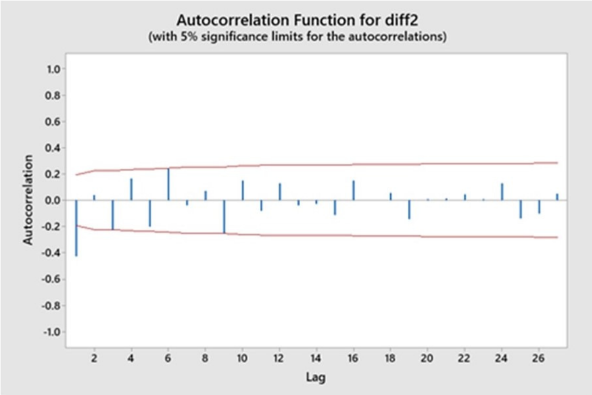


Figure 13. Time Series Plot of ICI (left) and Output Differencing of ICI (right)



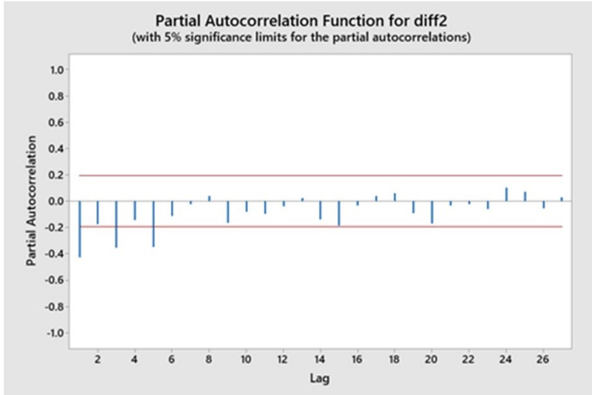


Figure 14. Plot ACF (left) Plot PACF (right)

Based on Figure 13, the data has been stationary in the second differencing process. The next step is to construct a tentative model based on the ACF and PACF in Figure 14. The tentative model based on ARIMA is shown in Table 16.

Table 16. Tentative Model-Based ARIMA Model

ARIMA Model	Koefisien	P-Value (parameter)	MSE	White Noise Assumptions
ARIMA (1,2,1)	AR 1 $= -0.6684$ MA 1 $= 0.97946$	0.000 0.000	123821	Lag 12 $= 0.000$ Lag 24 $= 0.001$ Lag 36 $= 0.000$ Lag 48 $= 0.000$
ARIMA (3,2,1)	AR 1 $= -0.923$ AR 2 $= 0.632$ AR 3 $= -0.3818$ MA 1 $= 0.987$	0.000 0.000 0.000 0.000	43383.6	Lag 12 $= 0.000$ Lag 24 $= 0.001$ Lag 36 $= 0.001$ Lag 48 $= 0.001$

Based on the tentative model in Table 16, it is concluded that the ARIMA model does not meet the white noise assumption (p-value ¡0.05), so the residuals still show autocorrelation patterns. This indicates that the obtained model is not fully accurate in measuring time series data. Furthermore, the ARIMA model is unable to explain other factors that influence the movement of the ICI. Therefore, an ARIMAX model was developed by adding other variables, in this case the BI rate and inflation.

Similar to the ARIMA stage, the ARIMAX process requires identifying data stationarity. After the second differencing process, the BI rate and inflation data are stationary. This is shown in Figure 15.

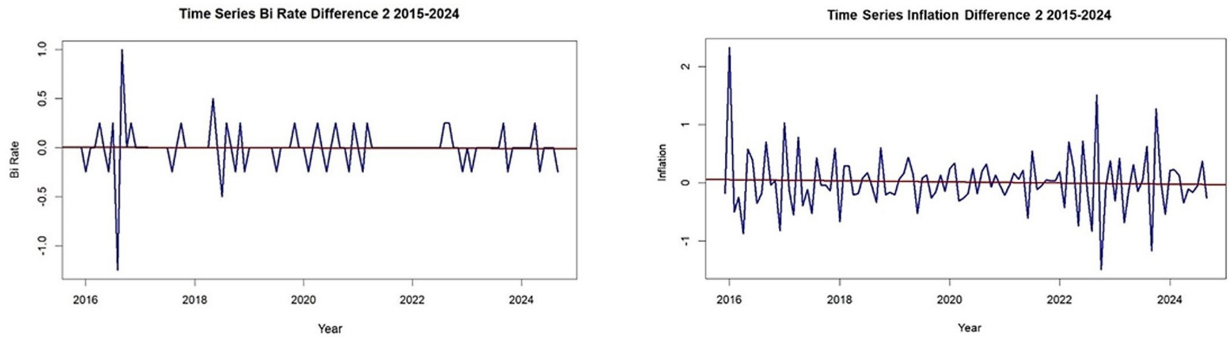


Figure 15. Output Differencing of BI Rate (left) and output Differencing of Inflation (right)

Next, Table 17 shows the possible ARIMAX models.

Table 17. Model Based ARIMAX

ARIMAX Model	Coefficient	p-value	Sig
	AR 1 $= -0.0395$	0.007	Significant
	MA 1 $= -1.000$	0.000	Significant
ARIMAX (1,2,1)	ICI_lag $= -0.416$	0.003	Significant
	BI Rate $= -5.491$	0.960	Not Significant
	Inflation $= -99.0217$	0.0324	Significant
	AR 1 $= -0.709$	0.000	Significant
	AR 2 $= -0.564$	0.000	Significant
	AR 3 $= -0.430$	0.000	Significant
ARIMAX (3,2,1)	MA 1 $= -0.99$	0.000	Significant
	ICI_Lag $= -0.376$	0.040	Significant
	BI Rate $= -37.944$	0.714	Not Significant
	Inflation $= -115.081$	0.015	Significant

The possible ARIMAX models in Table 17 show that there are insignificant parameter coefficients. Therefore, the testing process cannot be continued because there are assumptions in the time series that are not met. Based on the evaluation of the STSR model using LS-Spline, the results show that the MAPE and sMAPE values for the training, testing, and overall data are below 10%, while the MASE value is less than 1. These values confirm that the STSR model using LS-Spline has a very good and consistent level of prediction accuracy across all horizons, both short, medium, and long term. This advantage indicates that the LS-Spline approach is able to capture complex nonlinear patterns in the ICI data something that is difficult to explain by conventional linear models such as ARIMA and ARIMAX. In addition, the results of the residual diagnostic test through the ACF and PACF plots show that there are no bars outside the confidence interval limits, which means that the residuals meet the white noise assumption and are free from autocorrelation. This condition proves that the STSR model using LS-Spline has successfully captured the structure of temporal dependencies thoroughly, thereby leaving no systematic patterns in the residual errors. Thus, the STSR model using LS-Spline is proven to be superior not only in terms of prediction accuracy, but also in model stability and its ability to represent the dynamic relationship between macroeconomic variables and the movement of the Indonesian stock market.

5. Discussion

The estimation results of the STSR model using the LS-Spline method with four data-splitting schemes (6, 12, 18, and 24 months of testing data) show a consistent pattern across all forecast periods, including the short, medium, and long-term. The ICI variable in the previous period has a positive and significant influence on the current ICI value, indicating a momentum effect in the movement of the Indonesian stock market. This pattern reflects that the strengthening of the ICI in the previous period tends to be followed by further strengthening in the subsequent period, both in the short and long term. This finding aligns with the characteristics of financial markets that exhibit persistence or continuity in stock price behavior over time and is consistent with previous studies showing that the ICI in the prior period tends to positively influence the current ICI value [65, 76, 77].

The Bank Indonesia interest rate (BI Rate) exhibits a nonlinear relationship with the Indonesia Composite Index (ICI). In general, an increase in the BI Rate has a negative impact on the ICI because it raises borrowing costs, reduces liquidity, and encourages investors to shift their portfolios from risky assets such as stocks to safer fixed-income instruments like deposits and bonds ([80]). This finding is consistent with studies conducted by Winarko et al. [81] and Robiyanto [82], which also found that the BI Rate tends to negatively affect ICI movements. However, under certain conditions—particularly when the interest rate is above the average or exceeds 5.75%—the effect of the BI Rate on the ICI actually becomes positive. This indicates that the relationship between the BI Rate and the ICI is not direct but may be influenced by other factors such as market segmentation and investor expectations regarding economic prospects ([83, 84]). When market segmentation is favorable and investor expectations toward the stock market improve, the ICI can continue to strengthen despite high interest rates. This phenomenon reflects the presence of a threshold effect and shows that the dynamics of the Indonesian stock market are not solely determined by monetary policy but are also influenced by investor sentiment and behavior in the financial market [85, 86]. In a broader context, stable interest rate policies and positive market responses to controlled monetary conditions contribute to sustainable economic growth. The inflation variable also shows a negative effect on the ICI across all testing schemes. Rising inflation tends to suppress ICI movements because it reduces public purchasing power and increases production costs for companies [87]. In addition, it can decrease the attractiveness of investment in the stock market. Inflationary pressures often prompt monetary authorities to raise interest rates as a measure to control prices, which ultimately weakens stock market performance by increasing the cost of capital and reducing liquidity [88]. This finding is consistent with the results of studies by Hayati et al. [89] and Hermanto et al. [90], which indicate that price stability is an important factor in maintaining capital market performance. Price stability fosters a healthier investment climate, increases investor confidence, and sustains productive economic activities that form the foundation for decent employment and inclusive economic growth. Therefore, the negative relationship between inflation and ICI underscores the importance of monetary policies oriented toward economic stability to achieve sustainable economic development. Effective inflation control not only maintains macroeconomic stability but also supports the achievement of Decent Work and Economic Growth, one of the goals of the Sustainable Development Goals (SDGs).

Overall, the results of this study indicate that the STSR model using LS-Spline is capable of providing superior performance in modeling and predicting ICI movements. This model not only shows a high level of accuracy and good stability across various time horizons, but also successfully represents nonlinear relationships between macroeconomic variables and the stock market that cannot be adequately explained by conventional parametric models such as ARIMA and ARIMAX. With its adaptive ability to changes in data structure and market dynamics, this approach can be a more flexible and reliable alternative in financial time series analysis, especially in the context of emerging markets such as Indonesia.

6. Conclusion

The results of this study indicate that the STSR model using LS-Spline is capable of being a flexible and effective framework in the Semiparametric Time Series Regression (STSR) model. This model integrates the ICI variable in the previous period as an autoregressive parametric component, as well as the BI Rate and Inflation as nonparametric components, so that it can capture the structure of linear and nonlinear relationships in financial time series data. Empirical results from four data sharing schemes representing short, medium, and long-term prediction horizons indicate that the STSR model using LS-Spline has a very high level of prediction accuracy, with MAPE and sMAPE values below 10% and MASE values less than 1. Residual diagnostic tests through ACF and PACF plots also indicate that the model has met the white noise assumption, meaning that the temporal dependencies in the data have been well modeled. When compared with classical models such as ARIMA and ARIMAX, which do not meet the diagnostic criteria or have insignificant coefficients, the STSR model using Ls-Spline is proven to be more flexible, robust, and has superior predictive capabilities.

Although the results of this study confirm the potential of the LS-Spline approach in modeling and forecasting the ICI, further methodological development and broader validation are still needed to ensure empirical credibility and stronger generalizability. Future research is recommended to include larger data sets, the addition of other macroeconomic variables, or cross-country comparisons to strengthen the results. These improvements will position this study as a significant contribution to the literature on econometric forecasting and sustainable finance. Furthermore, the model's ability to generate accurate and adaptive predictions can also support data-driven decision-making in line with the SDGs, particularly in promoting Decent Work and Economic Growth.

Acknowledgement

This research was funded by:

1. Indonesia Endowment Fund for Education (LPDP), Ministry of Finance of Republic Indonesia.
2. Center for Higher Education Funding and Assessment, Ministry of Higher Education, Science, and Technology of Republic Indonesia, and
3. The Indonesian Education Scholarship (BPI) grant number: 00198/J5.2.3./BPI.06/9/2022.

REFERENCES

1. M. Ramli, I. N. Budiantara, V. Ratnasari, *A method for Parameter Hypothesis Testing in Nonparametric Regression with Fourier Series Approach*, MethodsX, vol. 11, no. 102468, pp. 1–13, 2023.
2. J. Yan, *Almost Sure Convergence for Weighted Sums of WNOD Random Variables and Its Applications to Nonparametric Regression Models*, Communications in Statistics - Theory and Methods, vol. 47, no. 16, pp. 3893–3909, 2018.
3. B. Lestari, Fatmawati, I.N. Budiantara, *Spline Estimator and Its Asymptotic Properties in Multiresponse Nonparametric Regression Model*, Songklanakarin Journal of Science & Technology, vol. 42, no. 3, pp. 533–548, 2020.
4. A. Caron, G. Baio, I. Manolopoulou, *Estimating Individual Treatment Effects Using Non-Parametric Regression Models: A Review*, Journal of the Royal Statistical Society Series A: Statistics in Society, vol. 185, no. 3, pp. 1115–1149, 2022.
5. A. T. Ampa, I. N. Budiantara, I. Zain, *Kernel Multivariable Semiparametric Regression Model in Estimating the Level of Open Unemption in East Java Province*, Journal of Physics: Conference Series, vol. 1899, no. 1, p. 012127, 2021.
6. N. Chamidah, B. Zaman, L. Muniroh, B. Lestari, *Multiresponse Semiparametric Regression Model Approach to Standard Growth Charts Design for Assessing Nutritional Status of East Java Toddlers*, Communications in Mathematical Biology and Neuroscience, vol. 2023, no. 30, pp. 1–23, 2023.
7. N. Chamidah, B. Lestari, A. Y. Wulandari, L. Muniroh, *Z-Score Standard Growth Chart Design of Toddler Weight Using Least Square Spline Semiparametric*, AIP Conference Proceedings, vol. 2329, no. 1, p. 060031, 2021.
8. N. Chamidah, B. Lestari, A. Massaid, T. Saifudin, *Estimating Mean Arterial Pressure Affected by Stress Scores Using Spline Nonparametric Regression Model Approach*, Communications in Mathematical Biology and Neuroscience, vol. 2020, no.72, pp. 1–12, 2020.
9. A. Ramadan, N. Chamidah, I.N. Budiantara, B. Lestari, D. Aydin, *Method for Modelling the Number of HIV and AIDS Cases Using Least Square Spline Biresponse Nonparametric Negative Binomial Regression*, MethodsX, vol. 13, no. 103336, pp. 1–11, 2025.
10. W. Ramadan, N. Chamidah, B. Zaman, L. Muniroh, B. Lestari, *Standard Growth Chart of Weight for Height to Determine Wasting Nutritional Status in East Java Based on Semiparametric Least Square Spline Estimator*, IOP Conference Series: Materials Science and Engineering, vol. 546, no. 5, p. 052063, 2019.
11. 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*, International Journal of Innovation, Creativity and Change, vol. 5, no. 3, pp. 1200–1216, 2019.
12. M. F. Hidayat, R. F. Adji Achmad, Solimun, *Estimation of Truncated Spline Function in Non-parametric Path Analysis Based on Weighted Least Square (WLS)*, IOP Conference Series: Materials Science and Engineering, vol. 546, no. 5, p. 052027, 2019.
13. Fatmawati, I.N. Budiantara, B. Lestari, *Comparison of Smoothing and Truncated Spline Estimators in Estimating Blood Pressure Models*, International Journal of Innovation, Creativity and Change, vol. 5, no. 3, pp. 1177–1199, 2019.

14. N. Chamidah, B. Lestari, I.N. Budiantara, T. Saifudin, R. Rulaningtyas, A. Aryati, P. Wardani, *Consistency and Asymptotic Normality of Estimator for Parameters in Multiresponse Multipredictor Semiparametric Regression Model*, Symmetry, vol. 14, no. 2, 336, 2022.
15. N. Chamidah, B. Lestari, H. Susilo, M.Y. Alsagaff, I.N. Budiantara, D. Aydin, *Spline Estimator in Nonparametric Ordinal Logistic Regression Model for Predicting Heart Attack Risk*, Symmetry, vol. 16, no.11, 1440, 2024.
16. M. Maharani, D.R.S. Saputro, *Generalized Cross Validation (GCV) in Smoothing Spline Nonparametric Regression Models*, Journal of Physics: Conference Series, vol. 1808, no. 1, p. 012053, 2021.
17. B. Lestari, N. Chamidah, D. Aydin, E. Yilmaz, *Reproducing Kernel Hilbert Space Approach to Multiresponse Smoothing Spline Regression Function*, Symmetry, vol. 14, no. 11, 2227, 2022.
18. B. Lestari, N. Chamidah, I.N. Budiantara, D. Aydin, *Determining Confidence Interval and Asymptotic Distribution for Parameters of Multiresponse Semiparametric Regression Model Using Smoothing Spline Estimator*, Journal of King Saud University - Science, vol. 35, no. 5, 102664, 2023.
19. D. Aydın, E. Yılmaz, N. Chamidah, B. Lestari, *Right-Censored Partially Linear Regression Model with Error in Variables: Application with Carotid Endarterectomy Dataset*, The International Journal Biostat., vol. 20, no. 1, pp. 245–278, 2024.
20. E. L. Montoya, N. Ulloa, V. Miller, *A Simulation Study Comparing Knot Selection Methods With Equally Spaced Knots in a Penalized Regression Spline*, International Journal of Statistics and Probability, vol. 3, no. 3, pp. 96–110, 2014.
21. N. Chamidah, B. Lestari, T. Saifudin, R. Rulaningtyas, P. Wardhani, I.N. Budiantara, *Estimating the Number of Malaria Parasites on Blood Smears Microscopic Images Using Penalized Spline Nonparametric Poisson Regression*, Communications in Mathematical Biology and Neuroscience, vol. 2024, 60, pp. 1–19, 2024.
22. N. Chamidah, B. Lestari, H. Susilo, T. K. Dewi, T. Saifudin, N.R.A.A. Siregar, D. Aydin, *Modeling Coronary Heart Disease Risk Based on Age, Fatty Food Consumption and Anxiety Factors Using Penalized Spline Nonparametric Logistic Regression*, MethodsX, vol. 14, no. 103320, pp. 1–11, 2025.
23. M. Hasyim, N. Chamidah, T. Saifudin, B. Lestari, *Nonparametric Bi-Response Ordinal Logistic Regression Model for Diabetes Mellitus and Hypertension Risks Based on Multivariate Adaptive Regression Spline*, Data and Metadata, vol. 4, no. 912, pp. 1–18, 2025.
24. G. Zhao, Y. Ma, *Robust Nonparametric Kernel Regression Estimator*, Statistics & Probability Letters, vol. 116, pp. 72–79, 2016.
25. J. Gao, I. Gijbels, *Bandwidth Selection in Nonparametric Kernel Testing*, Journal of the American Statistical Association, vol. 103, no. 484, pp. 1584–1594, 2008.
26. B. Lestari, Fatmawati, I. N. Budiantara, and N. Chamidah, *Estimation of Regression Function in Multi-Response Nonparametric Regression Model Using Smoothing Spline and Kernel Estimators*, Journal of Physics: Conference Series, vol. 1097, no. 1, p. 012091, 2018.
27. B. Lestari, Fatmawati, I. N. Budiantara, and N. Chamidah, *Smoothing Parameter Selection Method for Multiresponse Nonparametric Regression Model Using Smoothing Spline and Kernel Estimators Approaches*, Journal of Physics: Conference Series, vol. 1397, no. 1, p. 012064, 2019.
28. D. Aydin, E. Yilmaz, N. Chamidah, B. Lestari, and I.N. Budiantara, *Right-Censored Nonparametric Regression with Measurement Error*, Metrika, vol. 88, no. 3, p. 183–214, 2024
29. N. Chamidah, E. Tjahjono, A.R. Fadilah, B. Lestari, *Standard Growth Charts for Weight of Children in East Java Using Local Linear Estimator*, Journal of Physics: Conference Series, vol. 1097, no. 1, p. 012092, 2018.
30. E. Ana, N. Chamidah, P. Andriani, B. Lestari, *Modeling of Hypertension Risk Factors Using Local Linear of Additive Nonparametric Logistic Regression*, Journal of Physics: Conference Series, vol. 1397, no. 1, p. 012067, 2019.
31. L. R. Cheruiyot, *Local Linear Regression Estimator on the Boundary Correction in Nonparametric Regression Estimation*, Journal of Statistical Theory and Applications, vol. 19, no. 3, pp. 460–471, 2020.
32. N. Chamidah, B. Lestari, T.N. Larasati, L. Muniroh, *Designing Z-Score Standard Growth Charts Based on Height-for-Age of Toddlers Using Local Linear Estimator for Determining Stunting*, AIP Conference Proceedings, vol. 3083, no. 1, p. 030002, 2024.
33. A. Tohari, N. Chamidah, F. Fatmawati, B. Lestari, *Modelling the Number of HIV and AIDS Cases in East Java Using Biresponse Multipredictor Negative Binomial Regression Based on Local Linear Estimator*, Communications in Mathematical Biology and Neuroscience, vol. 2021, no. 73, pp. 1–17, 2021.
34. N. Chamidah, Y.S. Yonani, E. Ana, B. Lestari, *Identification the Number of Mycobacterium Tuberculosis Based on Sputum Image Using Local Linear Estimator*, Bulletin of Electrical Engineering and Informatics (BEEI), vol. 9, no. 5, pp. 2109–2116, 2020.
35. N. Chamidah, B. Zaman, L. Muniroh, B. Lestari, *Designing Local Standard Growth Charts of Children in East Java Province Using a Local Linear Estimator*, International Journal of Innovation, Creativity and Change, vol. 13, no. 1, pp. 45–67, 2020.
36. N. Chamidah, K.H. Gusti, E. Tjahjono, B. Lestari, *Improving of Classification Accuracy of Cyst and Tumor Using Local Polynomial Estimator*, Telkomnika, vol. 17, no. 3, pp. 1492–1500, 2019.
37. N. Chamidah, B. Lestari, *Estimation of Covariance Matrix Using Multi-response Local Polynomial Estimator for Designing Children Growth Charts: A Theoretically Discussion*, Journal of Physics: Conference Series, vol. 1397, no. 1, p. 012072, 2019.
38. D. Aydin, N. Chamidah, B. Lestari, S. Mohammad, E. Yilmaz, *Local Polynomial Estimation for Multi-response Semiparametric Regression Models with Right Censored Data*, Communications in Statistics Part B: Simulation and Computation, vol. 54, no. 3, pp. 1–32, 2025.
39. V. Fibriyani, N. Chamidah, *Prediction of Inflation Rate in Indonesia Using Local Polynomial Estimator for Time Series Data*, Journal of Physics: Conference Series, vol. 1776, no. 1, p. 012065, 2021.
40. V. Fibriyani, N. Chamidah, *Prediction of Inflation in Indonesia Using Nonparametric Regression Approach Based on Local Polynomial Estimator*, Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019), Yogyakarta, Indonesia: Atlantis Press, 2020.
41. R. I. U. Hidayat, J. K. Prasetyo, B. Larasati, M. Aisharezka, N. Chamidah, *JCI Modeling in Indonesia Based on Industrial Production Index with Local Polynomial Estimator Approach*, BAREKENG: Journal of Mathematics and Its Applications, vol. 17, no. 3, pp. 1277–1286, 2023.
42. S. E. Ahmed, D. Aydin, E. Yilmaz, *Semiparametric Time-Series Model Using Local Polynomial: An Application on the Effects of Financial Risk Factors on Crop Yield*, Journal of Risk Financial Management, vol. 15, no. 3, p. 141, 2022.

43. V. Fibriyani, N. Chamidah, T. Saifudin, *Modeling Case Fatality Rate of COVID-19 in Indonesia Using Time Series Semiparametric Regression Based on Local Polynomial Estimator*, Communications in Mathematical Biology and Neuroscience, vol. 2024, no. 9, pp. 1–22, 2024.
44. T.W. Utami, N. Chamidah, T. Saifudin, B. Lestari, D. Aydin, *Estimation of Biresponse Semiparametric Regression Model for Longitudinal Data Using Local Polynomial Kernel Estimator*, Symmetry, vol. 17, no.3, 392, pp. 1–22, 2025.
45. I.F. Amri, N. Chamidah, T. Saifudin, D. Purwanto, A. Fadlurrohman, A.F. Ningrum, S. Amri, *Prediction of Extreme Weather Using Nonparametric Regression Approach with Fourier Series Estimators*, Data and Metadata, vol. 3, no. 319, pp. 1–12, 2024.
46. A.A. Anandari, E. D. Supandi, M. W. Musthofa, *Fourier Series Nonparametric Regression Modeling in the Case of Rainfall in West Java Province*, IJID (International Journal on Informatics for Development), vol. 11, no. 1, pp. 142–151, 2022.
47. I.F. Amri, N. Chamidah, T. Saifudin, B. Lestari, D. Aydin, *Forecasting Temperature of Earth Surface in Sragen Regency Using Semiparametric Regression Based on Penalized Fourier Series Estimator*, Data and Metadata, vol. 4, no. 890, pp. 1–18, 2025.
48. I.F. Amri, N. Chamidah, T. Saifudin, B. Lestari, *Determining Relative Humidity in Surakarta City Using Semiparametric Regression Based on Fourier Series Penalized Least Square*, Communications in Mathematical Biology and Neuroscience, vol. 2025, no. 61, pp. 1–17, 2025.
49. N. Chamidah, B. Lestari, I.N. Budiantara, D. Aydin, *Estimation of Multiresponse Multipredictor Nonparametric Regression Model Using Mixed Estimator*, Symmetry, vol. 16, no. 4, 386, pp. 1–25, 2024.
50. H. Husain, P.I. Rahayu, M. R. Nisardi, M.A. Al-Fadhilah, A. Husain, *emiparametric Biresponse Regression Modeling Mixed Spline Truncated, Fourier Series, and Kernel in Predicting Rainfall and Sunshine*, Statistics, Optimization and Information Computing, Issue: Online First, 2025.
51. R. Li, J. Cui, D. Fu, *A Time Series Data Regression Based on a B-spline and Stacking Framework*, Journal of Physics: Conference Series, vol. 2813, no. 1, p. 012024, 2024.
52. R. Ospina, J. A. M. Gondim, V. Leiva, C. Castro, *An Overview of Forecast Analysis with ARIMA Models During the COVID-19 Pandemic: Methodology and Case Study in Brazil*, Mathematics, vol. 11, no. 14, p. 3069, 2023.
53. W. K. Adu, P. Appiahene, S. Afrifa, *VAR, ARIMAX and ARIMA Models for Nowcasting Unemployment Rate in Ghana Using Google Trends*, Journal of Electrical Systems and Information Technology, vol. 10, no. 12, pp. 1–16, 2023.
54. J. Gao, M. King, *Adaptive Testing in Continuous-Time Diffusion Models*, Econometric Theory, vol. 20, no. 5, pp. 844–882, 2004.
55. S. X. Chen, J. Gao, C. Y. Tang, *A test for model specification of diffusion processes*, Annals of Statistics, vol. 36, no. 1, pp. 167–198, 2008.
56. C. Dong, J. Gao, *Specification Testing Driven by Orthogonal Series for Nonlinear Cointegration with Endogeneity*, Econometric Theory, vol. 34, no. 4, pp. 754–789, 2018.
57. P. K. Dey, *Growth Analysis of Stock Market in India: An Empirical Study*, International Journal of Trend in Scientific Research and Development (IJTSRD), vol. 2, no. 3, pp. 386–392, 2018.
58. R.D.P. Aji, N. Abundanti, *The Effect of Asia Regional Stock Price Index on the Indonesia Composite Index (ICI) on the Indonesia Stock Exchange*, European Journal of Business & Management Research (EJBMR), vol. 7, no. 2, pp. 100–106, 2022.
59. A. Wibowo, H. Pujianto, D.R.S. Saputro, *Nonlinear Autoregressive Exogenous Model (NARX) in Stock Price Index's Prediction*, 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta: IEEE, pp. 26–29, 2017.
60. I. Saluza, E. Munarsih, F. Faradillah, L.D. Anggraini, *Prediction Indonesia Composite Index Using Integration Decomposition-Neural Network Ensemble During Vuca Era*, BAREKENG: Journal of Mathematics and Its Application, vol. 18, no. 4, pp. 2721–2736, 2024.
61. K.N. Khikmah, K. Sadik, I. Indahwati, *Transfer Function and ARIMA Model for Forecasting BI Rate in Indonesia*, BAREKENG: Journal of Mathematics and Its Application, vol. 17, no. 3, pp. 1359–1366, 2023.
62. R.P. Widhiarti, L. Anggraeni, S.H. Pasaribu, *Analysis of Investor Sentiment Impact in Indonesia Composite Stock Price Index Return Volatility*, Indonesian Journal of Business and Entrepreneurship, vol. 4, no. 3, pp. 239–248, 2018.
63. Z. Shao, *Research on the Interaction of Inflation, Stock Price and Development Trend of Trading Volume*, Advances in Social Science, Education and Humanities Research, Dalian, China: Atlantis Press, 2018.
64. P.D. Aryadi, W.P. Setyono, *Mutual Fund Performance: Impact of Stock Market, Interest Rates, and Inflation*, Academia Open, vol. 7, 2022.
65. A. T. Fitiryah, N. Chamidah, T. Saifudin. *Comparison Of Least Square Spline and ARIMA Models for Predicting Indonesia Composite Index*, BAREKENG: Journal Mathematics and Its Application, vol. 19, no. 3, p. 2169-2178, 2025.
66. M.M. Dhas, and N.S. Singh, *Optimized Haar Wavelet-based Blood Cell Image Denoising with Improved Multiverse Optimization Algorithm.*, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. Vol 11, Issue 4, 2023.
67. L. Li, Y. Kang, F. Petropoulos, F.Li, *Feature-based intermittent demand forecast combinations: accuracy and inventory implications*, International Journal of Production Research, 61(22), pp. 7557–7572, 2022.
68. J. Hauke, I. Lorscheid, M. Meyer, *Individuals and their interactions in demand planning processes: an agent-based, computational testbed*, International Journal of Production Research, 56(13), pp. 4644–4658, 2017.
69. R. J. Hyndman, A.B. Koehler, *Another look at measures of forecast accuracy*, International Journal of Forecasting, Vol. 22, Issue 4, pp. 679-688, 2006.
70. S. 1. D. Ruppert, M.P. Wand, and R. J. Carroll, *Semiparametric Regression. Cambridge University Press*, 2003.
71. 5. D. Ariesta, N. Gusriani, K. Parmikanti, *Estimasi Parameter Model Regresi Nonparametrik B-Spline Pada Angka Kematian Maternal*, Jurnal Matematika UNAND Vol. 10 No. 3, 2021.
72. J. Z. Huang, H. Shen, *Functional Coefficient Regression Models for Non-Linear Time Series: A Polynomial Spline Approach*, Journal of Statistics, Vol. 31, pp. 515-534, 2004.
73. S.K. Schnabel, *Expectile Smoothing: New Perspectives on Asymmetric Least Squares. An Application to Life Expectancy*, Proefschrift Universiteit Utrecht, Utrecht. 2011.
74. W. F. Velicer, P. C. Molenaar, *Time Series Analysis for Psychological Research*, Handbook of Psychology, Second Edition, John Wiley Sons, Inc, Hoboken, 2012.

75. B. A. Hakim, K. A. Notodiputro, Y. Angraini, L. Nissa, and A. Mualifah, *Time Series Model For Train Passenger Forecasting*, BAREKENG: Journal Mathematics and Its Application, vol. 19, no. 2, p. 755-766, 2025.
76. A. I. E. Nensi., M. Al Maida., K. A. Notodiputro., Y. Angraini., L. N. A. Mualifah, *Performance Analysis of ARIMA, LSTM, and Hybrid ARIMA-LSTM in Forecasting the Composite Stock Price Index*, Chaucy Jurnal Matematika Murni dan Aplikasi. Vol. 10 No. 2. 2025.
77. S.M. Ulyah, *AForecasting index and stock returns by considering the effect of Indonesia pre-presidential election 2019 using ARIMAX and VARX approaches*, Journal of Physics: Conference Series, vol. 1277, p. 012053, 2019.
78. S. Amilia, N. Herrhyanto, E. Puspita, *Application of Semiparametric Spline Regression Model in Modelling Human Development Index in 2016*, Journal of Physics: Conference Series, vol. 1280, no. 2, p. 022042, 2019.
79. J.J.M. Moreno, A.P. Pol, A.S. Abad, *Using the R-MAPE Index as a Resistant Measure of Forecast Accuracy*, Psicothema, vol. 25, no. 4, pp. 500–506, 2013.
80. D. Suselo, R. Agustina *The Analysis of Influence of Inflation, Exchange Rate and BI Rate Toward The Jakarta Composite Index in Indonesia Stock Exchange*, JPIM (Jurnal Penelitian Ilmu Manajemen), vol. 7. pp. 261-270, 2022
81. H.B. Winarko, J.C. Wahono, Y. Yuniningsih, S.T. Pannindriya, *How Do Inflation Rate, BI Rate, and Balance of Trade Directly Affect IDR to USD Exchange Rate and Indirectly Affect IDX Composite Index in Initial Stage of Covid-19 Outbreak?*, Research in World Economy, vol. 12, no. 1, pp. 56–67, 2021.
82. R. Robiyanto, *The Effect of Gold Price Changes, USD/IDR Exchange Rate Changes and Bank Indonesia (BI) Rate on Jakarta Composite Index (JCI)'s Return and Jakarta Islamic Index (JII)'s Return*, Jurnal Manajemen dan Kewirausahaan, vol. 20, no. 1, pp. 45–52, 2018.
83. D. I. Gojali, E. H. Juniwati, L. N. Pratiwi *The effect of narrow money (M1), interest rate of Bank Indonesia, inflation, exchange rate towards composite stock price index*, Indonesian Journal of Economics and Management, vol. 1, no. 3, pp. 561-577, 2021.
84. N. Rahayu, R. Sanjaya *Analisis Pengaruh BI-Rate, Kurs Rupiah dan Laba Akuntansi terhadap Indeks Harga Saham Gabungan*, Jurnal Rimba : Riset Ilmu manajemen Bisnis dan Akuntansi, vol. 3, no. 1, 2025.
85. D. Adrian, R.E. Tarigan, *The Effect of Revenue and Market Segmentation Level Towards Venture Capital Investment in Mobile Application Business*, COMMIT (Communication and Information Technology) J., vol. 8, no. 1, pp. 10–16, 2014.
86. R.M. Irani, D. Oesch, *Financial Constraints and Corporate Disclosure*, SSRN Electronic Journal, 2014, doi: 10.2139/ssrn.2530613.
87. H. Priyatna, I. Suryadi, Ersanti, *Facing Global Inflation: Economic Strategies to Strengthen People's Purchasing Power*, SJ : Majority Science Journal. vol. 3. 73-81, 2025.
88. I. Dridi, A. Boughrara, *Flexible Inflation Targeting and Stock Market Volatility: Evidence from Emerging Market Economies*, Economic Modelling, vol. 126, 106420, 2023.
89. R. Hayati, P. C. Jamil, Azmansyah, *Macroeconomics and Jakarta Composite Index*, Proc. Second Int. Conf. Soc., Econ., Edu. Humanity - ICoSEEH; ISBN 978-989-758-464-0, SciTePress, pp. 117–122, 2020.
90. T. Hermanto, H. Siregar, R. Sembel, R. Nurmalina, *Factors Affecting the Volatility of the Jakarta Composite Index before and after the Merger of Two Stock and Bond Markets in Indonesia*, Asian Social Science, vol. 10, no. 22, pp. 91–99, 2014.