CAUCHY – Jurnal Matematika Murni dan Aplikasi
Volume 10 (2) (2025), Pages 986-997
p-ISSN: 2086-0382; e-ISSN: 2477-3344

Forecasting Indonesia's Composite Stock Price Index with Semiparametric Cubic and Local Gaussian Polynomials

Mita Kornilia Dewi* and Nanang Susyanto

Department of Mathematics, Gadjah Mada University, Indonesia

Abstract

The Composite Stock Price Index (CSPI) serves as a crucial indicator for assessing the performance of the Indonesian capital market, reflecting both economic conditions and investor confidence. Its movements are influenced by macroeconomic factors such as exchange rates, inflation, interest rates, and commodity prices, including oil and gold. Parametric models often fail to capture nonlinear patterns, whereas nonparametric approaches lack efficiency and interpretability. To address this gap, this study develops a semiparametric regression model that integrates a cubic polynomial for parametric effects with local polynomial estimators using Gaussian kernels for nonparametric effects. The results show that the semiparametric model is effective, yielding an MSE of 0.569747, a MAPE of 8.60%, and an R^2 of 85%. This confirms its ability to capture nonlinear dynamics in the stock market. Moreover, the model provides accurate forecasting and practical insights for investors in portfolio strategies as well as for policymakers in managing financial market stability.

Keywords: Cubic polynomial; CSPI; Local Gaussian kernel; Semiparametric regression.

1 Introduction

The Composite Stock Price Index (CSPI) is a leading indicator that reflects the movements of share price of all companies listed on the Indonesia Stock Exchange [1]. This index is widely utilized by market participants as a benchmark to evaluate macroeconomic conditions and to gauge investor confidence in the stability of financial markets [2]. Several economic variables, including inflation, interest rates, exchange rates, and global commodity price such as oil and gold, are known to significantly influence fluctuations in the CSPI [3].

The CSPI and its relationship with various macroeconomic variables have been modeled in numerous previous studies using classical statistical approaches, such as multiple linear regression and the Partial Adjustment Model (PAM) [4]. While these methods effectively capture linear relationships, they exhibit limitations when addressing the nonlinear patterns inherent in capital market data. As an alternative, nonparametric approaches, such as kernel-based methods, have been adopted owing to their flexibility in modeling nonlinear relationships among economic variables. This is evident in studies that reveal varying relationships between inflation and sectoral stock indices across economies experiencing different inflationary pressures [5]. However, nonparametric models often encounter challenges related to the interpretability of coefficients and estimation efficiency, particularly when dealing with relatively small sample sizes [6].

* **Corresponding author**. E-mail: mitakorniliadewi@mail.ugm.ac.id

Submitted: August 20, 2025 Reviewed: September 12, 2025 Accepted: October 12, 2025
DOI: https://doi.org/10.18860/cauchy.v10i2.36180

To address the limitations of both approaches, semiparametric regression has emerged as a robust solution that combines the strengths of parametric and nonparametric models within a unified analytical framework. This method enables explicit interpretation of parameters through the parametric component, while simultaneously providing the flexibility to capture nonlinear relationships through the nonparametric component [7], [8]. The use of this approach has gained increasing attention, particularly through the application of combined estimators such as truncated splines and Fourier series, which have demonstrated improved estimation accuracy in the analysis of longitudinal data and complex time series [9]. In practical applications, semiparametric regression has also been successfully employed in various economic studies, including modeling life satisfaction as influenced by factors such as country, age, and gender; estimating property price; and forecasting inflation rates [10], [11], [12].

The effectiveness of semiparametric regression extends beyond economics, having also been demonstrated across diverse social and environmental contexts. For example, a time series semiparametric regression model based on least-squares spline estimation has been applied successfully to predict agricultural production in Indonesia [13]. Similarly, the multiresponse approach based on local polynomial methods has shown reliable performance in analyzing complex socio economic data [14]. Furthermore, semiparametric models have been utilized to model various environmental phenomena, including rainfall, air temperature, and COVID-19 related mortality [15], [16], [17].

However, previous semiparametric research has primarily focused on spline and Fourier series approaches. Although effective, these methods have limitations [18], [19], [20]. The spline approach is sensitive to the number and placement of knots, making it prone to overfitting and underfitting. Meanwhile, Fourier series are more suitable for seasonal or periodic data patterns but are less adept at capturing random and irregular market fluctuations. To date, the application of more adaptive combinations of parametric and nonparametric components remains limited, especially within the context of the Indonesian stock market [21], [22].

To the best of our knowledge, no previous research has specifically applied a semiparametric regression approach combining cubic and local polynomials with Gaussian weighting functions to predict the movement of the CSPI in Indonesia. Cubic polynomials were selected for their ability to capture complex nonlinear patterns more effectively than lower-order polynomials, while still maintaining interpretable coefficients [23]. Concurrently, local polynomials with a Gaussian kernel were utilized to generate smooth, adaptive estimates at each data point without imposing a specific global structure [24]. This integrated approach is hypothesized to offer advantages over traditional spline or Fourier methods by effectively balancing global interpretability with local flexibility [25].

2 Materials and Methods

This study utilizes secondary data comprising 4,813 daily observations retrieved from Yahoo Finance, covering the period from February 1, 2005, to November 25, 2024. With such extensive temporal coverage and high-frequency data, this research represents one of the more comprehensive studies in the context of the Indonesian stock market. The research data employed are summarized in Table 1.

Table 1: Research Variables

Symbol	Variable Type	Variable	Unit
y	Response	CSPI	Points
x_1	Parametric	Exchange Rate	IDR/USD
x_2	Nonparametric	Oil price	USD/Barrel
x_3	Nonparametric	Gold price	USD/Ounce

The selection of predictor variables is based on both theoretical considerations in economics and the statistical characteristics of their relationship with the Jakarta Composite Index (CSPI), the response variable. From an economic perspective, exchange rates play a crucial role in determining international trade competitiveness and influencing foreign capital flows, thereby directly impacting capital market performance. Given the tendency for a linear relationship between the exchange rate and the CSPI, the exchange rate is modeled as a parametric component.

In contrast, world oil price and gold price are dynamic global economic indicators that tend to exhibit nonlinear relationships with stock market movements. Oil price can influence both production costs and inflation rates, while gold price are often perceived as safe-haven assets during periods of economic uncertainty. Owing to the complexity and potential nonlinearity in their relationship with the CSPI, these variables are modeled as nonparametric components, allowing greater flexibility in capturing their patterns.

This study employs a semiparametric model that integrates cubic polynomial parametric regression with local polynomial nonparametric regression using a Gaussian kernel function. The model is specified as follows:

$$\begin{aligned} y_i = \beta_0 + \beta_1 x_{1(i-1)} + \beta_2 x_{1(i-1)}^2 + \beta_3 x_{1(i-1)}^3 + \sum_{j=0}^{10} \alpha_j (x_{2(i-1)} - x_2)^j K\left(\frac{x_{2(i-1)} - x_2}{h_2}\right) \\ + \sum_{k=0}^{10} \gamma_k (x_{3(i-1)} - x_3)^k K\left(\frac{x_{3(i-1)} - x_3}{h_3}\right) + \epsilon_i \end{aligned} \tag{1}$$

where:

y_i	:	response variable of the i-th observation,
$\beta_1, \beta_2, \beta_3$	:	parameters for the parametric component of x_1,
x_2	:	evaluation point of x_2 used to predict y_i,
x_3	:	evaluation point of x_3 used to predict y_i,
α_j	:	parameters for the nonparametric component of x_2,
γ_k	:	parameters for the nonparametric component of x_3,
K	:	kernel function,
h_2, h_3	:	bandwidths of the kernel function,
ϵ_i	:	error term of the i-th observation, assumed to follow $N(0, \sigma^2)$.

The general form of the semiparametric model is as follows [26]:

$$y_i = f(x_i; \beta) + m(z_i) + \epsilon_i, \quad i = 1, 2, \ldots, n, \tag{2}$$

where:

y_i	:	response variable of the i-th observation,
x_i	:	predictor variable associated with the parametric component,
z_i	:	predictor variable associated with the nonparametric component,
$f(x_i; \beta)$	:	parametric regression function of x_i with parameter vector β,
$m(z_i)$	:	nonparametric regression function of z_i,
ϵ_i	:	error term, assumed to follow $\mathcal{N}(0, \sigma^2)$.

The model comprises two main components to be estimated: the parametric parameters and the nonparametric regression function. Estimation proceeds sequentially, starting with the parametric component using the Least Squares Method, which minimizes the sum of squared errors:

$$S(\beta) = \sum_{i=1}^{n} \left(y_i - f(x_i; \beta)\right)^2. \tag{3}$$

The estimated parametric regression function:

$$\hat{y}_i = f(x_i; \hat{\beta}). \tag{4}$$

Next, the nonparametric regression function $m(\cdot)$ is estimated without assuming an explicit functional form, typically through kernel-based local polynomial regression. At a given point x^*, $m(x)$ is locally approximated by a degree-p polynomial, leading to the following estimator:

$$\hat{m}(x^*) = \hat{a}_0(x^*) = \mathbf{e}_1^T(\mathbf{X}^T\mathbf{K}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{K}\mathbf{y}, \tag{5}$$

where $K(\cdot)$ denotes the kernel function, commonly chosen as the Gaussian kernel.

A critical aspect of kernel regression is the selection of the bandwidth parameter (h), which governs the trade-off between bias and variance in the estimation. If h is too small, the estimator becomes excessively sensitive to noise, leading to overfitting, whereas a large h oversmooths the data and conceals important local patterns. Therefore, bandwidth selection plays a central role in ensuring reliable nonparametric estimation.

In this study, the model was implemented using 5-fold cross-validation, in which the dataset is divided into five nearly equal folds [27]. In each iteration, four folds are used for training to estimate model parameters, while the remaining fold serves as the test set, and the process is repeated so that each fold acts as a test set once. The optimal bandwidth is determined through this data-driven approach to minimize estimation error. Alternatively, a rule-of-thumb formula can be used:

$$h = 1.06 \cdot \hat{\sigma} \cdot n^{-1/5}, \tag{6}$$

where $\hat{\sigma}$ is the standard deviation of the predictor and n is the sample size [28]. To ensure robustness, a sensitivity analysis is conducted to evaluate how variations in h affect the stability and accuracy of the regression estimates.
Finally, the semiparametric regression model combines both components as follows:

$$\hat{y}_i = \hat{f}(x_i) + \hat{m}(z_i), \quad i = 1, 2, \ldots, n, \tag{7}$$

where $\hat{f}(x_i)$ derives from $\hat{\beta}$ and $\hat{m}(z_i)$ is estimated via local smoothing. This approach integrates the structural interpretability of parametric models with the flexibility of nonparametric techniques to capture nonlinear relationships.

3 Results and Discussion

Based on the model structure outlined in the previous section, the next step involves applying the estimation procedure to the research data. The results of the analysis are presented in this section, beginning with the descriptive statistics for each variable. A summary of the descriptive statistics is provided in Table 2.

Table 2: Descriptive Statistics of Variables

Variable	Minimum	Median	Mean	Maximum
CSPI (points)	995	4,832	4,491	7,905
Exchange Rate (IDR/USD)	888	12,023	12,023	16,505
Oil Price (USD/Barrel)	-38	72	72	145
Gold Price (USD/Ounce)	413	1,306	1,336	2,788

Based on Table 2, the movements of the CSPI, exchange rate, oil prices, and gold prices during 2005–2024 exhibit a relatively stable long-term trend, although they remain highly influenced by global turmoil. The CSPI followed a symmetric upward trajectory, while the rupiah exchange rate depreciated to 16,505 IDR/USD in March 2020 as a consequence of the COVID-19 pandemic. The sharpest volatility occurred in oil prices, which briefly turned negative in April 2020, reflecting the steep decline in global energy demand. In contrast, gold maintained a consistent upward trend, reaching USD 2,788 per ounce, thereby reinforcing its role as a safe-haven asset amid uncertainty. Overall, these dynamics highlight the sensitivity of the Indonesian market to global crises, particularly the COVID-19 pandemic.

To address differences in measurement scales across variables, this study applies data standardization using the Z-score method. This approach transforms each variable to have a mean of zero and a standard deviation of one, ensuring a balanced contribution of all variables in the model analysis [27]. Following normalization, the next step is to visualize each variable over the period from February 1, 2005, to November 25, 2024, as presented in Fig. 1.

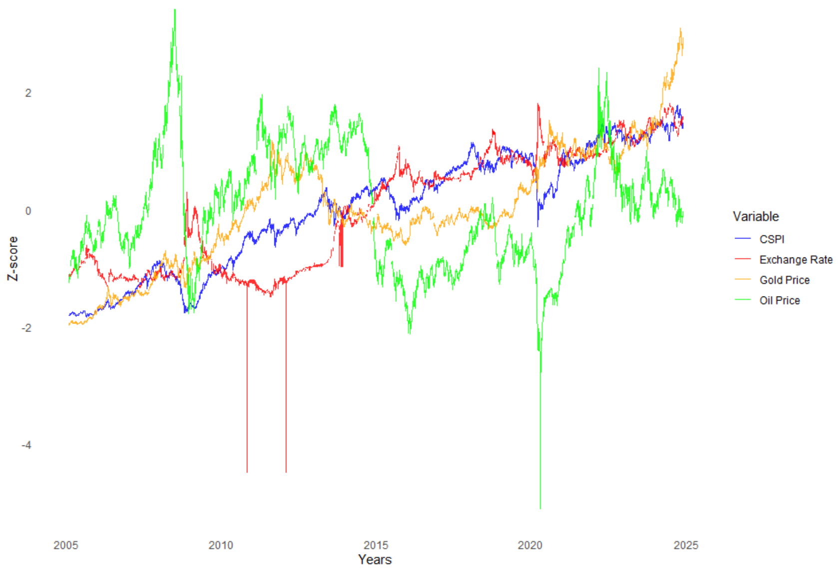

Figure 1: The relationship between variables and the Z-score

The Z-score visualization illustrates the relationship between the CSPI, exchange rate, oil, and gold over the period 2005–2024. The CSPI generally followed an upward trajectory in a relatively symmetric pattern, although it experienced notable declines during the 2008 global financial crisis and the 2019–2020 COVID-19 pandemic. The rupiah tended to depreciate against the US dollar and exhibited an inverse relationship with the CSPI through foreign capital flows and market responses to global uncertainty. Meanwhile, oil prices displayed considerably higher volatility than the CSPI, with sharp declines in 2008 and even negative values in April 2020 due to the collapse in global energy demand.

Gold prices, in contrast showed a consistent upward trend as a safe-haven instrument, particularly during crisis periods such as 2008 and 2020 when the CSPI declined. This pattern reflects a shift in investment allocation from risky assets to safer ones. Overall, the correlations between the CSPI, exchange rate, oil, and gold highlight the sensitivity of the Indonesian capital market to global dynamics and underscore the importance of considering external factors in investment analysis.

Subsequently, the variables are standardized using K-fold cross-validation to normalize their scales, ensuring stable and accurate estimation. The optimal polynomial order and bandwidth are then determined using the Generalized Cross-Validation (GCV) criterion. Table 3 summarizes the GCV values for polynomial orders at the 0.10 bandwidth.

Table 3: GCV Values for First-Order Polynomial

Bandwidth	Order	GCV
0.10	1	2.8262
0.10	2	2.8381
0.10	3	2.8500
0.10	4	2.8620
0.10	5	2.8740
0.10	6	2.8862
0.10	7	2.8984
0.10	8	2.9107
0.10	9	2.9231
0.10	10	2.9355

The minimum GCV is attained at order 1 with a bandwidth of 2.8262, indicating the best bias-variance tradeoff. Fig. 2 visualizes the GCV across bandwidths and orders, showing that smaller bandwidths and lower orders generally yield better performance.

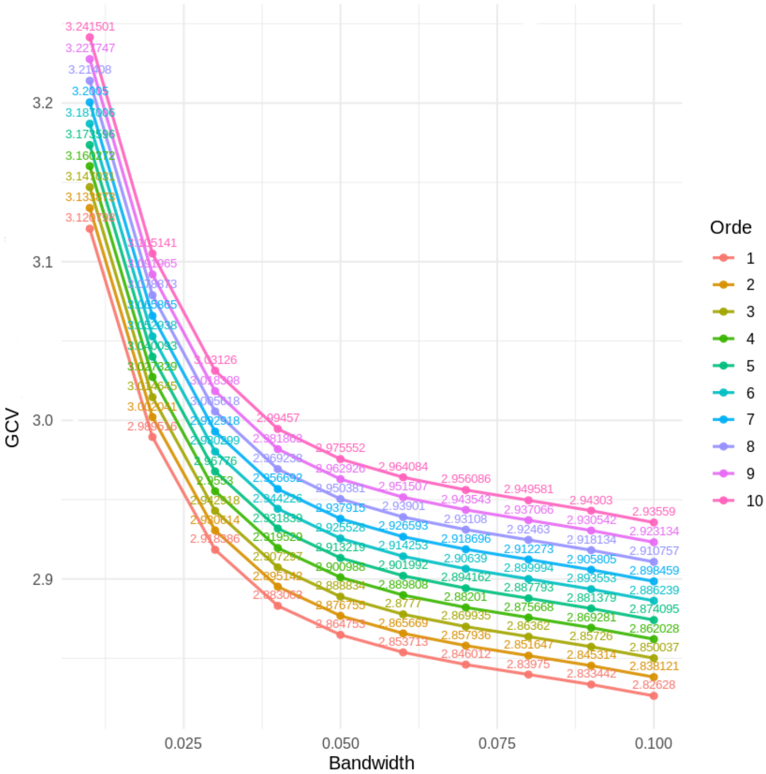

Figure 2: GCV Plot Across Polynomial Orders

Estimated parameters for the parametric and nonparametric components are shown in Tables 4 and 5, respectively.

Table 4: Parametric Parameters of Exchange Rate

Parameter	Estimate
$\hat{\beta}_0$	-0.089266
$\hat{\beta}_1$	0.904031
$\hat{\beta}_2$	0.087649
$\hat{\beta}_3$	-0.067659

Table 5: Nonparametric Parameters of Oil and Gold price

Oil (x_2)		**Gold** (x_3)	
j	$\hat{\alpha}_j$	k	$\hat{\gamma}_k$
0	−0.118143	0	0.325147
1	0.476209	1	0.772789
2	0.599326	2	−1.223115
3	−0.188590	3	−0.261291
4	−0.368967	4	0.928090
5	0.055631	5	−0.144415
6	0.067410	6	−0.273105
7	−0.007804	7	0.103602
8	−0.004835	8	0.019076
9	0.000388	9	−0.014647
10	0.000108	10	0.001889

While the primary focus of this study is a semiparametric model combining cubic and local polynomials, its performance is also benchmarked against the ARIMAX time series model. In addition, the model is compared with a Long Short-Term Memory (LSTM) network. This comparison evaluates the effectiveness of the proposed model against methods specifically designed to handle temporal dependencies.

The ARIMAX model (Autoregressive Integrated Moving Average with Exogenous Variables) is used in the form ARIMAX(1,1,1), which combines autoregressive, differentiation, and moving-average components, as well as involving exogenous variables [29]. The general form of the model used is

$$y_t = \mu + \phi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \varepsilon_t, \tag{8}$$

where μ is the constant, ϕ_1 and θ_1 are the first-order AR and MA parameters, respectively, β_j is the coefficient of the exogenous variable x_{jt}, and ε_t is the error term.

Additionally, the proposed model is benchmarked against a neural network-based approach, specifically the Long Short-Term Memory (LSTM) model [30]. The prediction output is formulated as:

$$\hat{y} = \mathbf{h}_T \cdot \mathbf{W}_{\text{dense}} + b_{\text{dense}}, \tag{9}$$

where $\mathbf{h}_T$ is the final hidden state (output) of the LSTM layer, $\mathbf{W}_{\text{dense}}$ is the weight matrix of the subsequent dense layer, and b_{dense} is its bias term.

The performance evaluation results of the three models are presented in Table 6.

Table 6: Model Performance Evaluation

Metric	**Semiparametric**	**ARIMAX (1,1,1)**	**LSTM**
MSE	0.569747	3.030235	1.60487
MAPE	8.60%	7.03%	8.44%
R^2	85%	70%	74%

Based on the performance evaluation, the semiparametric model achieved the lowest MSE of 0.569747, compared to 3.030235 for ARIMAX and 1.60487 for LSTM, along with a relatively low MAPE of 8.60%. Regarding R^2, the semiparametric model also outperformed the others, attaining 85%, compared to 70% for ARIMAX and 74% for LSTM, indicating a superior ability to explain data variation. This combination of low MSE and high R^2 demonstrates that the semiparametric model provides the best predictive performance, and therefore, further research will focus on this model without additional comparisons to ARIMAX or LSTM.

Next, an evaluation was performed on Fig. 3, which illustrates the variation in MAPE with bandwidth, highlighting the optimal elimination region.

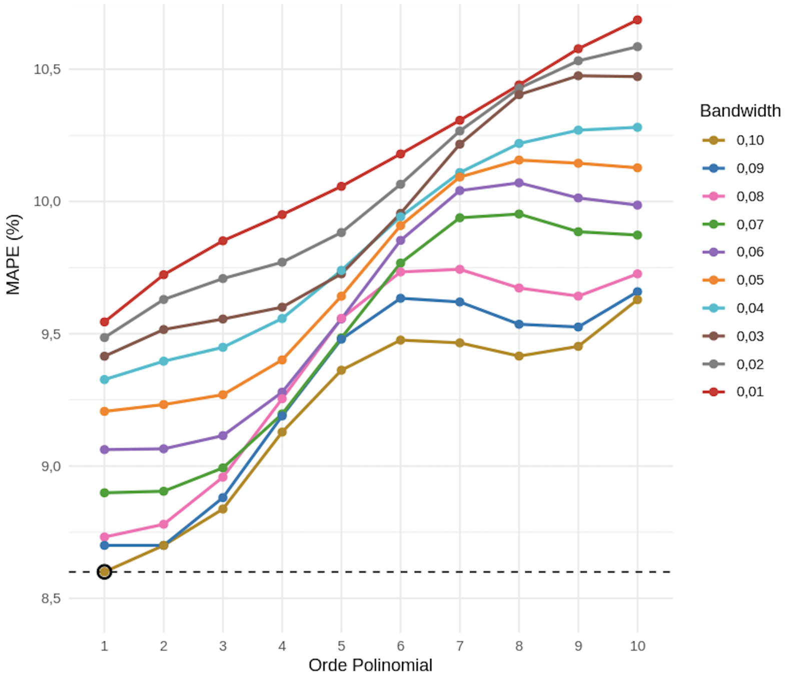

Figure 3: MAPE vs. Bandwidth

Forecasting was then performed for November 26, 2024, to July 28, 2025. Table 7 summarizes the test data.

Table 7: Test Data

No	Date	CSPI	Exchange Rate	Oil Price	Gold Price
4813	25-11-2024	7,314	15,940	68.90	2,616
4814	26-11-2024	7,206	15,901	68.77	2,620

The predicted value of the CSPI at data point 4,814 is estimated to be 7,206 (USD/IDR), with the *error* calculated as follows:

$$\begin{aligned} Error &= \text{CSPI}_{\text{Actual}} - \text{CSPI}_{\text{Predicted}} \\ &= y_{4813} - \hat{y}_{4814} \\ &= 7,246 - 7,206 \\ &= 0.040 \end{aligned}$$

Fig. 4 compares the predicted and actual CSPI values over the forecasting period. The blue line represents the actual CSPI data, while the dotted red line indicates the model's predictions. The two series closely track each other, particularly during the sharp decline in early March and the subsequent recovery, demonstrating that the model effectively captures CSPI movements. Despite minor deviations at certain points, the model overall provides reliable forecasts of CSPI dynamics.

After confirming the consistency between the actual CSPI and the model's prediction results during the observation period, the analysis proceeds with a visualization of the relationships between the predictor variables, namely the exchange rate, oil price, and gold price, and the actual CSPI. The semiparametric model exhibits high predictive accuracy for CSPI movements by effectively capturing nonlinear relationships with these macroeconomic variables.

Currency exchange rates negatively affect the CSPI, oil prices generally exert a positive influence, and gold prices display a nonlinear pattern, with extreme increases often triggering selling pressure. These findings underscore the significant role of macroeconomic factors and market sentiment, providing valuable insights for investors and policymakers. Moreover, they

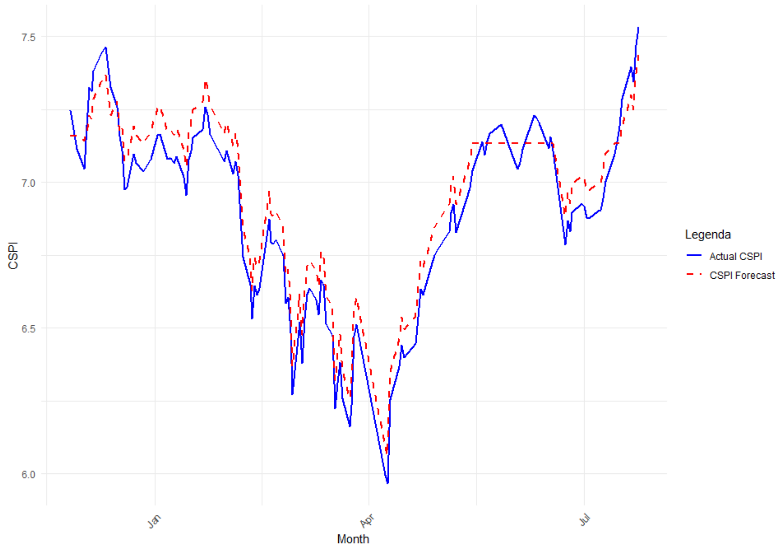

Figure 4: Actual vs. Forecasted CSPI

demonstrate that the semiparametric model can accurately predict the dynamics of the Indonesian stock market, as illustrated in Fig. 5 below.

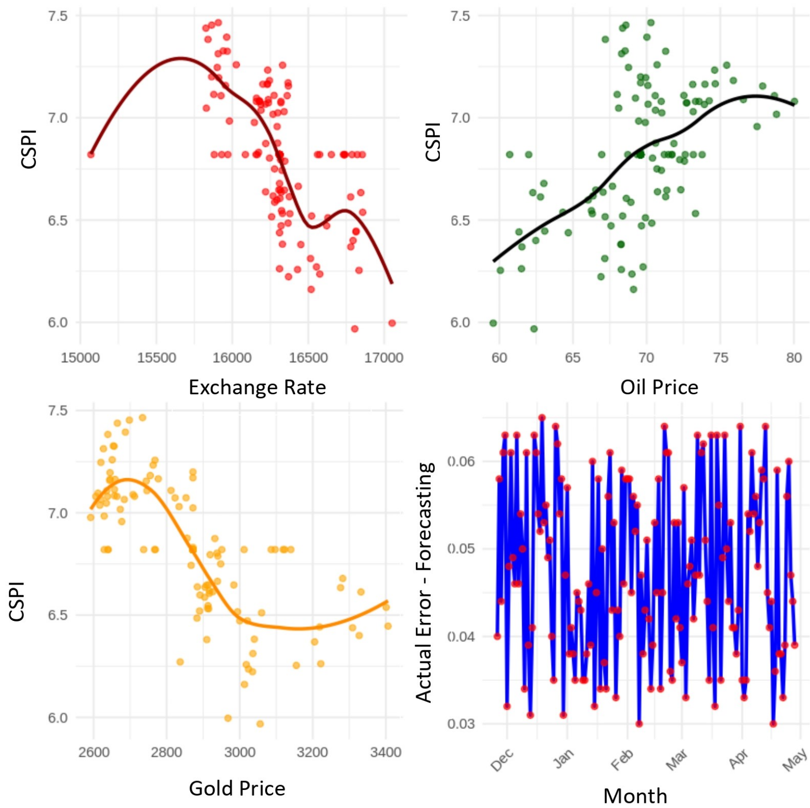

Figure 5: Variable and Error Relationship

4 Conclusion

Based on the implementation and evaluation of the semiparametric model for the Composite Stock Price Index (CSPI), several key findings emerge. The semiparametric model, which integrates cubic polynomials with local polynomials using Gaussian kernels, effectively captures the nonlinear relationships between the CSPI and major macroeconomic variables, namely the exchange rate, oil price, and gold price. The analysis reveals that an increase in the USD/IDR exchange rate tends to exert downward pressure on the CSPI, an increase in the oil price is generally associated with a rise in the CSPI, while a surge in the gold price is often accompanied by a decline in the CSPI due to investors reallocating assets to gold as a hedging instrument. Performance evaluation shows that the semiparametric model outperforms ARIMAX(1,1,1) and LSTM, with an MSE of 0.569747, a MAPE of 8.60%, and an R^2 of 85%, confirming its ability to accurately capture nonlinear patterns in the stock market.

These findings underscore the value of the semiparametric approach as a flexible method that combines the interpretability of parametric components with the adaptability of nonparametric components. The model also offers practical insights for investors and policymakers in formulating investment and risk management strategies. Nonetheless, this study is constrained by its reliance on only three predictor variables and its focus on the CSPI as a single case study. Future research is encouraged to incorporate additional macroeconomic variables, such as inflation, interest rates, or economic growth, and to apply the model to other stock indices or time periods to assess its generalizability and robustness under different market conditions.

CRediT Authorship Contribution Statement

Mita Kornilia Dewi: Data Curation, Conceptualization, Methodology, Validation, Visualization, and Project Administration. **Nanang Susyanto:** Supervisor, Review & Editing.

Declaration of Generative AI and AI-assisted technologies

Generative AI or AI-assisted technologies were not involved in the preparation of this manuscript, except for the use of Grammarly solely for grammar and spelling checking purposes. No other AI tools were used for writing, data analysis, or content generation.

Declaration of Competing Interest

The authors declare no competing interests.

Funding and Acknowledgments

The authors would like to thank MIPA UGM for the Lecturer Research Grant, Number 3954/UN1/FMIPA.1.3/KP/PT.01.03/2025, which has supported this research.

Data Availability

The datasets used in this study are publicly available from Yahoo Finance at the following links:

- Composite Stock Price Index (CSPI): Click here
- Exchange Rate (IDR/USD): Click here
- Oil Price (WTI Crude): Click here
- Gold Price: Click here

References

[1] D. B. Pinem, "Analysis of global stock exchange index, foreign exchange rate, interest rate and inflation rate influences cspi in indonesia stock exchange," *European Journal of Business and Management Research*, vol. 4, no. 6, 2019. DOI: 10.24018/ejbmr.2019.4.6.162.

[2] F. Ma, Y. Guo, J. Chevallier, and D. Huang, "Macroeconomic attention, economic policy uncertainty, and stock volatility predictability," *International Review of Financial Analysis*, vol. 81, p. 101 732, 2022. DOI: 10.1016/j.irfa.2022.102339.

[3] F. D. Sumaryana, N. Nugraha, M. Sari, and T. Heryawan, "The influence of macroeconomic fundamentals and investor sentiment on the indonesian stock exchange," *Proceedings of the 8th Global Conference on Business, Management, and Entrepreneurship (GCBME 2023)*, vol. 288, pp. 8–22, 2024. DOI: 10.2991/978-94-6463-443-3_2.

[4] M. Fuad and A. Imamudin, "Determinants of the composite stock price index (ihsg) on the indonesia stock exchange," *Journal of Economics Research and Social Sciences*, vol. 5, no. 1, 2021. DOI: 10.18196/jerss.v5i1.11002.

[5] L. R. A. Doho, S. M. Somé, and J. M. Banto, "Inflation and west african sectoral stock price indices: An asymmetric kernel method analysis," *Emerging Markets Review*, vol. 54, p. 100 987, 2023. DOI: 10.1016/j.ememar.2022.100987.

[6] B. Xu and R. Xu, "Assessing the carbon intensity of the heavy industry in china: Using a nonparametric econometric model," *Environmental Impact Assessment Review*, vol. 98, p. 106 925, 2023. DOI: 10.1016/j.eiar.2022.106925.

[7] D. Wied, "Semiparametric distribution regression with instruments and monotonicity," *Labour Economics*, vol. 90, p. 102 565, 2024. DOI: 10.1016/j.labeco.2024.102565.

[8] X. Chen and Y. Yi, "Information bounds for gaussian copula parameter in stationary semiparametric markov models," *Statistics & Probability Letters*, vol. 216, p. 110 254, 2025. DOI: 10.1016/j.spl.2024.110254.

[9] A. W. Wening, I. N. Budiantara, and I. Zain, "Semiparametric regression curve estimation for longitudinal data using mixed spline truncated and fourier series estimator," in *Journal of Physics: Conference Series*, vol. 1538, IOP Publishing, 2020, p. 012 061. DOI: 10.1088/1742-6596/1538/1/012061.

[10] J. L. Tobias and T. N. Bond, "Semiparametric bayesian estimation in an ordinal probit model with application to life satisfaction across countries, age and gender," *Journal of Econometrics*, vol. In Press, p. 105 917, 2025. DOI: 10.1016/j.jeconom.2024.105917.

[11] V. Fibriyani, N. Chamidah, and T. Saifudin, "Estimating time series semiparametric regression model using local polynomial estimator for predicting inflation rate in indonesia," *Journal of King Saud University - Science*, vol. 36, no. 1, 2024. DOI: 10.1016/j.jksus.2024.103549.

[12] G. M. Martin et al., "Bayesian forecasting in economics and finance: A modern review," *International Journal of Forecasting*, vol. 39, no. 3, 2023. DOI: 10.1016/j.ijforecast.2023.05.002.

[13] A. T. Fitriyah, N. Chamidah, and T. Saifudin, "Prediction of paddy production in indonesia using semiparametric time series regression least square spline estimator," *Data and Metadata*, vol. 4, 2025. DOI: 10.56294/dm2025527.

[14] D. Aydın, N. Chamidah, B. Lestari, S. Mohammad, and E. Yılmaz, "Local polynomial estimation for multi-response semiparametric regression models with right censored data," *Communications in Statistics - Simulation and Computation*, 2025. DOI: 10.1080/03610918.2025.2476595.

[15] N. M. Shimaponda-Mataa, E. Tembo-Mwase, M. Gebreslasie, T. N. Achia, and S. Mukaratirwa, "Reprint of "modelling the influence of temperature and rainfall on malaria incidence in four endemic provinces of zambia using semiparametric poisson regression"," *Acta Tropica*, vol. 175, pp. 60–70, 2017. DOI: 10.1016/j.actatropica.2017.08.014.

[16] B.-J. So, H.-H. Kwon, D. Kim, and S. O. Lee, "Modeling of daily rainfall sequence and extremes based on a semiparametric pareto tail approach at multiple locations," *Journal of Hydrology*, vol. 529, pp. 1442–1450, 2015. DOI: 10.1016/j.jhydrol.2015.08.037.

[17] M. Billio, R. Casarin, M. Costola, and M. Iacopini, "Covid-19 spreading in financial networks: A semiparametric matrix regression model," *Econometrics and Statistics*, vol. 29, pp. 113–131, 2024. DOI: 10.1016/j.ecosta.2021.10.003.

[18] Suliyanto, T. Saifudin, M. Rifada, and D. Amelia, "Statistical inferences and applications of nonparametric regression models based on fourier series," *MethodsX*, vol. 14, p. 103 217, 2025. DOI: 10.1016/j.mex.2025.103217.

[19] M. Zulfadhli, I. N. Budiantara, and V. Ratnasari, "Nonparametric regression estimator of multivariable fourier series for categorical data," *MethodsX*, vol. 13, p. 102 983, 2024. DOI: 10.1016/j.mex.2024.102983.

[20] W. Marbun, Suparti, and M. Maruddani, "Modeling of composite stock price index (cspi) using semiparametric regression truncated spline based on gui r," *Journal of Physics: Conference Series*, vol. 1524, 2020. DOI: 10.1088/1742-6596/1524/1/012096.

[21] W. Zhou and R. Harris, "Semi-parametric single-index predictive regression models with cointegrated regressors," *Journal of Econometrics*, 2024. DOI: 10.1016/j.jeconom.2023.105577.

[22] H. Zhu, Y. Zhang, Y. Li, and H. Lian, "Semiparametric function-on-function quantile regression model with dynamic single-index interactions," *Computational Statistics & Data Analysis*, vol. 182, p. 107 727, 2023. DOI: 10.1016/j.csda.2023.107727.

[23] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*. Springer, 2013. DOI: 10.1007/978-1-4614-7138-7.

[24] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*. London: Chapman and Hall, 1996, pp. 57–197. DOI: 10.1007/978-1-4899-0027-8.

[25] J. Friedman, T. Hastie, and R. Tibshirani, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009. DOI: 10.1007/978-0-387-84858-7.

[26] D. Y. Lin and Z. Ying, "Semiparametric and nonparametric regression analysis of longitudinal data," *Journal of the American Statistical Association*, vol. 96, no. 453, pp. 103–126, 2001. DOI: 10.1198/016214501750333018.

[27] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd. Morgan Kaufmann, an imprint of Elsevier, 2012, https://www.sciencedirect.com/book/9780123814791/data-mining.

[28] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986, pp. 37–43. DOI: 10.1201/9781315140919.

[29] D. Rosadi, *Analisa Runtun Waktu*. Yogyakarta: Gadjah Mada University Press, 2021.

[30] J. K. Saoudi, Y. Falloul, and M., "A comparison of lstm, gru, and xgboost for forecasting morocco's yield curve," *Mathematical Modeling and Computing*, vol. 11, no. 3, pp. 674–681, 2024. DOI: 10.23939/mmc2024.03.674.