Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 10, Number 1, 2026
DOI: 10.30829/zero.v10i1.29313
Page: 392-403

Robustness Evaluation of ANFIS, Hybrid GA-SVM, and SVM under Controlled Time Series Structures

[1] **Sri Kustiara**
Study Program of Statistics and Data Science, IPB University, Bogor, Indonesia

[2] **Kusman Sadik**
Study Program of Statistics and Data Science, IPB University, Bogor, Indonesia

[3] **Hari Wijayanto**
Study Program of Statistics and Data Science, IPB University, Bogor, Indonesia

Article Info

Article history:

Accepted 15 May 2026

Keywords:

ANFIS;
Factorial Simulation;
Hybrid GA-SVM;
Robustness Evaluation;
Time Series Forecasting.

ABSTRACT

This study evaluates the robustness of ANFIS, hybrid GA-SVM, and SVM under synthetic time-series structures using a factorial simulation framework combined with empirical validation. From a practical perspective, robust coal price forecasting is essential for supporting energy planning, trade management, and policy decision-making under uncertain market conditions. Empirical analysis of Indonesian coal prices reveals nonstationary behavior, high volatility, and nonlinear dynamics. Forecasting performance is assessed using walk-forward validation, where SVM and hybrid GA-SVM demonstrate comparable accuracy and outperform ANFIS on the empirical dataset. To systematically examine model sensitivity to structural variations, a $2 \times 3 \times 2$ factorial simulation design is implemented by varying seasonality, volatility, and predictor–response structure across 12 scenarios with 100 replications each. The results indicate that volatility is the most dominant factor affecting forecasting error, with significant interaction effects among structural factors. ANOVA and post hoc analysis further confirm that model performance depends more on data characteristics than on algorithmic complexity. These findings demonstrate that factorial simulation provides a systematic and robust framework for evaluating forecasting models beyond conventional empirical comparisons, while offering deeper insight into the relationship between data structure and model performance.

Corresponding Author:

Kusman Sadik,
Statistic and Data Science Program Study,
School of Data Science, Mathematics, and Informatics, IPB University, Bogor, Indonesia
Email: kusmans@apps.ipb.ac.id

1. INTRODUCTION

This study compares ANFIS, a hybrid GA-SVM, and SVM for time-series forecasting in both empirical and synthetic time-series settings. The empirical focus on Indonesian coal prices is important because Indonesia is one of the world's major coal exporters, making coal price dynamics economically relevant for energy trade, export performance, industrial activity, and policy formulation.

As a strategic commodity, coal prices are not determined solely by supply and demand, but are also affected by exchange rates, inflation, energy market interactions, trade conditions, and broader geopolitical disturbances. Recent evidence also shows that commodity markets are increasingly exposed to contagion effects and geopolitical spillovers that may alter price behavior over time [1], [2].

From a time-series perspective, commodity price data often violate simple assumptions such as stationarity, linearity, and constant variance. In practice, coal price series may exhibit level shifts, time-varying volatility, and complex relationships between the response and explanatory variables. Under such conditions, classical linear forecasting approaches may be insufficient to produce reliable predictions, especially when the data are influenced by structural change, market shocks, or evolving predictor interactions. [3], [4], [5], [6].

The development of computational forecasting methods has encouraged the use of more adaptive approaches, including Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machine (SVM), and optimization-based hybrid models. ANFIS can model nonlinear relationships by integrating neural networks and fuzzy inference systems into a unified learning framework [7]. SVMs are known for their strong generalization ability and effectiveness in handling complex nonlinear patterns through kernel-based learning [8],[9],[10]. Meanwhile, hybrid GA-SVM combines SVM with a Genetic Algorithm (GA) to optimize hyperparameters, which theoretically enhances model adaptability when the data structure is difficult, noisy, or unstable [11],[12].

Previous studies have shown that machine learning and hybrid methods can perform well in volatile time-series and commodity-related forecasting problems. Alameer et al. reported that hybrid GA-ANFIS outperformed ANFIS, SVM, GARCH, and ARIMA in forecasting copper prices [13]. Hendikawati et al. showed that ANFIS can yield reliable time-series predictions, although its performance is sensitive to preprocessing and input selection [14]. In reservoir inflow forecasting, Alquraish et al. found that hybrid ANFIS-GA and SVM-GA improved forecasting performance relative to benchmark models [15]. In the context of oil price forecasting, Kaymak and Kaymak emphasized that nonlinear and volatile price movements require flexible predictive models [16]. Sinaga et al., Oladipo and Sun, and Bakare et al. further demonstrated that optimization-based ANFIS hybrids can improve predictive performance in time-series and energy-related applications [17],[18],[19]. More broadly, Drachal and Pawłowski highlighted the increasing relevance of genetic algorithms and their hybrids in commodity price forecasting [20]. Related developments in support vector methods, including fuzzy and pinball-loss extensions, also suggest that SVM-based approaches remain highly relevant for complex predictive settings [21].

Recent developments in time-series learning also emphasize that multivariate forecasting becomes more challenging when temporal dependencies and inter-variable relations evolve over time [5],[22]. This is particularly relevant in data-rich forecasting environments, where performance depends not only on the forecasting algorithm itself, but also on predictor structure, dynamic dependence, and the complexity of the underlying data-generating process [4],[23],[24],[25],[6]. Recent surveys further show that forecasting research has expanded from statistical and machine learning approaches to multi-scale, self-supervised, and information-theoretic frameworks, reflecting the growing diversity of modern forecasting architectures [23],[24][26],[27],[22]. These perspectives are also consistent with broader machine learning arguments that predictive performance depends on the match between model structure and data complexity rather than methodological sophistication alone [28],[29].

Despite these developments, several gaps remain. First, many forecasting studies still focus on a single model or rely on static data splitting, which is less representative of real forecasting conditions than sequential validation [30],[4]. Second, direct comparisons among ANFIS, SVM, and hybrid GA-SVM on Indonesian coal price data remain limited. Third, empirical findings are rarely interpreted alongside controlled simulation evidence to explain why a particular method performs better under one structural condition but not under another. These gaps are important because modern time-series data are increasingly recognized as structurally dynamic, nonstationary, and rich in predictor information [4],[23],[24],[5],[6].

Based on this context, the present study makes two main contributions. First, it directly compares the forecasting performance of ANFIS, SVM, and hybrid GA-SVM on Indonesian coal price data using walk-forward validation, thereby making the empirical evaluation more consistent with real forecasting practice [30],[31]. Second, it complements the empirical findings with a factorial simulation framework that uses synthetic time-series structures to systematically vary seasonality, volatility, and predictor–response structure. This design makes it possible to evaluate model robustness more explicitly and to interpret empirical performance differences in light of underlying data characteristics. Such an approach is particularly relevant for commodity prices, which may be influenced by volatility transmission, spillover effects, and shifting structural conditions over time [1],[2],[20].

Therefore, this study evaluates the robustness of ANFIS, hybrid GA-SVM, and SVM across controlled time-series structures, with empirical validation using Indonesian coal price data.

From a practical perspective, accurate and robust coal price forecasting is essential for energy planning, export management, and policy decision-making under uncertain and volatile market conditions.

2. RESEARCH METHOD

2.1 Research Design

This study employed a quantitative computational framework consisting of two complementary stages: empirical analysis and factorial simulation. The empirical stage was conducted first to identify the actual characteristics of Indonesian coal price data, while the simulation stage was designed to evaluate model robustness under systematically controlled structural conditions. This sequence follows the rationale that empirical evidence should provide an initial understanding of the observed data-generating behavior before more structured simulation experiments are carried out.

The forecasting models evaluated in this study were ANFIS, hybrid GA-SVM, and SVM. These three models were selected because they represent different levels of modeling flexibility and optimization. ANFIS was included as a neuro-fuzzy approach capable of capturing nonlinear relationships [7], hybrid GA-SVM was included as an optimization-enhanced machine learning model [11],[12], and SVM was used as a benchmark model with strong generalization capability in nonlinear forecasting tasks [8],[9],[10].

The overall design was intended not only to compare forecasting accuracy but also to explain how model performance changes under different structural properties of time-series data. For this reason, the empirical analysis was complemented by a factorial simulation design and an ANOVA-based inferential analysis [11]. Although the main inferential focus of this study is factorial ANOVA, the formal comparison of predictive performance is also consistent with the broader literature on forecast evaluation [32].

2.2 Empirical Data and Preprocessing

The empirical dataset consisted of monthly Indonesian coal prices and several macroeconomic and energy-related variables assumed to influence coal price dynamics. These variables were selected for their economic relevance to coal price formation, since coal prices are not determined solely by their own historical movements but are also influenced by trade activity, inflation, exchange rates, and related energy commodity prices. The variables used in this study are presented in Table 1.

Table 1. Variables and Data Sources

Variable	Notation	Description	Data Source
Indonesian coal price	Y_t	Main response variable to be forecasted	World Bank
Coal exports	$X_{1,t}$	Volume of Indonesian coal exports	Ministry of Trade
Coal imports	$X_{2,t}$	Volume of coal imports	Ministry of Trade
Inflation	$X_{3,t}$	Indonesian inflation rate	Statistics Indonesia
Crude oil price	$X_{4,t}$	Indicator of global energy price dynamics	World Bank
Natural gas price	$X_{5,t}$	Complementary energy commodity price	World Bank
IDR/USD exchange rate	$X_{6,t}$	Exchange rate (IDR/USD)	Ministry of Trade

All variables were arranged as monthly time series. To reduce scale differences and stabilize variance, price-/related variables were transformed using the natural logarithm. This transformation is widely used in forecasting and commodity price modeling because it improves comparability across variables and reduces the effect of extreme magnitudes [3], [33]. After transformation, exploratory analysis was conducted the use of time-series plots, volatility inspections, and temporal dependence to identify nonstationarity, volatility shifts, and possible nonlinear behavior in the data.

The empirical findings were not treated as definitive evidence of model superiority. Instead, they served as an initial diagnostic stage from which the simulation framework could be designed more meaningfully. This is important because empirical performance on a single dataset may reflect only one realization of a broader structural environment [30],[4]. In addition, potential dependence among explanatory variables was considered conceptually because energy prices, exchange rates, and macroeconomic indicators may move together under changing market conditions. However, formal variable selection procedures such as LASSO or information criteria were not applied in this study, which is acknowledged as a limitation and a possible direction for further research.

2.3 Forecasting Models

2.3.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS is a hybrid intelligent system that combines the reasoning structure of fuzzy inference systems with the learning ability of neural networks. The model is designed to approximate nonlinear input–output relationships through a set of fuzzy if–then rules with adaptive membership functions [7]. In general, the ANFIS output is expressed in Equation (1).

$$f(x) = \sum_{i=1}^{R} w_i f_i(x) \tag{1}$$

where w_i denotes the normalized firing strength of the i-th rule and $f_i(x)$ represents the corresponding consequent function. This structure allows ANFIS to capture nonlinear and uncertain relationships more flexibly than conventional linear models.

In this study, ANFIS was implemented using a Takagi–Sugeno fuzzy inference system. The input variables were first scaled using min-max normalization to improve training stability. The model configuration, including the membership function structure, number of membership functions, training procedure, and stopping criteria, was fixed before evaluation to ensure comparability across empirical and simulation experiments. The training process followed the same walk-forward validation framework as the other forecasting models.

2.3.2 Hybrid GA-SVM

Hybrid GA-SVM combines Support Vector Machine with a Genetic Algorithm to optimize the hyperparameters of the SVM model. This hybridization is motivated by the fact that SVM performance is highly dependent on the selection of tuning parameters, particularly the penalty parameter C, the insensitive loss parameter ε, and the kernel parameter γ [9],[10]. In the hybrid framework, GA performs a global search over candidate parameter combinations, reducing the possibility of selecting suboptimal parameters [11],[12]. In addition to classical GA references, recent optimization-oriented studies also support the role of hybrid search strategies in improving predictive model configuration [17],[25].

The use of GA becomes particularly relevant under high-volatility conditions. When the data are noisier and structurally more complex, SVM performance may become more sensitive to hyperparameter choices. In this context, GA provides an adaptive search mechanism that allows the model to explore alternative parameter combinations and identify settings that are more suitable for unstable data structures.

To ensure reproducibility, the GA procedure was configured before the evaluation stage. In this study, GA was used to optimize the main SVM hyperparameters, namely C, ε, and γ. The GA configuration used a population size of 24, 20 generations, a crossover probability of 0.7, and a mutation probability of 0.3. Candidate solutions were represented as parameter chromosomes and evaluated using RMSE as the fitness criterion within the walk-forward validation framework. The optimization process consisted of population initialization, fitness evaluation, selection, crossover, mutation, and iterative updating until the stopping criterion was reached. The best parameter combination obtained from GA was then used to train the final SVM model at each walk-forward evaluation step.

2.3.3 Support Vector Machine (SVM)

SVM is a kernel-based machine learning approach with strong theoretical foundations in statistical learning theory [8],[9]. For nonlinear regression, the forecasting function is written in Equation (2).

$$f(x) = \omega^T \phi(x) + b, \tag{2}$$

where $\phi(x)$ maps the input vector into a higher-dimensional feature space, ω is the weight vector, and b is the intercept term. In practice, kernel functions allow SVM to model nonlinear relationships without explicitly constructing the feature space. Because of this capability, SVM is often considered a strong benchmark for nonlinear forecasting problems [10],[15].

In this study, an SVM was implemented with a radial basis function (RBF) kernel because it is well-suited for capturing nonlinear relationships in time-series forecasting. The main hyperparameters considered were C, ε, and γ. For the standard SVM model, the tuning procedure was kept fixed and comparable across all experiments, whereas in hybrid GA-SVM these parameters were optimized using the Genetic Algorithm. This distinction allowed the study to compare a standard SVM configuration with an optimization-enhanced SVM framework in a fair evaluation setting.

Classical benchmark models such as ARIMA and ETS were not included because the main objective of this study was to evaluate the robustness of ANFIS, hybrid GA-SVM, and SVM under empirical and controlled structural conditions. These benchmarks are acknowledged as useful comparative models and are recommended for future research to broaden the evaluation framework.

2.4 Empirical Evaluation Procedure

An empirical comparison among ANFIS, hybrid GA-SVM, and SVM was conducted using walk-forward validation with an expanding-window scheme. This strategy preserves the temporal order of observations and provides a more realistic evaluation than a static train-test split, especially for time-series forecasting [30]. It is also consistent with broader forecasting practice and applied forecasting guidelines in the literature [30],[31].

At each forecasting step, the model was trained on all available observations up to time t and then used to predict the next observation at time $t + 1$. The training sample was then expanded by one period, and the process was repeated until the end of the evaluation horizon. This procedure reduces information leakage and ensures that forecasting performance is assessed under genuinely out-of-sample conditions.

Model performance in the empirical stage was evaluated using RMSE, MAE, and MAPE. RMSE was then used as the main response variable in the inferential simulation stage because it is more sensitive to large forecasting errors and is widely used in comparative forecasting evaluation [34],[35]. Although formal predictive accuracy tests, such as the Diebold–Mariano test, are relevant in forecast comparison studies [32], this study focuses on descriptive empirical evaluation and on a factorial ANOVA-based robustness assessment. Therefore, the Diebold–Mariano test was not included in the main analysis and is recommended as a possible extension for future research.

2.5 Factorial Simulation Framework

To evaluate robustness under controlled structural conditions, this study employed a factorial simulation framework based on an additive time-series generation model. The general form of the simulated response series is shown in Equation (3).

$$Y_t = T_t + S_t + f(X_t) + \varepsilon_t \tag{3}$$

where T_t denotes the trend component, S_t denotes the seasonal component, $f(X_t)$ represents the relationship between predictors and the response, and ε_t is a random disturbance term.

In this study, the trend component was kept constant to focus the simulation on the three structural factors of interest: seasonality, volatility, and predictor–response structure. The seasonal component was generated using Equation (4).

$$S_t = A \sin\left(\frac{2\pi t}{s}\right), s = 12 \tag{4}$$

where A controls the intensity of seasonality. The random disturbance term followed the distribution in Equation (5).

$$\varepsilon_t \sim N(0, \sigma^2), \tag{5}$$

where σ determines the volatility level. The structural relationship between predictors and the response was defined as either linear or nonlinear, depending on the factor level being simulated.

This formulation was chosen because it provides a transparent decomposition of the time series into interpretable structural components while still allowing sufficient flexibility to examine model sensitivity under different conditions [13],[20],[16]. However, the additive data-generating process is a simplified representation of real commodity price dynamics. It may not fully capture more complex mechanisms such as regime shifts, policy shocks, market microstructure, or endogenous feedback. This limitation is acknowledged because the main purpose of the simulation is to isolate the effects of seasonality, volatility, and predictor–response structure in a controlled manner.

2.6 Factorial Design and Scenario Construction

The simulation adopted a full factorial design with three structural factors: seasonality, volatility, and predictor–response structure. Seasonality intensity (Factor A) consisted of two levels: weak seasonality ($A_1 = 0.5$) and strong seasonality ($A_2 = 3.0$). Volatility level (Factor B) consisted of three levels: low volatility ($B_1 = 0.5$), moderate volatility ($B_2 = 1.0$), and high volatility ($B_3 = 2.0$). Predictor–response structure (Factor C) consisted of two levels: linear structure (C_1) and nonlinear structure (C_2).

Accordingly, the simulation generated $2 \times 3 \times 2 = 12$ scenarios. Each scenario was replicated 100 times and evaluated using three forecasting models, producing $12 \times 100 \times 3 = 3600$ evaluation records.

The numerical values assigned to the factor levels were treated as controlled design levels rather than empirical estimates. These values were determined through preliminary calibration to ensure that the simulated conditions were interpretable, distinguishable, and capable of producing meaningful differences in model performance. This calibration served as a practical step in simulation design, not as an inferential procedure. The selected levels span low-to-high structural variation in seasonality and volatility, while the predictor–response structure was varied to distinguish between linear and nonlinear data-generating mechanisms.

To account for experimental variability, each factorial scenario was replicated 100 times. This repeated-simulation design allows forecasting performance to be summarized across replications and helps capture the variability of model errors under the same structural conditions. However, this study reports uncertainty through replication-based summaries rather than interval-based measures. Bootstrap confidence intervals were not included in the current analysis and are recommended as a possible extension for future research.

2.7 Forecasting Accuracy Measures

Forecasting performance was evaluated using three error measures. RMSE is defined in Equation (6).

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}, \tag{6}$$

MAE is defined in Equation (7).

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|, \tag{7}$$

MAPE is defined in Equation (8).

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t} \times 100\%. \tag{8}$$

Among these measures, RMSE was used as the primary metric because it is more sensitive to large forecasting errors and is therefore more suitable when substantial deviations are of primary concern [34],[35]. MAE and MAPE were retained as supporting measures to enrich interpretation, but they were not used as ANOVA response variables so that the inferential analysis remained focused and methodologically consistent. Forecast errors were reported as point estimates, while interval-based uncertainty assessment, such as bootstrap confidence intervals, was not included. This is acknowledged as a limitation and may be addressed in future work.

2.8 ANOVA and Post Hoc Analysis

The simulation results were analyzed using RMSE as the primary response variable, using two approaches: factorial ANOVA per method and combined factorial ANOVA with the forecasting method as an additional factor.

In the first approach, ANOVA was conducted separately for each forecasting method, namely ANFIS, hybrid GA-SVM, and SVM, using the full factorial model shown in Equation (9).

$$E_{rijk} = \mu + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + \varepsilon_{rijk} \tag{9}$$

where E_{rijk} denotes the forecasting error (RMSE) at replication r, and A, B, and C represent seasonality, volatility, and predictor–response structure, respectively.

In the second approach, all methods were analyzed jointly by including forecasting method as an additional factor M, as shown in Equation (10).

$$\begin{aligned} E_{lmijk} = \mu + M_l + A_i + B_j + C_k + (AB)_{ij} + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + (MA)_{li} \\ + (MB)_{lj} + (MC)_{lk} + \varepsilon_{lmijk} \end{aligned} \tag{10}$$

This formulation allows the assessment of differences in model sensitivity across structural conditions. Effect size was evaluated using partial eta squared [36], and Tukey's HSD was applied for post hoc analysis. The inclusion of effect size is important because statistical significance alone does not fully describe the practical magnitude of factor effects. Therefore, partial eta squared was used to interpret the practical significance of the ANOVA results.

2.9 Integration of Empirical and Simulation Analysis

The final stage integrated the empirical and simulation results. The empirical analysis showed how the models behaved under actual Indonesian coal price dynamics, whereas the simulation analysis explained how performance changed when structural properties of the data were varied in a controlled manner. This combined approach ensured that model comparison did not rely solely on a single empirical realization, but also reflected a broader assessment of robustness across alternative synthetic time-series structures.

The empirical findings should be interpreted in light of Indonesian coal prices and should not be generalized directly to all commodities or geographic markets. Market structure, institutional conditions, policy shocks, and commodity-specific mechanisms may influence the transferability of the results. Therefore, the findings provide methodological and empirical insight for Indonesian coal price forecasting, while further validation using other commodities and market settings remains an important direction for future research.

3. RESULT AND ANALYSIS

3.1 Characteristics of Empirical Data

The empirical dataset consisted of 262 monthly observations from January 2004 to October 2025 and contained no missing values, indicating strong temporal continuity for time-series analysis. After a natural logarithm transformation, the coal price series had an average of approximately 4.49 and a standard deviation of 0.48, suggesting that the transformation reduced scale differences while preserving the underlying dynamics of price movements [3], [33].

To provide an initial overview of the empirical behavior, the log-transformed coal price series is presented in Figure 1.

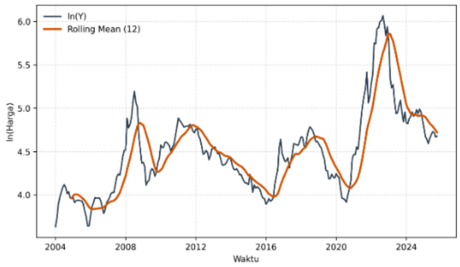

Figure 1. Log-transformed Indonesian Coal Price Series

Figure 1 shows several distinct phases of level change over time. A substantial increase occurred around 2008, followed by another upward movement in 2010–2011, a decline in the middle of the following decade, and a strong rise during 2021–2023, before weakening toward the end of the observation period. This pattern indicates that the series is not fully stationary in level and contains long-term structural changes. Such behavior is common in commodity and economic time-series data, where evolving market conditions alter the data-generating process over time [16],[4],[5].

To examine changes in local volatility, a 12-month rolling standard deviation analysis was conducted, as shown in Figure 2.

Figure 2. Dynamic Volatility of Indonesian Coal Prices Based on 12-Month Rolling Standard Deviation

Figure 2 indicates that volatility varies over time and does not remain constant throughout the observation period. Higher volatility is observed around 2008–2009, increases again around 2016–2017, and becomes more pronounced during 2021–2024. This confirms the presence of time-varying volatility in the empirical coal price data, consistent with broader findings in commodity market research [16], [1], [2].

Temporal dependence was further examined using ACF and PACF plots. The ACF shows strong positive autocorrelation at the initial lags and declines gradually, while the PACF indicates that direct dependence is concentrated in the first few lags. Overall, the empirical coal price data exhibit three key characteristics: nonstationarity in level, time-varying volatility, and strong short-term temporal dependence. These findings justify the use of flexible and adaptive forecasting models rather than approaches that rely on strict linearity and constant variance assumptions [8],[9],[10],[5].

3.2 Walk-Forward Empirical Performance

Model performance was evaluated using one-step-ahead walk-forward validation with an expanding window scheme. The empirical results are summarized in Table 2.

Table 2. Forecasting Performance under Walk-Forward Validation

Method	RMSE	MAE	MAPE
SVM	0.224775	0.148685	2.916890
Hybrid GA-SVM	0.224775	0.148685	2.916890
ANFIS	0.287135	0.205642	4.087419

Based on Table 2, SVM and hybrid GA-SVM produced identical performance across all evaluation metrics, whereas ANFIS yielded higher forecast errors. These results indicate that SVM-based approaches achieved higher predictive accuracy than ANFIS on Indonesian coal price data.

To clarify consistency across different error measures, the comparison of RMSE, MAE, and MAPE is presented in Figure 3.

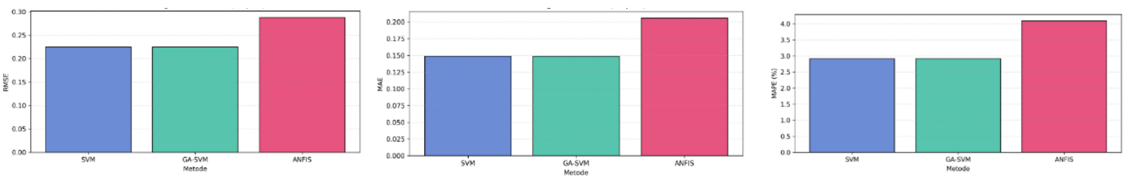

Figure 3. Comparison of RMSE, MAE, and MAPE across Forecasting Models

Figure 3 shows a consistent ranking across all three metrics. SVM and hybrid GA-SVM produced lower forecast errors than ANFIS, while the difference between SVM and hybrid GA-SVM remained negligible. This suggests that, under the observed empirical structure, GA-based optimization did not provide measurable improvement beyond the standard SVM model.

To ensure the validity of this result, the numerical outputs and model implementation were carefully rechecked. The identical performance values indicate that the baseline SVM configuration was already sufficient to capture the dominant data structure, so the additional optimization step in the hybrid GA-SVM did not yield further improvement. This finding is consistent with the broader forecasting principle that increased methodological complexity does not necessarily yield better performance when the underlying structure can already be adequately modeled [34],[13],[20],[28]. It is also aligned with general machine learning perspectives that emphasize the importance of matching model structure to data complexity rather than relying solely on algorithmic sophistication [28],[29].

3.3 Completeness of the Factorial Simulation Design

Following the empirical stage, robustness was further examined using a controlled factorial simulation framework based on synthetic time-series structures. The simulation employed a balanced design that incorporated seasonality, volatility, and a predictor–response structure, yielding 12 scenarios. Each scenario was replicated 100 times and evaluated using three forecasting methods, producing a total of 3600 evaluation records.

This balanced design ensures that the influence of each structural factor and its interactions can be assessed systematically. It also allows empirical findings to be interpreted in relation to a broader range of structural conditions beyond those directly observed in Indonesian coal price data.

3.4 Comparative Simulation Performance

The comparative performance of forecasting models under synthetic time-series structures was evaluated using mean RMSE across volatility levels, as presented in Table 3.

Table 3. Mean RMSE by Volatility Level

Volatility Level	ANFIS	Hybrid GA-SVM	SVM
Low	0.855919	0.920262	0.890638
Moderate	1.718657	1.664907	1.671510
High	2.819211	2.570143	2.607179

Table 3 shows that forecasting error increases consistently as volatility rises. Under low-volatility conditions, ANFIS achieves the lowest RMSE, indicating its strength in capturing stable patterns. However, under moderate to high volatility, SVM and hybrid GA-SVM demonstrate better performance.

Under moderate volatility, SVM achieves the lowest RMSE, followed closely by hybrid GA-SVM, suggesting that SVM-based approaches are effective in capturing moderately complex patterns. Under high-volatility conditions, hybrid GA-SVM produces the lowest RMSE, indicating that parameter optimization provides additional benefits when the data become more complex and noisier.

These findings confirm that model performance is not uniform across structural conditions. Instead, the relative advantage of each method depends on the level of volatility, which is consistent with theoretical expectations in time-series forecasting, where higher noise levels reduce predictability [3], [35].

3.5 ANOVA Main and Interaction Effects

A factorial ANOVA was conducted with RMSE as the response variable to evaluate the significance of structural factors and their interactions. The detailed ANOVA results are presented in Table 4.

Table 4. Factorial ANOVA Results (RMSE as Response Variable)

Factor	F-value	p-value	Partial Eta Squared
Seasonality (A)	1609.22	< 0.001	0.3096
Volatility (B)	4400.40	< 0.001	0.7104
Structure (C)	350.55	< 0.001	0.0890
A × B	244.88	< 0.001	0.1201
A × C	42.97	< 0.001	0.0118
B × C	9.27	< 0.001	0.0051

The results indicate that volatility is the most dominant factor affecting forecasting error, as reflected by the highest F-value and effect size. Seasonality and predictor–response structure also show significant effects, although with relatively smaller magnitudes.

The presence of significant interaction effects confirms that model performance is influenced by combined structural conditions rather than isolated factors. This implies that the impact of each factor depends on the levels of other factors, highlighting the importance of considering interaction effects when evaluating the robustness of forecasting models.

3.6 Post Hoc and Robustness Interpretation

Post hoc analysis using Tukey's HSD, as presented in Table 5, confirmed significant differences among volatility levels, indicating that forecasting error increased consistently from low to high volatility. This result reinforces the ANOVA finding that volatility is the dominant factor affecting RMSE.

In terms of robustness, ANFIS performed better under stable conditions, hybrid GA-SVM showed greater robustness under high volatility and structural complexity, and SVM maintained consistent performance under moderate conditions. These findings suggest that robustness should not be interpreted as universal superiority, but as the ability of a model to remain competitive under varying data structures.

From a practical perspective, these results can be translated into decision rules for model selection. ANFIS is more suitable for relatively stable data, SVM is effective under moderate conditions, and a hybrid GA-SVM is more appropriate for highly volatile, structurally complex data.

Table 5. Tukey HSD Post Hoc Results

Comparison	Mean Difference	p-value	Decision
Low vs Moderate	0.7961	< 0.001	Significant
Low vs High	1.7766	< 0.001	Significant
Moderate vs High	0.9805	< 0.001	Significant

3.7 Linking Empirical and Simulation Findings

The empirical and simulation results are complementary rather than contradictory. Empirically, SVM and the hybrid GA-SVM achieved identical or superior performance relative to ANFIS for Indonesian coal prices. In the simulation analysis, hybrid GA-SVM showed clearer advantages only under more complex structural conditions, particularly when volatility and nonlinearity increased.

This implies that the observed coal price data were closer to a moderate structural environment than to an extreme one. Under such conditions, the standard SVM was sufficient to capture the dominant dynamics, so the additional optimization step in the hybrid GA-SVM did not produce further gains. The simulation stage explains why the empirical ranking occurred, while the empirical stage provides real-world validation of the robustness interpretation.

4. CONCLUSION

This study evaluated the robustness of ANFIS, hybrid GA-SVM, and SVM using empirical analysis and factorial simulation based on synthetic time-series structures. Using empirical Indonesian coal price data, SVM and hybrid GA-SVM achieved identical and superior forecasting performance, whereas ANFIS produced higher forecast errors. The simulation results further showed that model performance was strongly influenced by structural data characteristics, with volatility emerging as the most dominant factor affecting forecasting error.

These findings indicate that greater algorithmic complexity does not automatically lead to better predictive accuracy. Under relatively moderate data structures, standard SVM may already be sufficient, whereas the advantage of hybrid GA-SVM becomes more visible under more volatile and structurally complex conditions. Therefore, forecasting model selection should be based on the compatibility between the model structure and the data characteristics rather than on methodological sophistication alone.

Future research may extend this study by incorporating more complex data-generating processes, interval-based uncertainty assessment, additional forecast-comparison methods, and broader applications across different commodities or regional markets.

5. REFERENCES

[1] B. Wu, D. Huang, and M. Chen, "Estimating contagion mechanism in global equity market with time-zone effect," *Financ. Manag.*, vol. 52, no. 3, pp. 543–572, 2023, doi: 10.1111/fima.12430.

[2] G. Palomba and M. Tedeschi, "Geopolitical Risks' Spillovers Across Countries and on Commodity Markets: A Dynamic Analysis," *Energy Res. Lett.*, vol. 6, no. 2, pp. 1–8, 2025, doi: 10.46557/001c.121262.

[3] J. D. Hamilton, *Time Series Analysis*. Princeton, New Jersey, USA: Princeton University Press, 1994. doi: 10.2307/j.ctv14jx6sm.

[4] A. Almeida, S. Brás, S. Sargento, and F. C. Pinto, "Time series big data: a survey on data stream frameworks, analysis and algorithms," *J. Big Data*, vol. 10, no. 1, p. 83, May 2023, doi: 10.1186/s40537-023-00760-1.

[5] B. Nguyen-Thai, V. Le, N. D. T. Tieu, T. Tran, S. Venkatesh, and N. Ramzan, "Learning evolving relations for multivariate time series forecasting," *Appl. Intell.*, vol. 54, no. 5, pp. 3918–3932, 2024, doi: 10.1007/s10489-023-05220-0.

[6] S. C. Huang and R. S. Tsay, "Time Series Forecasting with Many Predictors," *Mathematics*, vol. 12, no. 15, 2024, doi: 10.3390/math12152336.

[7] J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665–685, 1993, doi: 10.1109/21.256541.

[8] V. Vapnik, *The Nature of Statistical Learning Theory, Second Edition*, 2nd ed. New York Inc: Springer Berlin Heidelberg, 1995.

[9] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, pp. 273–297, 1995, doi: 10.1007/BF00994018.

[10] B. Schölkopf and A. J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2002.

[11] D. E. Goldberg, *Genetic algorithms in optimization, search and machine learning*, 1st editio. Reading, MA, USA: Addison-Wesley Publishing Company, 1989. [Online]. Available: https;//books.google.co.id/books?

[12] R. L. Haupt, S. E. Haupt, and A. J. Wiley, *Pratical Genetic Algorithms*, 2nd ed. Hoboken New Jersey: John Wiley & Sons, 2004.

[13] Z. Alameer, M. A. Elaziz, A. A. Ewees, H. Ye, and Z. Jianhua, "Forecasting Copper Prices Using Hybrid Adaptive Neuro-Fuzzy Inference System and Genetic Algorithms," *Nat. Resour. Res.*, vol. 28, no. 4, pp. 1385–1401, 2019, doi: 10.1007/s11053-019-09473-w.

[14] P. Hendikawati, Subanar, Abdurakhman, and Tarno, "ANFIS Performance Evaluation for Predicting Time Series with Calendar Effects," *IAENG Int. J. Appl. Math.*, vol. 51, no. 3, pp. 1–12, 2021, [Online]. Available: https://www.iaeng.org/IJAM/issues_v51/issue_3/index.html.

[15] M. . Alquraish, K. . Abuhasel, A. . Alqahtani, and M. A. Khadr, "Comparative Analysis of Hidden Markov Model, Hybrid Support Vector Machines, and Hybrid Artificial Neural Fuzzy Inference System in Reservoir Inflow Forecasting (Case Study: The King Fahd Dam, Saudi Arabia).," *Water*, vol. 13, no. 1236, 2021, doi: doi.org/10.3390/w13091236.

[16] Ö. Öztunç Kaymak and Y. Kaymak, "Prediction of crude oil prices in COVID-19 outbreak using real data," *Chaos, Solitons and Fractals*, vol. 158, 2022, doi: 10.1016/j.chaos.2022.111990.

[17] A. V. Sinaga, N. C. Biutarbutar, T. B. Simamora, and J. Amalia, "Adaptive Neuro Fuzzy Inference System Optimization by Genetic Algorithm pada Time Series," *DoubleClick J. Comput. Inf. Technol.*, vol. 6, no. 1, pp. 25–31, 2022, doi: 10.25273/doubleclick.v6i1.13368.

[18] S. Oladipo and Y. Sun, "Enhanced adaptive neuro - fuzzy inference system using genetic algorithm : a case study in predicting electricity consumption," *SN Appl. Sci.*, vol. 5, no. 6, p. 186, 2023, doi: 10.1007/s42452-023-05406-8.

[19] M. S. Bakare, A. Abdulkarim, A. N. Shuaibu, and M. M. Muhamad, "A hybrid long-term industrial electrical load forecasting model using optimized ANFIS with gene expression programming," *Energy Reports*, vol. 11, no. May, pp. 5831–5844, 2024, doi: 10.1016/j.egyr.2024.05.045.

[20] K. Drachal and M. Pawłowski, "A Review of the Applications of Genetic Algorithms to Forecasting Prices of Commodities.," *Economies*, vol. 9, p. 6, 2021, doi: 10.3390/economies9010006.

[21] R. Nayan, V. Rahul, D. Rakesh, N. Subbarao, and G. Pratap, "A new fuzzy support vector machine with pinball loss," *Discov. Artif. Intell.*, vol. 3, no. March, p. 14, 2023, doi: 10.1007/s44163-023-00057-5.

[22] D. Ullmann, O. Taran, and S. Voloshynovskiy, "Multivariate Time Series Information Bottleneck," *Entropy*, vol. 25, no. 5, p. 831, 2023, doi: 10.3390/e25050831.

[23] M. Cheng, Q. Liu, Z. Liu, Z. Li, Y. Luo, and E. Chen, *FormerTime: Hierarchical Multi-Scale Representations for Multivariate Time Series Classification*, vol. 1, no. 1. Association for Computing Machinery, 2023. doi: 10.1145/3543507.3583205.

[24] M. Cheng *et al.*, "TimeMAE: Self-Supervised Representations of Time Series with Decoupled Masked Autoencoders," *Proc. Ninet. ACM Int. Conf. Web Search Data Min. (WSDM '26), Febr. 22-26, 2026, Boise, ID, USA*, vol. 1, no. 1, pp. 498–508, 2026, doi: 10.1145/3773966.3778007.

[25] C.-M. Lin and Y.-S. Lin, "TPTM-HANN-GA: A Novel Hyperparameter Optimization Framework Integrating the Taguchi Method, an Artificial Neural Network, and a Genetic Algorithm for the Precise Prediction of Cardiovascular Disease Risk," *Mathematics*, vol. 12, no. 1303, p. 9, 2024, doi: 10.3390/math12091303.
[26] M. Hamoudia and S. and E. S. Makridakis, *Forecasting with artificial intelligence Theory and Artificial*, 1st ed. Switzerland: Palgrave Macmillan Cham, 2023. doi: 10.1007/978-3-031-35879-1.
[27] J. Kim, H. Kim, H. G. Kim, D. Lee, and S. Yoon, "A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges," *Artif. Intell. Rev.*, vol. 58, no. 7, p. 216, 2025, doi: 10.1007/s10462-025-11223-9.
[28] S. S. W. Fatima and A. Rahimi, "A Review of Time-Series Forecasting Algorithms for Industrial Manufacturing Systems," *Machines*, vol. 12, no. 6, p. 380, 2024, doi: 10.3390/machines12060380.
[29] M. Bishop,C, *Pattern Recognition and Machine Learning*. New York, USA: Springer US, 2006. doi: 10.53759/7669/jmc202404020.
[30] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Inf. Sci. (Ny).*, vol. 191, pp. 192–213, 2012, doi: 10.1016/j.ins.2011.12.028.
[31] R. J. Hyndman and George Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. Australia: OTexts, 2018.
[32] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253–263, 1995, doi: 10.1080/07350015.1995.10524599.
[33] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting: Methods and Applications*, 3rd Editio. New York, NY, USA: John Wiley & Sons, 1998.
[34] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.
[35] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.
[36] D. C. Montgomery, *Design and Analysis of Experiment*, 9th ed. Hoboken New Jersey: John Wiley & Sons, 2017.