JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index
Vol. 21, No. 3, May 2025, pp. 627-645
DOI: 10.20956/j.v21i3.42558

Forecasting Nickel Prices in Indonesia Using ARIMA, SVR, and Hybrid ARIMA-SVR Approach

Nashwa Carista[1], Arinda Mahadesyawardani[2], Adelia Sukma Dwiyanto[3], Elly Pusporani[4]*

[1234] Department of Mathematics, Airlangga University, Indonesia
***Email**: nashwa.carista-2022@fst.unair.ac.id[1], arinda.mahadesyawardani-2022@fst.unair.ac.id[2], adelia.sukma.dwiyanto-2022@fst.unair.ac.id[3], elly.pusporani@fst.unair.ac.id[4]*
[*]Corresponding author
Received: 17 December 2024, revised: 8 March 2025, accepted: 11 March 2025

Abstract

Nickel production plays a key role in reducing reliance on fossil fuels, supporting the 7th Sustainable Development Goals (SDGs) on clean and affordable energy. As the world's largest nickel ore producer, Indonesia significantly influences global market dynamics. This study evaluates the accuracy of ARIMA, SVR, and hybrid ARIMA-SVR models in forecasting Indonesia’s daily nickel futures prices for 2023 using historical data from official website investing.com. The results indicate that SVR outperforms the other models, achieving the lowest MAPE of 0.2532% with the Radial Basis Function (RBF) kernel and optimized parameters $\varepsilon = 0.1$, $C = 2^9$, and $\gamma = 2^5$ selected through grid search method which gives the minimum RMSE and MAE values as well. Accurate nickel price forecasting is essential for investors, mining companies, and policymakers to optimize production planning, manage risks, and enhance market stability. However, this study relies solely on historical price data, without considering external factors such as geopolitical events and market shocks, highlighting the need for future research incorporating broader economic indicators and alternative modeling approaches.

Keywords: Autoregressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), Hybrid ARIMA-SVR, Forecasting, Nickel.

1. INTRODUCTION

Indonesia is rich in natural resources and geographically spread across various parts of the country, making it an important asset for development. The Indonesian mining sector, known for its abundant resources, makes significant contributions each year through taxes and royalties. One of the strategic minerals is nickel, which plays a vital role in supporting industrial needs. Nickel, with its anticorrosive properties, is very valuable. In its pure form, nickel is easy to shape, but when combined with iron chromium, and other metals, it can produce strong and durable stainless steel [12].

The United States Geological Survey (USGS) report states that in 2019, Indonesia ranked as the largest nickel ore producer in the world, with a production of 800 thousands tons and reserves of 21 million tons [9]. According to the United States Geological Survey report in 2020, Indonesia managed to produce 760,000 tons of nickel from a total reserve of 21 million tons [19]. Data from the United States Geological Survey shows that Indonesia has positioned itself as the largest nickel producer in the world. In addition, Indonesia also plays an important role in supporting nickel exports in the international market.

This statement is supported by data from 2021 showing that Indonesia successfully produced 1 metric ton of nickel, equivalent to 37.04% of the total global nickel production [3]. Based on data from the United States Geological Survey in 2022, nickel ore production in Indonesia was around 1.6 million tons [11]. In 2023, total nickel production is estimated to reach 1.6 million metric tons, contributing 48.48% of the global nickel production [13]. Meanwhile, according to the Geological Agency, Indonesia’s nickel resources are even recorded at 11.7 billion tons. Most of these resources are concentrated in the regions of Central Sulawesi, South Sulawesi, Southeast Sulawesi, and North Maluku [18].

Based on supporting data sources, nickel production has been increasing year by year. This makes Indonesia a key player in determining the dynamics of the global market. However, the increase in production is not always accompanied by price stability. Nickel prices in the global market are greatly influenced by factors such as the demand for raw materials for electric vehicle batteries. For example, the high demand for nickel to support the energy transition towards environmentally friendly technologies has the potential to significantly increase nickel prices. This is related to nickel price forecasting, which is a key element for the government and industry players to formulate long-term strategies. By utilizing historical nickel futures price data, forecasting models can provide projections of future nickel prices. In this study, a comparison was made between three methods, namely ARIMA, Support Vector Regression, and hybrid ARIMA-SVR.

The previous studies that serve as references for the researchers are “Analysis of the Hybrid ARIMA-SVR Method on the Composite Stock Price Index”, which obtained a MAPE value of 70.8% on the testing data with the best model being hybrid ARIMA-SVR [14]. Furthermore, the study titled ”Forecasting Gold Prices during the Covid-19 Pandemic Using the hybrid Autoregressive Integrated Moving Average-Support Vector Regression Model”, which compared the implementation of the hybrid ARIMA-SVR method with ARIMA in forecasting gold prices with a total of 210 observations [10]. The best method obtained was hybrid ARIMA-SVR with a value of 0.355 on the training data and 4.001 on the testing data. Therefore, the researcher is interested in conducting a study comparing the models, namely ARIMA, SVR, and hybrid ARIMA-SVR, using daily nickel futures price data. This research supports the 7th Sustainable Development Goals (SDGs), which is clean and affordable energy, specifically point 7.2 that discusses substantially increasing the share of renewable energy in the global energy mix by 2030. This research is expected to be a first step for nickel production to contribute to the reduction of fossil fuel materials in transportation.

2. METHODS

2.1 Nickel

Nickel is a hard, silvery-white metal that is resistant to corrosion. This metal plays an important role in processing various metals into solutions capable of generating heat energy. Nickel ore consists of Ni-Sulfides (nickel sulfides) and Ni-Laterites (nickel laterites). Ni-Sulfide minerals are generally formed primarily and are associated with mafic and ultramafic rocks [7]. Some benefits of nickel in daily life include its use as a material for manufacturing stainless steel, as an alloying agent in steel, and in industries such as transportation, electric battery production, and others.

2.2 Time Series and Data Stationarity

Time series data is a type of data that is collected based on a specific period of time in a sequential manner. A process in time series analysis is said to be stationary if there are no changes in the trend of either the mean or the variance [4]. The stationarity test of the data can be conducted using the unit root test calculation, namely the Augmented Dickey-Fuller (ADF). Hypothesis testing uses the following ADF test [14].

$H_0 : \phi = 0$ (data is not stationary)

$H_1 : \phi \neq 0$ (data is stationary)

Calculate the test ADF statistical value using Equation (2.1).

$$ADF_{hit} = \frac{\hat{\phi}}{se(\hat{\phi})} \tag{2.1}$$

with $se(\hat{\phi})$ is the standard error of $\hat{\phi}$. Decision to reject H_0 if $|ADF_{hit}| > |t_{(\alpha,n)}|$ or p-value is less than the significance level of 0.05 where n is many observations. If the data is non-stationary in variance, then the data needs to be transformed using the Box-Cox transformation, which can be expressed in Equation (2.2).

$$Z_t = \begin{cases} \frac{Y_t^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \\ \ln(Y_t), & \lambda = 0 \end{cases} \tag{2.2}$$

with Y_t representing the t-th time series data and Z_t representing the transformed time series data, the corresponding transformations based on the parameter λ are shown in Table 2.1.

Table 2.1. Box-Cox Transformation

λ **Value**	-1.0	-0.5	0	0.5	1.0
Transformation	$\frac{1}{Y_t}$	$\frac{1}{\sqrt{Y_t}}$	$\ln(Y_t)$	$\sqrt{Y_t}$	Y_t

If the data is non-stationary in the mean, differencing is required. Differencing involves subtracting the observation at time Y_t from the observation at time $Y_t - 1$. This is expressed for the first order in Equation (2.3) and for the second order in Equation (2.4).

$$\Delta Y_t = Y_t - Y_{t-1} \tag{2.3}$$

$$\Delta^2 Y_t = Y_t - 2Y_{t-1} + 2Y_t Y_{t-2} \tag{2.4}$$

To perform differencing, the Backshift operator is used, as defined in Equation (2.5).

$$B^d Y_t = Y_{t-d} \tag{2.5}$$

Thus, the differencing equation can be expressed in Equation (2.6).

$$Y_t^d = (1-B)^d Y_t, \qquad d = 1,2,\ldots \tag{2.6}$$

where Y_t^d represents the time series data after differencing of order d.

2.3 Autoregressive Integrated Moving Average (ARIMA)

ARIMA model is a combination of the Autoregressive (AR) model of order p, Moving Average (MA), and differencing (I) [1]. The AR(p) model describes a prediction as a function of previous values and the current residual in a time series, with its general equation written in Equation (2.7).

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \cdots + \phi_i Y_{t-p} + \varepsilon_t \tag{2.7}$$

The MA (q) model describes a prediction as a function of the previous and current time series residuals with the general equation written in Equation (2.8).

$$Y_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \cdots + \theta_i \varepsilon_{t-q} + \varepsilon_t \tag{2.8}$$

So in general the ARIMA model (p,d,q) can be written in Equation (2.9).

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)\varepsilon_t \tag{2.9}$$

With $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is the coefficient of the AR(p) model, $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ is the coefficient of the MA(q) model, and ε_t is the residual of the *t* time series data. The ARIMA parameters (p,d,q) are selected based on ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) analysis after differencing which are functions that show the correlation between observations at a certain time and previous observations.

To get the best ARIMA model, model diagnostic tests are needed to test residual assumptions, including the residual white noise test, residual normality test, and residual homoscedasticity test. White noise testing is used to see whether the model residuals are identical and independent using the Ljung-Box test with the following hypothesis [2].

H_0: $\rho_1 = \rho_2 = \cdots = \rho_k = 0$ (residuals are normally distributed)

H_1: there is at least one $\rho_k \neq 0$ with $k = 1,2,\ldots,K$ (residuals are not normally distributed)

With the critical region rejecting H_0 if the p-value < a or $Q > \chi^2_{\alpha,K-p-q}$ with the Q test statistic written in Equation (2.10).

$$Q = n(n+2)\sum_{k=1}^{K}(n-k)^{-1}\hat{\rho}_k^2 \tag{2.10}$$

where $\hat{\rho}_k$ is the autocorrelation value at lag k.

Next, the normality test of the residuals is conducted to check whether the model's residuals follow a normal distribution using the Kolmogorov-Smirnov test, with the following hypotheses.

H_0: $F_n(a_t) = F_0(a_t)$ or residual is *white noise*

H_1: $F_n(a_t) \neq F_0(a_t)$ or residual is not *white noise*

With the critical region rejecting H_0 if the p-value < a or $D > \chi^2_{\alpha,K-p-q}$ with the D test statistic written in Equation (2.11).

$$D = sup\ |\mathrm{F_n}(a_t) - F_0(a_t)| \tag{2.11}$$

Here, $F_0(a_t)$ represents the cumulative distribution function (CDF) of the normal distribution, $F_n(a_t)$ is the CDF calculated from the sample data, and *sup* is the maximum difference between $F_n(a_t)$ and $F_0(a_t)$.

Next, the homoskedasticity test of residuals is performed to determine whether the residual variance is constant, using the ARCH-LM test with the following hypotheses.

H_0: $a_1 = a_2 = \cdots = a_i = 0$ (residual variance is constant)

H_1: there is at least one $a_i \neq 0$ with $i = 1,2,\ldots,q$ (residual variance is constant)

With the critical region rejecting H_0 if the p-value < a or $LM > \chi^2_{\alpha,q}$ with the LM test statistic written in Equation (2.12).

$$LM = nR^2 \tag{2.12}$$

Here, R^2 is the coefficient of determination of the model.

The selection of the best ARIMA model can be determined using Akaike's Information Criterion (AIC), which considers the number of parameters in the model. The formula for AIC is given in Equation (2.13).

$$AIC = n \ln \hat{\sigma}_{\alpha}^{2} + 2m \tag{2.13}$$

The best ARIMA model is the one with the smallest AIC value.

2.4 Linearity Test

In this study, after the data was analyzed using the ARIMA model, the residuals from the model were tested to determine its linearity using the terasvirta test. The Terasvirta test is a linearity testing method developed based on a neural network model. This test was developed using the Taylor series expansion approach and is included in the Lagrangian Multiplier (LM) test category. The following hypothesis is used in the Terasvirta test [16].

H_0 : $f(x)$ is a linear function in x (linear model)
H_1 : $f(x)$ is a nonlinear function in x (nonlinear model)
Calculate the test statistical value using Equation (2.14).

$$F_{hit} = \frac{\frac{SSR_0 - SSR_1}{m}}{\frac{SSR_1}{(n-p-1-m)}} \tag{2.14}$$

With SSR_0 is the total squared residual, while SSR_1 is the total squared residual calculated using the Taylor expansion approach. The variable *n* represents the amount of data, *p* indicates the number of predictor variables, and *m* is an additional predictor resulting from the Taylor expansion approach. Decision to reject H_0 if $F_{hit} > F_{(n-p-1-m)}$ or p-value is less than the significance level of 0.05.

2.5 Support Vector Regression (SVR)

Support Vector Regression (SVR) is an extension of the Support Vector Machine (SVM) method specifically designed to address and solve regression problems. The main concept of SVR involves dividing the data into two parts: training data and testing data. A regression function is then determined on the training data with a specified deviation boundary to produce predictions close to the actual values. Furthermore, SVR is based on the principle of Structural Risk Minimization (SRM), which aims to estimate the risk function by minimizing the upper bound of generalization error [9]. The general formula for the linear regression function in SVR can be expressed in Equation (2.15).

$$f(\boldsymbol{x}_t) = \boldsymbol{w}\boldsymbol{x}_t + b, \qquad i = 1,,\ldots,n \tag{2.15}$$

Let $\boldsymbol{w}$ represent the weight vector of the model, and *b* denote the bias. The risk function is expressed in Equation (2.16).

$$R\big(f(\boldsymbol{x}_t)\big) = \frac{1}{2}\left|\left|\boldsymbol{W}\right|\right|^2 + C\sum_{i=1}^{n} E_{\varepsilon}\big(y_t - f(\boldsymbol{x}_t)\big) \tag{2.16}$$

Here, $\left|\left|\boldsymbol{W}\right|\right|$ is the function minimized to achieve the flattest possible regression function. The constant $C > 0$ defines the trade-off between the tolerance for deviations exceeding ε and the flatness of f. Equation (2.16) is minimized to estimate the weight vector and bias, expressed in Equation (2.17).

$$min_{\boldsymbol{w}} \frac{1}{2}\left|\left|W\right|\right|^2 + C\sum_{t=1}^{n} E_{\varepsilon}\big(y_t - f(\boldsymbol{x}_t)\big) \tag{2.17}$$

where E_{ε} is the ε-insensitive loss function defined as

$$E_{\varepsilon}\big(y_i - f(\boldsymbol{x}_t)\big) = \begin{cases} |y_i - f(\boldsymbol{x}_t)| - \varepsilon, & \text{untuk } |y_t - f(\boldsymbol{x}_t)| \geq \varepsilon \\ 0, & \text{lainnya} \end{cases}$$

with the constraints

$$y_i \leq f(\boldsymbol{x}_t) + \varepsilon$$
$$y_i \geq f(\boldsymbol{x}_t) - \varepsilon$$

All points within the range $f \pm \varepsilon$ are considered feasible, while points outside this range are infeasible. To address the infeasible region and optimize predictions, slack variables ξ_t and ξ_t^* are introduced to handle outliers. Equation (2.17) is then rewritten as Equation (2.18).

$$\min_{\boldsymbol{w},\xi_t,\xi_t^*} \frac{1}{2}\left||\boldsymbol{W}|\right|^2 + C\sum_{t=1}^{n}(\xi_t - \xi_t^*) \tag{2.18}$$

subject to

$$y_i \leq f(\boldsymbol{x}_t) + \varepsilon + \xi_t$$
$$y_i \geq f(\boldsymbol{x}_t) - \varepsilon - \xi_t^*$$
$$\xi_t, \xi_t^* \geq 0$$

To solve Equation (2.18), the Lagrangian method is applied with Lagrange coefficients $\alpha_t, \alpha_t^*, \eta_t, \eta_t^* \geq 0$, yielding

$$L_p = \frac{1}{2}\left||\boldsymbol{W}|\right|^2 + C\sum_{t=1}^{n}(\xi_t - \xi_t^*) - \sum_{t=1}^{n}(\eta_t\xi_t + \eta_t^*\xi_t^*) - \sum_{t=1}^{n}\alpha_t(\varepsilon + \xi_t + f(\boldsymbol{x}_t) - y_t) - \sum_{t=1}^{n}a_t^*(\varepsilon + \xi_t - f(\boldsymbol{x}_t) + y_t)$$

The dual Lagrangian for optimizing Support Vector Regression (SVR) is expressed in Equation (2.19).

$$L_d(a_t, a_t^*) = \sum_{t=1}^{n} y_t(a_t - a_t^*) - \varepsilon\sum_{t=1}^{n}(a_t - a_t^*) - \frac{1}{2}\sum_{t=1}^{n}(a_t - a_t^*)(a_t - a_t^*)\boldsymbol{x}_i\boldsymbol{x} \tag{2.19}$$

The optimal solution for $\boldsymbol{w}$, derived from Lagrange multipliers α_i and α_i^*, is given in Equation (2.20).

$$w = \sum_{t=1}^{n}(a_t - a_t^*)\boldsymbol{x}_t \tag{2.20}$$

The linear SVR regression function is written in Equation (2.21).

$$f(\boldsymbol{x}_t) = \sum_{t=1}^{n}(a_t - a_t^*)\boldsymbol{x}_t\,\boldsymbol{x} + b \tag{2.21}$$

The bias b is estimated under the Karush-Kuhn-Tucker (KKT) conditions

$$a_t\big(\varepsilon + \xi_t - y_t + f(\boldsymbol{x}_t)\big) = 0$$
$$a_i\big(\varepsilon + \xi_t^* - y_t - f(\boldsymbol{x}_t)\big) = 0$$
$$(C - a_t)\xi_t = 0$$
$$(C - a_t)\xi_t^* = 0$$

Using KKT conditions, b is computed as $b = y_t - \boldsymbol{w}\boldsymbol{x}_t - \varepsilon$ for $0 \leq a_t \leq C$ or $b = -y_t + \boldsymbol{w}\boldsymbol{x}_t - \varepsilon$ for $0 \leq a_t^* \leq C$. For nonlinear cases, SVR utilizes a mapping function φ to transform $\varphi: \boldsymbol{x}_t \to \varphi(\boldsymbol{x})$ into higher dimensions, expressed as Equation (2.22) and Equation (2.23).

$$f(\boldsymbol{x}_i) = \boldsymbol{w}\varphi(\boldsymbol{x}_t) + b, \qquad i = 1,, \ldots, n \tag{2.22}$$

or equivalently

$$f(x_t) = \sum_{t=1}^{n}(a_t - a_t^*)\varphi(x_t)\varphi(x) + b \tag{2.23}$$

Since $\varphi: \boldsymbol{x}_t \to \varphi(\boldsymbol{x})$ is often unknown, a kernel function is used Equation (2.24).

$$K(x_t, x) = \varphi(x_t)\varphi(x) \tag{2.24}$$

Common kernel functions for SVR are shown in Table 2.2.

Table 2.2. Kernel Functions in SVR Model

Kernel Functions	Formula
Linear	$K(x_t, x) = x_t^T x$
Radial Basis Function (RBF)	$K(x_t, x) = \exp\left(-\gamma\|\|x_t - x\|\|^2\right)$
Sigmoid	$K(x_t, x) = \tanh\left(\gamma(x_t^T x) + r\right)$
Polynomial	$K(x_t, x) = (\gamma(x_t^T x) + r)^p, p = 1,2, \ldots$

Optimal parameter selection in the SVR model can be performed using grid search optimization, a method that tests various parameter combinations to minimize prediction error [17]. This process divides the parameter ranges $C > 0$, $\gamma > 0$, and ε into a grid and evaluates all combinations to identify the optimal parameters on the training data, resulting in an accurate model for testing data. For time series data, grid search employs the concept of walk-forward validation, which tests the model incrementally over time using an out-of-sample approach. The grid search procedure can be divided into two stages, such as loose grid which explores C and γ over a broad range, and finer grid which refines the search within narrower ranges to achieve the lowest error. The parameter ranges for the loose grid stage are presented in Table 2.3.

Table 2.3. Range of Parameter Values in the Grid Search Method

Parameter	Value Range
Cost (C)	$2^{-10}, 2^{-9}, 2^{-8}, \ldots, 2^{8}, 2^{9}, 2^{10}$
Gamma (γ)	$2^{-10}, 2^{-9}, 2^{-8}, \ldots, 2^{8}, 2^{9}, 2^{10}$
Espilon (ε)	$0, 0.01, \ldots, 0.09, 0.1$
$coef.0$	$0.1, 0.2, \ldots, 1.00$
$degree$	$1, 2, 3$

2.6 Hybrid ARIMA-SVR

The hybrid ARIMA-SVR model is a combination of the ARIMA model and the SVR model. In this hybrid approach, time series data is assumed to have two components, a linear component captured by ARIMA and a nonlinear component modeled using SVR. This method leverages the strength of ARIMA in modeling linear dependencies and the ability of SVR to capture nonlinear relationships in residual patterns. Based on Sihombing et al. (2022), time series data consisting of a combination of linear and nonlinear components can be represented systematically in Equation (2.25) [15].

$$Y_t = L_t + N_t \tag{2.25}$$

where

$$\phi_p(B)(1-B)^d L_t = \theta_q(B)\varepsilon_t \tag{2.26}$$

With Y_t is the t time series data, L_t is the t linear component of the ARIMA model, and N_t is the t nonlinear component of the SVR model [15]. The nonlinear component $\widehat{N_t}$ is assumed to be present in the residuals of the ARIMA model defined as $\varepsilon_t = Y_t - \widehat{L_t}$ that modeled with SVR written in Equation (2.27).

$$\widehat{N_t} = f(\varepsilon_t) = \sum_{t=1}^{n} (a_t - a_t^*)\varphi(\varepsilon_t)\varphi(\varepsilon) + b \tag{2.27}$$

Then, $\widehat{L_t}$ is the estimation result of the ARIMA model at time t. This residual series is then used as input for SVR modeling, aiming to capture nonlinear dependencies. Hence, the final hybrid ARIMA-SVR model can be written in Equation (2.28).

$$\widehat{Y_t} = \widehat{L_t} + \widehat{N_t} \tag{2.28}$$

with $\widehat{Y_t}$ is the estimation result of the hybrid ARIMA-SVR model at time t.

2.7 Model Validation

The hybrid ARIMA-SVR model is benchmarked against standalone ARIMA and SVR models using Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). MAPE was added to overcome the problem of the error size which can be distorted when the actual data is larger or smaller than the predicted data. MAPE is calculated by determining the absolute difference between predicted data and actual data for each time period, then dividing it by the actual value in that period, and finally taking the average of the absolute error percentage [8]. The formula for determining MAPE value is written in Equation (2.29).

$$MAPE = \frac{\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|}{n} \tag{2.29}$$

With Y_t = t-th observation and $\hat{Y}_t$ = t-th prediction. Based on MAPE research, the conclusion of the accuracy of the forecasting method is presented in Table 2.4.

Table 2.4. Criteria MAPE

MAPE Value	Criteria
MAPE ≤ 10%	Very Accurate
10% < MAPE ≤ 20%	Accurate
20% < MAPE ≤ 50%	Less Accurate
MAPE > 50%	Not Accurate

MAE is a common metric used to measure the accuracy of a forecasting model by calculating the average absolute difference between the actual and predicted values [6]. Unlike MSE, MAE does not square the errors, making it less sensitive to outliers. A lower MAE value indicates higher model accuracy. The formula of MAE is written in Equation (2.30).

$$MAE = \frac{\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right|}{n} \tag{2.30}$$

MSE is the average of the squared errors and RMSE measures the average magnitude of prediction errors, giving greater weight to larger errors due to squaring. A lower MSE and RMSE indicates better predictive accuracy. The formula for determining the MSE and RMSE value is written in Equation (2.31) and (2.32).

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2 \tag{2.31}$$

$$RMSE = \sqrt{MSE} \tag{2.32}$$

2.8 Research Data

The data used in this study are secondary data, specifically daily nickel futures price data in Indonesia for the period from January 3rd, to December 29th, 2023, encompassing 254 observations. These data were obtained from the official website https://id.investing.com/commodities/nickel-historical-data. The independent variable (X) is represented by the daily time index, while the dependent variable (Y) corresponds to the daily nickel prices. This research adopts a quantitative approach, focusing on comparing the predictive accuracy of ARIMA, SVR, and hybrid ARIMA-SVR models to determine the best approach in forecasting daily nickel prices in 2023.

2.9 Data Analysis Procedure

1. Conducting descriptive analysis to provide an overview of the data, including the mean, standard deviation, variance, maximum, and minimum values.

2. Developing a linear component model using the Autoregressive Integrated Moving Average (ARIMA) method through the following steps:
   a. Splitting the data into training and testing sets with a 90:10 ratio.
   b. Testing variance stationarity using the Box-Cox transformation as in Equation (2.2) and mean stationarity using the Augmented Dickey-Fuller (ADF) test as in Equation (2.1).
   c. Identifying the ARIMA model based on ACF and PACF plots.
   d. Conducting parameter significance testing for the ARIMA model. If multiple significant models are identified, the model with the smallest Akaike Information Criterion (AIC) value using Equation (2.13) will be selected.
   e. Performing diagnostic tests, including white noise testing using the Ljung-Box test according to Equation (2.10), normality test using Kolmogorov-Smirnov test as in Equation (2.11), and homoskedasticity testing using ARCH-LM test as in Equation (2.12).
   f. Formulating the ARIMA model equation based on Equation (2.9).
3. Constructing a hybrid ARIMA-SVR model by analyzing the nonlinear component of the residuals from the best ARIMA model using Support Vector Regression (SVR) through the following steps:
   a. The nonlinearity pattern in the residual data from ARIMA model was tested using the Terasvirta test according to Equation (2.14).
   b. Using the residuals from the best ARIMA model as input for SVR modeling.
   c. Identifying significant timelags from the ARIMA residuals based on PACF plots.
   d. Determining optimal parameters using the grid search method with the parameter space as in Table 2.3 through the tuning process for each kernel in Table 2.2 to obtain the optimal SVR model with the smallest MAPE value using Equation (2.30).
   e. Formulating the hybrid ARIMA-SVR model equation based on Equation (2.28), which integrates the linear ARIMA and nonlinear SVR components.
4. Conducting SVR analysis directly on the data through the following steps:
   a. The nonlinearity pattern in the data was tested using the Terasvirta test according to Equation (2.14).
   b. Identifying significant timelags in the data based on PACF plots.
   c. Splitting the data into training and testing sets with a 90:10 ratio.
   d. Determining optimal parameters using the grid search method with the parameter space as in Table 2.3 through the tuning process for each kernel in Table 2.2.
   e. Formulating the SVR model equation based on Equation (2.23).
5. Comparing the prediction results of the three models ARIMA, hybrid ARIMA-SVR, and SVR by comparing their MAPE using Equation (2.29), MAE using Equation (2.30), MSE using Equation (2.31), and RMSE using Equation (2.32) to identify the best approach for forecasting daily nickel prices in 2023.

3. RESULT AND DISCUSSION

3.1. Descriptive Statistics

This research uses secondary data, specifically daily nickel futures prices in Indonesia for the period from January 3rd, 2023, to December 29th, 2023, comprising 254 observations. A summary of the daily nickel futures prices is presented in Table 3.1.

Table 3.1. Descriptive Statistics

Data	Data Amount	Mean	Minimum	Maksimum
Train	229	1902.2	1413.5	2607.5
Test	25	1422.5	1400.8	1449.4
Overall	254	1855	1400.8	2607.5

Based on Table 3.1, it can be seen that the nickel price plummeted to 1400.8 on December 14, 2023. The decline was caused by operational disruptions at the ferronickel processing and refining business unit in Pomalaa, Southeast Sulawesi, influenced by reduced availability and declining nickel ore quality. Meanwhile, the maximum nickel price reached 2607.5 on January 4th, 2023. Nevertheless, nickel commodity prices were recorded to weaken at the beginning of 2023. The plot of daily nickel futures prices is shown in Figure 3.1.

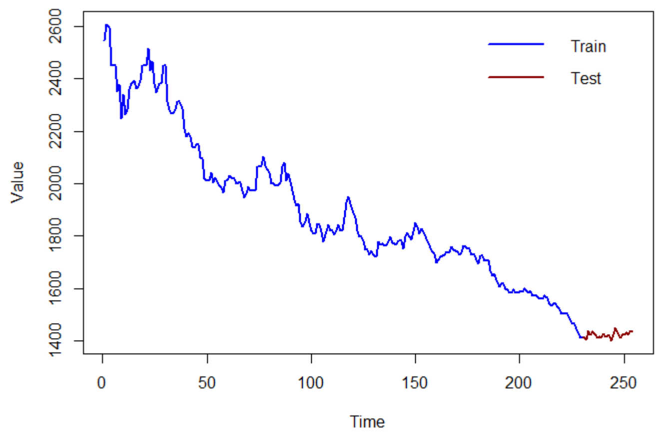


Figure 3.1. Plot of Daily Nickel Futures Price Data

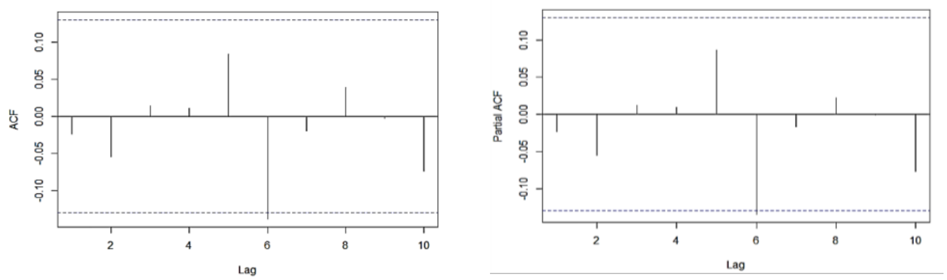
Based on Figure 3.1, it shows the model's performance on the training data. The graph appears to be fluctuating initially but tends to gradually decrease over time. The graph on the test data has lower values compared to the training data at the end of the period, with a more stable range.

3.2. Autoregressive Integrated Moving Average (ARIMA)

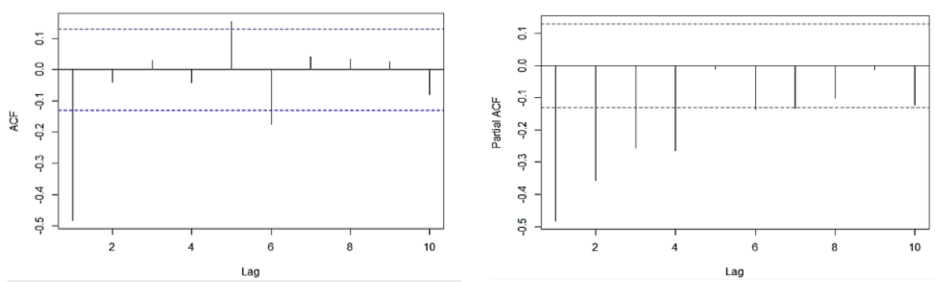
Before conducting the ARIMA analysis, it is necessary to check whether the data is stationary in mean and variance. The test for stationarity in variance is conducted by determining the lambda λ value on the nickel price training data. The obtained λ value is -1, thus data transformation is required. The data that has undergone the transformation process is referred to as data that is stationary in variance. After that, the testing continues to evaluate the stationarity of the data in the mean. This test was conducted using the Augmented Dickey-Fuller (ADF) method presented in Table 3.2.

Table 3.2. ADF Test Results

Test Statistics	Results		
	Before Differencing	Differencing 1	Differencing 2
Dickey-Fuller	-2.3303	-5.906	-9.8557
P-value	0.4372	0.01	0.01



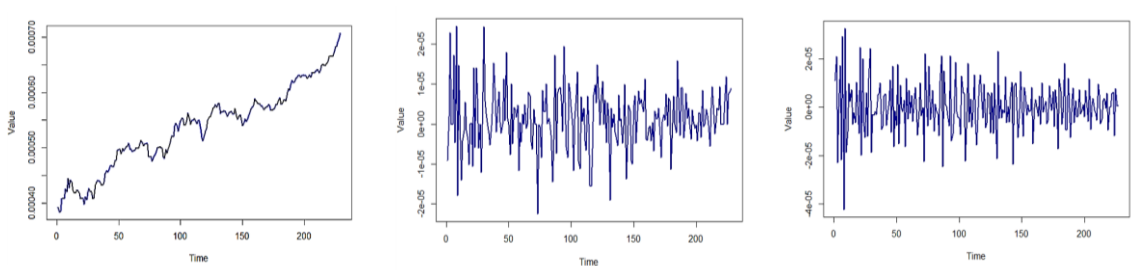
(a)



(b)

Figure 3.2. ACF and PACF plots (a) after differentiation 1 and (b) after differentiation 2

Based on Table 3.2, it can be seen that the nickel training data has a p-value of 0.4372 before differencing, which is greater than the significance level α of 0.05, thus failing to reject H_0 indicating that the data is not stationary in mean. Since the data is not stationary in mean, first differencing is required, which shows that the nickel training data has a p-value of 0.01 after first differencing, which is smaller than the significance level α of 0.05, thus rejecting H_0 indicating that the data is stationary in mean. However, it can be seen in Figure 3.2(a) that the ACF and PACF plots after differencing 1 do not show lag 1, so differencing 2 is necessary. After differencing 2, a p-value of 0.01, which is smaller than the significance level α of 0.05 was obtained leading to the rejection of H_0 indicating that the data is now stationary in mean. Additionally, lag 1 in the ACF and PACF plots in Figure 3.2(b) has appeared, allowing for the identification of the ARIMA model. The plots of the data after transformation and differencing are shown in Figure 3.3, where the plot after differencing 2 indicates that the data is now stationary in both mean and variance.



(a) (b) (c)

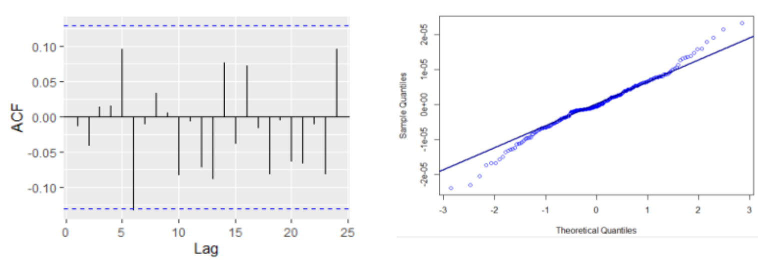
Figure 3.3. (a) Box-Cox Transformation Plot, (b) Differentiation 1, and (c) Differentiation 2

After the data is stationary in mean and variance, the next step is to identify the ARIMA model. Based on the ACF and PACF plots after differencing 2 in Figure 3.3(c), the ARIMA model that can be formed is presented in Table 3.3 along with the probability values $\Pr(>|z|)$ to test the significance of the parameters.

Table 3.3. Parameter Significance Testing

Model	Parameter	Estimation	Prob	Result	AIC
ARI (1,2)	$\hat{\phi}_1$	-0.4834	2.2 10^{-16}	**Significant**	-4617.84
ARI (2,2)	$\hat{\phi}_1$	-0.6708	2.2 10^{-16}	**Significant**	-4649.48
	$\hat{\phi}_2$	-0.3771	1.6 10^{-9}	**Significant**	
ARI (3,2)	$\hat{\phi}_1$	-0.7693	2.2 10^{-16}	**Significant**	-4664.19
	$\hat{\phi}_2$	-0.5582	5.5 10^{-14}	**Significant**	
	$\hat{\phi}_3$	-0.2709	3.1 10^{-5}	**Significant**	
IMA (2,1)	$\hat{\theta}_1$	-1.0000	2.2 10^{-16}	**Significant**	-4713.10
ARIMA (1,2,1)	$\hat{\phi}_1$	-0.0194	0.7712	Not Significant	-4711.19
	$\hat{\theta}_1$	0.0668	2.10^{-16}	Significant	
ARIMA (2,2,1)	$\hat{\phi}_1$	-0.0210	0.7525	Not Significant	-4709.78
	$\hat{\phi}_2$	-0.0515	0.4298	Not Significant	
	$\hat{\theta}_1$	-1.0000	2.10^{-16}	Significant	
ARIMA (3,2,1)	$\hat{\phi}_1$	-0.0203	0.7613	Not Significant	-4707.85
	$\hat{\phi}_2$	-0.0508	0.4462	Not Significant	
	$\hat{\phi}_3$	0.0169	0.8039	Not Significant	
	$\hat{\theta}_1$	-1.0000	2.10^{-16}	Significant	

Based on Table 3.3, it shows that all parameters in the ARI (1,2), ARI (2,2), ARI (3,2), and IMA (2,1) models are significant. The selection of the best ARIMA model can be seen from the smallest AIC value, where the IMA (2,1) model has the smallest AIC value of -4713.1. therefore, a model diagnostic test in the form of a white noise test on the residuals of the IMA (2,1) model was conducted, as shown in Figure 3.4.



(a) (b)

Figure 3.4. Model Diagnostic Test, (a) Residual ACF and (b) Residual Normality Test

Figure 3.4(a) shows that the lag in the ACF plot of the IMA (2,1) model residuals is within the significance line, indicating that the model residuals are not correlated with each other. Furthermore, Figure 3.4(b) shows that the model residuals are normally distributed because their spread approaches the diagonal line. This is reinforced by the Kolmogorov-Smirnov test with a p-value of 0.1006, indicating that the normality test for the residuals is met. In addition, the Ljung-Box test results with a p-value of 0.5411 indicate that the model residuals are white noise. Therefore, the IMA (2,1) model meets the white noise assumption. Next, a homoscedasticity test was conducted on the data using the ARCH-LM test to determine whether the variability between data groups is constant or not. The result of the ARCH-LM test with a p-value of 0.671 indicates that the residuals of the IMA (2,1) model have constant variance, thus fulfilling all model diagnostic tests.

Based on the parameter estimates in Table 3.3, the IMA (2,1) model equation can be written in Equation (3.1).

$$\begin{aligned}\widehat{L_t}^* &= 2Z_{t-1} - Z_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1}\\ &= 2Z_{t-1} - Z_{t-2} + \varepsilon_t - \varepsilon_{t-1}\end{aligned} \quad (3.1)$$

Because the Box-Cox transformation with a λ value of -1, the inverse of that transformation is $Z_t = \frac{1}{Y_t}$ so that the equation of the IMA (2,1) model is shown in the following Equation (3.2).

$$\widehat{L_t} = 2\frac{1}{Y_{t-1}} - \frac{1}{Y_{t-2}} + \varepsilon_t - \varepsilon_{t-1} \quad (3.2)$$

From Equation 4.2, the residual of the IMA (2,1) model is obtained as $\varepsilon_t = Y_t - \widehat{L_t}$.

3.3. Hybrid ARIMA-SVR

The results of testing the linearity of the residuals of the best IMA (2,1) model using the Terasvirta test are shown in Table 3.4.

Table 3.4. ARIMA Residual Linearity Testing

Data	Test Statistics χ^2	p-value
IMA (2,1) residuals	37.639	$6.712.10^{-9}$

Based on Table 3.4, it can be observed that the Terasvirta test statistic value $37.639 > F_{(0.05;1;250)} = 3.8789$ or p-value $6.712.10^{-9} < \alpha = 0.05$. Using $\alpha = 0.05$ corresponds to a 95% confidence level, meaning there is 95% confidence that the conclusion drawn is not due to random chance. Therefore, the decision is to reject H_0 indicating that the ARIMA residuals contain a nonlinear pattern, so the analysis can proceed with SVR modeling.

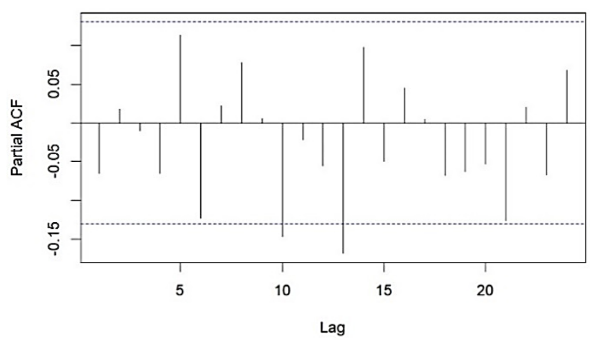


Figure 3.5. PACF Plot of ARIMA Residuals

The ARIMA residuals were transformed into time lags. Based on the PACF plot in Figure 3.5, the significant lags identified are lag 10 and lag 13, corresponding to Y_{t-10} and Y_{t-13}. These lags indicate that the values of the time series from 10 and 13 prior have a significant influence on the current value. Since the ARIMA model involved two differencing steps, the input data for the SVR model must start from the 16th observation, as the first two periods are removed during the differencing process.

The optimal parameter values for tuning the ARIMA residual model in SVR modeling used the grid search method by determining the same range of parameter values as in Table 2.3 Selection of the best kernel with optimal parameters is based on the minimum MAPE value. The optimal parameter evaluation of the tuning results of the four kernel functions on the ARIMA residual model is presented in Table 3.5.

Table 3.5. Grid Search Results and Kernel Function Evaluation in Residual Modeling with the SVR Method

Kernel	Parameter	MSE		MAPE	
		Training	Testing	Training	Testing
Linear	$\varepsilon = 0.7$ $C = 2^{-10}$	735.6188	2562.47	0.9864304	3.070923
Radial	$\varepsilon = 0.1$ $C = 2^{4}$ $\gamma = 2^{9}$	300.9595	2524.145	0.7657964	3.065773
Polynomial	$\boldsymbol{\varepsilon = 0.1}$ $\boldsymbol{C = 2^{-4}}$ $\boldsymbol{\gamma = 2^{-3}}$ $\boldsymbol{coef.0 = 0.1}$ $\boldsymbol{degree = 3}$	**703.6092**	**633.6779**	**0.9559922**	**1.441836**
Sigmoid	$\varepsilon = 0.7$ $C = 2^{-4}$ $\gamma = 2^{-4}$ $coef.0 = 1$	751.5647	631.2789	1.242527	1.452575

Through the grid search process in Table 3.5, the optimal values of the hybrid ARIMA-SVR model parameters are $\varepsilon = 0.1$, $C = 2^{-4}$, $\gamma = 2^{-3}, coef.0 = 0.1,$ and $degree = 3$ with the polynomial as the best kernel function. Based on the polynomial kernel model in Table 2.2, the SVR model of the ARIMA residuals is written in Equation (3.3).

$$\widehat{N}_t = \sum_{t=1}^{m} (a_t - a_t^*)(\gamma x_t x + r)^d + b \text{, with } 0 < a_t \le C, 0 < a_t^* \le C$$

$$\widehat{N}_t = \sum_{t=16}^{214} (a_t - a_t^*)(0.125\ x_t x + 0.1)^3 + 0.005395823 \tag{3.3}$$

with $a_t = 0.0625$.

After determining the best parameters of the ARIMA and SVR models, the next step is to combine the two component models. Based on Equations (3.2) and (2.2), the hybrid ARIMA-SVR model equation is written in Equation (3.4).

$$\hat{Y}_t = \hat{L}_t + \widehat{N}_t$$

$$\hat{Y}_t = 2\frac{1}{Y_{t-1}} - \frac{1}{Y_{t-2}} + \varepsilon_t - \varepsilon_{t-1} + \sum_{t=16}^{214} (a_t - a_t^*)(0.125\ x_t x + 0.1)^3 \tag{3.4}$$

with $a_t = 0.0625$.

Based on the hybrid ARIMA-SVR model in Equation (3.4), the linear component of the nickel price training data is modeled using ARIMA, where the nickel price at time t is influenced by price changes two days earlier and the residual of the nickel price one day earlier. The nonlinear component is modeled using SVR, which captures the residual patterns from ARIMA $t - 10$ and $t - 13$ with a degree 3 polynomial kernel by $\gamma = 0.125$ and a Lagrange coefficient between 0 and 0.0625. This nonlinear pattern, which cannot be explained by ARIMA is successfully captured by SVR. The residual linearity of the hybrid ARIMA-SVR model was then validated using the Teräsvirta test, as shown in Table 3.6.

Table 3.6. ARIMA-SVR Hybrid Residual Linearity Testing

Data	p-value
Hybrid ARIMA-SVR Residual	0.07586

Terasvirta test results on the hybrid model shown in Table 3.6 it can be seen that the p-value of the terasvirta test results is $0.07586 > \alpha = 0.05$, so the decision is to fail to reject H_0, which means that the residuals of the hybrid ARIMA-SVR are linear. It indicates that the residuals are now linear, suggesting that the hybrid model may has successfully captured both linear and nonlinear patterns in the data.

3.4. Support Vector Regression (SVR)

Linearity test using the Terasvirta test on Indonesian nickel price data during the period January 3rd, 2023 to December 29th, 2023 obtained the results presented in Table 3.7.

Table 3.7. SVR Linearity Testing

Data	Test Statistics χ^2	p-value
Nickel Price 2023	149.23	$2.2.10^{-16}$

Based on Table 3.7, it can be observed that the Terasvirta test statistic value $149.23 > F_{(0.05;1;252)} = 3.8786$ or p-value $2.2.10^{-16} < \alpha = 0.05$, Therefore, the decision is to reject H_0 indicating that the ARIMA residuals contain a nonlinear pattern.

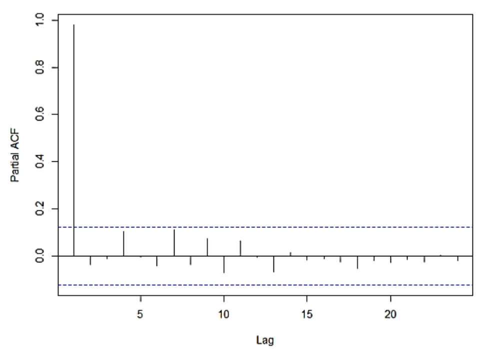


Figure 3.6. PACF Plot of SVR Data

Nickel price data shows that the significant lag only appears at lag 1 based on the PACF plot in Figure 3.6, meaning that the current nickel price value (Y_t) has a significant influence only from the nickel price value in one previous period (Y_{t-1}). Consequently, the input data for the SVR model begins from the second data point, as the first data point lacks sufficient lag information to serve as input for the model. The best SVR model parameters are determined by tuning the four kernel functions as presented in Table 2.2, using the grid search method within the parameter range specified in Table 2.3. The evaluation of model performance is shown in Table 3.8.

Table 3.8. Grid Search Results and Kernel Function Evaluation on the SVR Model

Kernel	Parameter	MSE		MAPE	
		Training	Testing	Training	Testing
Linear	$\varepsilon = 0.1$ $C = 2^8$	1020.632	135.6298	1.106037	0.6451849
Radial	$\boldsymbol{\varepsilon = 0.1}$	**774.7965**	**42.55313**	**1.058176**	**0.2532663**

Kernel	Parameter	MSE		MAPE	
		Training	Testing	Training	Testing
	$\boldsymbol{C = 2^9}$ $\boldsymbol{\gamma = 2^5}$				
Polynomial	$\varepsilon = 0.1$ $C = 2^4$ $\gamma = 2^2$ $coef.0 = 0.7$ $degree = 3$	1036.212	128.7259	1.099947	0.6130337
Sigmoid	$\varepsilon = 0.1$ $C = 2^9$ $\gamma = 2^{-8}$ $coef.0 = 0.3$	1005.683	137.0141	1.092687	0.64739

The tuning results in Table 3.8 show that the optimal values of the SVR model parameters for nickel price data are $\varepsilon = 0.1$, $C = 2^9$, and $\gamma = 2^5$ which produce minimum MSE and MAPE, both on training and testing data with radial as the best kernel function. Based on the radial kernel function SVR model in Table 2.2, the SVR model is shown in Equation (3.5).

$$K(x_t, x) = \sum_{t=2}^{79} (a_t - a_t^*) \exp\left(-32||x_t - x||^2\right) - 0.4288245 \qquad (3.5)$$

where $0 < a_t \leq 512,\ \ 0 < a_t^* \leq 512$.

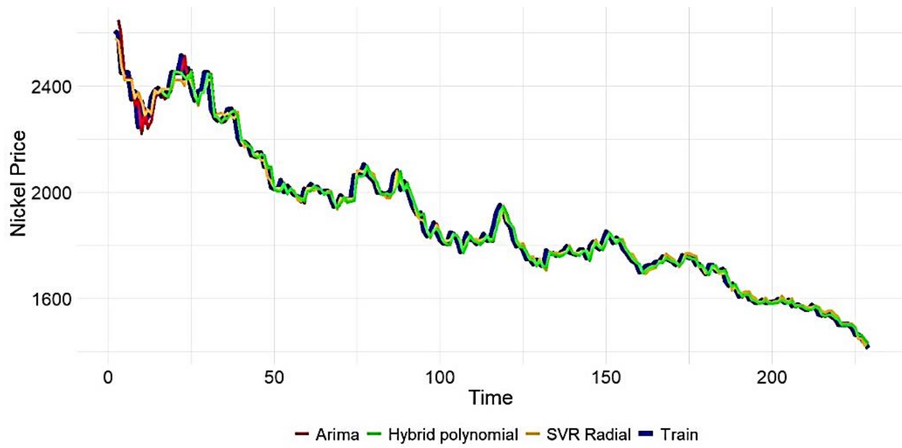
3.5. Comparison of Forecasting Accuracy of ARIMA, Hybrid ARIMA-SVR, and SVR Models

The results of the comparative evaluation of the three methods based on the forecasting accuracy values are presented in Table 3.9.

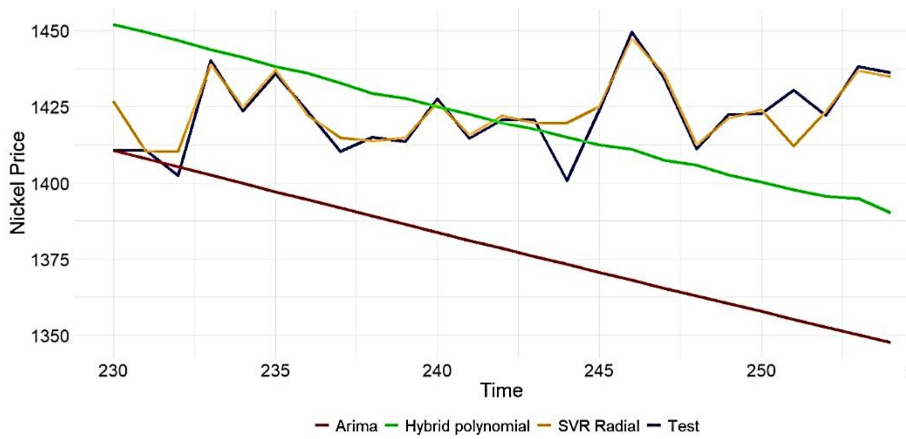
Table 3.9. Comparison of Accuracy Values

Model	MAPE		RMSE		MAE	
	Training	Testing	Training	Testing	Training	Testing
ARIMA	1.088411	3.076029	31.81666	50.70161	21.72896	43.90523
Hybrid ARIMA-SVR	0.955992	1.441836	26.42734	25.17296	18.49895	20.51660
SVR	**1.058176**	**0.253266**	**27.83517**	**6.523276**	**20.57020**	**3.586384**

Based on Table 3.9, it can be seen that the hybrid ARIMA-SVR and SVR models have better prediction accuracy than ARIMA on both training and testing data. hybrid ARIMA-SVR shows better performance on training data but produces performance that is not good enough in generalizing testing data. This can occur when the hybrid ARIMA-SVR model is too focused on learning the training data so that it has the potential for underfitting in testing, or it can be caused by the contribution of ARIMA which is too dominant in handling the linear component, so that SVR does not play a role enough to handle the nonlinear aspects of the testing data. Meanwhile, when using the SVR approach alone will provide better performance on testing data with a MAPE value of 0.2532% which the minimum among other models supported by the RMSE value 6.5232 and MAE value 3.5863 that also have the minimum testing prediction value where the SVR prediction plot in Figure 3.7(b) follows the actual testing data pattern so that it does not show symptoms of underfitting like the ARIMA or hybrid methods before. In accordance with the research objectives that focus on prediction results, that daily nickel price forecasting in 2023 using the SVR approach shows the best ability to handle testing data that has a more complex nonlinear component so as to provide more accurate predictions than ARIMA or hybrid ARIMA-SVR models. These results indicate that not always advanced methods such as the hybrid ARIMA-SVR method can produce better results [5].



(a)



(b)

Figure 3.7. Comparison of Data Plots, (a) Training Data and (b) Testing Data

Based on Figure 3.7(a), the model seems to follow the pattern of the training data quite well, showing similar predictive ability to the original data. While Figure 3.7(b) shows that Radial SVR appears superior in capturing the pattern of the testing data so that the training data pattern follows the original data pattern well, compared to ARIMA and hybrid ARIMA-SVR which are less successful in predicting this data, where the testing plot tends to be linear which may be due to the dominance of the ARIMA component which captures the main linear pattern in the data, so that hybrid SVR which is tasked with capturing nonlinear patterns from ARIMA residuals, may have a small contribution in the testing period, so that the nonlinear pattern does not look significant.

4. CONCLUSION

The purpose of this research focuses on the prediction results, so the best forecasting method for Indonesia's daily nickel prices is the SVR model that achieves the lowest MAPE of 0.2532% on the testing data, with parameters $\varepsilon = 0.1$, $C = 2^9$, and $\gamma = 2^5$ which gives the minimum RMSE and MAE values as well. These parameters were determined through a grid search optimization process, systematically evaluating different combinations to minimize prediction errors. The testing plot shows SVR closely follows the actual data. Although ARIMA and hybrid ARIMA-SVR perform well on training data, they are prone to underfitting, as seen in the testing plots. The ARIMA component captures the linear trend, while hybrid ARIMA-SVR's contribution to nonlinear patterns is minimal during testing, resulting in a less significant nonlinear pattern. The findings highlight the importance of accurate forecasting for market participants in optimizing production planning, managing risks, and making informed investment decisions, while policymakers can use these insights to ensure market stability, support domestic industries, and develop sustainable resource management strategies. However, since this study relies solely on historical price data, external factors such as policy changes and global market shocks were not considered, highlighting the need for future research to integrate broader economic indicators for improved forecasting accuracy. Technically, future researchers can develop a parameter space with a more detailed increment and adjust the data split for more accurate prediction results or other methods that can represent non-linear patterns, especially for time series data that are often affected by external factors.

REFERENCES

[1] Desi, D., Rizki, S., & Yundari, Y., 2022. Combined Model Time Series Regression – ARIMA on Stocks Prices. *Tensor: Pure and Applied Mathematics Journal*, Vol. 3, No. 2, 65-72.

[2] Farafilla, S. & Suharsono, A. 2023. Peramalan Penjualan Produk Baja dan Besi di PT MSU dengan Pendekatan Metode ARIMA dan Single Moving Average. *Jurnal Sains dan Seni ITS*, Vol. 12, No. 1, 88-95.

[3] Hutabarat, G, 2023. Sengketa Ekspor Nikel Indonesia dengan Uni Eropa di World Trade Organization. *Jurnal Ilmu Hubungan Internasional LINO*, Vol. 3, No. 2, 116–122.

[4] Jusmawati, J., Hadijati, M., & Fitriyani, N., 2020. Penerapan Model Vector Autoregressive Integrate Moving Average dalam Peramalan Laju Inflasi dan Suku Bunga di Indonesia. *Eigen Mathematics Journal*, Vol. 3, No. 2, 73–82.

[5] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. 2020. The M4 Competition: 100,000 Time Series and 61 Forecasting Method. *International Journal of Forecasting*, Vol. 36, No. 1, 54-74.

[6] Montesinos López, O. A., Montesinos López, A., & Crossa, J. 2022. Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer.

[7] Mubarak, A. M. N., Ariani, F., & Ridwan., 2022. Pengaruh Moisture Content Terhadap Zona Gradasi Tanah Tambang di Morowali Utara. *SAINTIS*, Vol. 3, No. 1, 50-57.

[8] Perdana, D. & Muklason, A. Machine Learning Untuk Peramalan Kualitas Indeks Standar Pencemar Udara DKI Jakarta Dengan Metode Hibrid ARIMAX-LSTM. *Journal of Computer Science and Applied Informatics*, Vol. 5, No. 3, 209-222.

[9] Purba, J. A. A, 2024. Batasan untuk Keamanan: Analisis Kebijakan Pembatasan Ekspor Nikel Indonesia Menggunakan Pendekatan Poskolonialisme. *Jurnal Hubungan Internasional*, Vol. 17, No. 1, 78–91.

[10] Purnama, D. I., 2021. Peramalan Harga Emas Saat Pandemi Covid-19 Menggunakan Model Hybrid Autoregressive Integrated Moving Average - Support Vector Regression. *Jambura Journal of Mathematics*, Vol. 3, No. 1, 52-65.

[11] Putri, N. N. & Abrar, H., 2023. Dampak Kebijakan Pembatasan Ekspor Nikel ke Uni Eropa: Global Trade Analysis Project (GTAP) Model. *SEIKO: Journal of Management & Business*, Vol 6, No. 2, 329-336.

[12] Safarudin, L., Syahrul, S., & Kumalasari, R., 2024. Estimasi Cadangan Nikel Laterit Menggunakan Metode Inverse Distance Weight dan Kriging di Blok X pada PT. Tambang Bumi Sulawesi, Kabaena Selatan Kabupaten Bombana, Sulawesi Tenggara. *Mining Science And Technology Journal*, Vol. 3, No. 1, 39-50.

[13] Salmon, D. L., Gelatan, L., Narew, I., & Murnitasari, 2024. Analisis Perbedaan Kinerja Keuangan Pada Perusahaan Tambang Nikel. *COSTING: Journal of Economic, Business, and Accounting*, Vol. 7, No. 4, 8212-8221.

[14] Salsabila, N. A., Wahyuningsih, S., & Purnamasari, I., 2022. Pemodelan Vector Autoregressive Exogenous (VARX) untuk Meramalkan Data Ekspor dan Impor Total di Indonesia. Jambura Journal of Probability and Statistics, Vol. 3, No. 2, 128-140.

[15] Sihombing, C. V. M., Martha, S., & Huda, N. M., 2022. Analisis Metode Hybrid ARIMA–SVR pada Indeks Harga Saham Gabungan. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, Vol. 11, No. 3, 413-422.

[16] Subiyanto, M. L., Amanada, Y., Fachrian, M. N., Afriani, Rohim, A. Y. B., & Chamidah, N., 2023. Peramalan Kasus Harian Monkeypox Dunia dengan Pendekatan Support Vector Regression. *Jurnal Aplikasi Statistika dan Komputasi Statistik*, Vol. 15, No. 1, 27-36.

[17] Sun, Y., Ding, S., Zhang, Z., & Jia, W., 2021. An Improved Grid Search Algorithm to Optimize SVR For Prediction. *Soft Computing*, Vol. 25, No. 7, 5633–5644.

[18] Syarifuddin, N., 2022. Pengaruh Industri Pertambangan Nikel Terhadap Kondisi Lingkungan Maritim di Kabupaten Morowali. *Journal Riset & Teknologi Terapan Kemaritiman*, Vol. 1, No. 2, 19-23.

[19] Wau, F. T., Kiton, M. A., Wau, M., & Fau, J. F., 2024. Analisis Strategis Kebijakan Hilirisasi Mineral: Implikasi Ekonomi dan Pengaruhnya terhadap Perekonomian Indonesia. Journal Publicuho, Vol. 7, No. 3, 1215-1224.