BAREKENG: Journal of Mathematics and Its Applications
March 2025 Volume 19 Issue 1 Page 0013-0024
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss1pp0013-0024

IMPLEMENTATION OF CROSS-VALIDATION ON HANG SENG INDEX FORECASTING USING HOLT'S EXPONENTIAL SMOOTHING AND AUTO-ARIMA METHOD

Christy Sheldy Sucipto[1], **Winita Sulandari**[2*], **Yuliana Susanti**[3]

[1,2,3]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret*
Jl. Ir. Sutami No. 36 Kentingan, Surakarta, 57126, Indonesia

*Corresponding author's e-mail: * **winita@mipa.uns.ac.id***

Article History:

Received: 1st January 2024
Revised: 13th October 2024
Accepted: 30th October 2024
Published: 13th January 2025

Keywords:

Exponential smoothing;
Multi-step forecasts;
Rolling window cross-validation;
Trend.

ABSTRACT

This study applies a rolling window cross-validation to evaluate the multi-step forecasts instead of using the traditional single split for Hang Sheng Index (HSI) forecasting. The forecasting methods discussed in this study are Holt's Exponential Smoothing and auto ARIMA, chosen because of their ability to model trend data as in the daily HSI. This research aims to evaluate up to five step forecast values obtained by the two forecasting methods built in the training data with rolling window cross-validation. In the experiment, each of the 21 auto ARIMA and Holt's models was constructed from 84 observations (as in-sample data) obtained from the rolling window cross-validation. The one to five step forecast values of daily HSI are then calculated using those models, and the accuracy of each forecast value is evaluated based on Mean Absolute Percentage Error (MAPE). The results show that the Auto ARIMA model produces a lower MAPE value than Holt's model, namely 2.9196%, 4.6553%, 6.4012%, 8.3083%, and 10.3781%, respectively, for one to five steps ahead. Therefore, auto ARIMA is more recommended for forecasting HSI values up to five steps ahead than Holt's method.

How to cite this article:

C. S. Sucipto, W. Sulandari, and Y. Susanti., "IMPLEMENTATION OF CROSS-VALIDATION ON HANG SENG INDEX FORECASTING USING HOLT'S EXPONENTIAL SMOOTHING AND AUTO-ARIMA METHOD", *BAREKENG: J. Math. & App.*, vol. 19, iss. 1, pp. 0013-0024, March, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*
Research Article · Open Access

1. INTRODUCTION

Forecasting accuracy can be obtained through the application of appropriate forecasting models on training and testing data. However, researchers generally agree that the accuracy of forecasting methods needs to be checked using testing data rather than the accuracy of the model on training data [1]. Evaluation of forecasting on testing data needs to be done to obtain a model that is not only good at training data but also good at generalizing to new data [2]. If the forecasting model is only good at training data, it can cause overfitting problems where the model is too precise in predicting training data so that it is less able to predict new data multiple steps ahead.

Time series cross-validation is a cross-validation technique used on time series data that maintains the time sequence to obtain the performance accuracy of the testing data in generalizing new data. There are several kinds of time series cross-validation that can be used, one of which is Rolling Window Cross-Validation. This technique maintains the size of the training data, by cutting one of the oldest observation data and adding one of the newest observation data at each model roll, resulting in a constant window length or training data period. The main purpose of truncating the oldest observation data and adding one new observation data in this technique is to update the model coefficients and equalize the comparison of forecasting accuracy between periods.

Updating the data tested in time series forecasting is important, especially in economic data. This is because economic data has erratic changes that can make previous data values less relevant to current conditions. Therefore, continuous data updates are required to produce more accurate forecasting values.

The time series data used in this study is the Hang Seng Index (HSI). The HSI is a stock index that is a leading indicator of Hong Kong's stock market performance and represents 65% of the stock exchange capitalization in Hong Kong [3]. In addition, the HSI is declared as one of the major stock market indices in Asia that has an influence on other countries. This is proven through research conducted by Aji and Abudanti [4] that the HSI positively influences the Composite Stock Price Index (JCI) on the Indonesia Stock Exchange. The annual report issued by the Hang Seng Index's Year-End Reports [5] states that the HSI decreased by 14.1% in 2021 and 15.5% in 2022. This statement indicates a downward trend pattern in the HSI data, which indicates that stock data tends to fluctuate and has a non-stationary pattern, so an appropriate forecasting method is needed to predict future HSI values.

Holt's Exponential Smoothing method, or what can be called Holt's method, is intended for forecasting with time series data that has a linear trend pattern and does not contain seasonality. Research that has been conducted by several researchers, such as Alias et al. [6] and Muchayan [7] prove that the forecasting error value obtained from the Holt's method is smaller than other methods for data with a linear trend pattern. The correlation between values in a time series and previous values in the same time series also needs to be known in order to know the right forecasting method in predicting. The ARIMA method can explain autocorrelation in a time series and is intended for data that does not meet the assumption of stationarity. The ARIMA method procedure in Ahmad and Ahmad [8] and Alias et al. [6] is done through the differencing examination stage and manual determination of model parameters. This requires a long time in determining the best ARIMA model because it requires checking the parameter significance test and selecting the smallest AIC value from each possible model obtained from several combinations of order values based on the identification of the model one by one.

In its development, Hyndman and Khandakar [9] created an algorithm to quickly obtain a model from the ARIMA method with optimal parameters using R software, called the Auto ARIMA method. Tiwari et. al [10] conducted research using the Auto ARIMA method on Nifty 50 stock price data with polynomial trend data patterns. The results show that the Auto ARIMA method is not good enough to be used in predicting future forecasting values. Research with the Auto ARIMA method has also been carried out by Kalyoncu et. al [11] to predict the stock market by applying fixed origin cross-validation. In Kalyoncu's research, it was found that Auto ARIMA has better prediction accuracy in short-term forecasting.

In the forecasting field, researchers certainly focus on model performance and accuracy resulting from forecasting results. The main goal is to obtain accurate and reliable forecasting values. Sulandari et al. [2] conducted research using time series cross-validation with a rolling window to evaluate the performance of a forecasting model applied to hourly electricity load data in Malaysia for one year. Numerous forecasting methods were implemented in the study, including ARIMA, Neural Network Autoregressive (NNAR), Exponential Smoothing, Singular Spectrum Analysis (SSA), and General Regression Neural Network

(GRNN). According to Sulandari et al. [2], the ARIMA, NNAR, and SSA models' Mean Absolute Percentage Error (MAPE) values are quite stable for forecasting up to seven steps ahead. In addition, it is known that the forecasting error values do not always increase as the forecasting step increases.

This research examines the performance evaluation of Holt's and Auto ARIMA methods in predicting HSI several periods in advance using Rolling Window Cross-Validation, in light of the aforementioned background. This type of cross-validation is chosen to evaluate the two methods objectively so that it can be ascertained thoroughly which one can produce better forecasts up to five steps ahead, measured based on the smallest MAPE value. Furthermore, accurate multi-step stock price forecasting results will help investors in making decisions related to better future economic growth prospects.

2. RESEARCH METHODS

2.1 Data Collection Methods

The data in this study uses secondary time series data taken from the *finance.yahoo.com* page. The time series data used is the weekly Hang Seng Stock Price Index in the period January 2021 to December 2022, totaling 105 data.

2.2 Time Series Cross Validation

In time series analysis, the validation technique for models is called time series cross-validation. The data are divided into two parts, training and testing datasets. The first part is used to build the model and the last to test the constructed model. In time series, the time sequence is an important thing to consider in the distribution of training and testing data on time series data. There are several time series cross-validation methods that can be used, including Fixed Origin Cross Validation, Rolling Origin Cross Validation and Rolling Window Cross Validation [12].

Fixed Origin Cross Validation is a common technique used by researchers in measuring forecasting accuracy by dividing data into two parts, namely training and testing using a single forecasting origin, in other words, using only one origin or one model in evaluating future forecasting values [13]. This results in the prediction results obtained will only be influenced by factors that occur in that one model, making it less reliable. The problem with Fixed Origin Cross-validation can be overcome using the Rolling Origin Cross-Validation technique, where the forecasting origin can be updated by adding one new observation data in each repetition of the model so that several models are obtained in evaluating the forecasting value [14]. Illustrations of how Fixed Origin Cross Validation and Rolling Origin Cross Validation work for one-step-ahead forecasting are shown in Figure 1 and Figure 2 respectively. Blue dots indicate training data and red dots indicate testing data.

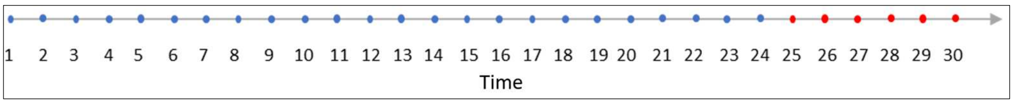


Figure 1. Illustration of Fixed Origin Cross Validation

By using Fixed Origin Cross Validation, some possible forecasting models are constructed from a training dataset, which in Figure 1 the models are estimated from the first twenty-four observations (blue dots) and the rest will be the testing data (red dots). Meanwhile, Rolling Origin Cross Validation provides several training datasets where each is updated by adding one new observation to the previous training data (see the blue dots in Figure 2, from line 1 to the last line).

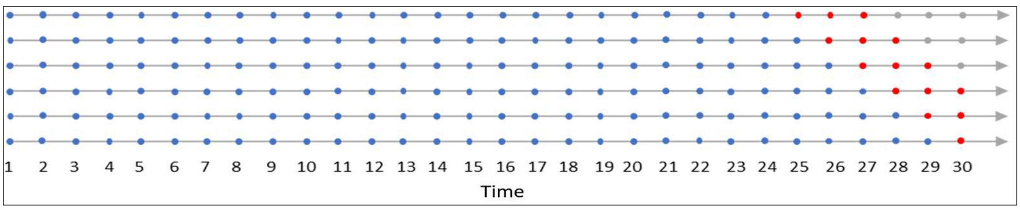


Figure 2. Illustration of Rolling Origin Cross Validation

In Rolling Origin Cross Validation, the accuracy comparison may be inconsistent because the second model is constructed from larger sample size of the training dataset than the first one, and so on [1]. There is an alternative to maintaining a constant size of the training data period, called Rolling Window Cross-Validation. Figure 3 shows an illustration of how Rolling Window Cross Validation works for multiple steps ahead forecasting. The blue dots show the training data while the red, black and green dots show the testing data which are the forecasting values one to three steps ahead sequentially in multiple steps ahead forecasting.

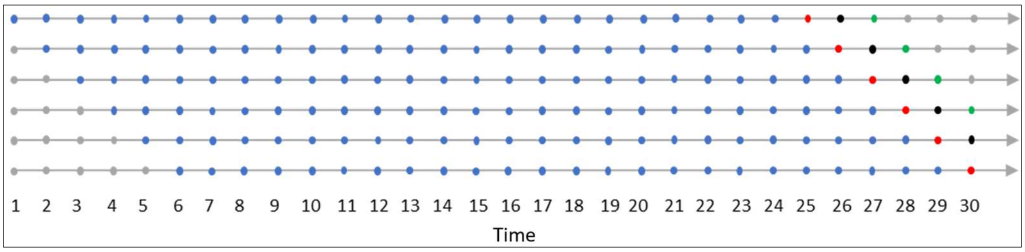


Figure 3. Illustration of Rolling Window Cross Validation

The Rstudio software has a function to simplify the calculation of forecasts that apply time series cross-validation, namely the '*tsCV*' function [15]. The output given from this function is the forecasting error by applying the forecasting function to a subset of the time series data $\boldsymbol{Y_t}$ using the rolling origin. The forecasting error output from the '*tsCV*' function is obtained through Equation (1).

$$e_{t+h} = Y_{t+h} - \hat{Y}_{t+h|t} \tag{1}$$

where $\boldsymbol{e_t}$ is residual value at time $\boldsymbol{t}$, $\boldsymbol{Y_t}$ is actual data value at time $\boldsymbol{t}$, $\boldsymbol{\hat{Y}_t}$ is forecast value at time $\boldsymbol{t}$, and $\boldsymbol{h}$ is forecast horizon.

2.3 Exponential Smoothing Holt

The Exponential Smoothing method is a relatively simple but effective forecasting method that produces reliable forecasts [16]. Exponential Smoothing Holt is a further development of the Single Exponential Smoothing method [17]-[18]. Exponential Smoothing Holt or Holt's method is specialized for time series data that has a linear trend pattern and can do multiple-step ahead forecasting. The smoothing equations for level and trend in this method are written in Equation (2) and Equation (3), while for the forecasting value h periods ahead is shown in Equation (4) [19].

$$l_t = \delta Y_t + (1-\delta)(l_{t-1} + b_{t-1}) \tag{2}$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \tag{3}$$

$$\hat{Y}_{t+h} = l_t + hb_t \tag{4}$$

where $\boldsymbol{l_t}$ is smoothing level value at time $\boldsymbol{t}$, $\boldsymbol{b_t}$ is smoothing trend value at time $\boldsymbol{t}$, $\boldsymbol{\delta}$ is smoothing parameter for level ($\boldsymbol{0 \leq \delta \leq 1}$), $\boldsymbol{\beta}$ is smoothing parameter for tren ($\boldsymbol{0 \leq \beta \leq 1}$).

In its implementation, Holt's parameters can be estimated quickly with the help of software. In R software, there are *HoltWinters*() and *ets*() functions that can estimate Holt's parameters with different estimation methods and provide different initial values. The differences are shown in Table 1.

Table 1. The Differences of *HoltWinters*() and *ets*() Functions

Item	***HoltWinters*()**	***ets*()**
1. Determining the initial states	minimising mean squared error (MSE)	maximizing the likelihood function
2. Estimating smoothing parameters	Based on the heuristic value	Maximizing the likelihood function

2.4 Auto ARIMA

The ARIMA method, also known as the Box-Jenkins approach, is a very common method used for nonstationary time series analysis. According to Box et al. [20], there are three main steps taken to build a time series model, namely model identification, parameter estimation, and model diagnostic test or model feasibility test. Before the model identification step, the stationarity of time series data in the ARIMA method needs to be checked, which is stationarity in variance and average [21]. Stationarity in variance is checked by Box-Cox transformation, while stationarity in the mean is checked through the ACF plot which is then subjected to differencing if the data is not yet stationary in the mean. The general equation of the ARIMA model is shown in Equation (5) [22].

$$\phi_p(B)(1-B)^d Y_t = c + \theta_q(B)\varepsilon_t \; ; \; \{\varepsilon_t\} \sim (0, \sigma^2) \tag{5}$$

where $\boldsymbol{\phi_p}$ is AR model parameters of order $\boldsymbol{p}$, $\boldsymbol{B}$ is backshift operator, $\boldsymbol{d}$ is amount of differencing, $\boldsymbol{c}$ is constant, $\boldsymbol{\theta_q}$is MA model parameters of order $\boldsymbol{q}$, and $\boldsymbol{\varepsilon_t}$ is error in period $\boldsymbol{t}$.

In the machine learning approach, Hyndman and Khandakar developed the ARIMA model into Auto ARIMA. It can help researchers by quickly trying more combinations of p, d, and q orders that might be missed when using the ARIMA method manually. The model can be constructed automatically using *auto.arima*() in the *'forecast'* package in R software. By this *auto.arima*() algorithm, the optimal order can be quickly selected. The auto ARIMA modeling algorithm is explained as follows [23].

1) The number of differences is $\boldsymbol{0 \leq d \leq 2}$
2) The selection of order $\boldsymbol{p}$ and $\boldsymbol{q}$ is determined by the smallest AICc value.
    a) The initial models defined by the *auto.arima*() algorithm are ARIMA(0,d,0), ARIMA(2,d,2), ARIMA(1,d,0) and ARIMA(0,d,1). A constant is included in the model for $\boldsymbol{d \leq 1}$. In fact, an additional model is also included in the initial model for $\boldsymbol{d \leq 1}$, namely ARIMA(0,d,0) without a constant.
    b) The best model determined in the initial model is designated as the "current model".
    c) There are variations considered in the current model by changing the $\boldsymbol{p}$ and $\boldsymbol{q}$ order values of the model by ±1 and by adding or removing constants from the current model. The limit of the order $\boldsymbol{p}$ and $\boldsymbol{q}$ in *auto.arima*() is 5.
    d) The final best model is chosen based on the lowest AICc value.

Furthermore, the forecasting value of the ARIMA model can be calculated according to the following three steps.

1) Explain the ARIMA equation in Equation (5) so that $\boldsymbol{Y_t}$ is in the left segment and the others are in the right segment.
2) Rewrite the equation in part a by changing $\boldsymbol{t}$ to $\boldsymbol{T+h}$.
3) In the right segment, replace the future observed value with the forecasted value, replace the future error with zero, and replace the previous period error with the appropriate residual.

2.5 Model Feasibility Test

A model is considered suitable for forecasting if the autocorrelation coefficient of the residuals is random and the residuals are normally distributed [23]. Statistically, the autocorrelation of the residuals can be checked using the Ljung-Box test [24] with '*Box.test*' function in RStudio software. Meanwhile, the normality of the residuals can be checked using the Anderson-Darling test [25] with '*ad.test*' function in RStudio software. Whether in the Ljung-Box test or the AD test, residuals are considered random and normally distributed if the p-value > the significance level $(\boldsymbol{\alpha})$ [26].

2.6 Analysis Methods

In this analysis, we use Microsoft Excel also Rstudio Software with the '*forecast*', and '*stats*' packages for all modeling and forecasting. The procedure of the research is explained in the following steps.

Step 1 : Plot HSI time series data and identify data patterns.

Step 2 : Implement Rolling Window Cross-Validation to the data.

a) Determine the window size $(\boldsymbol{w})$ which is the amount of data set as a model in each rolling window.
b) Determine the number of forward forecasting steps or forecast horizon $(\boldsymbol{h})$

c) Obtaining the number of modeling (subsample) as many as $\boldsymbol{M}$. The number of models based on the selection of window size and the amount of data available is formulated as follows $\boldsymbol{M} = \boldsymbol{n} - \boldsymbol{w}$. The first rolling window is the first model containing time series data from period 1 to $\boldsymbol{w}$, the second model contains time series data from period 2 to $\boldsymbol{w} + \boldsymbol{1}$, and so on.

Step 3 : Select models that pass the model feasibility test from the Holt's and Auto ARIMA methods for each window.

a) Determine the best parameters for each Holt's and Auto ARIMA model. In the Auto ARIMA model, Box-Cox transformation is performed first to check stationarity in variance.

b) Check the feasibility of the model through the normality residual test with AD test and the white noise residual assumption with the Ljung-Box test for each method.

Step 4 : Define the forecasting value for h steps ahead of each model and calculate the residuals.

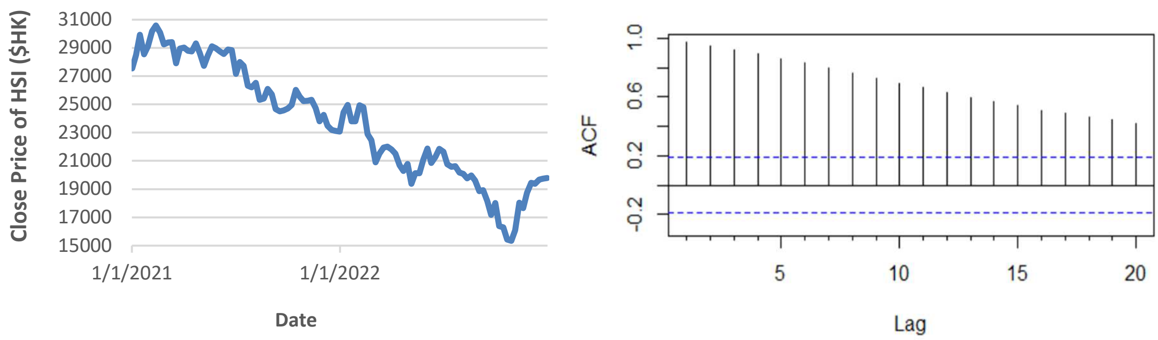
Step 5 : Evaluate the performance of the forecasting methods by calculating the MAPE values of the Holt's and Auto ARIMA methods for each h steps ahead. The formulas for MAPE are shown in **Equation (6)**. MAPE values are interpreted into four categories. MAPE values that are less than 10% are declared as highly accurate predictions, while MAPE values between 10-20% are interpreted as good predictions. MAPE values between 20-50% are categorized as feasible predictions and MAPE values above 50% are stated as inaccurate predictions **[27]**.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \widehat{Y}_t}{Y_t}\right| \times 100\% \tag{6}$$

3. RESULTS AND DISCUSSION

3.1 Identification of Data Patterns

The HSI data is visualized in **Figure 4** **(a)** while the ACF plot is in **Figure 4** **(b).**



(a) **(b)**

Figure 4. Daily HSI data
(a) Time series plot for period 1 January to 30 December 2022. (b) Its ACF

Based on **Figure 4** **(a)**, it can be seen that the closing price of HSI experienced several increases and decreases with a downward trend from the beginning to the end of the period data. Besides the time series plot, the data pattern can also be identified through the ACF plot. **Figure 4** **(b)** shows that the initial lag is significant above the confidence line and then drops slowly to zero. This indicates that there is a high correlation between the initial time interval and decreases at subsequent time intervals or lags, which means that the HSI data contains a trend data pattern.

3.2 Implementation of Rolling Window Cross Validation

The window size used in the application of rolling window cross-validation in this study is 84, which is obtained from 80% of the total data, that is 105 data. The forecast horizon determined is five steps ($h = 5$), so the number of models obtained is 21 models. An overview of the Rolling Window Cross-Validation procedure is shown in **Table 2**.

Table 2. Rolling Window Cross-Validation Procedure on HSI Data

Model	Window	h
1	1st – 84th data	85th – 89th data
2	2nd – 85th data	86th – 90th data
.	.	.
.	.	.
.	.	.
21	21st – 104th data	105th data

3.3 Modeling Using Holt's Method

The *HoltWinters*() estimate the smoothing parameter by minimizing the MSE and determining the initial values of states by heuristic values. In the experimental study, we get smoothing values, $\delta = 0.8824$ and $\beta = 0.1423$ for the first model. Meanwhile, the initial value of l_0 is obtained from the second observation values (at time $t = 2$), that is 28496.8594 and b_0 is the difference between the second and the first data, that is 948.3398.

On the other hand, the *ets*() determine the initial states and estimate smoothing parameters by maximizing the likelihood function. For the first dataset, we obtain smoothing parameter, $\delta = 0.7369$ and $\beta = 0.0001$, with initial values of $l_0 = 29193.5314$ and $b_0 = 104.3377$.

After obtaining the optimal parameters, the next step is to check the feasibility of the model through the AD normality test and residual autocorrelation with Ljung-Box. The residuals are said to be normally distributed and white noise if the p-value $>$ significance level ($\alpha = 0.05$).

In the *HoltWinters*() function, the AD test p-value given is 0.1644 and the Ljung-Box test p-value obtained is 0.9413. This indicates that the Holt's model with $\delta = 0.8824$ and $\beta = 0.1423$ in the *HoltWinters*() function can be used in predicting future HSI values because the p-value given from both tests is greater than the significance level.

The p-value of the AD test given in the *ets*() function is 0.3348 and the p-value of the Ljung-Box test obtained is 0.5706. This shows that the Holt's model with $\delta = 0.7369$ and $\beta = 0.0001$ in the *ets*() function can be used in predicting the future value of HSI because the p-value given from both tests is greater than the significance level.

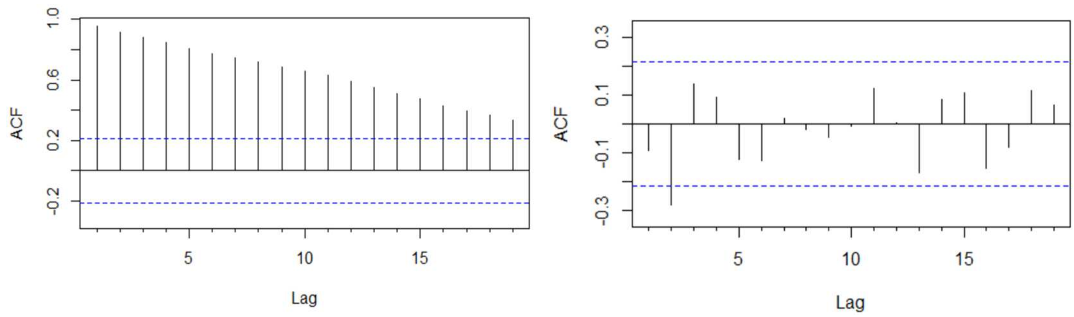
The Holt's method has 21 models in both the *HoltWinters*() and *ets*() functions. In the second to the 21st model, the same steps are taken to determine the initial values, estimate the parameters and check the model feasibility test. The results stated that all models, both from the *HoltWinters*() and *ets*() functions are feasible to be used in forecasting several steps ahead.

3.4 Modeling Using Auto ARIMA Method

The modeling process of the Auto ARIMA method starts with model identification, then parameter estimation and continued with diagnostic tests or model feasibility tests. The initial step taken before model identification is to check stationarity in variance using Box-Cox transformation. For example, in the first model, a value of $\lambda = 0.9500$ is obtained, which means that the data is not yet stationary in variance and needs to be transformed using Box-Cox transformation for each time series data contained in the first model. The value of Box-Cox transformation with $\lambda = 0.9500$ is obtained through **Equation (7)**.

$$X(Y_t) = \frac{sgn(Y_t)|Y_t|^{0{,}95} - 1}{0{,}95} \tag{7}$$

The HSI data that has been stationary in variance then checks for stationary on average using the ACF plot of the data. **Figure 5 (a)** shows that the lag decreases exponentially which means that the first model of HSI data is not stationary in the mean, so differentiation needs to be done. Stationary checks against the mean are performed until the data is stationary in the mean. Therefore, the first model of HSI data that has been differentiated once is checked again through the ACF plot. **Figure 5 (b)** shows that the lag in the ACF plot of the first model data that has been differentiated once does not decrease exponentially so that the data is stationary in the average and the order of differentiation (d) is determined to be 1.



(a) (b)

Figure 5. ACF plot of HSI for the first model
(a) ACF Plot before differentiation, (b) ACF Plot after differentiation ($d = 1$)

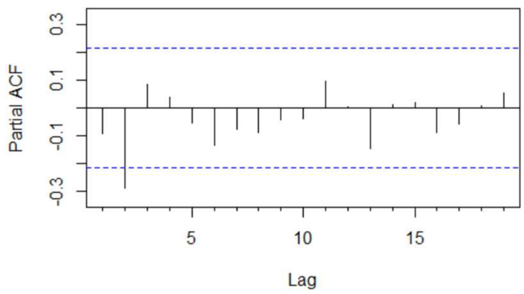


Figure 6. PACF plot of HSI for the first model after differentiation ($d = 1$)

Model identification then uses HSI data that are stationary in mean and variance. The determination of order p and q parameters is identified through ACF and PACF plots. Figure 5 **(b)** and Figure 6 show that lag 2 is significant past the significance limit on the ACF and PACF plots, so the order of p and q is determined to be 2. In addition, modeling will be carried out with a combination of orders p and q by 1, 2 and 3. The next step after identifying the order is to estimate the parameters of each model. The parameter estimation method used is maximum likelihood. This method determines the estimator value for the parameter by maximizing the likelihood function. Based on the determination of the smallest AICc value of possible ARIMA models, the best model is ARIMA (2,1,0) with drift. The parameter estimation values of the model are $\phi_1 = -0.1363$, $\phi_2 = -0.3134$, and drift $= -58.5307$.

The model feasibility test for the best ARIMA model was carried out with the AD and Ljung-Box tests with a p-value of 0.7098 and 0.4386 respectively. The p-value of both tests is greater than the significance level ($\alpha = 0.05$) which means that the model is feasible and can be used to predict future HSI values.

3.5 Multiple Steps Ahead Forecasting

The forecasting value of HSI multiple steps ahead was determined using RStudio software through the 'tsCV' function listed in the 'forecast' package. The output is the forecasting error value up to five steps ahead. The calculation of multiple steps ahead forecasting values can be obtained through **Equation (4)** for the Holt's method and **Equation (5)** for the Auto ARIMA method.

Holt's method with the *HoltWinters*() and *ets*() functions derive the equations for predicting the HSI value one step ahead at time $t = 84$, which are shown in **Equation (8)** and **Equation (9)** respectively.

$$\hat{Y}_{84+h} = \hat{Y}_{85} = l_{84} + (h)b_{84} = 20083{,}6922 + (h)(-124{,}2360) \tag{8}$$
$$\hat{Y}_{84+h} = \hat{Y}_{85} = l_{84} + (h)b_{84} = 20101{,}4354 + (h)(-104{,}3811) \tag{9}$$

In the first model of the ARIMA method, the best model is ARIMA(2,1,0) with drift. The general equation of the ARIMA model in **Equation (5)** then expanded into **Equation (10)** which is used to predict the HSI value multiple steps ahead.

$$\phi_p(B)(1 - B)^d Y_t = c + \theta_q(B)\varepsilon_t$$

Where $c = \mu(1 - \phi_1 - \cdots - \phi_p)$, orde $p = 2$, $d = 1$, and $q = 0$ are subtituted into the model.

$$\phi_2(B)(1-B)^1 Y_t = \mu(1-\phi_1-\phi_2) + \theta_0(B)\varepsilon_t$$
$$(Y_t - BY_t) - (\hat{\phi}_1 BY_t - \hat{\phi}_1 B^2 Y_t) - (\hat{\phi}_2 B^2 Y_t - \hat{\phi}_2 B^3 Y_t) = \mu(1-\hat{\phi}_1-\hat{\phi}_2) + \varepsilon_t$$
$$Y_t - (1+\hat{\phi}_1)Y_{t-1} + (\hat{\phi}_1 - \hat{\phi}_2)Y_{t-2} + \hat{\phi}_2 Y_{t-3} = \mu(1-\hat{\phi}_1-\hat{\phi}_2) + \varepsilon_t$$

Decomposes the ARIMA equation to let Y_t become on the left side.

$$Y_t = (1+\hat{\phi}_1)Y_{t-1} - (\hat{\phi}_1 - \hat{\phi}_2)Y_{t-2} - \hat{\phi}_2 Y_{t-3} + \mu(1-\hat{\phi}_1-\hat{\phi}_2) + \varepsilon_t$$
$$Y_{T+h} = (1+\hat{\phi}_1)Y_{(T+h)-1} - (\hat{\phi}_1 - \hat{\phi}_2)Y_{(T+h)-2} - \hat{\phi}_2 Y_{(T+h)-3} + \mu(1-\hat{\phi}_1-\hat{\phi}_2) + \varepsilon_{T+h}$$
$$\hat{Y}_{T+h} = (1+\hat{\phi}_1)Y_{(T+h)-1} - (\hat{\phi}_1 - \hat{\phi}_2)Y_{(T+h)-2} - \hat{\phi}_2 Y_{(T+h)-3} + \mu(1-\hat{\phi}_1-\hat{\phi}_2)$$

At time $t = 84$, with drift = $-58{,}5307$, $\hat{\phi}_1 = -0{,}1363$ and $\hat{\phi}_2 = -0{,}3134$, **Equation (10)** is obtained.

$$\hat{Y}_{84+h} = 0{,}8637Y_{(84+h)-1} - 0{,}1771Y_{(84+h)-2} + 0{,}3134Y_{(84+h)-3} - 84{,}852 \tag{10}$$

3.6 Model Performance Evaluation

Through selecting a model that is feasible for forecasting and finding the average forecasting error of each model, the MAPE value for each forward forecasting step is obtained as follows in **Table 3** and visualized in **Figure 7**.

Table 3. MAPE Values of *HoltWinters*(), *ets*() and *auto arima*()

Function	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
HoltWinters()	3.3834%	4.7903%	7.2524%	9.7318%	12.2149%
ets()	3.4088%	5.4025%	7.7144%	10.0780%	12.3298%
auto.arima()	2.9196%	4.6553%	6.4012%	8.3083%	10.3781%

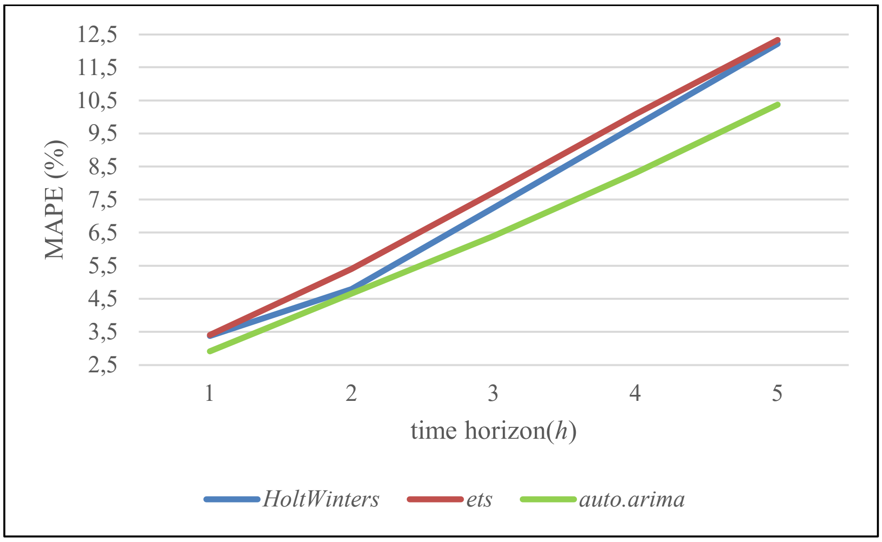


Figure 7. MAPE values of Holt's and Auto ARIMA methods

Based on **Figure 7**, for each forward forecasting step, MAPE produced by Holt's method is higher than those by Auto ARIMA method. The Holt's method with *ets*() function is stated to have the highest MAPE value, while the Auto ARIMA method has the lowest MAPE value compared to both Holt's method functions. However, both the Holt's method with *HoltWinters*() and *ets*() functions and the Auto ARIMA method have MAPE values that continue to increase along with the forecasting period increases.

Unlike the conventional cross-validation method that splits the data only into two parts so that it can produce forecast accuracy information more quickly, rolling window cross-validation has many training testing data pairs so that the modeling and evaluation process of forecast accuracy becomes longer. However, this process is feasible to be able to determine the best forecast model up to several steps ahead in more detail, namely the model with the smallest MAPE value [23].

4. CONCLUSIONS

This study evaluated the multi-step forecast values of HSI using auto ARIMA and Holt's exponential smoothing, which was built from the training data obtained by rolling window cross-validation with the length of window 84. Based on the experiment results to the daily HSI for January 2021 – December 2022, each auto ARIMA and Holt's method produces 21 models that meet the diagnostic test. Each model is then used to calculate the up to five steps ahead forecast values.

The forecasting value of HSI data in the future five steps was evaluated using Mean Absolute Percentage Error (MAPE). Holt's method with the *HoltWinters*() function has MAPE values of 3.3834%, 4.7903%, 7.2524%, 9.7318%, and 12.2149%, while the *ets*() function is 3.4088%, 5.4025%, 7.7144%, 10.0780% and 12.3298%. Meanwhile, the Auto ARIMA method has MAPE values of 2.9196%, 4.6553%, 6.4012%, 8.3083%, and 10.3781% for forecasting one to five-step-ahead, respectively. Based on the HSI forecasting error evaluation for up to five steps ahead, the Auto ARIMA method is considered the most appropriate method in forecasting the HSI.

ACKNOWLEDGMENT

We thank the two anonymous reviewers for their constructive suggestions that enables us to improve the quality of this manuscript.

REFERENCES

[1] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: an analysis and review," *International Journal of Forecast*, vol. 16, pp. 437–450, 2000, doi: 10.1016/S0169-2070(00)00065-0.

[2] W. Sulandari, Y. Yudhanto, S. Subanti, E. Zukhronah, and M. Z. Subarkah, "Implementing Time Series Cross Validation to Evaluate the Forecasting Model Performance," *International Conference on Mathematics and Science Education*, KnE Life Science, pp. 229-238, 2024, doi: 10.18502/kls.v8i1.15584.

[3] "Hang Seng Indexes," 2022. www.hsi.com.hk/eng/indexes/all-indexes/hsi (accessed May 13, 2022).

[4] R. D. P. Aji and N. Abundanti, "The Effect of Asia Regional Stock Price Index on the Indonesia Composite Index (ICI) on the Indonesia Stock Exchange," *European Journal of Business and Management Research*, vol. 7, no. 2, pp. 100–106, Mar. 2022, doi: 10.24018/ejbmr.2022.7.2.1337.

[5] "Hang Seng Indexes Year-End Reports," 2023. www.hsi.com.hk/eng/resources-education/reports (accessed May 03, 2023).

[6] A. R. Alias, N. Y. Zainun, and I. A. Rahman, "Comparison between ARIMA and DES Methods of Forecasting Population for Housing Demand in Johor," *ICTTE*, 2016, doi: 10.1051/matecconf/20168107002.

[7] A. Muchayan, "Comparison of Holt and Brown's Double Exponential Smoothing Methods in The Forecast of Moving Price for Mutual Funds," *Journal of Applied Science, Engineering, Technology, and Education*, vol. 1, no. 2, pp. 183–192, Dec. 2019, doi: 10.35877/454ri.asci1167.

[8] W. K. A. W. Ahmad and S. Ahmad, "Arima model and exponential smoothing method: A comparison," in *AIP Conference Proceedings*, 2013, pp. 1312–1321. doi: 10.1063/1.4801282.

[9] R. J. Hyndman and Y. Khandakar, "Automatic Time Series Forecasting: The forecast Package for R," *Journal of Statistical Software* , vol. 27, no. 3, pp. 1–22, 2008, doi: 10.18637/jss.v027.i03.

[10] S. Tiwari, A. Bharadwaj, and S. Gupta, "Stock Price Prediction Using Data Analytics," *International Conference on Advances in Computing, Communication and Control (ICAC3)*, pp. 1–5, 2017, doi: 10.1109/ICAC3.2017.8318783.

[11] S. Kalyoncu, A. Jamil, J. Rasheed, M. Yesiltepe, and C. Djeddi, "Machine Learning Methods for Stock Market Analysis," in *3rd International Conference on Data Science and Applications (ICONDATA)*, 2020.

[12] Z. Liu and X. Yang, "Cross-validation for uncertain autoregressive model," *Commun Stat Simul Comput*, vol. 51, no. 8, pp. 4715–4726, 2022, doi: 10.1080/03610918.2020.1747077.

[13] M. K. Okasha, "Using Support Vector Machines in Financial Time Series Forecasting," *Int J Stat Appl*, vol. 4, no. 1, pp. 28–39, 2014, doi: 10.5923/j.statistics.20140401.03.

[14] D. Koutsandreas, E. Spiliotis, F. Petropoulos, and V. Assimakopoulos, "On the selection of forecasting accuracy measures," *Journal of the Operational Research Society*, vol. 73, no. 5, pp. 937–954, 2022, doi: 10.1080/01605682.2021.1892464.

[15] H. Hewamalage, K. Ackermann, and C. Bergmeir, "Forecast evaluation for data scientists: common pitfalls and best practices," *Data Min Knowl Discov*, vol. 37, no. 2, pp. 788–832, Mar. 2023, doi: 10.1007/s10618-022-00894-5.

[16] W. Sulandari, Suhartono, Subanar, P. C. Rodrigues, "Exponential Smoothing on Modeling and Forecasting Multiple Seasonal Time Series: An Overview," Fluctuation and Noise Letters, vol. 20, no. 4, - 2130003-1 pp. 2130003-10, 2021, doi: 10.1142/S0219477521300032.

[17] A. Nazim and A. Afthanorhan, "A comparison between single exponential smoothing (SES), double exponential smoothing (DES), holt's (brown) and adaptive response rate exponential smoothing (ARRES) techniques in forecasting Malaysia population," *Global Journal of Mathematical Analysis*, vol. 2, no. 4, p. 276, Sep. 2014, doi: 10.14419/gjma.v2i4.3253.

[18] R.J. Hyndman, A.B. Koehler, R.D. Snyder, and S. Grose., "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods," 2002(18), pp. 439-454. doi: https://doi.org/10.1016/S0169-2070(01)00110-8.

[19] S. A. Shukor, S.F. Sufahani, K.Khalid, M.H.A. Wahab, S.Z.S Idrus, A. Ahmad, and T.S> Subramaniam, “Forecasting Stock Market Price of Gold, Silver, Crude Oil and Platinum by Using Double Exponential Smoothing, Holt’s Linear Trend and Random Walk,” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Jun. 2021. doi: 10.1088/1742-6596/1874/1/012087.

[20] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, “Time series analysis : forecasting and control,” J. Wiley & Sons, 2008. doi: 10.1002/9781118619193.

[21] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, and A. Lachhab, “Forecasting of demand using ARIMA model,” *International Journal of Engineering Business Management*, vol. 10, Oct. 2018, doi: 10.1177/1847979018808673.

[22] W. W. S. Wei, “Writing a Book on Multivariate Time Series Analysis and its Applications View project,” 2006.

[23] R. J. Hyndman and G. Athanasopoulus, “Forecasting: Principles and Practice (2nd ed),” 2018. https://otexts.com/fpp2/ (accessed Mar. 13, 2023).

[24] H. Hassani and M. R. Yeganegi, “Sum of squared ACF and the Ljung–Box statistics,” *Physica A: Statistical Mechanics and its Applications*, vol. 520, pp. 81–86, Apr. 2019, doi: 10.1016/j.physa.2018.12.028.

[25] M. Saculinggan and E. A. Balase, “Empirical Power Comparison of Goodness of Fit Tests for Normality in The Presence of Outliers,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, 2013. doi: 10.1088/1742-6596/435/1/012041.

[26] K. M. U. B. Konarasinghe, “Modeling COVID-19 Epidemic of USA, UK and Russia,” Journal of New Frontiers in Healthcare and Biological Sciences, vol. 1, no. 1, pp. 1–14, 2020, [Online]. Available: www.imathm.edu.lk/publications

[27] V. Yadav and S. Nath, “Forecasting of PM Models and Exponential Smoothing Technique,” *Asian Journal of Water, Environment and Pollution*, vol. 14, no. 4, pp. 109–113, 2017, doi: 10.3233/AJW-170041.