Department: Head
Editor: Name, xxxx@email

# Benchmark of Holt-Winters and SARIMA Methods in Predicting Jakarta Climate

**Amartya Natayu***
Mechanical Engineering, Sampoerna University, Jl. Raya Pasar Minggu No.Kav. 16 Pancoran, Jakarta, 12780, Daerah Khusus Ibukota Jakarta, Indonesia.
Email: amartya.natayu@my.sampoernauniversity.ac.id (corresponding author)

**Quincy Julian H. Clarke**
Information Systems, Sampoerna University, Jl. Raya Pasar Minggu No.Kav. 16 Pancoran, Jakarta, 12780, Daerah Khusus Ibukota Jakarta, Indonesia.
Email: quincy.clarke@my.sampoernauniversity.ac.id

**Muhamad Rausyan Fikri**
Information Systems, Sampoerna University, Jl. Raya Pasar Minggu No.Kav. 16 Pancoran, Jakarta, 12780, Daerah Khusus Ibukota Jakarta, Indonesia.
Email: muhamad.fikri@sampoernauniversity.ac.id

***Abstract—***
**As its capital, Jakarta plays a critical role in boosting Indonesia's economic growth and setting the precedent for broader change outside of the city. One crucial avenue of inquiry to better understand, and prepare for, the future of a country so heavily impacted by disastrous weather events is understanding the effects of climate change through data. This study investigates meteorological data collected from 1996 to 2021 and compares the application of the SARIMA and the Holt-Winters methods to predict the future influence of climatic parameters on Jakarta's weather. The performance of the SARIMA method is proven to provide better results than the Holt-Winter models and both methods showed the best performances when forecasting the humidity data. The results of the forecast are able to demonstrate the characteristic of the climate in Jakarta, with dry season ranging from May to October and wet season ranging from November to April.**

Keywords: Forecasting, SARIMA, Holt-Winters, Climate, Big Data

## THE INTRODUCTION

With the growing abundance and accessibility

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2022 doi:10.20944/preprints202204.0295.v1



of time series data , so too has the demand and use for data forecasting expanded. Forecasting is the process of predicting future trends based on previous historical data, and is utilized in many fields such as business and industry, government, environmental sciences, and the social sciences. Problems addressed using forecasting methods are divided into three classifications: short-term, medium-term, and long-term. Short-term forecasts involve making predictions of trends that are less than one year into the future, medium-term forecasts predict trends within periods that are one to two years into the future, and long-term forecasts predict trends beyond two years into the future [1, 1].

Forecasts are often estimated using time-series data: a sequence of data points chronologically indexed and recorded with an equal definitive time interval. Time-series data can be plotted and decomposed to reveal the components: the trend, the cycle, and the seasonal effects on the data. A trend is an overall increase or decrease within the time span of the observed data, which can be either linear or nonlinear. A The cycle and seasonal components of time-series data are, respectively, the irregular and regular fluctuations which occur within the time span of the observed data. The seasonal component is often measured weekly or yearly. These three components are used in time-series modeling to identify the underlying patterns in the data which assists in extrapolating the forecast. However, the identification of this pattern is complicated by the presence of a random error that is often present in time-series data [2]. Using the appropriate forecasting technique can minimize the random error or estimate to create a more accurate forecast result.

Techniques of forecasting can be generally divided into qualitative and quantitative methods [1, 3-4]. Qualitative methods are a subjective and interpretative approach which require the judgment of experts, and is used when there is little to no historical data. Quantitative methods use a more objective, systematic approach by processing historical data to create a forecasting model. This model captures the patterns based on the aforementioned components of the data and utilizes them to project future data.

There are three widely used quantitative forecasting models: regression, smoothing, and general time-series models. In the Regression models, the relationship between the observed variable with one or more related predictor variables is employed. Smoothing models apply a simple function based on previous observations to forecast the observed variables. Time series models apply the statistical component of the historical data to form a formal model and estimate the unknown parameter of the model. This present study focuses on using the forecasting method of a time series model defined as SARIMA and one of the most well-known ways of forecasting in the industrial and business field, which is exponential smoothing [3]; in particular, the Holt-Winters method. The results of the forecasts are then analyzed to determine which is the most effective in forecasting climate data in Jakarta based on their parametric errors.

Because of its proximity to the Equator, Indonesia has a tropical climate which significantly affects its temperature, humidity, and rainfall rate; all three of which are observed in this study.Tropical climates receive more intense sunlight than countries that are further away from the Equator; thus, such tropical countries will generally have warmer mean air temperatures [4]. Humidity is the concentration of water vapor content in the air, measured using hygrometers. The rainfall rate is the intensity of rainfall over a given time interval. Indonesia's rainfall pattern is mainly influenced by the monsoonal circulation pattern of wet and dry months [5]. Its rainfall rate peaks within the wet months from December to February [6], whereas during dry months—from June to August—there is a low rainfall rate.

This study aims to compare the forecasting results of predicting the climatic variables of Jakarta's weather for five years into the future using the SARIMA and Holt-Winters methods. The climate data that are observed are the mean temperature, humid ity, and rainfall rate. Both methods are done using the Python programming language with MATLAB and Microsoft Excel for data processing. The forecasting models are assessed and selected based on the models' errors, including normalized root mean squared error, mean absolute percentage error, and Nash-Sutcliffe efficiency. The results of the observed errors are also used to determine whichcan be used to more accurately forecast Jakarta's climate. According to similar previous research by Hansun and Kristanda [7], simply using exponential smoothing is not suitable to forecast Jakarta rainfall rate. Therefore, this study also hopes to find another more suitable method to forecast the Jakarta rainfall rate. In addition, if the forecastings are considered relatively accurate, the results may be considered to predict and analyze future Jakarta climate.

## 2 Materials and Methods

### 2.1 Case Study

The study focuses on Jakarta's monthly mean temperature, humidity, and rainfall rate observed between January 1996 and May 2021, . The mean temperature is measured in Celsius, the humidity is measured in percentage, and the rainfall rate is measured in millimeters. Data are measured from three different stations in Jakarta province - *Stasiun Halim Perdana Kusuma Jakarta, Stasiun Meteorologi Kemayoran, and Stasiun Meteorologi Maritim Tanjung Priok*. *Stasiun Halim Perdana Kusuma Jakarta* - referenced as Halim Station - is located at 6.27036 S and 106.88926 E. *Stasiun Meteorologi Kemayoran* - referenced as Kemayoran Station - is located at 6.15559 S and 106.84000 E. Stasiun Meteorologi Maritim Tanjung Priok - referenced as Tanjung Priok Station - is located at 6.10781 S and 106.88053 E. Each climate variable forecast is done individually per station, but the three stations' averaged results are also considered for further analysis.

Geographically, Jakarta is located at 6°12" S and 106°48" E, with an elevation of about 7 meters above sea level. According to the Government's Decree No.171 of 2007, the total area of Jakarta is 7,639.83 km$^2$ with a total landmass of 662.33 km$^2$ and sea area of 6,977.5 km$^2$ [8]. Jakarta is bordered by the Java Ocean and 13 river estuaries that run to the ocean in the north.

Jakarta possesses a tropical climate, and because of its archipelagic geography, sea wind also affects Jakarta's climate. The daily air temperature is relatively stable and spans from 26 - 28 °C. The air temperature difference between wet and dry months is also relatively small, which can be explained by the stable warm air temperature characteristic of tropical countries [9]. The city of Jakarta is located in the homonymous province of Jakarta, which touts a high population; Jakarta Central Statistic Center states that the total population of Jakarta reaches 10,562,088 people [10]. The dense population of Jakarta makes the province vulnerable to pollution. According to research on the impact of the population in Amman, Jordan, a high-density population increases dependence on cars and energy, which also increases environmental pollution by the use of gas and oil [11], which has been shown to be one of the leading causes of global warming and climate change [12][13][14]. Thus, the pollution level in Jakarta might also affect the climate variables observed in this study.

### 2.2 Data Explication and Seasonal Decomposition

Initially, the data are obtained as daily data from the official *Pusat Database Badan Meteorologi, Klimatologi, dan Geofisika* (BMKG) [15]. There exist missing values in the original data from the BMKG website, and based on the types of missing data listed by Kang [16], the missing values are identified to be *missing entirely at random* (MCAR) as the data are unmeasured by the stations. A listwise approach is used to handle the MCAR values. The listwise approach or case deletion omits the single missing variable in a case without removing the entire data tuple and analyzes the remaining data. This approach is the most frequently used to handle missing data, and if the sample size is large enough and the assumption of MCAR is fulfilled, the approach is known to produce unbiased estimates, and conservative results [17]. MATLAB is used to handle the missing data imputation and also averages the daily data to monthly data. The daily data in a month are summed and divided by the number of the daily data To average the daily data to monthly data.

Based on the data sets used in this study, the average, maximum, and minimum values of each monthly climate variable are described in Table 1, Table 2, and Table 3, respectively for mean temperature, humidity, and rainfall rate.

Data sets are also plotted with the help of *matplotlib* library [18] in Python and seasonal decomposed using the seasonal decompose function from the *statsmodel* Python library [19]. Seasonal decomposition is a procedure that decomposes a

Table 1: Description of Jakarta mean temperature from January 1996 - May 2021

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Average (°C) | 27.554 | 28.329 | 28.530 |
| Maximum (°C) | 31.126 | 29.861 | 29.970 |
| Minimum (°C) | 24.500 | 26.272 | 26.590 |

Table 2: Description of Jakarta humidity from January 1996 - May 2021

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Average (%) | 77.970 | 75.689 | 74.471 |
| Maximum (%) | 89.667 | 85.806 | 85.321 |
| Minimum (%) | 61.467 | 65.367 | 48.600 |

series into its seasonal components. The components observed are the trend, seasonal, and residual or error.

The seasonal trends can be explained by the relatively constant seasonal variation. To obtain the values of the observed component, the seasonally adjusted series are seasonal adjusted to remove the seasonal effect from the series to identify previously unseen characteristics; observations without seasonal variation are not adjusted [20]. Both the plotting and seasonal decomposition of the data are shown in Figure 1.

### 2.3 SARIMA Method

The Seasonal Autoregressive and Integrated Moving Average (SARIMA) method of forecasting is a product of the seasonal and non-seasonal polynomials of the components, represented by SARIMA $(p,d,q)x(P,D,Q)s$, [21]. The terms $p,d$, and $q$ are the trend elements, where $p$ is trend auto-regression order, $d$ is trend difference order, and $q$ is the trend moving average order. Meanwhile, $P,D,Q$, and $s$ are the seasonal elements, where $P$ is Seasonal autoregressive order, $D$ is seasonal difference order, $Q$ is seasonal moving average order, and $s$ is the number of time steps for a single seasonal period [22]. SARIMA method follows the following equation defined at Equation (1) [23][24][25].

$$\Phi(B^S)\varphi(B)(1\text{–}B^S)^D(1\text{–}B)^d yt = \Theta(B^S)\theta(B)\varepsilon_t \quad (1)$$

SARIMA requires the series to be univariate and stationary [26], which in this study is tested with the augmented Dickey-Fuller (ADF) parametric test [27] done using the *adfuller* function in the *statsmodels* Python library. The test returns the p-value and the critical values at 1%, 5%, and 10% confidence intervals. The null hypothesis is that the series has a unit root, meaning that the series is non-stationary. The alternate hypothesis is that the series does not contain a unit root; thus, the series is stationary [28][29]. If a series is non-stationary, the series requires to be differentiated in order to obtain stationary.

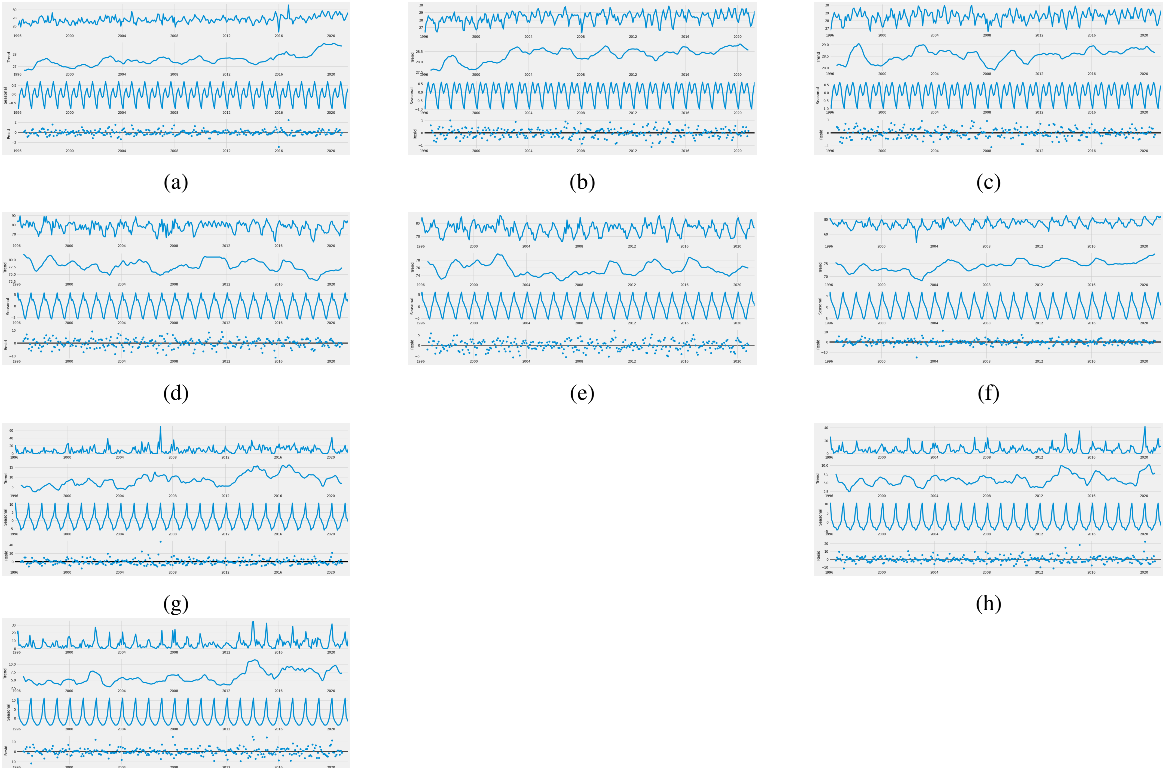
$$AIC = 2k - 2\ln(\hat{L}) \quad (2)$$

The method also uses the observation of the plotted *Autocorrelation Function* (ACF) and *Partial Autocorrelation Function* (PACF) of the data set. The ACF is a function that shows the correlation between the observed data at time $t$ and the observed data at previous times [30]. By observing the spikes that occur during certain lags, it can help to determine the parameter for the $q$ and $Q$ SARIMA parameters. The PACF is a function which summarizes the correlation between the observed data at previous times with the relationship of removing the intervening observed data [31]. This function also uses the spikes at certain lags to determine the parameter for the $p$ and $P$ SARIMA parameters. The plotting of ACF and PACF are both done using Python with functions from the *statsmodels* library. The ACF and PACF observations are used in determining the parameter of the SARIMA.

The SARIMA model is built using the-SARIMA parameters. A grid search for the optimal model parameters is conducted using a minimum value of zero and a maximum of 3 for [describe the parameters' utility here]. The s parameter, is set to 12 for each month of the yearly period. The models are then selected based on the Akaike Information Criterion (AIC) value.

Table 3: Description of Jakarta rainfa ll rate from January 1996 - May 2021

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Average (mm) | 5.988 | 6.056 | 5.988 |
| Maximum (mm) | 34.404 | 41.728 | 34.403 |
| Minimum (mm) | 0.000 | 0.000 | 0.000 |



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 1: This figure shows seasonal decompositions of (a) Mean temperature data Halim Station, (b) Mean temperature data Kemayoran Station, (c) Mean temperature data Tanjung Priok Station, (d) Humidity data Halim station, (e) Humidity data Kemayoran Station, (f) Humidity data Tanjung Priok Station, (g) Rainfall rate data Halim Station, (h) Rainfall rate data Kemayoran Station, and (i) Rainfall rate data Tanjung Priok Station.

The equation for the AIC is as defined in Equation (2) [32] [33], where $k$ is the number of estimated parameters in the model and $\hat{L}$ is the maximum value of the likelihood function for the model. AIC estimates the prediction error and, therefore, the relative quality of a given model for a data set. A better fit model will have lower AIC value [34], thus the model that has the lowest AIC value is selected for the next step of the study. The grid search and the AIC test are performed using the *statsmodels* Python library.

## 2.4 Holt-Winters Method

The Holt-Winters exponential smoothing method — an extension of the Holt smoothing method [35] - is commonly used for forecasting. This method captures seasonality and works on a single recursive system which makes the method simple, straightforward, and computationally efficient [36][37]. Holt's study on the use of exponentially-weighted moving averages (EWMA) on exhibiting variables' seasonal trends finds that a time series possesses additive, multiplicative, or multiplicative error structure components in a trend [36].

The Holt-Winters method's approaches depend on whether the seasonality is modeled additively or multiplicatively [38]. For this study, the Holt-Winters method for forecasting is done using the *SimpleExpSmoothing* and *ExponentialSmoothing* functions from the *statsmodels* Python library, in which both of the approaches are applied. By plotting both in an additive and multiplicative way, the best-fitted model of each data set is selected based on the best parametric error test results, from which the parameters for the model can be estimated.

#### 2.4.1 Additive Holt-Winters Method

The additive approach is defined as the following Equation (3) - 6 [38].

Level:

$$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (3)$$

Trend:

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad (4)$$

Seasonal:

$$s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1 - \gamma)s_{t-m} \quad (5)$$

Forecast:

$$\hat{y}_{t+h|t} = \ell_t + b_t h + s_{t-m+h_m^+} \quad (6)$$

Where $y_t$ is the actual data at time $t$ and $m$ is the period of seasonality. The $\ell_t$, $b_t$, $s_t$, and $\hat{y}_{t+h|t}$ are respectively the level, trend, seasonal factor at time $t$, and the forecast for $h$ periods ahead. As for $h_m^+$ , the variable is equal to $[(h - 1) mod\ m] + 1$. The smoothing parameters of the Holt-Winters model are defined as $\alpha$, $\beta$, and $\gamma$.

#### 2.4.2 Multiplicative Holt-Winters Method

For the Holt-Winters multiplicative, it is defined as the following Equation (7) - 10 [38]. The variable definition is as the same as the additive method (See Sect. **??**).

Level:

$$\ell_t = \alpha\frac{y_t}{s_{t-m}} + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (7)$$

Trend:

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \quad (8)$$

Seasonal:

$$s_t = \gamma(\frac{y_t}{\ell_{t-1} + b_{t-1}}) + (1 - \gamma)s_{t-m} \quad (9)$$

Forecast:

$$\hat{y}_{t+h|t} = (\ell_t + b_t h)s_{t-m+h_m^+} \quad (10)$$

#### 2.4.3 Numerical Instability

There is a possibility of numerical difficulty due to some of the combinations of trend, seasonality, and error. Models with a multiplicative error are only useful when the data are positive, but proven to be numerically unstable when the data contains zero or negative. Thus, when a time series is not strictly positive, only the additive models can be applied [38, 20-23]. This instability is caused by the need for logarithmic transformation of the purely multiplicative model. A logarithmic transformation may allow a model with parameters that combine multiplicatively to fit better [39]. This study, with the help of the *ExponentialSmoothing* function, tests both additive and multiplicative models; therefore, the data sets require adjustment before using the function. The adjustment works by first applying an increment to the data sets and then reducing the forecasted results by this increment to obtain the final results.

### 2.5 Error Parameters

The models' performances are evaluated based on the following error terms: mean squared error (MSE), root mean squared error (RMSE), normalized root mean squared error (NRMSE), mean absolute error (MAE), and the Nash-Sutcliffe efficiency (NSE). The error terms are sequentially defined by Equation (11 to 15), given $x_t$ as the observed and $\hat{x}_t$ as the estimated. Evaluations are done by using the previous one and two years of the training data set (full data set) as the test data set, which can be used to validate the forecast. The calculations of the MSE and MAE are done using functions from *sklearn* [40] Python library and the NSE is calculated using the *hydroeval* [41] Python library.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}[x_t - \hat{x}_t]^2 \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}[x_t - \hat{x}_t]^2} \quad (12)$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}[x_t - \hat{x}_t]^2}}{(\max value - \min value)} \quad (13)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|x_t - \hat{x}_t| \quad (14)$$

Table 4: ADF test results of Jakarta mean temperature from January 1996 - May 2021

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Test statistic | -2.068 | -3.384 | -4.422 |
| p-value | 0.258 | 0.012 | 0.000 |
| Critical value (1%) | -3.453 | -3.453 | -3.453 |
| Critical value (5%) | -2.871 | -2.871 | -2.871 |
| Critical value (10%) | -2.572 | -2.572 | -2.572 |
| Stationary | No | Yes | Yes |

Table 5: ADF test results of Jakarta humidity from January 1996 - May 2021

| | Halim Station | Kema yoran Station | Tanjung Priok Station |
|---|---|---|---|
| Test statistic | -3.322 | -4.386 | -2.756 |
| p-value | 0.013 | 0.000 | 0.065 |
| Critical value (1%) | -3.453 | -3.453 | -3.453 |
| Critical value (5%) | -2.871 | -2.872 | -2.872 |
| Critical value (10%) | -2.572 | -2.572 | -2.572 |
| Stationary | Yes | Yes | Yes |

Table 6: ADF test results of Jakarta rainfall rate from January 1996 - May 2021

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Test statistic | -3.250 | -3.920 | -2.502 |
| p-value | 0.017 | 0.002 | 0.115 |
| Critical value (1%) | -3.453 | -3.453 | -3.453 |
| Critical value (5%) | -2.871 | -2.872 | -2.872 |
| Critical value (10%) | -2.572 | -2.572 | -2.572 |
| Stationary | Yes | Yes | No |

$$NSE = 1 - \frac{\sum_{t=1}^{n} (\hat{x}_t - x_t)^2}{\sum_{t=1}^{n} (x_t - \overline{x_t})^2} \qquad (15)$$

## 3 Results and Discussion

### 3.1 SARIMA

**3.1.1 Augmented Dickey-Fuller Test** The augmented Dickey-Fuller (ADF) test results for stationarity are described in Table 4 to 6. Based on the results, all data sets are stationary except for the mean temperature data from Halim Station and the rainfall rate data from Tanjung Priok Station. Therefore, the $d$ and $D$ SARIMA parameters (See Sect. ) of the forecasting model for mean temperature Halim Station and rainfall rate Tanjung Priok Station cannot be zero due to data sets necessitating differentiation. After performing a first-order differentiation on the non-stationary data, the data is already stationary; thus, the $d$ and $D$ parameter values for the Halim Station mean temperature SARIMA model and the Tanjung Priok Station rainfall rate SARIMA model are 1. The new p-values of both data after first order differentiation are 5.786948e-13 and 4.450709e 30, respectively for the Halim Station mean temperature and Tanjung Priok Station rainfall rate. Both of these p-values are below the significance value, thus rejecting of the ADF Test null hypothesis. Regardless of the ADF test results, there are other possibilities of $d$ and $D$ parameter values for the models, which will be tested using the grid search.

**3.1.2 ACF and PACF Plotting** Figures 2 and 3 display the ACF and PACF plots. By observing the positive and negative significant spikes of the ACF plots, the SARIMA parameters $q$ and $Q$ (See Sect. ) can be estimated. Out of all of the ACF plots, the highest lag value that drops around the confidence interval (blue area) is 3. To estimate the $p$ and $P$ parameters, the PACF plots are observed. Based on the PACF observations, the highest lag value that drops around the confidence interval is also 3.

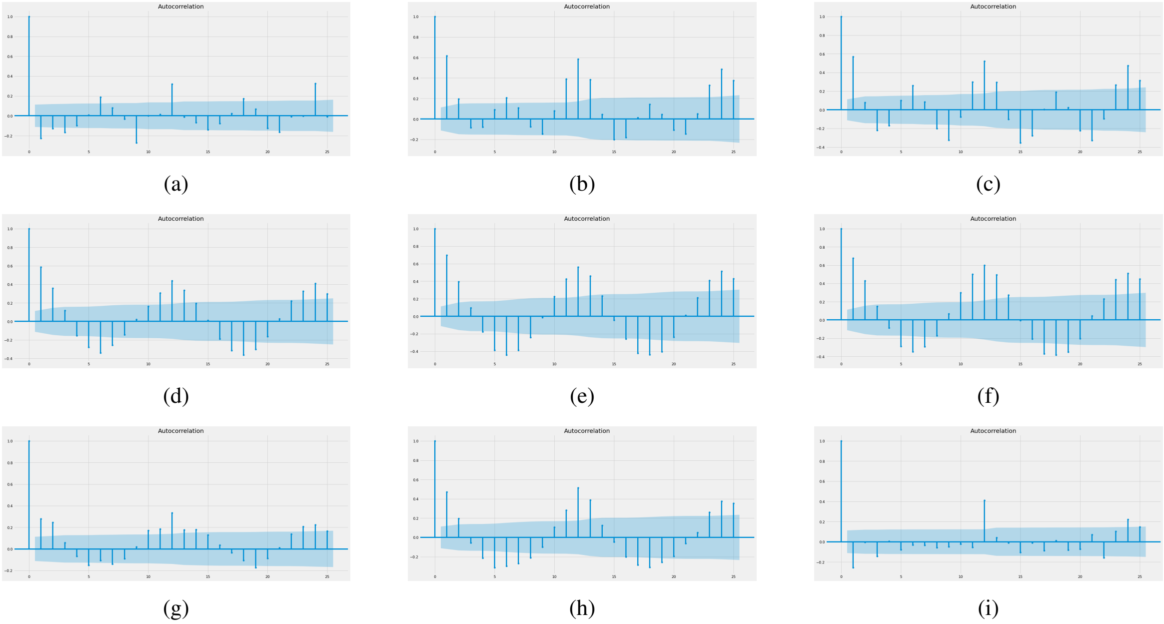
**3.1.3 Model Assessment** Based on the ADF test results, and the observations of the ACF and PACF plots, the grid search for every SARIMA parameter will test values between [0,3]. The models are selected based on the lowest AIC result from the grid search using the non-adjusted, undifferentiated data sets. The models and their AIC values are stated in Table 7. The plotting of the forecasting results are described in Figure 4 and the forecasted values will be examined in Section . The tables show that the non-stationary data sets indeed have a $d$ and $D$ parameter values of 1, following the differentiation order. Although, some data sets that are already stationary are shown to have better results when the $d$ and $D$ parameter values are above zero. By observing the plots, it is shown that the forecasting pattern does not feature many variations after the first period of forecasting.

### 3.2 Holt-Winters

**3.2.1 Zero-value Adjustment** Before undergoing the Holt-Winters forecasting, the rainfall rate data sets need to be adjusted to contain zero values. The adjustment is made by adding a 0.001

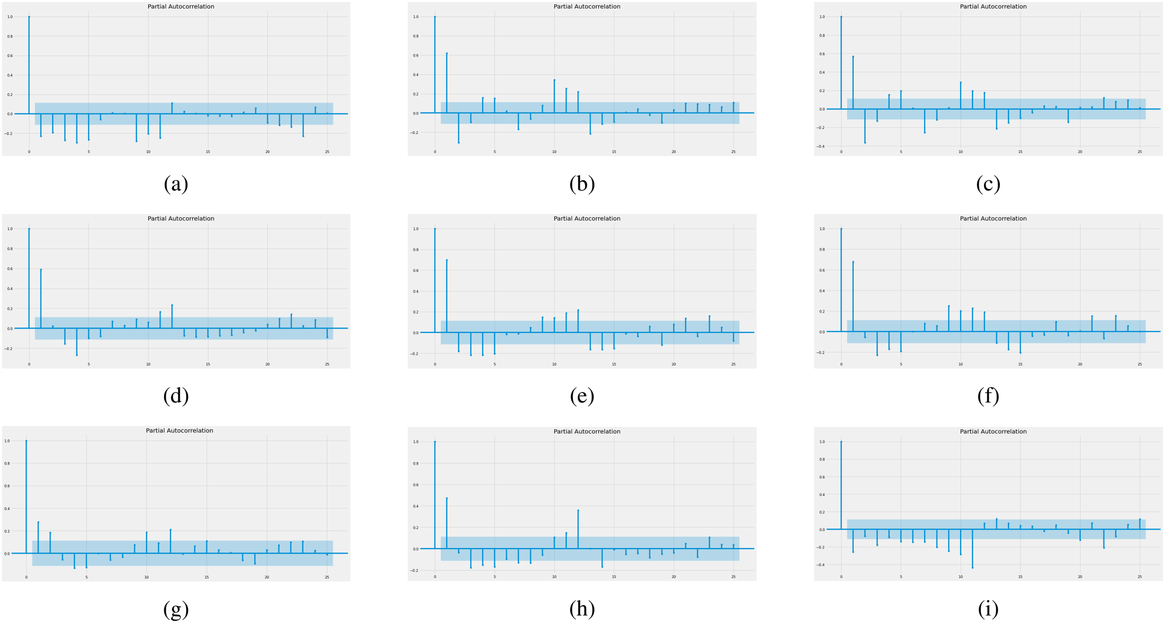
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2022 doi:10.20944/preprints202204.0295.v1

Department Head




(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 2: This figure shows ACF plots of (a) Mean temperature data Halim Station, (b) Mean temperature data Kemayoran Station, (c) Mean temperature data Tanjung Priok Station, (d) Humidity data Halim station, (e) Humidity data Kemayoran Station, (f) Humidity data Tanjung Priok Station, (g) Rainfall rate data Halim Station, (h) Rainfall rate data Kemayoran Station, and (i) Rainfall rate data Tanjung Priok Station.



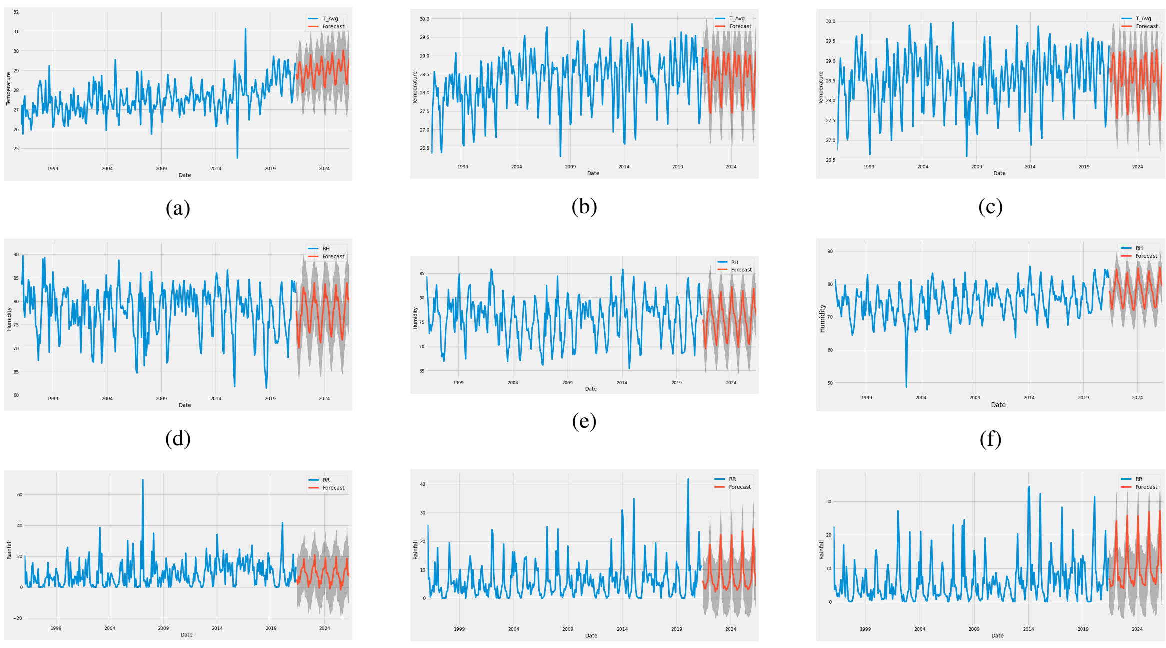

(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 3: This figure shows PACF plots of (a) Mean temperature data Halim Station, (b) Mean temperature data Kemayoran Station, (c) Mean temperature data Tanjung Priok Station, (d) Humidity data Halim station, (e) Humidity data Kemayoran Station, (f) Humidity data Tanjung Priok Station, (g) Rainfall rate data Halim Station, (h) Rainfall rate data Kemayoran Station, and (i) Rainfall rate data Tanjung Priok Station.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2022 doi:10.20944/preprints202204.0295.v1

Table 7: SARIMA models for Jakarta climate forecasting

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Mean temperature data | | | |
| SARIMA (p,d,q) x (P,D,Q)s | SARIMA (0,1,3) x (2,1,3)12 | SARIMA (2,0,1) x (2,1,3)12 | SARIMA (3,0,2) x (2,1,3)12 |
| AIC | 470.702 | 251.887 | 238.041 |
| Humidity data | | | |
| SARIMA (p,d,q) x (P,D,Q)s | SARIMA (0,0,3) x (2,1,3)12 | SARIMA (3,0,3) x (2,1,3)12 | SARIMA (1,1,2) x (3,1,3)12 |
| AIC | 1377.845 | 1172.544 | 1241.914 |
| Rainfall rate data | | | |
| SARIMA (p,d,q) x (P,D,Q)s | SARIMA (2,1,3) x (1,2,3)12 | SARIMA (0,0,3) x (1,2,3)12 | SARIMA (3,1,3) x (1,2,3)12 |
| AIC | 1741.553 | 1458.619 | 1460.059 |



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 4: This figure shows SARIMA forecast plots of (a) Mean temperature data Halim Station, (b) Mean temperature data Kemayoran Station, (c) Mean temperature data Tanjung Priok Station, (d) Humidity data Halim station, (e) Humidity data Kemayoran Station, (f) Humidity data Tanjung Priok Station, (g) Rainfall rate data Halim Station, (h) Rainfall rate data Kemayoran Station, and (i) Rainfall rate data Tanjung Priok Station.

increment to the data because the forecasting results are observed up to 3 decimal places, thus making 0.001 the smallest feasible increment. After the data sets have been adjusted, the data sets can then be used for forecasting. Before noting the forecasting results, the pre-forecast data sets are reduced by the same increment.

**3.2.2 Model Assessment** The Holt-Winters forecasting models are selected based on the additive and multiplicative models' lowest error terms. The seasonal period of the data sets determines the type of Holt-Winters model that is the best suited for the data set. The parameters and the AIC values of each data set's best-fitted model are described in Tables 8 to 10. The plots of the models' forecasts are described in 5 and the forecast estimations are later stated in Section . As shown in the plots, the forecasting pattern does not show many variations after the first period of forecasting. It should be noted that the models are built using adjusted data that have not been readjusted by reducing the increment for the rainfall rate data.

 doi:10.20944/preprints202204.0295.v1



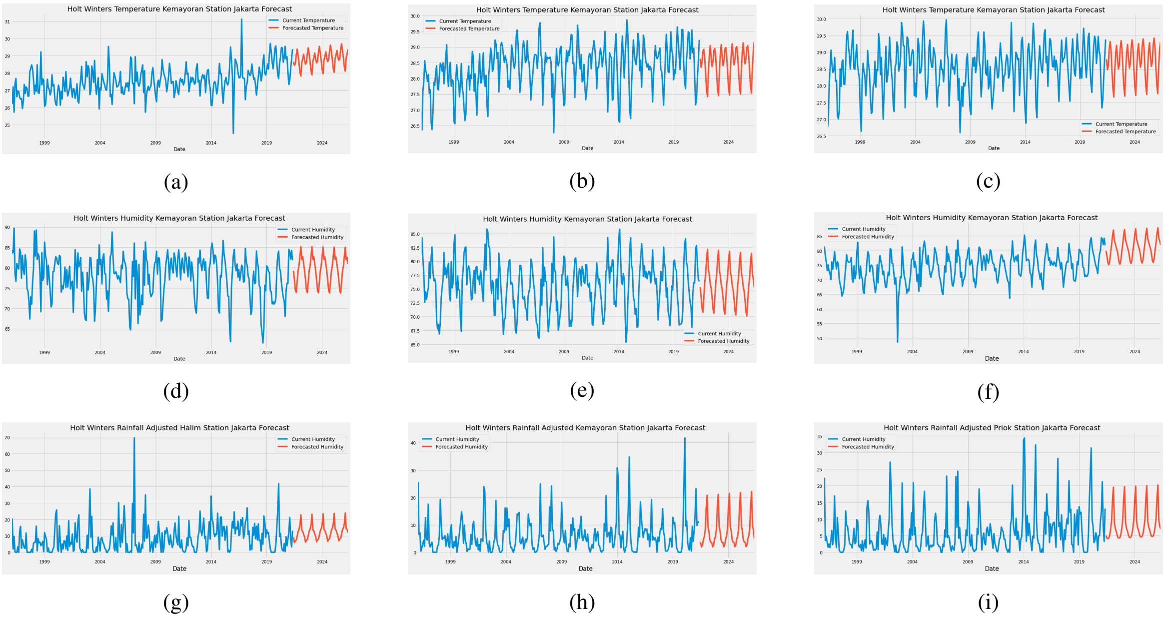



Figure 5: This figure shows Holt-Winters forecast plots of (a) Mean temperature data Halim Station, (b) Mean temperature data Kemayoran Station, (c) Mean temperature data Tanjung Priok Station, (d) Humidity data Halim station, (e) Humidity data Kemayoran Station, (f) Humidity data Tanjung Priok Station, (g) Rainfall rate data Halim Station, (h) Rainfall rate data Kemayoran Station, and (i) Rainfall rate data Tanjung Priok Station.

Table 8: Holt-Winters model assessment of Jakarta mean temperature

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Model | Multiplicative | Multiplicative | Multiplicative |
| $\alpha$ | 0.1829279 | 0.3868695 | 0.4463127 |
| $\beta$ | 7.1839e-15 | 1.8292e-13 | 2.8837e-12 |
| $\gamma$ | 1.0705e-14 | 7.5104e-14 | 2.9607e-12 |
| AIC | -312.732 | -524.473 | -539.434 |

Table 9: Holt-Winters model assessment of Jakarta humidity

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Model | Additive | Multiplicative | Additive |
| $\alpha$ | 0.3222196 | 0.5258200 | 0.3123139 |
| $\beta$ | 9.71e-08 | 2.9623e-14 | 4.3526e-08 |
| $\gamma$ | 0.000000 | 2.9224e-14 | 2.3477e-10 |
| AIC | 823.665 | 588.218 | 627.418 |

### 3.3 Results of Forecasting and Parametric Error Tests

The results of the SARIMA and Holt-Winters forecasts are described in Tables 11 to 13. The averages of the forecasts are also averaged. The error terms are described in Tables 14 to 16.

Table 10: Holt-Winters model assessment of Jakarta rainfall rate

| | Halim Station | Kemayoran Station | Tanjung Priok Station |
|---|---|---|---|
| Model | Additive | Multiplicative | Additive |
| $\alpha$ | 1.4914e-08 | 1.4949e-08 | 4.9918e-06 |
| $\beta$ | 1.02e-09 | 1.4947e-08 | 2.627e-07 |
| $\gamma$ | 1.0295e-06 | 2.2751e-11 | 9.6416e-13 |
| AIC | 1230.569 | 905.537 | 627.418 |

Based on the values of the model errors, the best forecasting method for each data set is noted in Table 17.

For NRMSE and MAE values, the lower the value, the more accurately a model is able to forecast. Conversely, the higher the NSE value, the more accuracte a model is able to forecast. Except for the SARIMA forecasts of humidity data from Tanjung Priok Station, the results of the one-year parametric error tests are better than those for a two-year period (See Table 14 to 16.

Based on Table 17, it is shown that the SARIMA method is preferred in all data sets except for the rainfall rate data from Kemayoran Station, in which the Holt-Winter method

Table 11: Forecasting results of Jakarta mean temperature in Celsius

| Date | Halim Station | | Kemayoran Station | | Tanjung Priok Station | | Total Average |
|---|---|---|---|---|---|---|---|
| | SARIMA | Holt-Winters | SARIMA | Holt-Winters | SARIMA | Holt-Winters | |
| 6/1/2021 | 28.816 | 28.650 | 28.955 | 28.622 | 28.807 | 28.787 | 28.807 |
| 7/1/2021 | 28.532 | 28.442 | 28.542 | 28.347 | 28.470 | 28.476 | 28.470 |
| 8/1/2021 | 28.621 | 28.533 | 28.783 | 28.507 | 28.660 | 28.590 | 28.660 |
| 9/1/2021 | 29.274 | 29.053 | 29.168 | 28.869 | 29.117 | 29.099 | 29.117 |
| 10/1/2021 | 29.434 | 29.392 | 28.969 | 28.924 | 29.195 | 29.321 | 29.195 |
| 11/1/2021 | 29.141 | 29.012 | 28.657 | 28.631 | 28.993 | 29.047 | 28.993 |
| 12/1/2021 | 28.679 | 28.564 | 28.117 | 28.130 | 28.198 | 28.419 | 28.198 |
| 1/1/2022 | 27.866 | 28.047 | 27.532 | 27.618 | 27.858 | 27.911 | 27.858 |
| 2/1/2022 | 27.961 | 27.823 | 27.439 | 27.419 | 27.542 | 27.662 | 27.542 |
| 3/1/2022 | 28.573 | 28.611 | 28.177 | 28.211 | 28.268 | 28.435 | 28.268 |
| 4/1/2022 | 28.898 | 28.824 | 28.730 | 28.724 | 28.936 | 28.968 | 28.936 |
| 5/1/2022 | 29.248 | 29.083 | 29.108 | 29.045 | 29.222 | 29.208 | 29.222 |
| 6/1/2022 | 28.976 | 28.722 | 28.601 | 28.650 | 28.755 | 28.812 | 28.755 |
| 7/1/2022 | 28.581 | 28.514 | 28.332 | 28.376 | 28.349 | 28.501 | 28.349 |
| 8/1/2022 | 28.768 | 28.606 | 28.518 | 28.536 | 28.439 | 28.615 | 28.439 |
| 9/1/2022 | 29.201 | 29.127 | 28.858 | 28.898 | 29.002 | 29.124 | 29.002 |
| 10/1/2022 | 29.789 | 29.466 | 28.955 | 28.953 | 29.247 | 29.347 | 29.247 |
| 11/1/2022 | 29.004 | 29.086 | 28.638 | 28.660 | 28.938 | 29.073 | 28.938 |
| 12/1/2022 | 28.606 | 28.636 | 28.239 | 28.158 | 28.462 | 28.444 | 28.462 |
| 1/1/2023 | 28.135 | 28.118 | 27.751 | 27.645 | 27.829 | 27.936 | 27.829 |
| 2/1/2023 | 28.030 | 27.893 | 27.585 | 27.447 | 27.639 | 27.686 | 27.639 |
| 3/1/2023 | 28.738 | 28.683 | 28.264 | 28.239 | 28.274 | 28.460 | 28.274 |
| 4/1/2023 | 29.087 | 28.897 | 28.806 | 28.753 | 28.780 | 28.993 | 28.780 |
| 5/1/2023 | 29.342 | 29.157 | 29.102 | 29.074 | 29.068 | 29.234 | 29.068 |
| 6/1/2023 | 29.046 | 28.795 | 28.704 | 28.679 | 28.635 | 28.838 | 28.635 |
| 7/1/2023 | 28.747 | 28.586 | 28.472 | 28.404 | 28.433 | 28.526 | 28.433 |
| 8/1/2023 | 28.793 | 28.678 | 28.579 | 28.564 | 28.511 | 28.640 | 28.511 |
| 9/1/2023 | 29.344 | 29.200 | 28.935 | 28.926 | 28.887 | 29.150 | 28.887 |
| 10/1/2023 | 29.718 | 29.540 | 29.059 | 28.982 | 29.255 | 29.373 | 29.255 |
| 11/1/2023 | 29.398 | 29.159 | 28.767 | 28.689 | 29.003 | 29.098 | 29.003 |
| 12/1/2023 | 28.892 | 28.708 | 28.250 | 28.186 | 28.246 | 28.469 | 28.246 |
| 1/1/2024 | 28.187 | 28.189 | 27.705 | 27.673 | 27.850 | 27.960 | 27.850 |
| 2/1/2024 | 28.093 | 27.964 | 27.442 | 27.474 | 27.481 | 27.710 | 27.481 |
| 3/1/2024 | 28.807 | 28.756 | 28.301 | 28.267 | 28.347 | 28.485 | 28.347 |
| 4/1/2024 | 29.121 | 28.970 | 28.770 | 28.781 | 28.875 | 29.019 | 28.875 |
| 5/1/2024 | 29.450 | 29.231 | 29.107 | 29.103 | 29.022 | 29.260 | 29.022 |
| 6/1/2024 | 29.127 | 28.868 | 28.745 | 28.708 | 28.694 | 28.863 | 28.694 |
| 7/1/2024 | 28.796 | 28.658 | 28.432 | 28.432 | 28.283 | 28.551 | 28.283 |
| 8/1/2024 | 28.991 | 28.750 | 28.642 | 28.593 | 28.473 | 28.666 | 28.473 |
| 9/1/2024 | 29.533 | 29.274 | 29.008 | 28.955 | 29.123 | 29.176 | 29.123 |
| 10/1/2024 | 29.899 | 29.615 | 28.979 | 29.011 | 29.198 | 29.399 | 29.198 |
| 11/1/2024 | 29.232 | 29.233 | 28.668 | 28.717 | 28.920 | 29.124 | 28.920 |
| 12/1/2024 | 28.848 | 28.781 | 28.195 | 28.214 | 28.396 | 28.494 | 28.396 |
| 1/1/2025 | 28.241 | 28.260 | 27.658 | 27.700 | 27.842 | 27.985 | 27.842 |
| 2/1/2025 | 28.295 | 28.035 | 27.515 | 27.501 | 27.647 | 27.735 | 27.647 |
| 3/1/2025 | 28.919 | 28.829 | 28.238 | 28.295 | 28.232 | 28.510 | 28.232 |
| 4/1/2025 | 29.270 | 29.043 | 28.777 | 28.810 | 28.839 | 29.044 | 28.839 |
| 5/1/2025 | 29.558 | 29.304 | 29.114 | 29.132 | 29.192 | 29.286 | 29.192 |
| 6/1/2025 | 29.298 | 28.941 | 28.671 | 28.736 | 28.684 | 28.889 | 28.684 |
| 7/1/2025 | 28.939 | 28.731 | 28.414 | 28.461 | 28.456 | 28.577 | 28.456 |
| 8/1/2025 | 29.010 | 28.823 | 28.567 | 28.621 | 28.481 | 28.691 | 28.481 |
| 9/1/2025 | 29.468 | 29.348 | 28.917 | 28.984 | 28.841 | 29.201 | 28.841 |
| 10/1/2025 | 30.024 | 29.690 | 29.012 | 29.040 | 29.280 | 29.424 | 29.280 |
| 11/1/2025 | 29.535 | 29.307 | 28.707 | 28.746 | 29.003 | 29.150 | 29.003 |
| 12/1/2025 | 29.043 | 28.854 | 28.247 | 28.243 | 28.312 | 28.519 | 28.312 |
| 1/1/2026 | 28.471 | 28.332 | 27.728 | 27.728 | 27.841 | 28.010 | 27.841 |
| 2/1/2026 | 28.275 | 28.106 | 27.521 | 27.529 | 27.501 | 27.759 | 27.501 |
| 3/1/2026 | 29.040 | 28.901 | 28.285 | 28.324 | 28.357 | 28.535 | 28.357 |
| 4/1/2026 | 29.360 | 29.117 | 28.794 | 28.839 | 28.831 | 29.070 | 28.831 |
| 5/1/2026 | 29.654 | 29.379 | 29.112 | 29.161 | 28.966 | 29.311 | 28.966 |

is shown to be a better fit. Variables that are more accurately predicted using SARIMA are humidity, mean temperature, and rainfall rate respectively. The SARIMA forecasting model for humidity data has the best parametric error values , with the data from the Kemayoran Station being the best overall. The Holt-Winters forecasting model for the rainfall rate data has the most worst parametric error values, with the data from the Halim Station being the least optimal overall.

### 3.4 Climate Analysis

The average forecasting results of the climatic variables show a slight trend, although the values are generally stable as there is no massive change in the maximum and minimum point, or a shift in the seasonality. The climate variables are observed using the total average value. The

Department Head

Table 12: Forecasting results of Jakarta humidity in percentage

| Date | Halim Station | | Kemayoran Station | | Tanjung Priok Station | | Total Average |
|---|---|---|---|---|---|---|---|
| | SARIMA | Holt-Winters | SARIMA | Holt-Winters | SARIMA | Holt-Winters | |
| 6/1/2021 | 77.917 | 79.221 | 75.463 | 75.307 | 77.873 | 79.841 | 77.873 |
| 7/1/2021 | 75.289 | 77.062 | 73.956 | 73.630 | 76.629 | 78.116 | 76.629 |
| 8/1/2021 | 71.443 | 74.462 | 71.190 | 71.529 | 73.780 | 75.861 | 73.780 |
| 9/1/2021 | 70.009 | 73.874 | 69.449 | 70.777 | 72.251 | 75.106 | 72.251 |
| 10/1/2021 | 74.025 | 76.573 | 73.027 | 73.490 | 75.016 | 76.414 | 75.016 |
| 11/1/2021 | 76.944 | 79.271 | 76.525 | 76.181 | 77.624 | 79.423 | 77.624 |
| 12/1/2021 | 76.891 | 80.494 | 77.215 | 77.691 | 79.540 | 81.927 | 79.540 |
| 1/1/2022 | 81.182 | 82.951 | 79.687 | 80.531 | 81.808 | 85.081 | 81.808 |
| 2/1/2022 | 82.961 | 85.131 | 81.611 | 82.162 | 84.404 | 87.021 | 84.404 |
| 3/1/2022 | 81.508 | 82.053 | 78.825 | 78.491 | 81.526 | 83.429 | 81.526 |
| 4/1/2022 | 81.758 | 82.682 | 76.659 | 77.519 | 80.503 | 82.350 | 80.503 |
| 5/1/2022 | 79.485 | 80.953 | 76.057 | 75.941 | 79.697 | 81.040 | 79.697 |
| 6/1/2022 | 78.557 | 79.188 | 74.829 | 75.137 | 77.994 | 80.042 | 77.994 |
| 7/1/2022 | 75.861 | 77.028 | 73.367 | 73.464 | 75.845 | 78.316 | 75.845 |
| 8/1/2022 | 73.568 | 74.429 | 71.353 | 71.368 | 73.452 | 76.061 | 73.452 |
| 9/1/2022 | 73.211 | 73.841 | 70.140 | 70.617 | 72.415 | 75.307 | 72.415 |
| 10/1/2022 | 75.093 | 76.539 | 72.496 | 73.324 | 73.009 | 76.614 | 73.009 |
| 11/1/2022 | 78.064 | 79.238 | 75.351 | 76.010 | 76.117 | 79.623 | 76.117 |
| 12/1/2022 | 79.619 | 80.460 | 77.181 | 77.516 | 78.355 | 82.128 | 78.355 |
| 1/1/2023 | 81.157 | 82.918 | 79.884 | 80.350 | 82.162 | 85.281 | 82.162 |
| 2/1/2023 | 83.897 | 85.097 | 81.244 | 81.977 | 83.564 | 87.222 | 83.564 |
| 3/1/2023 | 79.636 | 82.020 | 77.413 | 78.314 | 79.652 | 83.630 | 79.652 |
| 4/1/2023 | 80.283 | 82.649 | 77.620 | 77.344 | 78.879 | 82.551 | 78.879 |
| 5/1/2023 | 79.254 | 80.920 | 75.270 | 75.770 | 76.828 | 81.241 | 76.828 |
| 6/1/2023 | 76.901 | 79.154 | 74.363 | 74.968 | 76.303 | 80.243 | 76.303 |
| 7/1/2023 | 74.857 | 76.995 | 72.419 | 73.299 | 74.629 | 78.517 | 74.629 |
| 8/1/2023 | 72.062 | 74.396 | 70.612 | 71.207 | 72.370 | 76.262 | 72.370 |
| 9/1/2023 | 71.126 | 73.808 | 70.479 | 70.458 | 71.962 | 75.507 | 71.962 |
| 10/1/2023 | 73.791 | 76.506 | 72.595 | 73.159 | 73.144 | 76.815 | 73.144 |
| 11/1/2023 | 76.855 | 79.205 | 75.512 | 75.839 | 76.438 | 79.824 | 76.438 |
| 12/1/2023 | 78.503 | 80.427 | 77.419 | 77.341 | 79.553 | 82.329 | 79.553 |
| 1/1/2024 | 81.443 | 82.885 | 80.100 | 80.169 | 81.823 | 85.482 | 81.823 |
| 2/1/2024 | 83.676 | 85.064 | 82.200 | 81.792 | 84.791 | 87.423 | 84.791 |
| 3/1/2024 | 80.721 | 81.986 | 78.223 | 78.138 | 81.181 | 83.831 | 81.181 |
| 4/1/2024 | 80.997 | 82.616 | 76.800 | 77.170 | 79.944 | 82.751 | 79.944 |
| 5/1/2024 | 79.548 | 80.887 | 76.287 | 75.599 | 79.494 | 81.442 | 79.494 |
| 6/1/2024 | 78.074 | 79.121 | 75.281 | 74.799 | 77.480 | 80.443 | 77.480 |
| 7/1/2024 | 75.601 | 76.962 | 73.615 | 73.134 | 76.188 | 78.718 | 76.188 |
| 8/1/2024 | 72.980 | 74.362 | 71.145 | 71.046 | 73.797 | 76.463 | 73.797 |
| 9/1/2024 | 72.368 | 73.774 | 69.677 | 70.300 | 72.387 | 75.708 | 72.387 |
| 10/1/2024 | 75.123 | 76.473 | 73.149 | 72.994 | 74.333 | 77.016 | 74.333 |
| 11/1/2024 | 78.016 | 79.172 | 75.882 | 75.668 | 77.168 | 80.025 | 77.168 |
| 12/1/2024 | 78.562 | 80.394 | 76.624 | 77.167 | 78.854 | 82.529 | 78.854 |
| 1/1/2025 | 81.024 | 82.852 | 79.613 | 79.988 | 82.523 | 85.683 | 82.523 |
| 2/1/2025 | 83.451 | 85.031 | 81.389 | 81.608 | 83.972 | 87.623 | 83.972 |
| 3/1/2025 | 80.198 | 81.953 | 78.056 | 77.962 | 80.643 | 84.031 | 80.643 |
| 4/1/2025 | 80.798 | 82.583 | 77.062 | 76.996 | 79.981 | 82.952 | 79.981 |
| 5/1/2025 | 79.241 | 80.853 | 74.998 | 75.429 | 77.956 | 81.642 | 77.956 |
| 6/1/2025 | 77.288 | 79.088 | 74.448 | 74.630 | 77.441 | 80.644 | 77.441 |
| 7/1/2025 | 75.076 | 76.928 | 72.737 | 72.969 | 75.267 | 78.918 | 75.267 |
| 8/1/2025 | 72.491 | 74.329 | 70.664 | 70.886 | 72.999 | 76.663 | 72.999 |
| 9/1/2025 | 71.734 | 73.741 | 70.317 | 70.141 | 72.701 | 75.909 | 72.701 |
| 10/1/2025 | 73.885 | 76.439 | 72.271 | 72.830 | 73.101 | 77.216 | 73.101 |
| 11/1/2025 | 76.982 | 79.138 | 75.571 | 75.497 | 76.515 | 80.225 | 76.515 |
| 12/1/2025 | 79.140 | 80.361 | 77.762 | 76.993 | 79.791 | 82.730 | 79.791 |
| 1/1/2026 | 81.487 | 82.818 | 79.891 | 79.808 | 82.361 | 85.883 | 82.361 |
| 2/1/2026 | 83.924 | 84.997 | 81.820 | 81.424 | 85.077 | 87.824 | 85.077 |
| 3/1/2026 | 80.342 | 81.920 | 78.034 | 77.786 | 81.121 | 84.232 | 81.121 |
| 4/1/2026 | 80.672 | 82.549 | 77.456 | 76.823 | 79.896 | 83.153 | 79.896 |
| 5/1/2026 | 79.530 | 80.820 | 76.129 | 75.259 | 79.198 | 81.843 | 79.198 |

observations prove the characteristics of a tropical country, in which the climate is stable throughout the year. In forecasting over a one-year period, the diurnal pattern of the dry and wet seasons can also be distinguished. By observing the rainfall rate forecast using the median value of 7.724 mm as the criterion that differentiates the dry and wet season, the dry season spans from May toOctober and the wet season spans from November to April. The rainfall rate peaks during the wet season, especially in February, while the rainfall rate drops during the dry season, most notably in July. There are anomalies in the rainfall rate forecast from Halim Station, in which several values are negative. The humidity forecast shows an exact pattern as the rainfall rate forecast rises and drops during the same month. This result also states the direct relationship between rainfall rate

Table 13: Forecasting results of Jakarta rainfall rate in mm

| Date | Halim Station | | Kemayoran Station | | Tanjung Priok Station | | Total Average |
|---|---|---|---|---|---|---|---|
| | SARIMA | Holt-Winters | SARIMA | Holt-Winters | SARIMA | Holt-Winters | |
| 6/1/2021 | 4.932 | 9.147 | 6.119 | 3.557 | 6.929 | 5.048 | 5.955 |
| 7/1/2021 | 1.740 | 5.913 | 4.834 | 3.446 | 5.795 | 4.184 | 4.319 |
| 8/1/2021 | 6.636 | 7.074 | 3.331 | 1.927 | 4.418 | 4.123 | 4.585 |
| 9/1/2021 | 3.077 | 7.418 | 3.217 | 2.592 | 4.792 | 4.280 | 4.229 |
| 10/1/2021 | 7.269 | 9.975 | 5.005 | 4.680 | 4.858 | 5.225 | 6.169 |
| 11/1/2021 | 11.296 | 12.581 | 5.383 | 5.666 | 7.995 | 6.994 | 8.319 |
| 12/1/2021 | 12.428 | 13.838 | 6.978 | 8.221 | 10.279 | 9.474 | 10.203 |
| 1/1/2022 | 12.552 | 16.277 | 13.845 | 16.047 | 15.811 | 15.747 | 15.047 |
| 2/1/2022 | 18.387 | 22.874 | 18.735 | 20.810 | 24.046 | 19.570 | 20.737 |
| 3/1/2022 | 9.725 | 13.359 | 8.647 | 9.141 | 11.858 | 9.032 | 10.294 |
| 4/1/2022 | 9.458 | 12.764 | 7.539 | 7.520 | 6.264 | 7.139 | 8.447 |
| 5/1/2022 | 8.410 | 10.896 | 4.918 | 4.476 | 7.510 | 6.353 | 7.094 |
| 6/1/2022 | 6.315 | 9.407 | 4.873 | 3.613 | 6.635 | 5.212 | 6.009 |
| 7/1/2022 | -0.121 | 6.173 | 4.183 | 3.500 | 4.166 | 4.348 | 3.708 |
| 8/1/2022 | 3.813 | 7.334 | 2.180 | 1.958 | 4.722 | 4.287 | 4.049 |
| 9/1/2022 | 2.186 | 7.678 | 3.450 | 2.633 | 4.121 | 4.444 | 4.085 |
| 10/1/2022 | 5.606 | 10.235 | 3.113 | 4.755 | 3.956 | 5.389 | 5.509 |
| 11/1/2022 | 10.774 | 12.841 | 5.286 | 5.756 | 9.438 | 7.158 | 8.542 |
| 12/1/2022 | 13.472 | 14.098 | 7.641 | 8.352 | 10.314 | 9.638 | 10.586 |
| 1/1/2023 | 13.621 | 16.537 | 15.841 | 16.302 | 16.289 | 15.911 | 15.750 |
| 2/1/2023 | 20.839 | 23.134 | 22.261 | 21.140 | 25.724 | 19.734 | 22.139 |
| 3/1/2023 | 9.250 | 13.619 | 8.532 | 9.286 | 11.702 | 9.196 | 10.264 |
| 4/1/2023 | 10.490 | 13.024 | 7.462 | 7.640 | 8.208 | 7.303 | 9.021 |
| 5/1/2023 | 7.021 | 11.157 | 3.385 | 4.547 | 7.437 | 6.517 | 6.677 |
| 6/1/2023 | 6.190 | 9.667 | 4.212 | 3.671 | 6.276 | 5.376 | 5.899 |
| 7/1/2023 | -0.584 | 6.433 | 3.796 | 3.556 | 5.457 | 4.512 | 3.862 |
| 8/1/2023 | 4.152 | 7.594 | 2.770 | 1.989 | 6.269 | 4.451 | 4.538 |
| 9/1/2023 | 2.112 | 7.938 | 3.333 | 2.675 | 4.529 | 4.608 | 4.199 |
| 10/1/2023 | 6.499 | 10.495 | 4.050 | 4.830 | 5.349 | 5.553 | 6.129 |
| 11/1/2023 | 10.735 | 13.101 | 5.281 | 5.848 | 9.548 | 7.322 | 8.639 |
| 12/1/2023 | 12.284 | 14.358 | 7.157 | 8.484 | 10.072 | 9.802 | 10.360 |
| 1/1/2024 | 12.438 | 16.797 | 15.345 | 16.560 | 17.074 | 16.075 | 15.715 |
| 2/1/2024 | 19.069 | 23.394 | 22.123 | 21.476 | 25.606 | 19.898 | 21.928 |
| 3/1/2024 | 8.907 | 13.879 | 8.747 | 9.433 | 12.263 | 9.360 | 10.432 |
| 4/1/2024 | 9.383 | 13.284 | 7.882 | 7.761 | 7.736 | 7.467 | 8.919 |
| 5/1/2024 | 7.162 | 11.417 | 4.126 | 4.619 | 7.956 | 6.681 | 6.993 |
| 6/1/2024 | 5.731 | 9.927 | 4.485 | 3.729 | 7.168 | 5.540 | 6.097 |
| 7/1/2024 | -1.198 | 6.693 | 3.960 | 3.612 | 5.365 | 4.676 | 3.851 |
| 8/1/2024 | 3.440 | 7.854 | 2.745 | 2.021 | 5.242 | 4.615 | 4.319 |
| 9/1/2024 | 1.310 | 8.198 | 3.357 | 2.718 | 4.877 | 4.772 | 4.205 |
| 10/1/2024 | 5.446 | 10.755 | 3.851 | 4.907 | 4.809 | 5.717 | 5.914 |
| 11/1/2024 | 10.375 | 13.361 | 5.239 | 5.940 | 9.967 | 7.486 | 8.728 |
| 12/1/2024 | 12.525 | 14.618 | 7.134 | 8.619 | 11.162 | 9.966 | 10.671 |
| 1/1/2025 | 12.483 | 17.057 | 15.851 | 16.823 | 17.103 | 16.239 | 15.926 |
| 2/1/2025 | 19.432 | 23.654 | 23.456 | 21.817 | 26.897 | 20.062 | 22.553 |
| 3/1/2025 | 8.367 | 14.139 | 8.829 | 9.583 | 12.953 | 9.524 | 10.566 |
| 4/1/2025 | 9.452 | 13.544 | 8.102 | 7.884 | 8.792 | 7.631 | 9.234 |
| 5/1/2025 | 6.436 | 11.677 | 3.954 | 4.693 | 8.207 | 6.845 | 6.968 |
| 6/1/2025 | 5.367 | 10.187 | 4.383 | 3.788 | 7.243 | 5.704 | 6.112 |
| 7/1/2025 | -1.825 | 6.953 | 3.903 | 3.670 | 5.985 | 4.840 | 3.921 |
| 8/1/2025 | 3.249 | 8.114 | 2.968 | 2.053 | 7.075 | 4.779 | 4.706 |
| 9/1/2025 | 0.862 | 8.458 | 3.325 | 2.761 | 5.220 | 4.936 | 4.260 |
| 10/1/2025 | 5.403 | 11.015 | 4.108 | 4.985 | 5.780 | 5.881 | 6.195 |
| 11/1/2025 | 10.117 | 13.621 | 5.212 | 6.035 | 10.567 | 7.650 | 8.867 |
| 12/1/2025 | 11.907 | 14.878 | 6.925 | 8.756 | 10.764 | 10.130 | 10.560 |
| 1/1/2026 | 11.789 | 17.317 | 15.954 | 17.091 | 17.851 | 16.403 | 16.067 |
| 2/1/2026 | 18.553 | 23.914 | 24.199 | 22.163 | 27.179 | 20.226 | 22.706 |
| 3/1/2026 | 7.868 | 14.399 | 8.965 | 9.735 | 13.238 | 9.688 | 10.649 |
| 4/1/2026 | 8.803 | 13.804 | 8.402 | 8.010 | 8.649 | 7.795 | 9.244 |
| 5/1/2026 | 6.124 | 11.937 | 4.148 | 4.767 | 8.783 | 7.009 | 7.128 |

and humidity. For the mean temperature forecast, the temperature peaks during the beginning and the end of the dry season, which are May and October, and drops during the peak of the wet season, which is February.

## END SECTIONS

## CONCLUSION

SARIMA and the Holt-Winter method are both commonly used for their effectiveness in forecasting tasks. For the SARIMA method, the grid search parameters are limited using observations from the ADF test for stationarity, ACF plots, and PACF plots. By observing the data prior to using a grid search, hardware limitation can

Table 14: Parametric error test results of Jakarta mean temperature forecasting

| | SARIMA | | | Holt-Winters | | |
|---|---|---|---|---|---|---|
| | Halim Station | Kemayoran Station | Tanjung Priok Station | Halim Station | Kemayoran Station | Tanjung Priok Station |
| MSE (1 year) | 0.223 | 0.223 | 0.089 | 0.253 | 0.207 | 0.201 |
| MSE (2 year) | 0.183 | 0.183 | 0.074 | 0.167 | 0.179 | 0.211 |
| RMSE (1 year) | 0.472 | 0.323 | 0.298 | 0.503 | 0.454 | 0.448 |
| RMSE (2 year) | 0.428 | 0.314 | 0.272 | 0.409 | 0.423 | 0.460 |
| NRMSE (1 year) | 0.071 | 0.090 | 0.088 | 0.076 | 0.127 | 0.133 |
| NRMSE (2 year) | 0.065 | 0.087 | 0.081 | 0.062 | 0.118 | 0.136 |
| MAE (1 year) | 0.351 | 0.291 | 0.255 | 0.426 | 0.387 | 0.379 |
| MAE (2 year) | 0.315 | 0.278 | 0.226 | 0.332 | 0.340 | 0.367 |
| NSE (1 year) | 0.516 | 0.808 | 0.808 | 0.450 | 0.621 | 0.567 |
| NSE (2 year) | 0.551 | 0.787 | 0.812 | 0.591 | 0.612 | 0.465 |

Table 15: Parametric error test results of Jakarta humidity forecasting

| | SARIMA | | | Holt-Winters | | |
|---|---|---|---|---|---|---|
| | Halim Station | Kemayoran Station | Tanjung Priok Station | Halim Station | Kemayoran Station | Tanjung Priok Station |
| MSE (1 year) | 4.894 | 2.613 | 4.433 | 8.265 | 2.235 | 4.836 |
| MSE (2 year) | 4.916 | 2.980 | 3.556 | 13.832 | 6.155 | 11.328 |
| RMSE (1 year) | 2.212 | 1.616 | 2.105 | 2.875 | 1.495 | 2.199 |
| RMSE (2 year) | 2.217 | 1.726 | 1.886 | 3.719 | 2.481 | 3.366 |
| NRMSE (1 year) | 0.079 | 0.079 | 0.057 | 0.102 | 0.073 | 0.060 |
| NRMSE (2 year) | 0.079 | 0.085 | 0.051 | 0.132 | 0.121 | 0.092 |
| MAE (1 year) | 1.943 | 1.403 | 1.780 | 2.360 | 1.099 | 1.726 |
| MAE (2 year) | 1.934 | 1.451 | 2.974 | 2.943 | 2.173 | 2.911 |
| NSE (1 year) | 0.830 | 0.841 | 0.744 | 0.713 | 0.864 | 0.721 |
| NSE (2 year) | 0.808 | 0.880 | 0.839 | 0.459 | 0.751 | 0.486 |

Table 16: Parametric error test results of Jakarta rainfall rate forecasting

| | SARIMA | | | Holt-Winters | | |
|---|---|---|---|---|---|---|
| | Halim Station | Kemayoran Station | Tanjung Priok Station | Halim Station | Kemayoran Station | Tanjung Priok Station |
| MSE (1 year) | 26.637 | 14.749 | 10.847 | 38.822 | 11.661 | 9.051 |
| MSE (2 year) | 57.432 | 37.475 | 17.229 | 64.277 | 43.754 | 20.715 |
| RMSE (1 year) | 5.161 | 3.841 | 3.293 | 6.231 | 3.415 | 3.008 |
| RMSE (2 year) | 7.578 | 6.122 | 4.151 | 8.071 | 6.615 | 4.551 |
| NRMSE (1 year) | 0.074 | 0.092 | 0.096 | 0.090 | 0.092 | 0.087 |
| NRMSE (2 year) | 0.109 | 0.147 | 0.121 | 0.115 | 0.147 | 0.132 |
| MAE (1 year) | 4.154 | 3.315 | 2.518 | 5.452 | 2.934 | 2.338 |
| MAE (2 year) | 5.792 | 4.237 | 3.318 | 6.930 | 4.279 | 3.585 |
| NSE (1 year) | 0.230 | 0.617 | 0.630 | -0.123 | 0.697 | 0.691 |
| NSE (2 year) | 0.397 | 0.607 | 0.733 | 0.325 | 0.541 | 0.679 |

be minimized as running a full grid search can be computationally inefficient. The Holt-Winter method is relatively straightforward, especially when the data requires no adjustment. For the purposes of this study, SARIMA models are proven to provide better results than the Holt-Winter models. The methods have also shown that the most optimal variable to be forecasted using SARIMA and Holt-Winter is the humidity, while the least optimal is the rainfall rate. These statements are obtained from observing the parametric error tests. Despite the results of the tests, all models from both methods show some somewhat iterative patterns after the first period of forecasting. Therefore, using SARIMA and Holt-Winter to forecast beyond one year is not recommended. This result might affect the method's inability to capture the component precisely, and there is a possibility of unpredictable future events, such as Covid-19.

The results of the forecast can be utilized to observe Jakarta's climate. The results demonstrate the characteristics of dry and wet seasons in Jakarta from May 2020 to May 2021. As shown by the study, the dry season will occur from May to October, while the wet season will

Table 17: Forecasting method assessment for each data sets

| | SARIMA AVG NRMSE | SARIMA AVG MAE | SARIMA AVG NSE | Holt-Winters AVG NRMSE | Holt-Winters AVG NSE | Holt-Winters AVG MAE | Better-fitted method |
|---|---|---|---|---|---|---|---|
| Mean temperature data | | | | | | | |
| Halim Station | 0.068 | 0.333 | 0.534 | 0.069 | 0.379 | 0.521 | SARIMA |
| Kemayoran Station | 0.089 | 0.285 | 0.798 | 0.123 | 0.364 | 0.617 | SARIMA |
| Tanjung Priok Station | 0.085 | 0.241 | 0.810 | 0.135 | 0.373 | 0.516 | SARIMA |
| Humidity data | | | | | | | |
| Halim Station | 0.079 | 1.934 | 0.819 | 0.117 | 2.652 | 0.586 | SARIMA |
| Kemayoran Station | 0.082 | 1.427 | 0.861 | 0.097 | 1.636 | 0.808 | SARIMA |
| Tanjung Priok Station | 0.054 | 2.377 | 0.792 | 0.076 | 2.319 | 0.604 | SARIMA |
| Rainfall rate data | | | | | | | |
| Halim Station | 0.092 | 4.973 | 0.314 | 0.103 | 6.191 | 0.101 | SARIMA |
| Kemayoran Station | 0.120 | 3.776 | 0.612 | 0.120 | 3.607 | 0.619 | Holt-Winter |
| Tanjung Priok Station | 0.109 | 2.918 | 0.682 | 0.110 | 2.962 | 0.689 | SARIMA |

occur from November to April. The results also generally show stable values for all three climatic variables, in line with what can be expected from the climate of a tropical country. Future environmental studies can use this study's results as a foundation to perform further climate analysis on the meteorological conditions of Jakarta, and other such regions heavily impacted by climate change.

## ACKNOWLEDGMENT

We would like to thank CRCS for partially funding this project with the internal research grant 2020 - 2021.

## REFERENCES

1. Montgomery, D.C., Jennings, C.L., Kulahci, M., Introduction to Time Series Analysis and Forecasting, John Wiley & Sons. Inc., New Jersey, 2008.
2. Gerbing, D., Time Series Components, School of Business Administration Portland State of University, http://web.pdx.edu/~gerbing/515/Resources/ts.pdf, (2016).
3. Gardner Jr., E.S., Exponential smoothing: The state of the art—part II, International Journal of Forecasting, 22 (2006) 637-666.
4. Stevens, A., Introduction to the Basic Drivers of Climate, Nature Education Knowledge, 3(10) (2010) 10.
5. Analysis and Projection of Temperature and Rainfall, Indonesia Climate Change Sectoral Roadmap, www.bappenas.go.id/files/2013/5229/9917/analysis-and-projection-of-temperature-and-rainfall__20110217130224__0.pdf, (2010).
6. Loo, Y. Y., Billa, L., Singh, A., Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia, Geoscience Frontiers, 6(6) (2015) 817-823, https://doi.org/10.1016/j.gsf.2014.02.009.
7. Hansun, S., Kristanda, M.B., Central Jakarta Rainfall Intensity Forecast using Single Exponential Smoothing, International Journal of Recent Technology and Engineering, 8(4) (2019).
8. Keputusan Gubernur No. 171 Tahun 2007 tentang Penataan, Penetapan Batas Dan Luas Wilayah Kelurahan Di Provinsi Daerah Khusus Ibukota Jakarta, Jaringan Dokumentasi & Informasi Hukum Pemprov Jakarta, https://jdih.jakarta.go.id/himpunan/produkhukum_detail/885, n.d..
9. Fajri, K., Trilistyo, H., Hermanto, E., Rusunami di Jakarta Timur (Undergraduate Thesis), Universitas Diponegoro, Semarang, Indonesia. http://eprints.undip.ac.id/45023/, (2015).
10. Jumlah Penduduk Provinsi Jakarta Menurut Kelompok Umur dan Jenis Kelamin 2018-2020, Badan Pusat Statistik Provinsi Jakarta, https://jakarta.bps.go.id/indicator/12/111/1/jumlah-penduduk-provinsi-dki-jakarta-menurut-kelompok-umur-dan-jenis-kelamin.html, n.d..
11. Dawodieh, E., The Impact of High Population Density on the Built Environment and the Behavior of Individuals in Amman, Global Journal of Researches in Engineering: J, 17(1) (2017) 13-21.
12. Larr, A.S., Neidell, M., Pollution and Climate Change, The Future of Children, 26(1) (2016) 93-113, https://www.jstor.org/stable/43755232.
13. D'Amato, G., Bergmann, K.C., Cecchi, L. et al, Climate change and air pollution, Allergo Journal International, 23 (2014) 17–23, https://doi.org/10.1007/s40629-014-0003-7.
14. Sario, M.D., Katsouyanni, K., Michelozzi, P., Climate change, extreme weather events, air pollution and respiratory health in Europe, European Respiratory Journal, 42 (2013) 826-843.
15. Data Online: Pusat Database, Badan Meteorologi, Klimatologi, and Geofisika, https://dataonline.bmkg.go.id/, n.d..

16. Kang, H., The prevention and handling of the missing data, Korean Journal of Anesthesiology, 64(5) (2013) 402-406, https://doi.org/10.4097/kjae.2013.64.5.402.
17. Donner, A., The Relative Effectiveness of Procedures Commonly Used in Multiple Regression Analysis for Dealing with Missing Values, The American Statistician, 36(4) (1982) 378-381, https://www.jstor.org/stable/2683092.
18. Hunter, J.D., Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9(3) (2007) 90-95, https://doi.org/10.1109/MCSE.2007.55.
19. Seabold, S., Perktold, J., statsmodels: Econometric and statistical modeling with python, Proceedings of the 9th Python in Science Conference, 2010.
20. Seasonal Decomposition, International Business Machines, https://www.ibm.com/docs/de/spss-statistics/24.0.0?topic=option-seasonal-decomposition, (n.d.).
21. Tadesse, K.B., Dinka, M.O., Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa, Journal of Water and Land Development, 35 (2017) 229–236, https://doi.org/10.1515/jwld-2017-0088.
22. Brownlee, J., A Gentle Introduction to SARIMA for Time Series Forecasting in Python, Machine Learning Mastery, https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/, 2018.
23. Box, G.E.P., Jenkins, G.M., Reinsel, G.C., Time series analysis, forecasting and control (4th ed.), John Wiley & Sons. Inc., New Jersey, 2008.
24. Chang, X., Gao, M., Wang, Y., Hou, X., Seasonal autoregressive integrated moving average model for precipitation of time series, Journal of Mathematics and Statistics, 8(4) (2012) 500-505.
25. Cryer, J.D., Chan, K.S., Time series analysis: With applications in R, Springer, New York, 2008.
26. Rosyid, H.A., Aniendya, M.W., Herwanto, H.W., Comparison of Indonesian Imports Forecasting by Limited Period using SARIMA Method. Knowledge Engineering and Data Science, 2(2) (2019) 90-100.
27. Dickey, D.A., Fuller, W.A., Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74(366) (1979) 427-431.
28. statsmodels.tsa.stattools.adfuller, statsmodels, https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html, n.d..
29. Bronwlee, J., How to Check if Time Series Data is Stationary with Python, Machine Learning Mastery, https://machinelearningmastery.com/time-series-data-stationary-python/, 2016.
30. Teusch, A., Introduction to the Spectral and Time Series Analysis with Examples from Geodesy, Verlag der Bayerischen Akademie der Wissenschaften, München, 2006.
31. Tinungki, G.M., The analysis of partial autocorrelation function in predicting maximum wind speed, IOP Conf. Series: Earth and Environmental Science, 235 (2019), https://iopscience.iop.org/article/10.1088/1755-1315/235/1/012097.
32. Akaike, H., A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19(6) (1974) 716-723, https://doi.org/10.1109/TAC.1974.1100705.
33. Burnham, K.P., Anderson, D.R., Model Selection and Multimodel Inference: A practical information-theoretic approach (2nd ed.), Springer, New York, 2002. https://doi.org/10.1007/b97636.
34. Arnold, T.W., Uninformative Parameters and Model Selection Using Akaike's Information Criterion, The Journal of Wildlife Management, 74(6) (2010) 1175-1178 , https://www.jstor.org/stable/40801110.
35. Cipra, T., Trujillo, J., Rubio, A., Holt-Winters Method with Missing Observations, Management Science, 41(1) (1995) 174-178, https://www.jstor.org/stable/2632910.
36. Makatjane, K.D., Moroke, N.D., Comparative study of holt-winters triples exponential smoothing and seasonal Arima: Forecasting short term seasonal car sales in South Africa. *Risk governance & control: financial markets & institutions*, 6(1) (2016) 71-82, https://doi.org/10.22495/rgcv6i1art8.
37. Chatfield, C., Yar, M., Holt-Winters forecasting: some practical issues, The Statician, 129-140 (1988).
38. Hyndman, R.J., Koehler, A.B., Ord, J.K., Snyder, R.D., Forecasting with Exponential Smoothing: The State Space Approach, Springer, Heidelberg, 2008. https://doi.org/10.1007/978-3-540-71918-2.
39. Cialdella, L., When do we log transform the response variable? Model assumptions, multiplicative combinations and log-linear models, Casual Inference, https://lmc2179.github.io/posts/multiplicative.html, 2020.
40. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12 (2011) 2825 - 2830.
41. Hallouin, T., HydroEval: Streamflow Simulations Evaluator (Version 0.0.2) (v0.0.2-1), Zenodo, https://doi.org/10.5281/zenodo.2591218, 2019.

**Amartya Natayu** is a student at Sampoerna University, Indonesia and the University of Arizona Global Campus, US. She is currently pursuing B.Sc degree in Mechanical Engineering from Sampoerna University and University of Arizona. She is a member of Cyber-physical systems research group from Sampoerna University. She is available for contact at amartyanatayu@gmail.com.

**Quincy Clarke** is a researcher in cyber-physical systems research group SU. His research interest is semantic analysis and machine learning in social science.

**Muhamad Rausyan Fikri** is a lecturer at Sampoerna University, Indonesia, and a Global lecturer at the University of Arizona, US. He is a coordinator at Cyber-physical systems research group Sampoerna University. He received a B.Sc in Computer Science and Electronics from Universitas Gadjah Mada (UGM) and an M.Eng degree in Systems and Control Engineering from Tokyo Institute of Technology (Tokyo Tech), Japan. He is also a member of IEEE, AIAA, and IFAC Affiliated.