JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI e-ISSN: 2614-8811
Published by Departement of Mathematics, Hasanuddin University, Indonesia
https://journal.unhas.ac.id/index.php/jmsk/index

Vol. 19, No. 3, May 2023, pp. 484- 497
DOI: 10.20956/j.v19i3.24807

Forecasting Time Series Data Using Haar Discrete Wavelet Transformation

Hartina Husain[1], Amran[2]

[1] *Institut Agama Islam Negeri Parepare,* [2]*Universitas Hasanuddin*
Email address: [1]*hartinahusain@iainpare.ac.id,* [2]*amranihsani@gmail.com*

Received: 28 December 2022; Accepted: 3 March 2023; Published: 5 May 2023

Abstract

Discrete Wavelet Transform is a data transformation method that represents data in the time domain and frequency domain. This transformation appears to overcome the weakness of the Fourier transform which is only able to provide one domain information and is limited to certain windowing . The type of wavelet used is the Haar Wavelet. Identification of data periodicity using Periodogram analysis with Fisher's Test statistics. The transformed data is decomposed into two components, namely the Approximation Coefficient and the Detail Coefficient. Both components are predicted using the Box-Jenkins ARIMA method. Model selection was carried out using the Akaike Information Criterion (AIC) and Mean Square Error (MSE) methods. This study uses secondary data, namely Air Humidity from September 2006 to Desember 2011 obtained from the Meteorology, Climatology and Geophysics Agency of Makassar City. The result shows that forecasting on the Approximation Coefficient obtained by the ARIMA model (0,0,3) with AIC = 112.2142 and MSE = 29.673. While forecasting on Detailed Coefficients is obtained by the ARIMA model (2,1,0) with AIC = 89.2 and MSE = 15,989.

Keywords: Discrete Wavelet Transform, Fourier Transform, Haar Wavelet, Periodogram, Fisher's Test.

1. INTRODUCTION

Forecasting is a way to predict future conditions by taking into account past data and current data. Forecasting is a study of historical data to find relationships, trends and systematic data patterns to predict future values [1]. Future circumstances that are full of uncertainty often make someone try to predict something so that later they can prepare and make the right decision.

The forecasting process is done using time series data. In its implementation, time series data is a series of observational data that occurs based on the time index sequentially with fixed time intervals [2]. Time series analysis in general aims to study or create a stochastic model mechanism that is capable of analyzing observation series and predicting future time series values based on the history of the series itself [3]. The most frequently used forecasting method is the Autoregressive Integrated Moving Average (ARIMA) developed by Box and Jenkins.

Analysis based on the domain is divided into two, namely time series analysis in the time domain and time series analysis in the frequency domain. The time domain uses autocorrelation and partial autocorrelation functions to study changes in time series data using parametric models. Meanwhile for the frequency domain, time series analysis is considered as a result of the presence of cycle components at different frequencies which are difficult to obtain in the time domain. The data can then be represented in the time and frequency domain using transformations.

Transformation is the process of changing data or signal into another form to make it easier to analyze, like the Fourier transform which turns the signal inward some sine or cosine wave with that frequency different, while the wavelet transform transforms the signal into various basic wavelet forms with various shifts and scaling [4].

Wavelets can be used as a transformation tool that automatically dissects data into different components and studies each component at a resolution appropriate to its scale [5]. The data will be described based on a predetermined scale so that the movement of the amplitude of the data can be seen over time. The wavelet transform that is considered more suitable for time series data is the Discrete Wavelet Transform (DWT) because at each decomposition level there are wavelet coefficients and scales as much as the length of the data [6] . DWT can reduce filtering weaknesses that can be performed at any sample size. DWT has advantages including being easy to implement and efficient in terms of computation time[7] [8].

The wavelet transform that is considered more suitable for time series data is the Discrete Wavelet Transform because at each decomposition level there are wavelet coefficients and scales as much as the length of the data. DWT can reduce filtering weaknesses that can be performed on several sample sizes. Data analysis with DWT was performed at different frequencies with different resolutions by decomposing the data into detail components and approximation components. One type of wavelet is the Haar wavelet which is easier to use in calculations, especially in graph smoothing[7].

The development of wavelets has progressed rapidly, namely the discovery of various types of existing wavelets. One of them is the Haar wavelet where it is easier to use in calculations, especially in graph smoothing .This transformation is generally used to solve broader issues such as rainfall, climate, or other national scale data.

Previous research has discussed forecasting using time series data in the time domain or frequency domain [9]. While the time series data in this study were first transformed into the time and frequency domain, then Box-Jenkins forecasting methods were used. This is to find out how to get wavelet coefficients and their application to Makassar City air humidity data which tends to change every month.

2. PRELIMINARIES

2.1 Periodogram

The periodogram is used to determine the hidden periodicity of time series data that is difficult to find in the time domain by looking at the periodogram plot. The period obtained from the periodogram plot will then be tested for the periodic component of the periodogram. This can be done using Fisher's test with test statistics, namely [10]:

$$T = \frac{I^{(1)}\left(\omega_{(1)}\right)}{\sum_{k=1}^{\left[\frac{n}{2}\right]} I_{(\omega_k)}} \tag{2.1}$$

Where $I^{(1)}\left(\omega_{(1)}\right) = \max\{(\omega_k)\}$ spectrum value for the Fourier frequency which has the maximum value of the periodogram ordinate.

With the following hypothesis:

H_0 : The periodogram does not have a periodic component

H_1 : The periodogran has a periodic component.

Decision Criteria is if the T value calculated from the time series is larger than the table value g_α, then H_0 is rejected which means that the Y_t time series contains a periodic component.

2.2 Discrete Wavelet Transform

Wavelet is a data model with the form of time and frequency. Wavelet transform is a function of the real variable t which is used to localize a function in space and scale $L^2(R)$, denoted $\psi(t)$ as a mother wavelet [11] . Doughter wavelet $\psi_{a,b}(t)$ produced by the parameters dilation a and translation/contraction b , which are expressed in the equation [5][12] :

$$\psi_{a,b}(t) = a^{-1/2}\psi\left(\frac{t-b}{a}\right); a > 0, b \in R \tag{2.2}$$

Where :
a = Parameter of dilation or contraction
b = Parameter of translation
R = Conditional values of a and b in integer values

The transformation can be divided into two forms namely low pass filter $\mathcal{H} = \{h_k\}$ dan high pass filter $\mathcal{G} = \{g_k\}$, where h_k and g_k is the coefficient of scaling function $\phi(x)$ dan wavelet function $\psi(x)$ which is shown in equation [13] :

$$\phi(x) = \sum_{k \in \mathbb{Z}} h_k \sqrt{2}\phi(2x-k) \tag{2.3}$$

$$\psi(x) = \sum_{k \in \mathbb{Z}} g_k \sqrt{2}\psi(2x-k) \tag{2.4}$$

One type of wavelet that was first recognized and very simple to use, especially in quantitative calculations, is the Haar wavelet. The Haar wavelet function is given [14]:

$$\psi(t) = \begin{cases} 1 & , 0 \le t < 0.5 \\ -1 & , 0.5 \le t < 1 \\ 0 & , \text{untuk yang lain} \end{cases} \tag{2.5}$$

As well as the Haar scale function is defined :

$$\phi(t) = \begin{cases} 1 & , 0 \le t \le 1 \\ 0 & , \text{untuk yang lain} \end{cases} \tag{2.6}$$

2.3 Forecasting

One of the time series models is ARIMA. The general form of the ARIMA equation (p,d,q) is as follows [15][9]:

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1-B)^d Y_t = (1 - \theta_1 B - \cdots - \theta_q B^q)a_t \tag{2.7}$$

Where

Y_t = Observation data
a_t = Residual value (error) at time t
d = Differencing order
ϕ_P= The coefficient of the p -parameter autoregressive (AR) model
θ_q = The coefficient of the parameter of the qth moving average (MA) model
The forecasting process consists of five stages, namely:

1. Data Identification Stage

As for Model Identification the steps taken to identify the model are as follows[15]:

a. Plot the time series data and if the time series data is not stationary in terms of variance, a Box-Cox transformation is performed.
b. Calculating and examining the Autocorrelation function (ACF) and partial autocorrelation function (PACF) samples from the original series and if the time series data is indicated to be not stationary in the mean, a differencing process is carried out.
c. Calculating and examining sample ACF and PACF from time series data resulting from Box-Cox transformation and differencing to identify the order of the ARIMA model (p, d, q)
d. Test the deterministic trend if $d > 0$

2. Parameter Estimation Stage

Hypothesis:

$H_0 : \theta = 0$ (estimator not significant)

$H_1 : \theta \neq 0$ (significant estimator)

Test Statistics

$$t = \frac{\hat{\theta}}{SE(\hat{\theta})} \quad (2.8)$$

Decision Criteria is reject H_0 if $|t| > t_{\frac{\alpha}{2};(n-n_p)}$; n_p = the number of parameter

3. Model Diagnostic Stage

Model Diagnostics include model adequacy (test whether the rest is white noise) and normal distribution assumption test.

4. Best Model Selection Stage

The criteria for selecting the best model are carried out through the Akaike's Information Criterion (AIC) and Mean Square Error (MSE) equations:

$$AIC = T \ln \hat{\sigma}_a^2 + 2M \quad (2.9)$$

Where

M = The number of parameters in the model, the

N = Number of observations, the

$\hat{\sigma}_a^2$ = Maximum Likelihood estimation of σ_a^2

and to calculate the Mean Square Error (MSE) value used equation:

$$MSE = \frac{\sum_{h=1}^{H} e_h^2}{(H - M)} \quad (2.10)$$

Where

M = The number of parameters estimated to be

e = The rest of the out sample

5. Forecasting Stage

The forecasting results obtained are then compared with the validation data.

2.4 Humidity

Air humidity is the amount of water vapor present in the air [16]. There are 2 types of air humidity as follows:

a. Absolute humidity is a number indicating the amount of water vapor in grams in one cubic meter of air.
b. Relative humidity (relative humidity), which is a number in percent that shows the ratio between the amount of water vapor actually contained in the air at a certain temperature. Relative humidity is calculated using the following formula:

$$K = T/P \text{ x } 100\% \quad (2.11)$$

Where
K = relative humidity.
T = water vapor contained in the air at a certain temperature.
P = maximum water vapor content capacity.

3. MAIN RESULTS

3.1 Data Stationarity Test

Stationarity of data is one of the assumptions that must be met before forecasting time series data. The stationarity of the data can be seen by looking at the data plots, correlograms, and unit root tests. Data plots, correlograms, and unit root test results are as follows

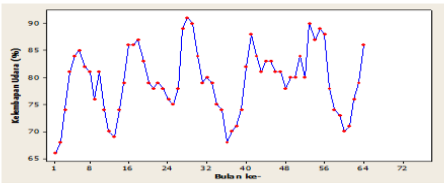


Figure 3.1 Air Humidity Data Plot

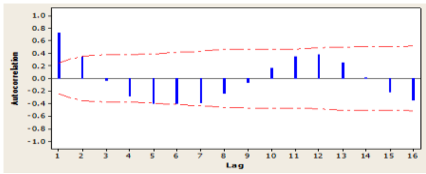


Figure 3.2 Air Humidity Data ACF Diagram

Based on figure 3.1 shows that the data plots are stationary because data changes from time to time do not increase or decrease significantly (does not have a slope or trend). Figure 3.2 shows the autocorrelation value fluctuates or does not decrease linearly close to zero. The results of the unit root test get the output as in $p - \text{value} = 0.01 < 0{,}05$ then H_0 being rejected means there is no unit root or the data is stationary.

3.2 Periodogram Analysis

Periodic data will form a seasonal pattern, meaning that the SARIMA forecasting process will be carried out and conversely, non-periodic data or non-seasonal patterns are used by ARIMA in its forecasting.

$$T = \frac{0.078125}{2376.734} = 0.0000328 \text{ and } g_{0.05} = 0.19784$$

Because of the value $T = 0.0000328 < 0.19784$ then the decision accept H_0 or reject H_1. Thus it can be concluded that the air humidity data does not contain a periodic component or non-seasonal data, so ARIMA is used in forecasting.

3.3 Discrete Wavelet Transform

This transformation uses a scaling filter (h_k) and a wavelet function (g_k) on the type wavelet Haar. This transformation decomposes the data into two coefficients, namely the approximation coefficient and the detail coefficient.

Table 3.1 Approximation Coefficient Value

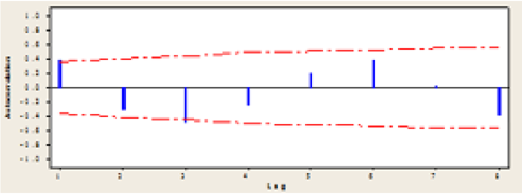
No	Approximate Coefficient	No	Approximate Coefficient	No.	Approximate Coefficient
1	94.7523	12	108.8944	23	115.9655
2	109.6016	13	108.1873	24	112.4300
3	119.5011	14	127.2792	25	113.1371
4	115.2584	15	123.0366	26	115.9655
5	111.0158	16	112.4300	27	125.1579
6	101.8234	17	108.8944	28	125.1579
7	101.1163	18	100.4092	29	107.4802
8	116.6726	19	99.7021	30	101.1163
9	122.3295	20	110.3087	31	103.9447
10	114.5513	21	121.6224	32	116.6726
11	111.0157	22	115.9655		

Table 3.2 Detailed Coefficient Values

No	Detail Coefficient	No	Detail Coefficient	No.	Detail Coefficient
1	1.4142	12	-1.4142	23	-1.4142
2	4.9497	13	2.1213	24	-2.1213
3	0.7071	14	1.4142	25	0
4	-0.7071	15	-4.2426	26	-2.8284
5	3.5355	16	0.7071	27	-2.1213
6	-2.8284	17	-2.8284	28	-0.7071
7	3.5355	18	-4.2426	29	-2.8284
8	4.9497	19	0.7071	30	-2.1213
9	0.7071	20	5.65685	31	3.5355
10	-2.8284	21	-2.8284	32	4.9497
11	0.7071	22	1.4142		

3.4 ACF and PACF Diagrams

Before making the ACF and PACF digrams, the approximate coefficient of approximation data is first tested for stationarity with DF. The value is obtained $\mathrm{p-value} = 0.02136 < \alpha(0.05)$, which means that there is no unit root or the data is stationary in both the mean and variance.



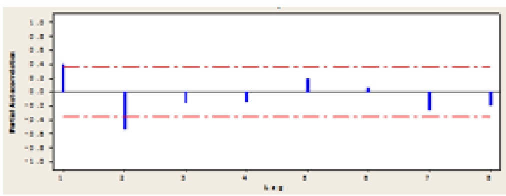
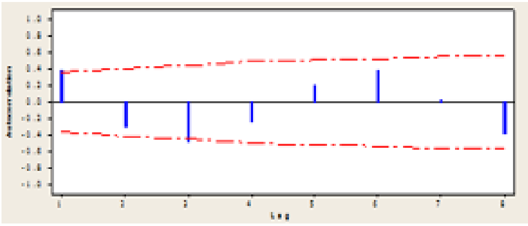


Figure 3.3 ACF Coefficient Approximation Diagram, **Figure 3.4** PACF diagram of approximation coefficients

An ARIMA model (p,d,q) each consists of an order of Autoregressive (p) determined based on the truncated lag on the PACF diagram, the order of Integrated (d) shows how much differencing is done on the data, and the order of the Moving Average (q) determined based on the truncated lag on the ACF diagram. Figure 3.3 shows a truncated ACF diagram at lag 1 and lag 3, and decreases exponentially. While in Figure 3.4 the PACF diagram is truncated at lag 1 and lag 2 and also decreases exponentially. Therefore the possible ARIMA models are ARIMA (0,0,3), ARIMA(0,0,1), ARIMA(1,0,0), ARIMA(2,0,0), ARIMA (1,0,3), ARIMA (1,0,1), and ARIMA (2,0,1).

Tests carried out on the Coefficient of Detail data based on the Dickey Fuller test obtained a value $p - \text{value} = 0.4822$ greater than 0.05. It means reject H_0, that is, there is a unit root or the data is not stationary. The non-stationary of the data is overcome by differencing once.



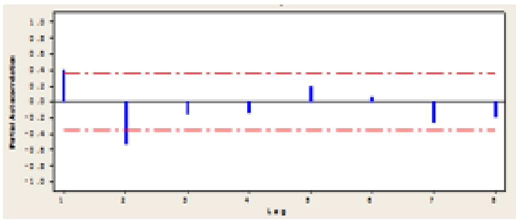


Figure 3.5 Detail Coefficient ACF Diagram Coefficient Diagram

Figure 3.6 PACF Detail

Based on figure 3.5 shows that the data is significant/truncated at lag 1 which shows the possible order of the Moving Average (q). While in Figure 3.6 the data is significant or truncated in lag 1 and lag 2 which shows the order of Autoregressive (p). And the Integrated order (d) shows the number of differencing done on the data, namely once. Thus, several possible models are ARIMA (0,1,1), ARIMA (1,1,0), ARIMA (1,1,1), ARIMA (2,1,0), and ARIMA (2,1,1).

3.5 Parameter Estimation

The results of estimation of possible model parameters for forecasting the Coefficient of Approximation were obtained using Minitab Software to obtain the following results:

Table 3.3 Estimation of ARIMA Model Parameters (0,0,3)

Parameter	Estimation	$p-$ value
MA(1)	-0.5932	0.000
MA(2)	0.5511	0.005
MA(3)	0.8523	0.000
Constant	112,733	0.000

Table 3.4 Estimation of ARIMA Model Parameters (0,0,1)

Parameter	Estimation	$p-$ value
MA(1)	-0.9795	0.000
Constant	111,672	0.005

Table 3.5 Estimation of ARIMA Model Parameters(1,0,0)

Parameter	Estimation	$p-$ value
AR(1)	0.4581	0.008
Constant	60,642	0.005

Table 3.6 Estimation of ARIMA Model Parameters (2,0,0)

Parameter	Estimation	$p-$ value
AR(1)	0.7909	0.000
AR(2)	-0.7431	0.000
Constant	107,106	0.005

Table 3.7 Estimation of ARIMA Model Parameters (1,0,3)

Parameter	Estimation	$p-$ value
AR(1)	0.1290	**0.572**
MA(1)	-0.5806	0.002
MA(2)	0.5489	0.009
MA(3)	0.8536	0.000
Constant	98.1837	0.005

Table 3.8 Estimation of ARIMA Model Parameters (1,0,1)

Parameter	Estimation	$p-$ value
AR(1)	-0.0189	**0.921**
MA(1)	-0.9800	0.000
Constant	113,782	0.000

Table 3.9 Estimation of ARIMA Model Parameters (2,0,1)

Parameter	Estimation	p − value
AR(1)	0.0320	**0.868**
AR(2)	-0.1505	**0.436**
MA(1)	-0.9413	0.000
Constant	125,408	0.000

Table 3.3 shows the model where all the parameters are significant, the remaining four models are ARIMA(0,0,3), ARIMA(0,0,1), ARIMA(1,0,0) and ARIMA(2,0,0) .

The estimation results of possible model parameters for Detail Coefficient forecasting are obtained using Minitab Software to obtain the following results:

Table 3.10 Estimation of ARIMA Model Parameters (0,1,1)

Parameter	Estimation	p − value
MA(1)	0.9427	0.001
Constant	0.0457	**0.794**

Table 3.11 Estimation of ARIMA Model Parameters (1,1,0)

Parameter	Estimation	p − value
AR(1)	-0.6240	0.000
Constant	0.1387	**0.890**

Table 3.12 Estimation of ARIMA Model Parameters (1,1,1)

Parameter	Estimation	p − value
AR(1)	-0.4688	0.014
MA(1)	0.9367	0.000
Constant	0.07494	**0.434**

Table 3.13 Estimation of ARIMA Model Parameters (2,1,0)

Parameter	Estimation	p − value
AR(1)	-1.0428	0.000
AR(2)	-0.7090	0.000
Constant	0.3440	**0.641**

Table 3.14 Estimation of ARIMA Model Parameters (2,1,1)

Parameter	Estimation	$p-\text{value}$
AR(1)	-1.5734	0.000
AR(2)	-0.7317	0.000
MA(1)	-0.9419	0.000
Constant	0.517	**0.778**

In the next stage, a diagnostic examination is carried out for the five forecasting models, namely ARIMA(0,1,1), ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(2,1,0), and ARIMA (2,1,1,) because it has fulfilled the assumption of significance for each of its parameters.

3.6 Test the Residual White Noise

White noise is a process that is independent and identical if the forms of successive random variables are not correlated with each other and follow a certain distribution. Table 3.15 show the result of testing the residual white noise from the approximation coefficient:

Table 3.15 Testing the Ljung Box Approximation Coefficient

Model	lag	
	12	24
ARIMA(0,0,3)	0.115	0.297
ARIMA(0,0,1)	0.000	0.000
ARIMA(1,0,0)	0.000	0.000
ARIMA(2,0,0)	0.09	0.174

Errors in the lags of the model are said to be uncorrelated or H_0 rejected if the p-value for each error in each lag is smaller than the error level of 0.05. Based on Table 3.15 the value is smaller than 0.05, namely the ARIMA(0,0,1) and ARIMA(1,0,0) models so that it does not meet the white noise error assumption. In other side, ARIMA (0,0,3) and ARIMA (1,0,0) that model meet the white noise error assumption.

In other side, the results of testing the remaining white noise from the detail coefficient shown in Table 3.16.

Table 3.16 Ljung Box Test for Detail Coefficients

Model	lag	
	12	24
ARIMA(0,1,1)	0.000	0.000
ARIMA(1,1,0)	0.000	0.000
ARIMA(1,1,1)	0.000	0.000
ARIMA(2,1,0)	0.100	0.090
ARIMA(2,1,1)	0.000	0.000

Based on Table 3.16, the model that meets the white noise error assumption is the ARIMA (2,1,0) model, namely the Ljung-Box value is 0.1 at lag 12 and 0.09 at lag 24.

3.7 Normality Test

The normality test was carried out to find out whether the residuals were normally distributed or not. The model that was tested for normality, namely the model meet the white noise error assumption. Table 3.17 show the result of approximation coefficient normality test using Kolmogorov Smirnov test of ARIMA(0,0,3) and ARIMA (2,0,0).

Table 3.17 Kolmogorov Smirnov Test

Model	$\mathbf{p-value}$
ARIMA(0,0,3)	0.307
ARIMA(2,0,0)	0.208

Table 3.17 shows that the ARIMA(0,0,3) model has $\mathrm{p-value}\ 0.307$ and the ARIMA(2,0,0) model has 0.208. Obtained that value$\mathrm{p-value}$ Both models are greater than 0.05 so they $H_0$are accepted, which means that both models are said to have normal distribution errors/residuals.

Furthermore, the ARIMA(2,1,0) model in detail coefficient was tested for normality in using the Kolmogorov Smirnov.

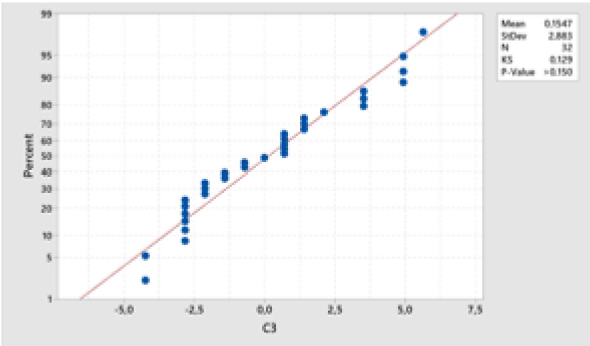


Figure 3.7 Kolmogorov Smirnov Normality Test Plot

From the Kolmogorov Smirnov test in Figure 3.7, a value is obtained $\mathrm{p-value} = 0.150 > 0.05$ or H_0 accepted, which means that the model is said to have a normally distributed error.

3.8 Selection of the Best Model

Selection of the best model was carried out using the smallest Akaike Information Criterion (AIC) and Mean Square Error (MSE).

Table 3.18 Value of AIC and MSE in Approximation Coefficient Model

Model	AIC	MSE
ARIMA(0,0,3)	112.2142	29,673
ARIMA(2,0,0)	113.2602	31,511

Table 3.18 shows that the smallest AIC and MSE values are in the ARIMA(0,0,3) model.

$$\hat{Y}_t = 112.733 + 0.5932a_{t-1} - 0.5511a_{t-2} - 0.8523a_{t-3}$$

In the ARIMA (0,0,3) model it can be seen that the change in the value of $\hat{Y}_t$ is equal to 112,733 when the effects of a_{t-1}, a_{t-2}, and a_{t-3} are zero. For an increase of one unit a_{t-1} , it gives an effect of adding value of 0.5932 to the average variable $\hat{Y}_t$. Meanwhile, a one-unit increase in a_{t-2} and a_{t-3} has a decreasing effect of 0.5511 and 0.8523 on the average $\hat{Y}_t$ variable.

In detail coefficent model, namely the ARIMA model (2,1,0) has a MSE of 15,989 and also an AIC value of 89.2. The ARIMA(2,1,0) model is obtained as follows:

$$\widehat{Y}_t = 0.3440 - 0.0428y_{t-1} - 1.7518y_{t-2} + 0.7090y_{t-3}$$

The forecasting model on ARIMA (2,1,0) shows that if the values y_{t-1}, y_{t-2}, and y_{t-3} are constant, then the value of $\widehat{Y}_t$ is equal to 0.3340. Meanwhile, a one-unit increase in y_{t-1} gives a decrease in value of 0.0428 to the average response variable $\widehat{Y}_t$. Whereas for each increase of one unit the value y_{t-2} and the value y_{t-3} each decreased by 1.7518 for y_{t-2} and increased by 0.7090 for y_{t-3}.

3.9 Forecasting

For example $c'_{5,i}$, the results of forecasting the coefficient of approximation level 1 ($c_{5,i}$) and $d'_{5,i}$ the results of forecasting the coefficient of detail level 1 ($d_{5,i}$). Obtained forecasting results are:

Table 3.19 Forecasting Results

No	Approximate Coefficient$(\mathbf{c'_{5,i}})$	Detail Coefficient$(\mathbf{d'_{5,i}})$
1	111,582	2.67301
2	114,060	4.71229
3	115,457	2.03737
4	112,733	3.72485
5	112,733	4.20566
6	112,733	2.85192

The results of the forecasting in Table 3.19 in the frequency domain are first transformed back to the time domain (inverse). The process of transforming back to the time domain can be used the following equation: If i is even $c'_{j,i} = \frac{\left(c'_{j-1,i}+d'_{j-1,i}\right)\sqrt{2}}{2}$, if i is odd$c'_{j,i} = \frac{(c'_{j-1,i}-d'_{j-1,i})\sqrt{2}}{2}$

4. CONCLUSION

From this research, several conclusions were obtained, namely:

1. Coefficients - wavelet coefficients consist of two, namely the approximation coefficient and the detail coefficient. Coefficient approximation $(c_{j-1,i})$ and detail coefficient $\left(d_{j-1,i}\right)$ respectively are obtained by using the following formula:

$$c_{j-1,i} = \sum_n h_{n-2i}c_{j,n}$$

$$d_{j-1,i} = \sum_n g_{n-2i}c_{j,n}$$

2. Models using discrete wavelet transforms applied to air humidity data obtained the best model for each component, namely

The forecasting model for the coefficient of approximation is obtained by ARIMA(0,0,3) as follows:

$$\hat{Y}_t = 112.733 + 0.5932a_{t-1} - 0.5511a_{t-2} - 0.8523a_{t-3}$$

From this model, the Akaike Information Criterion value = 112.2142 and Mean Square Error = 29.673 are obtained.

The Detail Coefficient forecasting model is obtained by ARIMA(2,1,0) as follows:

$$\hat{Y}_t = 0.3440 - 0.0428y_{t-1} - 1{,}7518y_{t-2} + 0.7090y_{t-3}$$

From this model, the value of Akaike Information Criterion = 89.2 and Mean Square Error = 15,989 is obtained.

REFERENCES

[1] Aditya Satrio, C. B., Darmawan, W. B., Nadia, U. and Hanafiah, N., 2020. Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET," *Procedia Comput. Sci.*, vol. 179, no. pp. 524–532, 2021, doi: 10.1016/j.procs.2021.01.036.

[2] Amin, R. , Shah, K., Asif, M., Khan, I. and Ullah, F., 2020. An efficient algorithm for numerical solution of fractional integro-differential equations via Haar wavelet. *J. Comput. Appl. Math.*, vol. 381, p. 113028, doi: 10.1016/j.cam.2020.113028.

[3] Apelgren, P. *et al.*, 2019. Study of wavelet-based denoising and a new shrinkage function for low-dose CT Scans. *Mater. Today Proc.*, vol. 27, no. xxxx, pp. 0–31, [Online].

[4] Burrus et al, C. S., 1998. *Introduction to Wavelets and Wavelet Transforms a Primer*. United State of America: Prentice Hall.

[5] Daubechies, I., 1992. *Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics*. SIAM.

[6] Friadi, R. and Junadhi, J., 2019. Sistem Kontrol Intensitas Cahaya, Suhu dan Kelembaban Udara Pada Greenhouse Berbasis Raspberry PI. *J. Technopreneursh. Inf. Syst.*, vol. 2, no. 1, pp. 30–37, doi: 10.36085/jtis.v2i1.217.

[7] Guo, T., Zhang, T., Lim, E., Lopez-Benitez, M., Ma, F. and Yu, L., 2022. A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities. *IEEE Access*, vol. 10, pp. 58869–58903, doi: 10.1109/ACCESS.2022.3179517.

[8] Katsavrias, C., Papadimitriou, C., Hillaris, A. and Balasis, G., 2022. Application of Wavelet Methods in the Investigation of Geospace Disturbances: A Review and an Evaluation of the Approach for Quantifying Wavelet Power. *Atmosphere (Basel).*, vol. 13, no. 3, doi: 10.3390/atmos13030499.

[9] Khan, S. and Alghulaiakh, H., 2020. ARIMA model for accurate time series stocks forecasting. *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 524–528, doi: 10.14569/IJACSA.2020.0110765.

[10] Lim, B. and Zohren, S., 2021. Time-series forecasting with deep learning: A survey. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 379, no. 2194, doi: 10.1098/rsta.2020.0209.

[11] Olsavszky,V., Dosius, M., Benecke, J., and Vladescu, C., 2020. Time series analysis and forecasting with automated machine learning on a national ICD-10 database. *Int. J. Environ. Res. Public Health*, vol. 17, no. 14, pp. 1–17, doi: 10.3390/ijerph17144979.

[12] Othman, G. and Zeebaree, D.Q., 2020. The Applications of Discrete Wavelet Transform in Image Processing: A Review. *J. Soft Comput. Data Min.*, vol. 1, no. 2, pp. 31–43,

[Online].

[13] Rhif, M., Ben Abbes, A., Farah, I. R., Martínez, B. and Sang, Y., 2019. Wavelet transform application for/in non-stationary time-series analysis: A review. *Appl. Sci.*, vol. 9, no. 7, pp. 1–22, doi: 10.3390/app9071345.

[14] Solichin,Z., 2011. *Program Aplikasi Keamanan Citra Dengan Algoritma Des dan Transformasi Wavelet Diskrit*. [Online].

[15] WAHYUNI, N. P. M. S., SUMARJAYA, I. W. and SRINADI, I. G. A. M., 2016 Peramalan Curah Hujan Menggunakan Metode Analisis Spektral. *E-Jurnal Mat.*, vol. 5, no. 4, p. 183, doi: 10.24843/mtk.2016.v05.i04.p139.

[16] Wei, W. W., 1994. *Time Series Analysis Univariate and Multivariate Methods*. California: Addison-Wesley Publishing Company.