Multiscale Autoregressive (MAR) Models with MODWT Decomposition on Non-Stationary Data

Melda Juliza\textsuperscript{1*}, Umu Sa’adah\textsuperscript{1} and Adjri A R Fernandes\textsuperscript{1}
\textsuperscript{1}Statistics Department – Universitas Brawijaya, Malang

*Corresponding Author: melda_juliza@yahoo.com

Abstract. Time series forecasting often shows behavior that is non-stationary, so it is necessary to do a forecasting method that can predict non-stationary data in order to obtain good forecast results, including Autoregressive Integrated Moving Average (ARIMA) and Multiscale autoregressive (MAR). The characteristics of this model do not include predictor variables in the model. The MAR model is a model that performs the transformation process using wavelet. The MAR Model adopts an autoregressive (AR) time series model with predictors used are wavelet and scale coefficients. The wavelet and scale coefficients are obtained by decomposing using Maximal Overlap Discrete Wavelet Transformation (MODWT). MODWT functions to decipher data based on the level of each wavelet filter. This research aims to determine the best forecasting model using the ARIMA and MAR models. Testing performed on non-stationary data, so the ARIMA and MAR models can be used in this research. This research is expected to be able to obtain the best time series model and the most suitable to be used in predicting on non-stationary data.

Keywords: Non-stationary, ARIMA, MAR, MODWT.

1. Introduction

Wavelet transformation is a transformation function that divide data into different frequency components, then analyzes for each component using a resolution that is in accordance with the scale [1]. This transformation is very suitable to be used for non-stationary data, because it is automatically able to separate trend from a data pattern. Wavelet transform can make data simpler without changing the shape of the initial pattern of data [2]. Several research have been conducted relating to the wavelet transformation method for time series analysis, i.e. [3] and [4], development of wavelet transformation for prediction of time series data based on an autoregressive model i.e. [5], [6], [2] and [7]. Wavelet transformation with DWT decomposition is carried out by [8].

Wavelet transformation is done by forming a multiscale autoregressive (MAR) model. The MAR model adopt the AR time series model with the input used are wavelet and scale coefficients. The wavelet and scale coefficients are obtained by decomposing using the maximal overlap discrete wavelet transform (MODWT). MODWT functions to decipher data based on the level of each wavelet filter. MODWT is a development of discrete wavelet transform (DWT), where DWT has a weakness that data must qualify of $n = 2^j$, where n is a measure of the number of time series data and j is a positive integer, whereas generally time series data does not always qualify of $2^j$ [4].

This research aims to apply the ARIMA and MAR models with the decomposition of MODWT on the daily data on the exchange rate of Rupiah to US dollar in 2014-2019, then analyze the model that is best used for forecasting exchange rate data in the future. this method is expected to be an alternative in predicting the exchange rate of Rupiah to US dollar in the future.
2. Literature Review

2.1. ARIMA

The non-stationary time series model is called the ARIMA process \((p, d, q)\) where \(p\) is the order of the autoregressive parameter, \(d\) is the quantity that states how many times differencing is done to reach stationary, and \(q\) the order of moving average parameters \([4]\). Mathematically the ARIMA \((p, d, q)\) model can be written as follows \([5]\):

\[
\phi(B)(1 - B)^d X_t = \theta(B)a_t
\]

Suppose \(d = 1\), then \(W_t = (1 - B)X_t = X_t - X_{t-1}\). So that:

\[
W_t = \phi W_{t-1} + \ldots + \phi p W_{t-p} + a_t - \theta a_{t-1} - \ldots - \theta q a_{t-q}
\]

The establishment of the ARIMA model was carried out with a Box-Jenkins procedure consisting of 4 stages, i.e. model identification, parameter estimation, diagnostic test and best model selection \([6]\). Model identification is the stage for knowing stationarity in mean and varians. If the data is not stationary in mean, then it is differencing and if it is not stationary in varians then Box-Cox transformation is carried out.

2.2. MAR

A modification of the AR model is proposed by \([2]\) known as the multiscale autoregressive (MAR) model. The MAR model is a model with a transformation process using wavelet, which assumes each scale of a wavelet transformation follows an AR process. Determination of lags on input variables for MAR model using wavelet and scale coefficients obtained from the results of wavelet transformation \([2]\). wavelet and scale coefficients from wavelet transformation through MODWT decomposition are considered to have an effect for predictions at time \(t+1\) to be in the form \(w_{j,t-2^j(k-1)}\) of and \(v_{j,t-2^j(k-1)}\), or can be written as:

\[
\hat{X}_{t+1} = \sum_{j=1}^{J} \sum_{k=1}^{A_j} \hat{a}_{j,k} w_{j,t-2^j(k-1)} + \sum_{k=1}^{A_{j+1}} \hat{a}_{j+1,k} v_{j,t-2^j(k-1)} + \varepsilon_t
\]

with \(J\) is the level of decomposition MODWT, \(A_j\) is the order of the MAR model \((A_j)\), \(a_{j,k}, \hat{a}_{j+1,k}\) is the value of the coefficients of the MAR model, \(w_{j,t-2^j(k-1)}\) is the wavelet coefficient and is \(v_{j,t-2^j(k-1)}\) the scale coefficient.

Determination of input MAR models on data forecasting to \((t+1)\) is shown in Figure 1, namely the first input on each scale is the \(t\)-data, and the second input on each scale is \((t-2^j)\)-data.

![Figure 1. Illustration of the Wavelet MAR Model](image-url)
wavelet coefficients at t = 20 and t = 16, level 3 wavelet coefficients at t = 20 and t = 12, level 4 wavelet coefficients at t = 20 and t = 4, and scale coefficients at t = 20 and t = 4.

Suppose J = 6 and \( A_j = 2 \), the MAR model based on equation (3) can be expressed as:

\[
\hat{X}_{t+1} = \sum_{j=1}^{6} \sum_{k=1}^{2} \hat{a}_{j,k} w_{j,2^j-2^j(k-1)} + \sum_{k=1}^{2} \hat{a}_{j,k} v_{j+1,2^j-2^j(k-1)} \\
= \hat{a}_{1,1} w_{1,t} + \hat{a}_{1,2} w_{2,t-2} + \hat{a}_{2,1} w_{2,t-4} + \hat{a}_{2,2} w_{2,t-8} + \hat{a}_{3,1} w_{3,t-16} + \hat{a}_{3,2} w_{3,t-32} + \hat{a}_{4,1} w_{4,t-32} + \hat{a}_{4,2} w_{4,t-64} + \hat{a}_{5,1} w_{5,t-64} + \hat{a}_{5,2} w_{5,t-96} + \hat{a}_{6,1} w_{6,t-96} + \hat{a}_{6,2} w_{6,t-128} + \hat{a}_{7,1} w_{7,t-128} + \hat{a}_{7,2} w_{7,t-192}
\]

or can be written as \( \hat{a}_t = A_t\alpha \).

The estimation of \( \alpha \) vector parameters can be solved by the least squares method which minimizes the squared error [7], namely

\[
\hat{\alpha} = (A'A)^{-1}A'y
\]

Calculation of the estimated value of the MAR model, used the value of \( \alpha \) which has been estimated using equation (5). The expected \( \alpha \) value and the wavelet and scale coefficient obtained through decomposition are entered into the MAR model as in equation (3).

2.3. MODWT

Wavelet transformations which are considered more suitable for time series data are MODWT because in each decomposition level there are wavelet and scale coefficients as much as the length of the data [3]. Determination of level (J) for MODWT decomposition depends on the width of the decomposition filter (L) and the number of data (n), with the formula:

\[
J < \ln \left( \frac{n}{L-1} + 1 \right)
\]

The pyramid algorithm for MODWT is an algorithm to calculate the scale and wavelet coefficient of MODWT at the \( j \)th level. If a data \( x \) is decomposed with a wavelet and scale filter, it will produce a wavelet coefficient and scale coefficient. Figure 2 below is a pyramid algorithm for MODWT [8].

\[
\begin{align*}
X & \rightarrow v_1 = A_1x, w_1 = B_1x \\
v_2 = A_2A_1x, w_2 = B_2A_1x \\
v_3 = A_3A_2A_1x, w_3 = B_3A_2A_1x \\
\vdots \\
v_j = A_j \ldots A_3A_2A_1x, w_j = B_j \ldots A_3A_2A_1x
\end{align*}
\]

**Figure 2.** MODWT Pyramid Algorithm Scheme

3. Research Methods

The time series data used in this research is the daily data on the exchange rate of Rupiah to US Dollar in January 2014-February 2019, which is 1385 observations. The data is a non-stationary time series
data. The type of wavelet used are Haar and Daubechies wavelet. The research steps taken in this research are:

a) Determine the level according to the filter used, that is Haar with a filter width of 2 and Daubechies with a filter width of 4.
b) Check stationary data.
c) If the data is stationary, decompose with MODWT at each level.
d) Selecting the wavelet and scale coefficient according to [2].
e) Selecting a wavelet and scale coefficient based on a significant PACF coefficient.
f) Perform a stepwise method.
g) Checking the assumptions of normality and white noise in the MAR model. If the assumption of normality is not fulfilled, the model is not used. Whereas if the white noise assumption is not fulfilled, MAR modeling is done again by adding wavelet and scale coefficients based on a significant lag on the residual ACF plot.
h) Perform MAR modeling on wavelet coefficients and significant scale.
i) Form a forecasting model MAR.
j) Selecting the MAR forecasting model with the smallest RMSE, then comparing it with ARIMA.

4. Results and Discussion

In this research, the MAR model is applied to daily data on the exchange rate of Rupiah to US dollar from 01 January 2014 - 28 February 2019, which consist of 1385 observations. The data used is divided into 2 parts, namely 1365 observations as training data and 20 observations as testing data. Description of exchange rates is used to determine the general description of the data, namely how much the mean value, distribution of data, maximum and minimum values, and the amount of exchange rate data used in this research.

Table 1. Descriptive Statistics

| Variable     | N   | Mean   | std deviation | Minimum | Maximum |
|--------------|-----|--------|---------------|---------|---------|
| Exchange rates| 1365| 13285  | 872           | 11288.5 | 15235   |

Based on Table 1. it can be seen that the amount of exchange rate data used for modeling is 1365 observations. The mean value of the exchange rate is 13285. The distribution of exchange rate data that is equal to 872. The minimum value of exchange rates occurs on April 8, 2014 amounting to 11288.5. While the maximum value of the Exchange occurs on October 11, 2018 amounting to 15235.

The first step is testing data stationarity. Stationarity test using the augmented dickey-fuller unit root test (ADF) can be seen in Table 2. The following:

Table 2. Test of ADF Exchange Rate Data.

| Data           | P-value before differencing | P-value after differencing | Conclusion          |
|----------------|----------------------------|---------------------------|---------------------|
| Exchange rate  | 0.559                      | 0.01                      | Stationary after differencing to 1 |

Based on Table 2, it is known that after differencing 1 time is obtained p-value = 0.01. This means that stationarity in the mean is achieved after differing 1 time. Plot data for exchange rates after differencing 1 can be seen in Figure 3. The following:
The Box-Cox transformation shows that the Exchange rate data produces a value of $\lambda = 1$. It can be said that the exchange rate data 1st differencing has been stationary in varians. Thus the results of stationary test show that the exchange rate on 1st differencing has been stationary in mean and varians.

4.1. The ARIMA Model

Determination of the best ARIMA model for Exchange Rate data uses software R with the function `auto.arima()`. The best ARIMA model for Exchange Rate data is presented in Table 3.

| Data       | Model       | RMSE | Parameter | s.e   | p-value    | Shapiro-Wilk |
|------------|-------------|------|-----------|-------|------------|--------------|
| Exchange   | ARIMA (0,1,1)| 50.63| MA 1      | 0.117 | 0.0269     | <1.309e-05   | <2.2e-16    |

So, the ARIMA (0,1,1) model formed in the Exchange rate data is:

$$X_t = a_t - 0.117a_{t-1}$$

To know the accuracy of the ARIMA model, the fitted test of the model by using the Ljung-Box (Q) test with $\alpha = 5\%$. With the help of software R, obtained p-value = 0.9785 a greater than $\alpha$. However, in testing the assumptions of model residuals only the assumptions of normality are fulfilled. This can be seen in the Shapiro_Wilk test value which is smaller than $\alpha$. While the assumption of white-noise is not fulfilled. So, it can be concluded that the ARIMA model (0,1,1) is not suitable to be used in this research. So that modeling will be done using the MAR model with MODWT decomposition.

4.2. The MAR Model

Determination of the level on the Exchange Rate data for the wavelet Daubechies filter based on equation (6) obtained a value $J < 6,122$, then the level used is the level 1,2,3,4,5,6. Before decomposing using the MODWT method, the first step taken is checking the exchange rate data stationarity. For example, the MAR (2) level 6 in the Exchange Rate data has been differencing stationary, it is known that the filter width ($L_j$) = 4, MAR order ($A_j$) = 2 and level ($J$) = 6. MODWT decomposition can be done with the help of software R.

Best on the Renaud et al [2] predictors for modeling MAR (2) level 6 are $w_{1,1}, w_{1,1-2}, w_{2,1}, w_{2,1-4}, w_{3,2}, w_{3,2-8}, w_{4,3}, w_{4,3-16}, w_{5,4}, w_{5,4-32}, w_{6,5}, w_{6,5-64}, v_{6,6}$ dan $v_{6,6-64}$. Then a stepwise method is used to obtain the right predictor. Next is the residual checking of the MAR model (2), which is the assumption of normality and white noise. Input of the MAR model is in accordance with Renaud et al. presented in Table 4.
Based on Table 4, the input of the MAR model is in accordance with Renaud et al. obtained that the MAR (2) model fulfilled the assumption of normality. While the assumption of white noise is not fulfilled. So that additions are made to the MAR model, namely the addition of the PACF lag from the wavelet and significant scale coefficient.

Table 4. Input MAR According to Renaud with the Daubechies Filter.

| Level (j) | Variable input | Normality | White noise | RMSE   |
|-----------|----------------|-----------|-------------|--------|
| 1         | $v_{1,t}$      | yes       | -           | 50.91740 |
| 2         | $w_{1,t-2}; v_{2,t}$ | yes       | -           | 50.88288 |
| 3         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.82072 |
| 4         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.93443 |
| 5         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.84304 |
| 6         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.79404 |

Table 5. Input MAR Based on Renaud and PACF Wavelet Coefficient with the Daubechies Filter.

| Level (j) | Variable input | Normality | White noise | RMSE   |
|-----------|----------------|-----------|-------------|--------|
| 1         | $w_{1,t-2}; v_{1,t}$ | yes       | -           | 50.87269 |
| 2         | $w_{1,t-2}; v_{2,t}$ | yes       | -           | 50.82966 |
| 3         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.82072 |
| 4         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.93443 |
| 5         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.84304 |
| 6         | $w_{1,t-2}; w_{3,t-8}$ | yes       | -           | 50.79404 |

Based on Table 5. Input MAR models based on Renaud et al. and a significant PACF coefficient is obtained that the MAR model that fulfilled the assumptions of normality. While the assumption of white noise is not fulfilled. So that MAR modeling is done again by adding wavelet and scale coefficients based on a significant lag on the residual ACF plot. However, the formation of a fixed model does not fulfilled the white noise assumption. So, the MAR model with MODWT decomposition using Daubechies filters is not suitable to be applied to Exchange Rate data.

Furthermore, determining the level of the Exchange Rate data for the Haar wavelet filter based on equation (6) obtained value $J < 7.22$, the levels used are 1, 2, 3, 4, 5, 6, and 7. Predictors for MAR modeling based on Renaud et al. and PACF from wavelet and scale coefficients that are significant with Haar filter are presented in Table 6.

Table 6. MAR Inputs Based on Renaud et al. and PACF Wavelet Coefficients with Haar Filter.

| Level (j) | Variable input | Normality | White noise | RMSE   |
|-----------|----------------|-----------|-------------|--------|
| 1         | $w_{1,t-1}; v_{1,t}$ | yes       | -           | 50.69195 |
| 2         | $w_{1,t}; v_{2,t}$ | yes       | -           | 50.74439 |
| 3         | $w_{1,t}; w_{2,t-1}; v_{3,t}$ | yes       | -           | 50.74443 |
| 4         | $w_{1,t}; w_{3,t-1}; w_{3,t-4}$ | yes       | -           | 50.83645 |
| 5         | $w_{1,t}; w_{3,t-1}; w_{3,t-4}$ | yes       | -           | 50.72204 |
| 6         | $w_{3,t-1}; w_{3,t-4}$ | yes       | -           | 50.61472 |
| 7         | $w_{1,t-1}; w_{3,t}; w_{3,t-1}; w_{3,t-4}$ | yes       | yes         | 50.65370 |
Based on Table 6. Input of the MAR model based on Renaud et al. and a significant PACF coefficient is obtained that the MAR (2) model fulfilled the assumption of normality. While the assumption of white noise is not fulfilled. So that MAR modeling is done again by adding wavelet and scale coefficients based on a significant lag on the residual ACF plot. The result of the analysis state that only 7th-level in the formation of the MAR (2) model with the Haar filter fulfilled the assumption of white noise with the addition of lag-20. So, the MAR model used in forecasting exchange rates is the MAR(2)-level7-Haar-lag20. The coefficients of the MAR model are shown in Table 7.

Table 7. Coefficients of MAR Model.

| Variable  | Estimate | s.e  | p-value  |
|-----------|----------|------|----------|
| $w_{1,t-19}$ | 0.11166 | 0.04222 | 0.00828 |
| $w_{3,t}$ | 0.59525 | 0.12077 | 9.41e-07 |
| $w_{3,t-1}$ | -0.35436 | 0.10853 | 0.00112 |
| $w_{3,t-4}$ | 0.31326 | 0.08872 | 0.00043 |

So, MAR (2)-level7-Haar-lag20 model formed in the Exchange rate data is:

$$\hat{X}_t = 0.11166w_{1,t-19} + 0.59525w_{3,t} - 0.35436w_{3,t-1} + 0.31326w_{3,t-4}$$

Table 8 below results of forecasting the Exchange Rate data for the next 20 periods based on the MAR(2)-level7-Haar-lag20, the RMSE value of the testing data obtained is 128.7749.

Table 8. Forecasting Results of the MAR Model

| No | Testing Data | Forecast Data | No | Testing Data | Forecast Data |
|----|--------------|---------------|----|--------------|---------------|
| 1  | 14040.0      | 13971.78      | 11 | 14017.5      | 13982.89      |
| 2  | 14068.5      | 13967.75      | 12 | 13991.0      | 13985.46      |
| 3  | 14057.5      | 13972.28      | 13 | 14030.0      | 13986.01      |
| 4  | 14090.0      | 13973.48      | 14 | 14065.0      | 13975.25      |
| 5  | 14142.5      | 13977.70      | 15 | 14115.0      | 13981.53      |
| 6  | 14106.5      | 13986.86      | 16 | 14130.0      | 13983.76      |
| 7  | 14102.0      | 13986.08      | 17 | 14120.0      | 13982.25      |
| 8  | 14040.0      | 13985.49      | 18 | 14140.0      | 13979.86      |
| 9  | 14063.0      | 13986.19      | 19 | 14140.0      | 13985.45      |
| 10 | 14057.5      | 13982.23      | 20 | 14310.0      | 13982.16      |

5. Conclusion

Based on the results of the analysis it can be concluded that for data problems the model rate used is ARIMA (0,1,1) with RMSE values in the sample 50.63 and RMSE out sample 143.96. But this model is not suitable for use because the assumption of white noise is not fulfilled. So that the model is formed using Model MAR. The best MAR model is the MAR model according to Renaud et al. and based on PACF the wavelet and scale coefficients are significant with RMSE values of training data 50.65 and the RMSE of testing data 128.77. So, the model that is suitable to be used in the exchange rate data problem is MAR(2)-level7-Haar-lag20 based on Renaud et al. and significant PACF.
References

[1] Graps, A. 1995. An Introduction to Wavelets. *IEEE Computational Science and Engineering*, 2 (2).

[2] Renaud, O., Starck, J. L and Murtagh, F. 2003. Prediction Based on a Multiscale Decomposition. *International Journal of Wavelets, Multiresolution and Information Processing*, 1(2), 217-232.

[3] Nason, G. P. and von Sachs, R. (1999). Wavelets in Time Series Analysis. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 357(1760), 2511-2526.

[4] Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.

[5] Bjorn, V. (1995). Multiresolution Methods for Financial Time Series Prediction. *Proc. IEEE/IAFE 1995 Conf. Computational Intelligence for Financial Eng.*, 97.

[6] Soltani, S., Boichu, D., Simard, P. and Canu, S. (2000). The Long-term Memory Prediction by Multiscale Decomposition. *Signal Processing*, 80, 2195-2205.

[7] Sa’adah, U., Subanar, Guritno, S. and Suhartono (2010). Multiscale Seasonal Autoregressive for Forecasting Trend and Seasonal Time Series. *Proceedings of the Tenth Islamic Countries Conference on Statistical Sciences (ICCS-X)*. 2, 1079–1086.

[8] Ulfiati, L and Sugiman. (2015). Peramalan Data Saham dengan Transformasi Wavelet Haar. *Unnes Journal of Mathematics*, 4 (2).

[9] Montgomery, D. C., Jennings, C. L. and Kulahci, M. 2008. *Introduction Time Series Analysis and Forecasting*. John Wiley & Sons, Inc., New Jersey.

[10] Cryer, J. D., and Chan, K. S. 2008. *Time Series Analysis with Application in R*. Second Edition. Spinger Science dan Business Media., USA.

[11] Box, G. E. P. and Jenkins, G. M. 1994. *Time Series Analysis: Forecasting and Control*. Third Edition. Prentice Hall, New Jersey.

[12] Makridakis, S., Wheelwright, S. C. and McGee, V. 1999. *Metode dan Aplikasi Peramalan*. Translation Hari Suminto. Revised Edition. Binarupa Aksara, Jakarta.

[13] Popoola, A. O. 2007. *Fuzzy-Wavelet Method for Time Series Analysis*. Submitted for the degree of doctor of philosophy in Surey University, England.