Model Selection in Feedforward Neural Networks for Forecasting Inflow and Outflow in Indonesia

Suhartono(✉), Prilyandari Dina Saputri, Farah Fajrina Amalia, Dedy Dwi Prastyo, and Brodjol Sutijo Suprih Ulama

Department of Statistics, Institut Teknologi Sepuluh Nopember, Kampus ITS Keputih Sukolilo, Surabaya 60111, Indonesia
{suhartono,dedy-dp,brodjol_su}@statistika.its.ac.id, prildisa@gmail.com, farahfajrinaamalia@gmail.com

Abstract. The interest in study using neural networks models has increased as they are able to capture nonlinear pattern and have a great accuracy. This paper focuses on how to determine the best model in feedforward neural networks for forecasting inflow and outflow in Indonesia. In univariate forecasting, inputs that used in the neural networks model were the lagged observations and it can be selected based on the significant lags in PACF. Thus, there are many combinations in order to get the best inputs for neural networks model. The forecasting result of inflow shows that it is possible to testing data has more accurate results than training data. This finding shows that neural networks were able to forecast testing data as well as training data by using the appropriate inputs and neuron, especially for short term forecasting. Moreover, the forecasting result of outflow shows that testing data were lower accurate than training data.

Keywords: Forecasting · Inflow · Outflow · Nonlinear · Neural network

1 Introduction

In recent years, neural networks are one of the most popular methods in forecasting. Neural networks applied in vary fields such as energy, financial and economics, environment, etc. [1–4]. The interest in study using neural networks models has increased as they are able to capture non linear pattern and have a great accuracy.

In financial and economics, neural networks are used to predict the movement of stock price index in Istanbul [5]. This method was compared to Support Vector Machine (SVM). They found that neural networks model were significantly better than SVM model. This result shows that neural networks have a great performance in financial data.

Forecasting inflow and outflow plays an important role to achieve the stability of economics in Indonesia. The currency that exceeded the demand will lead to inflation. On the other hand, the currency that less than the demand will lead to the declining of economic growth. In Indonesia, the suitability of the currency was maintained by Bank Indonesia. In order to guarantee the availability of currency, Bank Indonesia needs to plan the demand and supply of currency. The forecasting results of inflow and outflow will be used as the indicator in determining the demand of currency in the next period.

A. Mohamed et al. (Eds.): SCDS 2017, CCIS 788, pp. 95–105, 2017.
https://doi.org/10.1007/978-981-10-7242-0_8

Thus, an accurate forecast of inflow and outflow in Indonesia will lead to the suitability of the demand and supply currency.

The study about forecasting inflow and outflow has been done by using many methods, such as ARIMAX, SARIMA, time series regression, etc. [6–9]. By using classical methods, there are many assumptions that must be fulfilled. One of them is the homoscedasticity of the variance residuals. Unfortunately, the previous study found that the residuals from ARIMAX model didn't fulfill the assumption [6, 8]. Thus, using neural networks for forecasting inflow and outflow in Indonesia seems to have promising opportunity, since this method was free from assumption.

The selection of the input were one of the most important decisions for improve the forecast accuracy [10]. In univariate forecasting, inputs that used in the model were the lagged observations. In this study, selections of inputs are based on the significant lags in PACF [11]. Thus, there are many combinations in order to get the best inputs for neural networks model. The model selection will be done by using cross validation. Cross validation is the most generally applicable methods for model selection in neural networks [12]. The forecasting results from FFNN model will be compared with the widely used classical methods, i.e. ARIMA and ARIMAX.

2 Methods

2.1 Data

The data used in this study is secondary data obtained from Bank Indonesia. The data used are inflow and outflow data from January 2003 to December 2016. The data will be divided into training data and testing data. The training data are inflow and outflow data from January 2003 to December 2015, while testing data are inflow and outflow data from January 2016 to December 2016.

2.2 Autoregressive Integrated Moving Average (ARIMA)

The ARMA model is a combined model of the Autoregressive and Moving Average process. The AR process is a process that describes Y_t that influenced by the previous condition $\left(Y_{t-1}, Y_{t-2}, \ldots Y_{t-p}\right)$ and has a white noise a_t. The MA process is a process which shows that the estimated value of Y_t influenced by error at the time t and the previous error $\left(a_{t-1}, a_{t-2}, \ldots a_{t-q}\right)$. Non-stationary time series data can be differenced on a certain order to produce a stationary data. The general equation of the ARIMA model (p, d, q) can be written as follows [13].

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B) a_t, \tag{1}$$

where:

$\phi_p(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)$
$\theta_q(B) = (1 - \theta_1 B - \ldots - \theta_q B^q)$
$(1-B)^d$= differencing order,
a_t= error at the time t.

2.3 Autoregressive Integrated Moving Average with Exogenous Variable (ARIMAX)

ARIMAX model is a development of the ARIMA model. In the ARIMAX model, there is used an additional variable known as exogenous variable. The exogenous variables used can be dummy variables (non-metric) or other time series variables (metrics). In this study, exogenous variables used are dummy variables i.e. trend, monthly seasonal, and calendar variations effects. The general equation of the ARIMAX model can be written as follows [14].

$$Y_t = \beta_0 + \beta_1 V_{1,t} + \beta_2 V_{2,t} + \ldots + \beta_h V_{h,t} + N_t, \tag{2}$$

$$N_t = \frac{\theta_q(B)}{\phi_p(B)} a_t, \tag{3}$$

where:

$V_{h,t}$ = dummy variable,
$\phi_p(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)$,
$\theta_q(B) = (1 - \theta_1 B - \ldots - \theta_q B^q)$,
N_t = residual of the time series regression process,
a_t = residual of the ARIMAX process.

2.4 Feedforward Neural Network

Neural networks are one of nonlinear regression methods and widely applied in pattern recognition. The idea in building neural networks models is motivated by their similarity to working biological system, which consist of large number of neurons that work in parallel and have the capability to learn. Neural networks are able to process vast amounts of data and make accurate predictions [15].

In time series forecasting, the most popular neural networks model are Feedforward Neural Network. In FFNN, the process starts from inputs that are received by the nodes, where these nodes are grouped in input layers. Information received from the input layer proceeds to the layers in the FFNN up to the output layer. Layers between input and output are called hidden layers. The input that used in neural network for forecasting are the previous lagged observations and the output describing the forecasting results. The selection of input variables are based on the significant lags in PACF [11].

The accuracy of neural networks model is determined by three components, i.e. the network architecture, training methods or algorithms, and activation functions. FFNN with p input and one hidden layer that consist of m neuron can be described as the figure below.

The model of FFNN in Fig. 1 can be written as follows:

$$f(\mathbf{x}_t, \mathbf{v}, \mathbf{w}) = g_2 \left\{ \sum_{j=1}^{m} v_j g_1 \left[\sum_{i=1}^{p} w_{ji} x_{it} \right] \right\}, \tag{4}$$

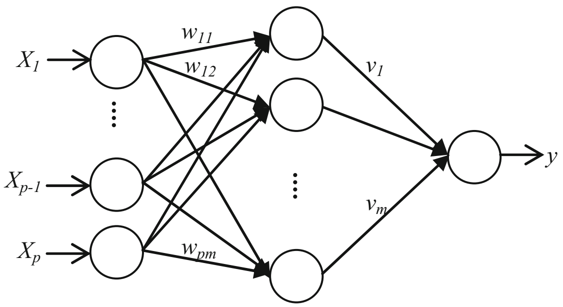

Fig. 1. Neural networks architecture

where **w** is the weights that connect the input layer to the hidden layer, **v** is the weights that connect the hidden layer to the output layer, $g_1(\cdot)$ and $g_2(\cdot)$ is the activation function, while w_{ji} and v_j are the weights to be estimated. The widely used activation function is tangent hyperbolic with the function below:

$$g(x) = \tanh(x) \tag{5}$$

2.5 Model Evaluation

The purpose of model evaluation is to know the performance of model in forecasting future period. Model evaluation will be based on the accuracy of forecasting result using *Root Mean Square Error* (RMSE). Model with the smallest RMSE will be selected as the best model in neural network. RMSE can be calculated by using the following formula [13]:

$$RMSE = \sqrt{\frac{1}{L}\sum_{l=1}^{L} e_l^2} \tag{6}$$

where $e_l = Y_{n+l} - \widehat{Y_n}(l)$.

3 Results

3.1 Forecasting Inflow and Outflow in Indonesia

The growth of inflow and outflow in Indonesia can be shown at the time series plot in Fig. 2. It shows that both inflow and outflow in Indonesia tends to increase every year. In other words, inflow and outflow in Indonesia have a trend pattern. In general, inflow and outflow data in Indonesia has a seasonal pattern [16]. However, the time series plot can't captured this pattern clearly. Thus, to determine the monthly pattern of the inflow and outflow, it will be used PACF plot in the next step.

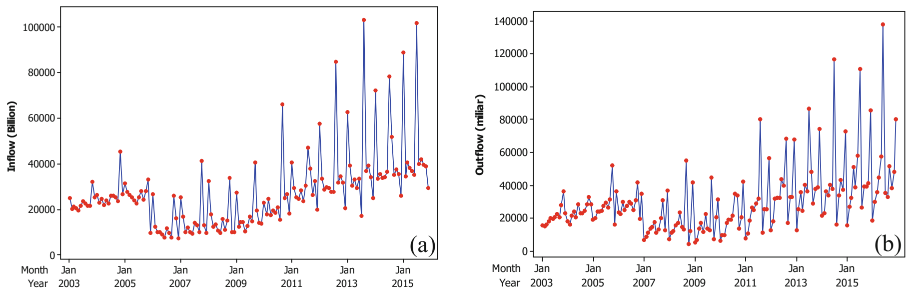

Fig. 2. Time series plot data of inflow (a) and outflow (b)

Selection of inputs that used in the neural network can be performed using lag plot. If nonlinear pattern is formed at a certain lag, then this lag will be used as input to the neural network models. Lag plot of inflow and outflow in Indonesia can be seen at the following figure.

Based on the lag plot in Fig. 3, the correlations with lagged observation until lag 15 tend to have the same pattern. Therefore, input determination based on the lag plot would be difficult to do. Thus, input determination will be done by using PACF from stationary data. If at certain lag the partial autocorrelation was significant, then this lag will be used as input to the neural network models. First, we will check the stationary of the data by using Augmented Dickey Fuller test, as shown in Table 1.

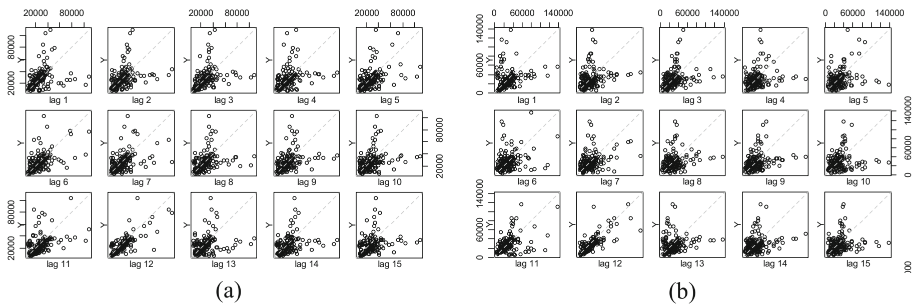

Fig. 3. Lag plot of inflow (a) and outflow (b)

Table 1 shows that the inflow data is not stationary in mean. Thus, the determination of the input for the inflow data will be based on the PACF of the differenced inflow data, while for outflow data, Augmented Dickey Fuller test shows that outflow data has been stationary in mean. Thus, the determination of inputs for outflow data

Table 1. Stationary test for inflow and outflow data

Data	Dickey-Fuller	p-value	Conclusion
Inflow	−2.238	0.476	Not stationary
Outflow	−3.824	0.019	Stationary

will be done based on PACF from outflow data. The PACF of differenced inflow data and PACF from outflow data can be seen at the Fig. 4.

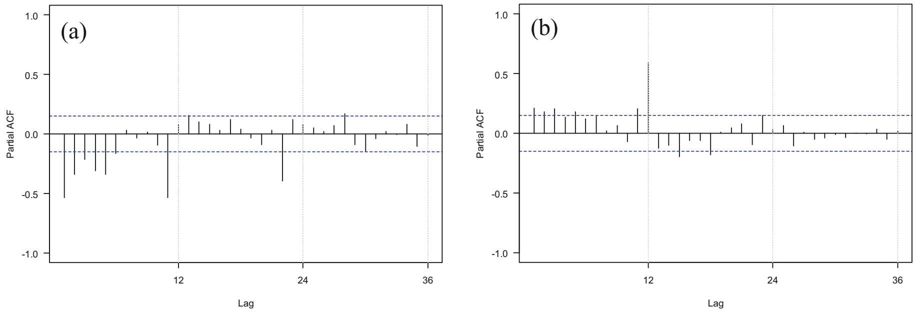

Fig. 4. PACF of stationary inflow data (a) and outflow data (b)

Figure 4 shows that data inflow and outflow in Indonesia have a seasonal pattern. In the PACF plot of differenced inflow data, lag 11 has the highest partial correlation. Thus, the high correlation of the inflow data is at lag 11 and 12, while in the outflow data, lag 12 has the highest partial correlation.

Significant lags in differenced inflow data are at lags 1, 2, 4, 5, and 11. Thus, the input combinations to be used are lag 1, 2, 3, 4, 5, 6, 11, and 12. In outflow data, the input combinations to be used are lag 1, 2, 3, and 12. Having known some input combinations, these combinations will be checked in linearity using the terasvirta test. Forecasting with neural network will be done if there is a nonlinear relationship between the input and the output. The results of the linearity testing for all possible inputs are shown at Table 2.

Table 2. Linearity test for all possible inputs for inflow and outflow models

Inflow			Outflow		
Input's lag	χ^2	p-value	Input's lag	χ^2	p-value
1	6.43	4.01×10^{-2}	12	9.91	7.03×10^{-3}
12	16.68	2.39×10^{-4}	1 and 12	56.71	6.82×10^{-10}
1 and 2	14.12	4.91×10^{-2}	2 and 12	39.72	1.43×10^{-6}
11 and 12	67.74	4.21×10^{-12}	1, 2, and 12	91.49	1.33×10^{-12}
1, 2, 11 and 12	159.77	0.00	1, 3, and 12	73.08	2.86×10^{-9}
1, 2, 3, 4, 5, 6, 11, and 12	Inf	0.00	1, 2, 3, and 12	102.04	8.84×10^{-10}

Based on Terasvirta test, all possible combinations input have a nonlinear relationship with the output. This is indicated by p-value for all combinations are less than α (α = 0.05). Under null hypotheses that the input has linear relationship with the output, this null hypothesis was rejected. Thus, all combinations of inputs will be analyzed using the neural network in order to obtain the best input with the optimal number of neurons.

In forecasting using neural networks models, preprocessing of the data also plays an important role. One of the most widely used types of preprocessing is transformation. In general, transformations can improve the accuracy of forecasting especially for short-term forecasting. The better forecasting results are generally obtained by using logarithm natural transformation [17]. Thus, this transformation will be used in order to improve the accuracy of forecasting results. The comparison of RMSE obtained for training data and testing data shown at Table 3. In order to make it easier to compare, the RMSE value in Table 3 then be illustrated as Fig. 5.

The selections of optimal input are based on the input that has the smallest RMSE on the testing data. Figure 5 shows that for inflow models, the input with smallest RMSE is combination of lag 11 and 12 with 4 number of neuron. Meanwhile for outflow data, the input with smallest RMSE is combination of lag 1, 2, and 12 with 5 number of neuron. However, Fig. 5 also shows that the increasing number of neurons does not always produce smaller RMSE, especially on testing data. Therefore, in the selection of the number of neurons, trial and error need to be done to determine the optimal number of neurons that produce the smallest RMSE. Thus, the forecasting with neural network will be done by using the transformed data using natural logarithm, and different input and neuron for inflow and outflow. The comparison of actual and forecast data for the inflow and outflow in Indonesia shown as Fig. 6.

Figure 6 shows that the forecast for training data can follow the actual data for inflow and outflow data. Moreover, for inflow data, the accuracy for testing data has smaller RMSE than training data. RMSE for training data is 10409.87 billion, while RMSE for testing data is 6560.62 billion. Whereas, forecasting results for testing data in outflow tend to be less able to capture the actual data patterns. RMSE for the training data is 9042.16 billion, while RMSE of testing data is 21566.85 billion. The value of RMSE in testing data is larger than the training data. It is in line with the previous study that found neural network models which can capture patterns well in training data can also produce forecasting less accurate in testing data [18].

3.2 Comparison of FFNN Methods with Other Classical Methods

The accuracy of FFNN in forecast the data of Inflow and Outflow in Indonesia will be compared with classical method which is widely used in forecasting, i.e. ARIMA and ARIMAX. ARIMA model that fulfilled white noise assumption for inflow data is ARIMA (2, 1, [23, 35]) $(1, 0, 0)^{12}$, while ARIMA model for outflow data is ARIMA ([2, 3, 23], 0 ,[35]) $(1, 0, 0)^{12}$. In the ARIMAX model, dummy variables used are trend, seasonal, and calendar variations effects. The comparison of RMSE for FFNN, ARIMA, and ARIMAX models can be shown as follows.

Table 4 shows that the best method for forecasting inflow data is FFNN, whereas for outflow data, the smallest RMSE is produced by ARIMAX method. This can be caused by the use of different inputs on the FFNN model and ARIMAX model. Based on Fig. 6(b), it can be seen that the FFNN model can't capture the effect of calendar variation so that the value of RMSE becomes large.

In ARIMAX model there is dummy variable that is calendar variations effects so that ARIMAX model will be able to capture the pattern of calendar variation in outflow data. However, compared to the ARIMA method, with the parsimonious inputs, the

Table 3. Comparison of RMSE for all possible inputs

Inflow				Outflow			
Input lag	Neuron	RMSE training	RMSE testing	Input lag	Neuron	RMSE training	RMSE testing
1	1	14143.91	28898.81	12	1	11703.26	28241.51
	2	22215.66	27086.60		2	11766.28	28979.40
	3	14266.64	29339.61		3	12806.41	29151.20
	4	14133.84	29068.68		4	12287.39	29129.95
	5	14563.87	26951.62		5	11567.62	28283.98
	10	15038.23	25980.72		10	11292.59	28334.56
	15	14154.05	28997.70		15	12751.28	29225.10
12	1	10871.00	14148.57	1 and 12	1	14653.29	30294.51
	2	11015.59	10919.43		2	10660.71	27789.08
	3	10516.49	11755.03		3	9598.81	25075.43
	4	11249.39	11739.50		4	10615.72	26766.65
	5	11275.52	9915.43		5	10100.57	27372.06
	10	11855.57	10518.17		10	11838.60	23831.50
	15	10732.50	11644.34		15	12612.10	26941.21
1 and 2	1	13562.32	27175.61	2 and 12	1	11526.31	26101.39
	2	13451.37	26923.92		2	11753.04	27320.93
	3	21500.65	22702.19		3	11600.49	24242.59
	4	15872.27	23861.09		4	11714.06	28441.15
	5	13404.75	26736.95		5	11495.97	27781.79
	10	13624.41	26655.08		10	11167.16	26931.99
	15	13469.46	26412.88		15	11884.34	28023.63
11 and 12	1	11160.52	9351.40	1, 2 and 12	1	12006.61	27205.56
	2	10577.98	18359.11		2	11266.77	26221.86
	3	9560.15	15585.81		3	11636.90	26548.47
	4	10409.87	**6560.62**		4	13738.92	29349.27
	5	9668.54	13062.22		5	9042.16	**21566.85**
	10	9637.45	10465.48		10	9549.50	25354.33
	15	8232.85	10779.40		15	9473.79	25289.16
1, 2, 11, and 12	1	10743.57	19329.36	1, 3 and 12	1	12810.72	27453.98
	2	11344.87	11356.93		2	13024.16	29289.44
	3	11997.43	8759.97		3	12016.07	28383.20
	4	10262.76	15117.49		4	10615.99	28469.27
	5	9173.02	10138.90		5	10295.00	26808.46
	10	9571.04	10089.72		10	11249.07	26378.39
	15	8754.30	11735.84		15	10363.29	25116.26
1, 2, 3, 4, 5, 6, 11, and 12	1	12432.70	22895.15	1, 2, 3 and 12	1	12203.34	27810.09
	2	9929.39	11359.77		2	9660.77	25247.26
	3	9891.63	8493.18		3	9240.43	22234.38
	4	8539.83	13215.84		4	10447.53	24038.35
	5	11106.36	9379.35		5	9463.79	23241.48
	10	10175.12	7727.62		10	9407.18	25092.63
	15	8462.65	8659.24		15	8264.03	22848.71

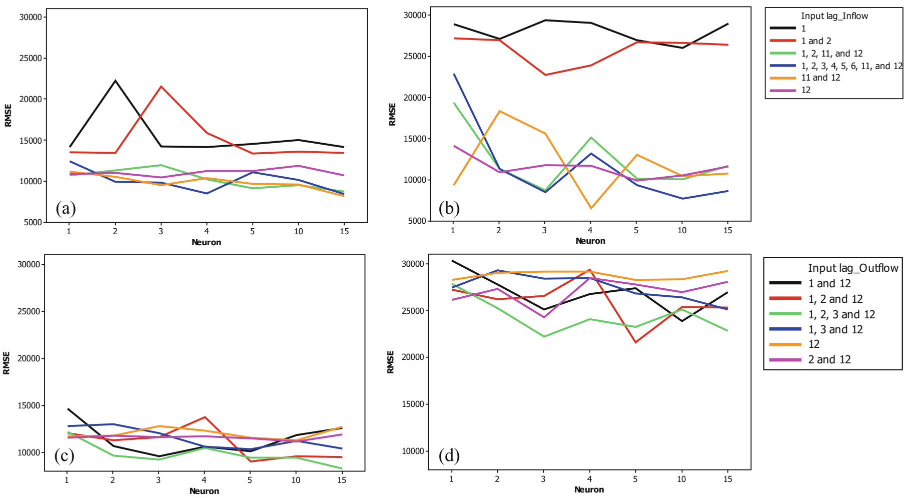

Fig. 5. RMSE comparison for selection input and neuron in training inflow (a), Testing inflow (b), Training outflow (c), and Testing outflow (d) Models

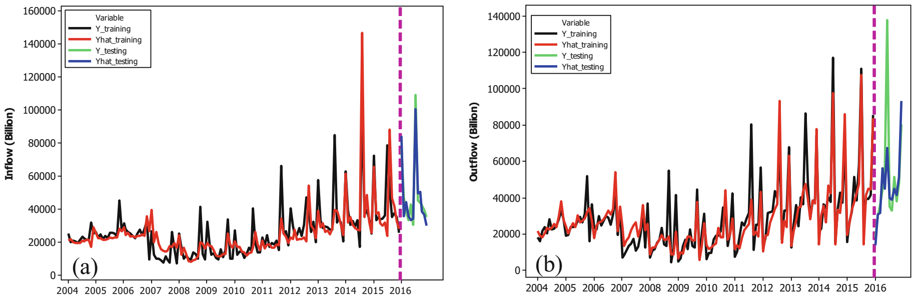

Fig. 6. Comparison of actual and forecast data for inflow (a) and outflow (b)

Table 4. Comparison of RMSE for FFNN, ARIMA, and ARIMAX models

Model	Inflow		Outflow	
	RMSE training	RMSE testing	RMSE training	RMSE testing
FFNN	10409.87	**6560.62**	9042.16	21566.85
ARIMA	6531.24	11697.04	7975.92	27212.73
ARIMAX	6551.84	15577.45	7969.11	**19381.40**

FFNN model has smaller RMSE values for testing data, both for inflow and outflow data. While for training data, RMSE of ARIMA model is lower than FFNN.

4 Discussion

Model selection in forecasting inflow and outflow in Indonesia are based on the significant lags in PACF. Thus, there are many combinations in order to get the best inputs for neural networks model. The model selection has been done by using cross validation. The forecasting results of inflow shows that it is possible that testing data have more accurate results than training data. This finding shows that neural networks were able to predict testing data as well as training data by using the suitable inputs and neuron, especially for short term forecasting. Meanwhile, by using FFNN the forecasting result of outflow shows that testing data were lower accurate than training data. Some previous studies showed that a neural network models can't capture the trend and seasonal patterns well [19, 20]. The best way to overcome these problems are to do detrend and deseasonal. So that, preprocessing by using detrend and deseasonal might prompting to improve the accuracy of forecasting.

All possible inputs that used in this study only based on the significant lag in PACF. There are many other methods to get the combination of inputs, such as by using stepwise linear regression [21], using lag from ARIMA models [18], and based on the increment of R^2 of FFNN model when added an input variable [22]. Additional research might examine the comparison between selection inputs by using those methods.

Acknowledgments. This research was supported by DRPM-DIKTI under scheme of "Penelitian Berbasis Kompetensi", project No. 532/PKS/ITS/2017. The authors thank to the General Director of DIKTI for funding and to anonymous referees for their useful suggestions.

References

1. Maier, H.R., Dandy, G.C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ. Model Softw. **15**, 101–124 (2000)
2. Kandil, N., Wamkeue, R., Saad, M., Georges, S.: An efficient approach for short term load forecasting using artificial neural networks. Electr. Power Energy Syst. **28**, 525–530 (2006)
3. Azad, H.B., Mekhilef, S., Ganapathy, V.G.: Long-term wind speed forecasting and general pattern recognition using neural networks. IEEE Trans. Sustain. Energy **5**, 546–553 (2014)
4. Claveria, O., Torra, S.: Forecasting tourism demand to Catalonia: neural networks vs. time series models. Econ. Model. **36**, 220–228 (2014)
5. Kara, Y., Boyacioglu, M.A., Baykan, O.K.: Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul stock exchange. Expert Syst. Appl. **38**, 5311–5319 (2014)
6. Guler, H., Talasli, A.: Modelling the daily currency in circulation in Turkey. Central Bank Repub. Turkey **10**, 29–46 (2010)
7. Nasiru, S., Luguterah, A., Anzagra, L.: The efficacy of ARIMAX and SARIMA models in predicting monthly currency in circulation in Ghana. Math. Theory Model. **3**, 73–81 (2013)
8. Rachmawati, N.I., Setiawan, S., Suhartono, S.: Peramalan Inflow dan dan Outflow Uang Kartal Bank Indonesia di Wilayah Jawa Tengah dengan Menggunakan Metode ARIMA, Time Series Regression, dan ARIMAX. Jurnal Sains dan Seni ITS, 2337–3520 (2015)

9. Kozinski, W., Swist, T.: Short-term currency in circulation forecasting for monetary policy purposes – the case of Poland. Financ. Internet Q. **11**, 65–75 (2015)
10. Hill, T., Marquezb, L., O'Connor, M., Remusa, W.: Artificial neural network models for forecasting and decision making. Int. J. Forecast. **10**, 5–15 (1994)
11. Crone, S.F., Kourentzes, N.: Input-variable specification for neural networks - an analysis of forecasting low and high time series frequency. In: International Joint Conference on Neural Networks, pp. 14–19 (2009)
12. Anders, U., Korn, O.: Model selection in neural networks. Neural Netw. **12**, 309–323 (1999)
13. Wei, W.W.S.: Time Series Analysis Univariate and Multivariate Methods. Pearson, New York (2006)
14. Lee, M.H., Suhartono, S., Hamzah, N.A.: Calendar variation model based on ARIMAX for forecasting sales data with Ramadhan effect. In: Regional Conference on Statistical Sciences, pp. 349–361 (2010)
15. Sarle, W.S.: Neural networks and statistical models. In: Proceedings of the Nineteenth Annual SAS Users Group International Conference (USA 1994), SAS Institute, pp. 1538–1550 (1994)
16. Apriliadara, M., Suhartono, Prastyo, D.D.: VARI-X model for currency inflow and outflow forecasting with Eid Fitr effect in Indonesia. In: AIP Conference Proceedings, vol. 1746, p. 020041 (2016)
17. Proietti, T., Lutkepohl, H.: Does the Box-Cox transformation help in forecasting macroeconomic time series? Int. J. Forecast. **29**, 88–99 (2013)
18. Faraway, J., Chatfield, C.: Time series forecasting with neural networks: a comparative study using the airline data. Appl. Stat. **47**, 231–250 (1998)
19. Zhang, G.P., Qi, M.: Neural network forecasting for seasonal and trend time series. Eur. J. Oper. Res. **160**, 501–514 (2005)
20. Suhartono, S., Subanar, S.: The effect of decomposition method as data preprocessing on neural networks model for forecasting trend and seasonal time series. Jurnal Keilmuan dan Aplikasi Teknik Industri 27–41 (2006)
21. Swanson, N.R., White, H.: Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models. Int. J. Forecast. **13**, 439–461 (1997)
22. Suhartono: New procedures for model selection in feedforward neural networks. Jurnal Ilmu Dasa, **9**, 104–113 (2008)