

Automatic time series forecasting using nonlinear autoregressive neural network model with exogenous input

Hermansah¹, Dedi Rosadi², Abdurakhman³, Herni Utami⁴

^{1,2,3,4}Department of Mathematics, Universitas Gadjah Mada, Indonesia

¹Department of Mathematics Education, Universitas Riau Kepulauan, Indonesia

Article Info

Article history:

Received Feb 9, 2021

Revised Jun 14, 2021

Accepted Jul 5, 2021

Keywords:

Automatic forecasting

Ensemble operator

NARX model

Neural network

Stepwise algorithm

Time series

ABSTRACT

This study aims to determine an automatic forecasting method of univariate time series, using the nonlinear autoregressive neural network model with exogenous input (NARX). In this automatic setting, users only need to supply the input of time series. Then, an automatic forecasting algorithm sets up the appropriate features, estimate the parameters in the model, and calculate forecasts, without the users' intervention. The algorithm method used include preprocessing, tests for trends, and the application of first differences. The time series were tested for seasonality, and seasonal differences were obtained from a successful analysis. These series were also linearly scaled to $[-1, +1]$. The autoregressive lags and hidden neurons were further selected through the stepwise and optimization algorithms, respectively. The 20 NARX models were fitted with different random starting weights, and the forecasts were combined using the ensemble operator, in order to obtain the final product. This proposed method was applied to real data, and its performance was compared with several available automatic models in the literature. The forecasting accuracy was also measured by mean squared error (MSE) and mean absolute percent error (MAPE), and the results showed that the proposed method outperformed the other automatic models.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Dedi Rosadi

Department of Mathematics

Universitas Gadjah Mada

Sekip Utara Bulaksumur, Mlati District, Sleman Regency, Yogyakarta Special Region, Indonesia

Email: dedirosadi@gadjahmada.edu

1. INTRODUCTION

Time series data forecasting is often referred to as a periodical forecasting, which involves the study of existing historical patterns and estimation of future values. This concept is classified into linear and nonlinear methods, with the popular univariate types being exponential smoothing and autoregressive integrated moving average (ARIMA) models. These methods have been reported to be successful in forecasting linear time-series data, and very poor based on designing nonlinear and complex parameters [1]. Meanwhile, the nonlinear forecasting generally provides irregular function specification requirements. Furthermore, an artificial neural network (ANN) is being introduced as a universal approach for this method, as several studies have confirmed its excellent performance in long-term forecasting, based on monthly, and quarterly time series of nonlinear data [2]-[5].

Two ANN architectural models are reported to have been widely applied to several time series forecasts, including the time-lagged feed-forward and dynamically-driven recurrent network methods. Both are

found to use a time-lagged and feed-forward network architecture, with the second model also using a feedback approach [6], [7]. Based on several applications, obtaining ANN models with accurate time series forecasting performance requires intervention to set up several optimal setting parameters, including the number of lag inputs and those associated with neurons in the hidden layer. However, this situation is not practical for some applications, such as obtaining (near) real-time forecast. Based on other cases, it is possible that no one understands how to use the ANN model to forecast, as one possible solution to this issue is the automatic forecasting [8]-[10]. This only requires the supply of input, as an automatic forecasting algorithm automatically sets up the appropriate data, estimates the parameters in the model, and calculates forecasts without intervention.

Based on multilayer perceptron parallel architecture without feedback (NAR model), the automatic ANN time series forecasting was first discussed in [8], and the result was further extended in [11]-[13], by implementing ensemble/combination operators. The studies in [8], [11]-[13] used the NAR model with sigmoid and linear activation functions at the hidden and output layers, with the backpropagation learning algorithm applied to update the parameters. Specifically, the non-automatic version of the forecasting algorithm based on the NAR model was observed in [1]. Forecasting time series with NAR is also found to be possible, by using multiple univariate models. The univariate aspect applies the past data from those predicted as the input, as external factors with possible effects are not allowed in the model. However, these are accommodated in the multiple univariate models, in order to ensure improved accuracy. Furthermore, the nonlinear autoregressive neural network model with exogenous input (NARX) is used as multiple univariate methods, as it is also considered a variant of NAR that utilizes external/exogenous inputs in the learning process. It generally has a more accurate forecasting capability compared to NAR, when the utilized external inputs have a strong relationship with the predicted data [14]-[16].

According to this research, an extension of the previous results in [8], [11]-[13] is being proposed on two folds. Firstly, an automatic forecasting algorithm is considered for a more general class of model, i.e., the NARX with parallel architecture without feedback. Secondly, the ensemble operators (mean, median, and mode are usable) with both logistic and tangent hyperbolic functions are also considered to activate the hidden layer of the NARX model. The descriptions and implementation details of each ensemble operator is clearly observed in [11]-[13], [17]-[20]. Also, several learning algorithms are being considered to update the parameters, namely the backprop (backpropagation), rprop+ and rprop- (resilient backpropagation with and without weight backtracking), as well as the grprop sag, and grprop slr (globally resilient backpropagation without weight backtracking and smallest absolute gradient or learning rate). The descriptions and implementation details for each learning algorithm are further shown in [21]-[25]. Based on the empirical study, the proposed automatic method is applied to forecast two real data, i.e., the Indonesian inflation and exchange rates between the Rupiah and US Dollar. Furthermore, the performance of this proposed model is compared with several available automatic methods in the literature, namely exponential smoothing (see discussion in [26]-[29]), ARIMA (see discussion in [30]-[33]), and NAR parallel architecture without feedback. The forecasting accuracy is also being measured by mean squared error (MSE) and mean absolute percent error (MAPE). This research is organized such that section 1 and 2 introduces the background of the study, as well as provide some necessary concepts while introducing the automatic NARX modeling for time series forecasting, respectively. Also, section 3 and 4 discusses the empirical results and conclusions of the study, respectively.

2. RESEARCH METHOD

2.1. NARX Model

The NARX model with exogenous input was reported to be very essential to the discrete-time nonlinear systems, and defined using the following mathematical relationship [34],

$$y(t+1) = f(y(t), y(t-1), \dots, y(t-n_y+1); u(t), u(t-1), \dots, u(t-n_u+1); w) \quad (1)$$

where $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$ indicates the input and output of the model at time t , $n_u \geq 1$ and $n_y \geq 1$ ($n_y \geq n_u$) represents the input and output-memory orders, w is the weights matrix, and f is the nonlinear function expected to be estimated through multilayer perceptron [35].

The NARX network was basically trained under one of two models [36]. The first model was the series-parallel architecture (or parallel architecture without feedback), where formation of the regressors was only obtained through the use of the output actual values.

$$\hat{y}(t+1) = \hat{f}(y(t), y(t-1), \dots, y(t-n_y+1); u(t), u(t-1), \dots, u(t-n_u+1); w) \quad (2)$$

The second model was the parallel architecture (or parallel architecture with feedback), where the output was the feedback for the feed-forward neural input, which was part of the standard architecture,

$$\hat{y}(t+1) = \hat{f}(\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t-n_y+1); u(t), u(t-1), \dots, u(t-n_u+1); w) \quad (3)$$

Based on being previously stated, the NARX network include the regressors of the system (inputs and outputs), with the time-series used as the output channels without measured input. Therefore, the forecasting ability of the model is limited in its application for time-series data without input regressors, due to the elimination of the tapped-delay line over the signal. This further leads to the reduction of NARX to plain time-delay neural network architecture [37]-[39], as shown in,

$$\hat{y}(t+1) = f(y(t), y(t-1), \dots, y(t-n_y+1); w) \quad (4)$$

According to [20], a simple strategy was proposed in line with the embedding theorem of Takens. This served as a solution to the problem, by providing the opportunity for the full exploitation of the actual NARX network computational abilities, towards predicting nonlinear time-series. Furthermore, the input signal regressors ($u(t)$) were initially defined through the delay-embedding coordinates, as shown in,

$$u(t) = (y(t), y(t-\tau), \dots, y(t-(d_E-1)\tau)) \quad (5)$$

where $d_E = n_u$ and τ are the embedding dimension and delay, respectively.

Secondly, the output signal regressors ($y(t)$) are presented as shown in the following relationships, due to the possibility of training the NARX network in two different architectures,

$$y_{sp}(t) = (y(t), y(t-1), \dots, y(t-n_y+1)) \quad (6)$$

$$y_p(t) = (\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t-n_y+1)) \quad (7)$$

where the output regressor ($y(t)$) for the parallel architecture without and with feedback in (6) have previous n_y actual and estimated time values, respectively.

These outputs were values of $y(t+1)$, which were previously estimated for a network that had been effectively trained. They were also required to follow the prognostic relationships applied, by using the NARX network. This is represented as being as [15], [16],

$$\hat{y}(t+1) = f(y_{sp}(t); u(t); w) \quad (8)$$

$$\hat{y}(t+1) = f(y_p(t); u(t); w) \quad (9)$$

Therefore, the NARX networks trained in line with (8) and (9) were represented as NARX-SP and NARX-P, respectively. However, this research focused on the model that had parallel architecture and without feedback (NARX-SP).

2.2. Automatic NARX modeling

Based on the NARX model, this section describes the automatic forecasting procedure, with the assumptions that $y(t)$ and $u(t)$ are the main series and external/exogenous variables to be predicted, respectively. Furthermore, several key steps contained in the algorithm were explained as being as,

- Step 1: Preprocessing of the series ($y(t)$) started with trend, seasonality check, and seasonal difference application. The Cox-Stuart test was used to determine the trend in a time series data, based on a 12-period centred moving average. This test was carried out in order to smoothen effects, due to irregularities. Furthermore, the de-trended time series data were used to calculate and identify seasonal indices, which were analyzed to determine their significant deviations, through the use of a Friedman test. These data were also linearly scaled between -1 and $+1$, in order to facilitate the NARX training.
- Step 2: The following strategy was used to obtain the autoregressive lags for the $y(t)$ and $u(t)$ series. When the frequency of $y(t)$ is equal to m , all products from 1 to m were considered possible lag numbers. For example, 1-4 and 1-12 were considered as quarterly and monthly data, respectively. Furthermore, the order for $u(t)$ was assumed equal to $y(t)$, as the significant lags were selected using the stepwise algorithm. The model also included seasonal dummy variables, due to the identification of periodical patterns. Creating additional features in form of dummy variables was also one of the methods used to

capture deterministic seasonal components. The conventional approach used to model these periodical patterns were $s - 1$ binary dummy variables with t , which denoted the seasonal length. Meanwhile, long input vectors used additional $s - 1$ for high frequency data ($s > 12$). Two inputs including $x_{s,1}$ and $x_{s,2}$ were also used to encode seasonality with variables created, by using $\sin(t)$ and $\cos(t)$ for an explicit representation of the point within an identified seasonal length (s). Furthermore, $x_{s,1}$ and $x_{s,2}$ were used to determine the sine-cosine-pairs for each s , as well as the input vector for long and multiple seasonalities.

- Step 3: There was also procedural division into training and testing parts, where 80% and 20% of the data processes were used respectively. The mean squared error (MSE) of training data was checked, when the value of neurons was set at 1 towards the maximum lag observed (as the total input in step 2 plus 2). Furthermore, the optimal number of neurons in the hidden layer was defined to be the value providing the minimum MSE. The number of neurons in this layer was also experimentally identified for each time series. Also, the maximum lag (as the total input in step 2 plus 2) neurons in the hidden layer were evaluated for each time series, as the values that minimized the MSE validation were selected.
- Step 4: The NARX model obtained in step 3 were fitted 20 times using different random starting weights, and the forecasts obtained were combined using the ensemble operator approach (mean, median, and mode are usable), in order to produce the final product. Based on avoiding local minima and providing an adequate error distribution using sufficient results, each NARX candidate was initialized 20 times with random starting weights in the interval of $[-1, +1]$.
- Step 5: The recursive or iterative strategy was used for multi-step ahead forecasts.

Further details that also showed the applicability of the above algorithm for the automatic NAR model, were provided in [8], [11]-[13]. Meanwhile, the non-automatic version of the forecasting algorithm based on the model was observed in [1].

3. RESULT AND DISCUSSION

This research was conducted using two real cases, with the first being the inflation rate data in Indonesia, with the external/exogenous variable being the interest of the Indonesian Central Bank. The second case was the exchange rate data for the Indonesian rupiah against the US dollar, with the external variable being the composite stock price index.

3.1. Inflation rate data

The Indonesian monthly inflation rate data from January 2007 to February 2018 that contained 134 observations, were used, with the initial 129 and final 5 applied for training, and testing, respectively. Based on simple terms, inflation is understood as a persistent, and continuous rise across a broad spectrum of prices. The investigation on forecasting inflation in a specific country had received significant attention from several macroeconomics experts. Based on most central banks, one of the monetary policy objectives was inflation. Monetary policy also needs to consider future inflation, due to the occurrence of typical time lags. Furthermore, the current inflation levels that were the result of past policies, should provide only incomplete information. Therefore, the forecasts that linked future inflation to current developments were found to bridge this gap. This study attempts to develop an inflation forecasting model for Indonesia, which serves as an input for policy setting in Bank Indonesia (BI). Based on evaluating the accuracy of the models, two forecast error measurements were used, namely the mean squared error (MSE) and mean absolute percent error (MAPE). The MSE and MAPE are further defined as being as,

$$MSE = \sum_{t=1}^N \frac{(A_t - F_t)^2}{N} = \sum_{t=1}^N \frac{e_t^2}{N}; MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{A_t - F_t}{A_t} \right|$$

where A_t and F_t are actual and forecast values at data time t , e_t is the error at data time t , and N is the number of data.

Thirty models were considered based on the combination of ensemble operators, activation functions, and algorithm types, in order to calculate weight networks applied in the automatic algorithms. The ensemble operators were mean, median, and mode, while the two activation functions compared were logistic and hyperbolic-tangent. Furthermore, the five utilized algorithm types included backpropagation (backprop),

resilient backpropagation with weight backtracking (rprop+), resilient backpropagation without weight backtracking (rprop-), globally resilient backpropagation without weight backtracking and smallest absolute gradient (grprop sag), and globally resilient backpropagation without weight backtracking and smallest learning rate and (grprop slr). The summary of the forecasting performance was shown in Table 1, and the automatic NARX model that combined the median ensemble operator, hyperbolic-tangent activation function, and rprop+ was observed to have produced the smallest MSE and MAPE values.

Table 1. Summary of the forecasting performance of the automatic NARX model for the inflation rate data and the exchange rate data

2*Prediction method			Inflation rate data		Exchange rate data	
			MSE	MAPE	MSE	MAPE
10*mean	5*logistic	backprop	0.757364	0.086202	73345.44	0.016121
		rprop+	0.364176	0.069689	72433.13	0.016063
		rprop-	0.319551	0.066502	72056.55	0.016041
		grprop sag	0.358235	0.069661	73387.39	0.016169
		grprop slr	0.349365	0.069136	71971.57	0.015998
		backprop	0.539312	0.075055	71649.17	0.016045
	5*tanh	rprop+	0.285420	0.064619	67087.11	0.015678
		rprop-	0.288407	0.063821	68718.11	0.015846
		grprop sag	0.283124	0.066306	67233.85	0.015736
		grprop slr	0.283034	0.065415	69153.00	0.015828
		backprop	0.738433	0.085848	72653.74	0.016012
		rprop+	0.296109	0.065271	72430.28	0.016063
10*median	5*logistic	rprop-	0.316643	0.066174	72630.40	0.016087
		grprop sag	0.321834	0.065195	69711.36	0.016011
		grprop slr	0.312252	0.067723	72464.98	0.016035
		backprop	0.442624	0.074001	72220.73	0.016102
		rprop+	0.271799	0.062696	70385.47	0.015981
		rprop-	0.282693	0.064414	70839.00	0.016042
	5*tanh	grprop sag	0.283691	0.066704	67937.07	0.015895
		grprop slr	0.286132	0.063374	70899.88	0.015952
		backprop	0.723456	0.083285	72507.50	0.015999
		rprop+	0.332833	0.066525	72172.29	0.016060
		rprop-	0.312909	0.067373	73040.55	0.016088
		grprop sag	0.333010	0.071677	71941.49	0.015779
10*mode	5*logistic	grprop slr	0.352009	0.069826	72408.98	0.016186
		backprop	0.445159	0.072309	72403.39	0.016105
		rprop+	0.312847	0.065484	71571.02	0.016033
		rprop-	0.310376	0.067916	71763.39	0.015974
		grprop sag	0.324551	0.074489	70864.37	0.016323
		grprop slr	0.484687	0.076009	71721.46	0.016033

There was also a performance comparison of the proposed methods with the other forecasting models, including automatic exponential smoothing (see discussion in [26]-[29]), automatic ARIMA (see discussion in [30]-[33]), and automatic NAR parallel architecture without feedback (see discussion in [8], [11]-[13]), respectively. Furthermore, the plot of in-sample fitting and out-sample forecasts for inflation rate data were shown in Figure 1, by using considered automatic algorithms. It was also observed that all the considered methods relatively performed accurately for modeling the data, as differences were hardly detected. However, the numerical summary presented in Table 2 showed that the proposed automatic NARX method outperformed other available methods, both in the training and testing data.

Table 2. The performance of four automatic methods for the inflation rate data

2*Automatic prediction method	Training		Testing	
	MSE	MAPE	MSE	MAPE
Exponential Smoothing	1.738208	0.105420	1.409523	0.363102
ARIMA	1.579745	0.088359	1.469312	0.370713
NAR	0.860586	0.084574	0.341019	0.150416
NARX	0.271799	0.062696	0.287479	0.134564

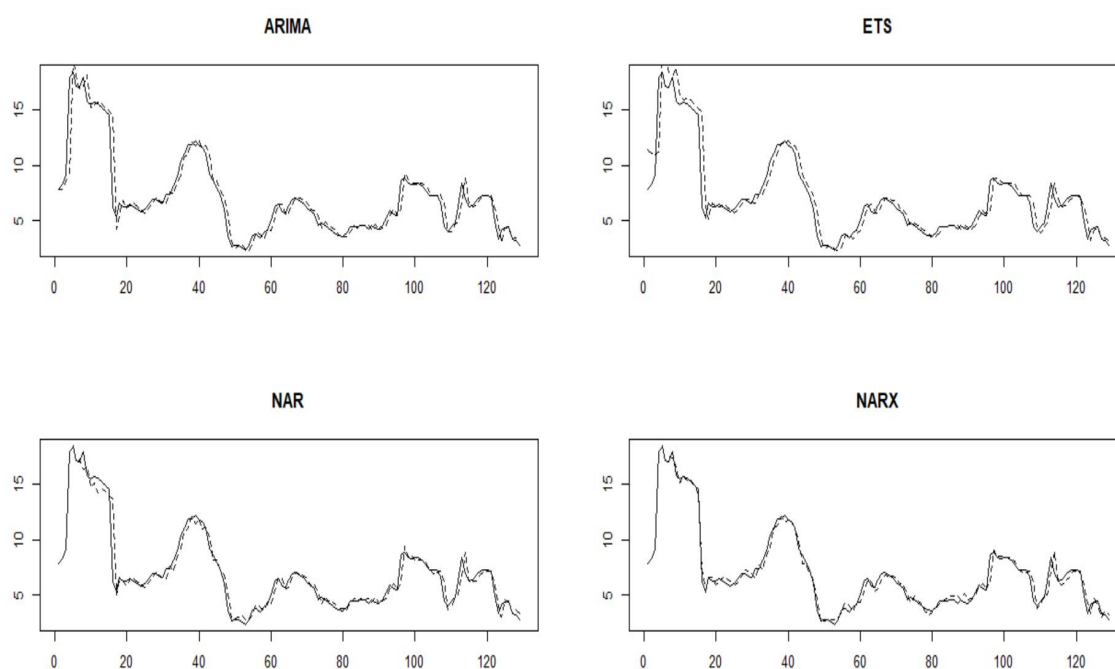


Figure 1. Plot of real data, in-sample fitting, and out-sample forecast using several automatic algorithms

3.2. Exchange rate data

There were 198 monthly observations for the exchange rate data from January 2003 to June 2019, with the initial 188 and final 10 applied for training and testing, respectively. Based on finance, an exchange rate is the value at which one currency is transformed for another. It is also regarded as the value of one country's currency with another. Also, the exchange rate describes the price of one currency in terms of another. This price is found to be essential for the government or company, especially when the business extends over different countries or firms. Furthermore, the exchange rate forecasting is an essential input for the decision-making management of exposure or hedging strategies.

This structure was similar as observed in the first cases. The summary of the forecasting performance was shown in Table 1, as the automatic NARX model with mean ensemble operator, hyperbolic-tangent activation function, and rprop+ produced the smallest MSE and MAPE values. This method was further compared with other automatic models, and the numerical performance was summarized in Table 3. Furthermore, the plots of in-sample fitting and out-sample forecasts that used all considered automatic algorithms were omitted. However, they showed that all methods were reasonably performed. Based on the results in Table 3, the automatic NARX method was superior to other models considered in the study.

Table 3. The performance of four automatic methods for the exchange rate data

2*Prediction method	Training		Testing	
	MSE	MAPE	MSE	MAPE
Exponential Smoothing	88053.32	0.017740	57775.10	0.021914
ARIMA	88051.71	0.017739	57773.90	0.021914
NAR	83522.71	0.017040	48219.30	0.021845
NARX	67087.11	0.015678	28558.70	0.016348

4. CONCLUSION

This study already proposed an automatic forecasting method of univariate time series, by using the nonlinear autoregressive neural network model with exogenous input (NARX). The automatic algorithm only allowed the supply of input data, as a forecasting algorithm automatically sets up the appropriate data, estimated the parameters in the model, and calculated forecasts without intervention. Furthermore, the empirical studies conducted showed that the automatic NARX models outperformed the other available methods in the literature,

by using two monthly series data. However, further research needs to be conducted based on checking and improving the effectiveness of the method, in order to forecast different types of time series data.

ACKNOWLEDGEMENT

The authors appreciate the UGM RTA 2019 and PDD 2020 Program from Deputy of Research and Development, Ministry of Research, and Technology/National Agency for Research and Innovation of Republic of Indonesia program for providing financial support for this research. The authors would like to thank anonymous referees for their constructive comments and valuable suggestions to improve the quality of this manuscript.

REFERENCES

- [1] Hermansah, D. Rosadi, A. Abdurakhman, and H. Utami, "Selection of input variables of nonlinear autoregressive neural network model for time series data forecasting," *Media Statistika*, vol. 13, no. 2, pp. 116-124, December 2020, doi: 10.14710/medstat.13.2.116-124.
- [2] S. F. Crone, M. Hibon, and K. Nikolopoulos, "Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction," *International Journal of Forecasting*, vol. 27, no. 3, pp. 635-660, July–September 2011, doi: 10.1016/j.ijforecast.2011.04.001.
- [3] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218-230, January 2016, doi: 10.1016/j.apenergy.2015.09.087.
- [4] D. Barrow and N. Kourentzes, "The impact of special days in call arrivals forecasting: A neural network approach to modelling special days," *European Journal of Operational Research*, vol. 264, no. 3, pp. 967-977, February 2018, doi: 10.1016/j.ejor.2016.07.015.
- [5] G. P. Herrera, M. Constantino, B. M. Tabak, H. Pistori, Jen-Je Su, and A. Naranpanawa, "Long-term forecast of energy commodities price using machine learning," *Energy*, vol. 179, pp. 214-221, July 2019, doi: 10.1016/j.energy.2019.04.077.
- [6] A. H. Ribeiro and L. A. Aguirre, "Parallel training considered harmful? Comparing series-parallel and parallel feedforward network training," *Neurocomputing*, vol. 316, pp. 222-231, November 2018, doi: 10.1016/j.neucom.2018.07.071.
- [7] S. M. Cabaneros, J. K. Calautit, and B. R. Hughes, "A review of artificial neural network models for ambient air pollution prediction," *Environmental Modelling and Software*, vol. 119, pp. 285-304, September 2019, doi: 10.1016/j.envsoft.2019.06.014.
- [8] S. F. Crone and N. Kourentzes, "Feature selection for time series prediction-A combined filter and wrapper approach for neural networks," *Neurocomputing*, vol. 73, no. 10-12, pp. 1923-1936, June 2010, doi: 10.1016/j.neucom.2010.01.017.
- [9] S. Maldonado, A. González, and S. Crone, "Automatic time series analysis for electric load forecasting via support vector regression," *Applied Soft Computing*, vol. 83, p. 105616, October 2019, doi: 10.1016/j.asoc.2019.105616.
- [10] D. Barrow, N. Kourentzes, R. Sandberg, and J. Niklewski, "Automatic robust estimation for exponential smoothing: Perspectives from statistics and machine learning," *Expert Systems with Applications*, vol. 160, pp. 113637, December 2020, doi: 10.1016/j.eswa.2020.113637.
- [11] N. Kourentzes, D. K. Barrow, and S. F. Crone, "Neural network ensemble operators for time series forecasting," *Expert Systems with Applications*, vol. 41, no. 9, pp. 4235-4244, July 2014, doi: 10.1016/j.eswa.2013.12.011.
- [12] D. K. Barrow and S. F. Crone, "Cross-validation aggregation for combining autoregressive neural network forecasts," *International Journal of Forecasting*, vol. 32, no. 4, pp. 1120-1137, October–December 2016, doi: 10.1016/j.ijforecast.2015.12.011.
- [13] D. K. Barrow and S. F. Crone, "A comparison of AdaBoost algorithms for time series forecast combination," *International Journal of Forecasting*, vol. 32, no. 4, pp. 1103-1119, October–December 2016, doi: 10.1016/j.ijforecast.2016.01.006.
- [14] A. Sihabuddin, Subanar, D. Rosadi, and E. Winarko, "A second correlation method for multivariate exchange rates forecasting," *International Journal of Advanced Computer Science and Applications*, vol. 5, no. 7, pp. 30-33, 2014.
- [15] A. Sihabuddin, Subanar, D. Rosadi, and E. Winarko, "Exchange rates forecasting using variable length moving average-NARX," *International Journal of Computer Science and Information Security*, vol. 13, no. 9, pp. 15-18, September 2015.
- [16] Y. Amellas, O. El Bakkali, A. Djebli, and A. Echhelh, "Short-term wind speed prediction based on MLP and NARX networks models," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 1, pp. 150-157, April 2020, doi: 10.11591/ijeecs.v18.i1.pp150-157.

- [17] F. Petropoulos and N. Kourentzes, "Forecast combinations for intermittent demand," *Journal of the Operational Research Society*, vol. 66, no. 6, pp. 914-924, 2015, doi: 10.1057/jors.2014.62.
- [18] D. Barrow and N. Kourentzes, "Distributions of forecasting errors of forecast combinations: Implications for inventory management," *International Journal of Production Economics*, vol. 177, pp. 24-33, July 2016, doi: 10.1016/j.ijpe.2016.03.017.
- [19] N. Kourentzes, D. Barrow, and F. Petropoulos, "Another look at forecast selection and combination: Evidence from forecast pooling," *International Journal of Production Economics*, vol. 209, pp. 226-235, March 2019, doi: 10.1016/j.ijpe.2018.05.019.
- [20] Hermansah, D. Rosadi, Abdurakhman, and H. Utami, "Time series forecasting with trend and seasonal patterns using NARX network ensembles," *Mathematics and Statistics*, vol. 9, no. 4, pp. 511-520, 2021, doi: 10.13189/ms.2021.090410.
- [21] A. D. Anastasiadis, G. D. Magoulas, and M. N. Vrahatisb, "New globally convergent training scheme based on the resilient propagation algorithm," *Neurocomputing*, vol. 64, pp. 253-270, March 2005, doi: 10.1016/j.neucom.2004.11.016.
- [22] F. Gunther and S. Fritsch, "Neuralnet: Training of neural networks," *The R Journal*, vol. 2, no. 1, 30-38, 2010.
- [23] J. H. Lee, T. Delbruck, and M. Pfeiffer, "Training deep spiking neural networks using backpropagation," *Frontiers in Neuroscience*, vol. 10, pp. 508-520, November 2016, doi: 10.3389/fnins.2016.00508.
- [24] A. Poole and A. Kotsialos, "Traffic flow model validation using METANET, ADOL-C and RPROP," *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 291-296, 2016, doi: 10.1016/j.ifacol.2016.07.049.
- [25] R. P. S. Hermanto and A. Nugroho, "Waiting-time estimation in bank customer queues using RPROP neural networks," *Procedia Computer Science*, vol. 135, pp. 35-42, 2018, doi: 10.1016/j.procs.2018.08.147.
- [26] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods," *International Journal of Forecasting*, vol. 18, no. 3, pp. 439-454, July–September 2002, doi: 10.1016/S0169-2070(01)00110-8.
- [27] R. J. Hyndman, M. Akram, and B. C. Archibald, "The admissible parameter space for exponential smoothing models," *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 2, pp. 407-426, 2008, doi: 10.1007/s10463-006-0109-x.
- [28] M. Akram, I. Bhatti, M. Ashfaq, and A. A Khan, "New approach to forecasting agro-based statistical models," *Journal of Statistical Theory and Applications*, vol. 15, no. 4, pp. 387-399, 2016, doi:10.2991/Jsta.2016.15.4.6.
- [29] Y. Shi, S. Tang, and J. Li, "A two-population extension of the exponential smoothing state space model with a smoothing penalisation scheme," *Risks*, vol. 8, no. 3, p. 67, June 2020, doi: 10.3390/risks8030067.
- [30] X. Wang, K. Smith, and R. Hyndman, "Characteristic-based clustering for time series data", *Data Mining and Knowledge Discovery*, vol. 13, no. 3, pp. 335-364, 2006, doi: 10.1007/s10618-005-0039-x.
- [31] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for R," *Journal of Statistical Software*, vol. 27, no. 3, pp. 1-22, 2008, doi: 10.18637/jss.v027.i03.
- [32] R. Sukmana and H. Kuswanto, "Assessment on the islamic banking market share projection by bank indonesia and proposed methods," *Journal of Islamic Monetary Economics and Finance*, vol. 1, no. 1, pp. 107-133, August 2015, doi: 10.21098/jimf.v1i1.485.
- [33] J. Arlt and P. Trcka, "Automatic SARIMA modeling and forecast accuracy," *Communications in Statistics-Simulation and Computation*, vol. 48, pp. 1-22, June 2019, doi: 10.1080/03610918.2019.1618471.
- [34] A. E. Filali and M. Zazi, "A nonlinear TSNN based model of a Lead acid battery," *Bulletin of Electrical Engineering and Informatics*, vol. 7, no. 2, pp. 169-175, June 2018, doi: 10.11591/eei.v7i2.675.
- [35] Y. Boujoudar, H. Elmoussaoui, and T. Lamhamdi, "Lithium-ion batteries modeling and state of charge estimation using artificial neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, 3415-3422, October 2019, doi: 10.11591/ijece.v9i5.pp3415-3422.
- [36] A. Y. Hatata and M. Eladawy, "Prediction of the true harmonic current contribution of nonlinear loads using NARX neural network," *Alexandria Engineering Journal*, vol. 57, no. 3, pp. 1509-1518, September 2018, doi: 10.1016/j.aej.2017.03.050.
- [37] Z. Yusof, N. A. Wahab, S. Ibrahim, S. Sahlan, and M. C. Razali, "Modeling of submerged membrane filtration processes using recurrent artificial neural networks," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 1, pp. 155-163, March 2020, doi: 10.11591/ijai.v9.i1.pp155-163.
- [38] S. M. Guzman, J. O. Paz, and M. L. M. Tagert, "The use of NARX neural networks to forecast daily groundwater levels," *Water Resources Management*, vol. 31, no. 5, pp. 1591-1603, 2017, doi: 10.1007/s11269-017-1598-5.
- [39] W. M. Z. W. Yahaya, F. H. K. Zaman, and M. F. A. Latip, "Prediction of energy consumption using recurrent neural networks (RNN) and nonlinear autoregressive neural network with external input (NARX)," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 3, pp. 1215-1223, March 2020, doi: 10.11591/ijeecs.v17.i3.pp1215-1223.

BIOGRAPHIES OF AUTHORS



Hermansah is a lecturer and researcher at Universitas Riau Kepulauan, Batam, Indonesia. He obtained a bachelor's degree in Mathematics Study Program from the State Islamic University of Sunan Kalijaga in 2010 and his master's degree in Mathematics Study Program from Universitas Gadjah Mada in 2012. He was a doctoral student in Mathematics Study Program from Universitas Gadjah Mada in 2017. His researches are in the fields of Computational Statistics, Financial Statistics, and Time Series Forecasting.



Dedi Rosadi is a lecturer and researcher at Universitas Gadjah Mada, Yogyakarta, Indonesia. He obtained a bachelor's degree in Statistics Study Program from Universitas Gadjah Mada. He completed his master's degree in Applied Probability from Faculty of Applied Mathematics, Department of Stochastics and Operation Research (STOR), University of Twente, Netherlands, and his Ph.D. degree in Econometrics and Time Series from the Institute of Mathematical Methods in Economics, Department of Econometrics (E119), Vienna University of Technology (TU Wien), Austria. He is active in conducting research, teaching, and scientific publications in several fields of Statistics, such as Econometrics, Time Series Analysis, Statistics Computing, and its applications. He currently works as the full professor at the research group Computational Statistics, Department of Mathematics, Universitas Gadjah Mada.



Abdurakhman is a lecturer and researcher at Universitas Gadjah Mada, Yogyakarta, Indonesia. He obtained a bachelor's degree in Statistics Study Program from Universitas Gadjah Mada. He completed his master's degree and his Ph.D. degree at Mathematics Study Program from Universitas Gadjah Mada. His researches are in fields of Mathematical Sciences, Applied Mathematics, Applied Statistics, Financial and Actuarial Mathematics Statistics, and Financial Mathematics. He was a trainee in training and course at READI in 2018 about modeling and loss modeling (risk, economic sustainability, and actuarial science development in Indonesia).



Herni Utami is a lecturer and researcher at Universitas Gadjah Mada, Yogyakarta, Indonesia. She obtained a bachelor's degree in Statistics Study Program from Universitas Gadjah Mada. She completed her master's degree and her Ph.D. degree at Mathematics Study Program from Universitas Gadjah Mada. Her researches are in the fields of Mathematical Sciences, Applied Statistics, Stochastic Analysis and Modelling, Probability Theory, Time Series, and Forecasting.