Contemporary Engineering Sciences, Vol. 8, 2015, no. 32, 1535 - 1550
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2015.510283

Fractional Integrated Recurrent Neural Network (FIRNN) for Forecasting of Time Series Data in Electricity Load in Java-Bali

Walid[1], Subanar[2], Dedi Rosadi[3] and Suhartono[4]

[1]Department of Mathematics, State University of Semarang, Indonesia

[2]Department of Mathematics, University of Gadjah Mada, Indonesia

[3]Department of Mathematics, University of Gadjah Mada, Indonesia

[4]Department of Statistics, State Institute of Technology, Indonesia

Abstract

The Increasing demand for electricity causes problems for State Electricity Company (SEC), which is called PLN in Indonesia in providing services to the public. Neural network (NN) is one of the methods that is often used in forecasting electricity load in different countries. Another form of neural network which is widely used for the analysis of issues that have a repeating pattern is the model of Recurrent Neural Network (RNN). Particularly, the problem in this paper is how to how to develop a model of Fractional Integrated Recurrent Neural Networks (FIRNN) in forecasting time series data on National Electricity Load and how are the forecasting results of time series data on national electricity load using Fractional Integrated Recurrent Neural Networks (FIRNN). Furthermore, RNN models in long memory nonlinear models in this study will be called Fractional Integrated Recurrent Neural Networks (FIRNN). This research was conducted with literature studies, simulations and applications to real cases in memory of long time series data, by taking the case of the burden of the use of electricity in Indonesia. The previous studies show that most of the time series on consumption patterns of electrical load in Semarang city shows the pattern of long memory, because it has a fractional difference parameter which can be seen in [26].

This study is aimed at assessing and developing a model of Integrated Fractional Recurrent Neural Networks (FIRNN). This study is a form of renewal of other studies, considering study long memory models using Neural Network has not been done by other researchers. RNN models used in this study is a model of Elman recurrent neural network (Elman-RNN). The results show that the forecasting by using FIRNN is much better in comparison with ARIMA models. It can be seen on the value of Mean Absolute Percentage Error (MAPE) and the Root of Mean Square Error (RMSE) that the results of prediction accuracy using FIRNN is better than forecast results using ARIMA.

Keywords: Double Seasonal, FIRNN, Long Memory

1. Introduction

Electricity demand in Indonesia continues to increase every year, along with the increase of population and human activities. The increase in demand for electricity stirs problems for the State Electricity Company (SEC) in providing services to the public. There is dissatisfaction experienced by Indonesian consumers in using SEC service.

Analysis of time series has several objectives, namely forecasting, modeling, and control [5]. Forecasting issues are related to the establishment of models and methods that can be used to generate an accurate forecast. Neural Networks (NN) is a nonparametric models that has a flexible functional form, contains several parameters that cannot be interpreted as in the parametric model. In its application, the NN contains a number of limited parameters (weight). Recent developments related to the application of NN for time series forecasting indicate that several studies related to the development of recurrent neural networks (RNN) has reviewed the problems that contains case seasonal. RNN is an artificial neural network architecture (NN-artificial) which is inspired by the connectivity of neurons in brain function cycle repeated use to store lap information [9].

In a further development, the NN models have been widely used for prediction or forecasting of time series data is real [20],[22],[19], [6], [16], [17], [29], and [15]. In addition, the application of NN models for time series analysis, especially in the field of time series econometrics, also encourages the development of several tests to test nonlinearity [28] and [14]. RNN has some good properties in sequence labeling.

RNN is flexible in terms of its usage in of information context, and is able to receive many types of information and different data representations, to recognize the shape of a sequential pattern in data presentation, and sequential patterns with a sequential distortion. However, RNN has some disadvantages. It has limited application in the problem of labeling sequence in real world. The serious drawback of RNN is its difficulty to store information for long periods of time [11]. This drawback limits RNN to access important labelling sequence.

During the recent decades, long short-term memory (LSTM) has been proven capable in various synthesis tasks requiring long-term memory (long memory), including learning context-free language, given the amount of real high-precision on the noise line up and a variety of tasks that require time and exact calculation. LSTM is the redesigned version of the special cell memory unit architecture [12]. In the synthesis of various tasks, LSTM has been proven capable of storing and accessing information in a very long time [8]. In addition, LSTM has also been applied to various real world problems, such as the prediction of secondary protein structure [3]. The use of LSTM as network architectures in a recurring RNN produces bidirectional LSTM [10], [3], and [25]. This bidirectional LSTM gives access to the context of long memory in the input direction.

There is another type of NN which is Elman RNN. Elman RNN can explain the sequential effect of the AR model and MA simultaneously to forecast multiple seasonal time series, and compare the accuracy of forecasting by using seasonal ARIMA models [30]. Several previous studies in many countries around the world, including in Indonesia show that the ARIMA model for electricity consumption data tends to have MA sequence, [26] and [18]. Several studies related to the short-term electric power can be seen in [4], [13], and [23].

Based on the open issues related to the RNN models by focusing the study on a model of time series data that have a long-term dependence, the research was conducted with the general goal to learn about the RNN models in forecasting modeling framework, especially on the major problems faced by this nation in terms of anticipating changes in the electricity load from time to time. In particular, forecasting modeling which will be discussed will focus more on time series models with application state electrical loads using RNN models. Further models of RNN on the model fractional integrated in this study will be called the term Fractional Integrated Recurrent Neural Networks (FIRNN).

Based on the background described above, the problems of this study can be formulated as follows:

1) How to how to develop a model of Fractional Integrated Recurrent Neural Networks (FIRNN) in forecasting time series data on National Electricity Load?

2) how are the forecasting results of time series data on national electricity load using Fractional Integrated Recurrent Neural Networks (FIRNN)?

2. Theory of Time Series Analysis and Neural Network

2.1. Time Series Analysis

Time series is a sequence (row) of Y_t observations on a variable Y, in which each observation is recorded at a certain time $t \in T$ [1] . In this case T is the set of the time at which the observation is made. If T is a discrete set, then Y_t is a discrete time series. Y_t is a notation whole time series, which is the observation of Y_t in time to t. In the case of discrete time series, observations are usually taken at the same time interval.

A Y_t process can be considered as a strong stationary if $(y_{t_1}, y_{t_2}, .., y_{t_k})'$ and $(y_{t_{1+l}l}, y_{t_{2+l}}, .., y_{t_{k+1}})'$ 'has the similar joint distribution function for all integers $k \geq 1$ dan for all $t_1, t_2, \dots, t_k,\ \ k \in N$ in [2]. Processes with the first and second moments are independent of time is also a concern in the analysis of time series. The following definitions is related to the concept of weakly stationary or stationary to second order. A sequence $\{Y_t\}$ can be included into the model Autoregressive Integrated Moving Average when the d difference which is $W_t = \nabla^d Y_t$ is a stationary ARMA process. If W_t is ARMA (p, q) then Y_t is ARIMA (p, d, q). In practice, the value of d is used generally worth 1 or at most 3 [27].

in general SARIMA models $(P, D, Q)^S$ [1]

$$\Phi_p(B^S)(1-B^S)^D Z_t = \Theta_Q(B^S)a_t \tag{2.1}$$

With $\Phi_p(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} \dots - \Phi_P B^{PS}$

$\Theta_Q(B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} \dots - \Theta_Q B^{QS}$

Multiplicative SARIMA models can be written in the following equation,

$$\phi_p(B)\Phi_p(B^S)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\Theta_Q(B^S)a_t \tag{2.2}$$

This model is often called by the SARIMA model (p, d, q) $(P, D, Q)^S$. This model will be reduced to the ARIMA model (p, d, q) when there is no seasonal effects as well as being the ARMA (p, q) when the time series is stationary. The generalized Additive SARIMA models can be written in the following equation

$$\begin{aligned}(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p - \Phi_1 B^S - \Phi_2 B^{2S} \dots \\ - \Phi_P B^{PS})(1-B)^d(1-B^S)^D Z_t \\ = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q - \Theta_1 B^S - \Theta_2 B^{2S} \dots - \Theta_Q B^{QS})a_t\end{aligned}$$

ARFIMA model is able to model both the short-term and long-term dependence. Observations generated by the ARMA structure shows the short-term dependence, while the fractional differentiation parameter d, which causes values to fall hyperbolic ACF showed long-term dependence. ARFIMA (p, d, q) models developed [7] is written as:

$$\phi(B)(1-B)^d(Y_t-\mu)=\theta(B)a_t \tag{2.3}$$

2.2. Neural Network (NN)

RNN has some good properties in sequence labeling, RNN is flexible in the use of information context, RNN is able to receive many types of information and different data representations and can recognize the shape of a sequential pattern in data presentation can recognize sequential patterns with a sequential distortion. However, RNN also has some disadvantages which are their limited application in the problem of labeling in real world sequence. The possibility of a very serious shortage of RNN is difficult to store information for long periods of time [11] The limit to be able to access RNN labeling sequence is very important.

In a study that used activation function is the sigmoid function in the form of bipolar

$$f(x)=\frac{1-e^{-x}}{1+e^{-x}} \tag{2.4}$$

With the value of the derivative$f'(x)=\frac{\sigma}{2}[1+f(x)][1-f(x)]$. This function is very close to the hyperbolic tangent function. Both have a range between -1 to 1. Hyperbolic tangent function, formulated as $f(x)=\frac{e^x-e^{-x}}{e^x+e^{-x}}$ or $f(x)=\frac{1-e^{-2x}}{1+e^{-2x}}$ With the value of derivatives is $f'(x)=[1+f(x)][1-f(x)]$.

In addition to the bipolar sigmoid function, activation function which is used is a linear function of the form$f(x)=x$. This linear function has an output value which is equal to the value of output.

Both of the activated functions which are used in the analysis of the real case study are ARIMA models and RNN-Elman. The figure of Elman RNN-or which is often called the ARMA (2,1)–NN anonymous has 4 units of neurons in layer hidden show in figure 1. The main difference between the RNN and NN Elman other types is the display of the feedback process that presents its output into the next input.

Weights and biases in model-Elman RNN are estimated using the back propagation algorithm. Generally RNN with one hidden layer, input units q and p units in the hidden screen is

$$\hat{X}_t=\left\{\beta_0+\sum_{j=1}^{p}\beta_j f^h\left(\gamma_{j0}+\sum_{i=1}^{q}\gamma_{ji}X_i\right)\right\}f^0 \tag{2.5}$$

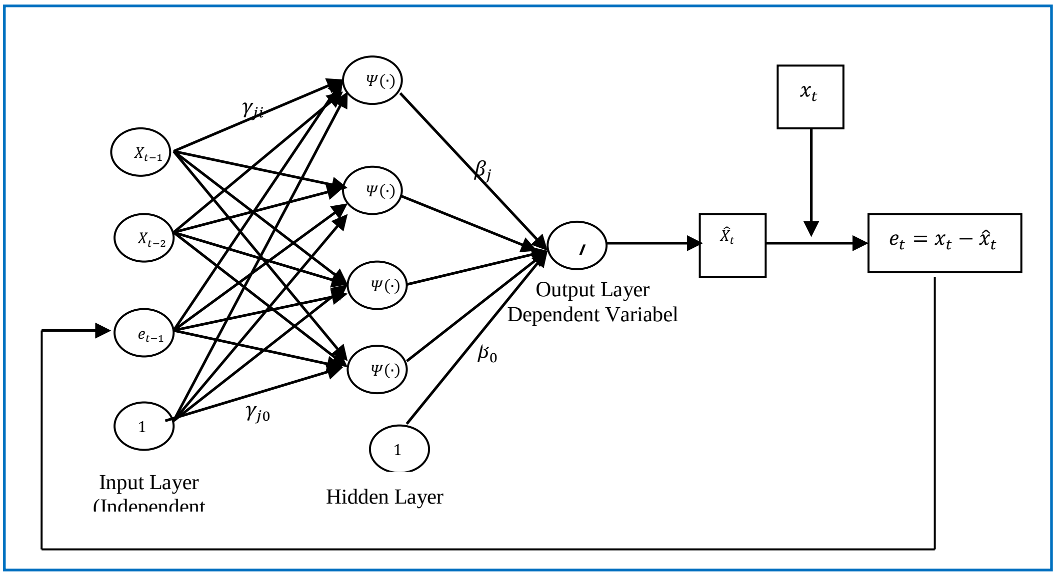

Figure 1. RNN-Elman Architecture

With β_j is weight on the j-th unit in the hidden layer, γ_{ji} is the weight of the i-th input to the j-th unit in the hidden layer, $f^h(x)$ is the activation function in the hidden layer, and $f^0(x)$ is function in the output layer. The above weights and biases can be estimated by minimizing the value of e, which is described in the following equation.

$$e = \frac{1}{2}\sum_{t=1}^{n}\left[X_t - \hat{X}_t\right]^2 \tag{2.6}$$

3. Methods

This study was conducted with the study of literature, simulations and applications to real cases of long time series data in energy consumption load in Indonesia. The study of literature and in-depth theoretical study were carried out to obtain the following results:

1) Retrieving studies of RNN model forecast on the time series and the seasonal pattern of long memory.
2) Retrieving forecasting ARFIMA modeling studies on time series pattern of long memory.
3) Developing theoretical models in the RNN and time series models and other related.

The study of the application was conducted by taking the load electricity consumption data the in Java and Bali every half an hour, taken from December 1st to December 31st, 2010. Furthermore, the data were divided into two groups, as the incoming data in the sample was observed during the period of December 1st until December 25th, 2010 and the out coming data in the sample was observed on 26th up to 31st December 2010. The following steps were carried out to analyze the data.

1. modeling ARIMA using the Box-Jenkins procedure
2. modeling RNN using the inputs based on the sequence of the best ARIMA model in the first step
3. Forecasting the outcoming data sample using both RNN and ARIMA models
4. Comparing the accuracy of forecasting results by using the Mean Absolute Percentage Error (MAPE) and Root of Mean Square Error (RMSE) to determine the best forecasting model.

4. Results

As described in the research methods above, the development and application of RNN – based modeling procedures as a predictive model to model time series data with the application of the national electricity load will be conducted in the following steps. Procedures for the RNN modeling of time series was conducted with the following steps.

4.1. Identification of Long Memory Models

Based on the original data plots from the program, the original data can be considered as stationary. Consequently, the time series analysis which is used to determine the Hurst value can be conducted in order to determine the long memory identification.

Box-Jenkins method based - ARFIMA modeling was conducted through several steps, which are model identification, parameter estimation, verification and forecasting. Basically, the method has the same phase with ARIMA models but each stage has its own distinction. In order to have a better look on the difference, a sample application will be given using the power of data which was created every hour. Based on analysis of the program and the study of theory, ARFIMA models were obtained.

H values in the regression equation can be estimated through ordinary Least Square method (OLS). If $| H | = 0.5$, then $\{Y_t\}$ shows a pattern of short memory, whereas if $0 <| H | <0.5$, then $\{Y_t\}$ shows an intermediate pattern memory, and if $0.5 <| H | <1$ then $\{Y_t\}$ shows a pattern of long memory. The $\hat{d}$

fractional differentiation parameter estimation was given with the estimated value of $d = H - 0.05$.

Results of analysis using Matlab program to determine Hurst value can be seen in the following figure.

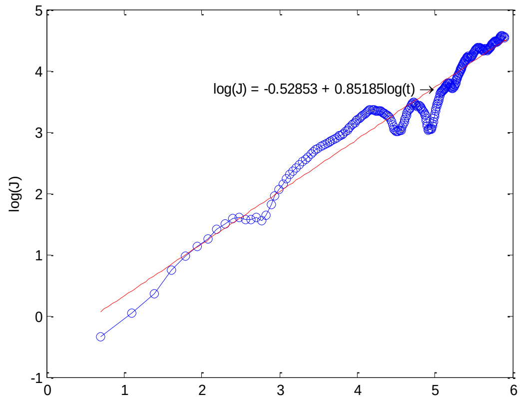

Figure 2: Results of analysis with Matlab program for Hurst value

Based on the calculations, $\bar{y} = 14487{,}6$ was obtained and Ordinary Least Square method with X = log(t) and Z = log (R/S) obtained estimated values H = 0.85185 so that d = 0.35185. These results indicate that the data of electricity consumption in Java and Bali follow the pattern of long memory.

D.2. Testing Stationary and Nonlinearity

The first section has explained that the original data plot shows that the data is in a stationary condition. Therefore, it is necessary to analyze whether the data is also stationary at the first difference of the data plot (differential level 1). Based on the results of the analysis with the program, the first plot in the data difference can be seen below.

The results above indicate that the data as the first difference is stationary, it can be shown by the degree of variance values which are smaller than the value of the plot of the original data variance.

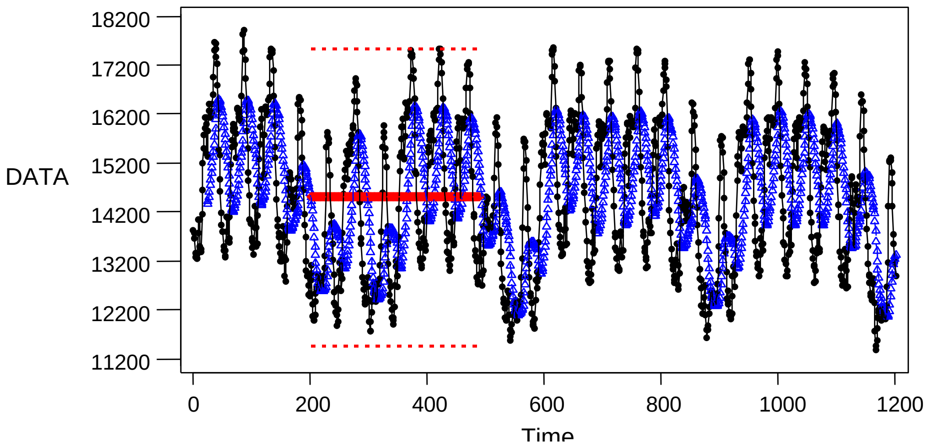

Figure 3: Plot of original data of electricity consumption load per half hour

Based on calculations using Terasvirta Neural Network Test (can be seen in the attachment by using the program R), the lag-1: X-squared = 61.2218, df = 2, p-value = 5.085e-14. These results indicate that the data show a nonlinear pattern, because the P-Value <0,005. Based on calculations using Terasvirta Neural Network Test using R program for deriving the lag-2: X-squared = 164.6963, df = 7, p-value <2.2e-16. These results indicate that the data show a nonlinear pattern, because the P-Value <0,005. Based on calculations, Terasvirta Neural Network Test using R program for deriving the lag-2: X-squared = 164.6963, df = 7, p-value <2.2e-16. These results indicate that the data show a nonlinear pattern, because the P-Value <0,005.

4.3. Data Analysis and Discussion

Based on the results of the descriptive analysis of the data, the electricity consumption per month and per half hours on Bali - Java System Load (MW) can be seen in table 1 below.

Table 1 above illustrates that the highest electricity load occurs in November is 17992 MW at 19:00 on Wednesday the 4^{th}, and the lowest on the same day ranging from 8940 MW. This condition illustrates that at 07.00 most customers turn off the lights, when the community is ready to work and go to the office. Generally in Indonesia, customers' working hours usually start at 08.0 and finish at 16.00.

Table 1. Descriptive Load Power Consumption in 2010

MONTH	DATE/TIME		DAY	MAXIMUM LOAD	MINIMUM LOAD	DURING PEAK LOAD	THE AVERAGE MAXIMUM LOAD
January	**22**	**19:00**	Thursday	**16896**	**9445**	**15911**	**14773**
February	**27**	**19:00**	Friday	**17157**	**11059**	**16262**	**15074**
March	**24**	**18:30**	Tuesday	**17503**	**11711**	**16668**	**15415**
April	**28**	**19:30**	Wednesday	**17776**	**11285**	**16824**	**15562**
May	**8**	**19:00**	Thusday	**17896**	**11661**	**16865**	**15656**
June	**17**	**19:00**	Tuesday	**17802**	**11454**	**16695**	**15449**
July	**31**	**18:30**	Thursday	**17645**	**11322**	**16469**	**15570**
August	**12**	**19:00**	Tuesday	**17840**	**9779**	**16428**	**15714**
September	**18**	**19:00**	Thursday	**17860**	**7101**	**16470**	**15431**
October	**22**	**19:00**	Thursday	**18100**	**11558**	**16769**	**15658**
November	**4**	**19:00**	Wednesday	**17992**	**8940**	**16655**	**15582**
December	**3**	**19:30**	Thursday	**17919**	**11334**	**16424**	**15368**

Furthermore, household electricity consumption in a period is less or more than the overall average power consumption. At 18:00, the customer turns on the light in the evening and at 19:00, most customers come back from work, and do a lot of activities in their houses, so they use large amounts of electricity such as the use of electronic devices. The load of average power consumption ranges from 15582 MW and the highest occurred in November, while the highest peak load during the day occurred in May which has reached about 16865 MW. This condition suggests that the usage of load power consumption at any time shows a different trend, so that there is a difference between the peak load during the day, night and in the morning. That becomes the characteristic of use loads of electricity consumption in Indonesia and it is always different from the rest of the world.

4.4. ARIMA Models and Model Results Elman RNN

The construction of the ARIMA model based on the procedure Box-Jenkins (Box, Jenkins 1994) begins with the identification of the sequence of the data stationary models. Figure 1 and Figure 2 show that the data is stationary.

(PACF) results can be seen in the following figure.

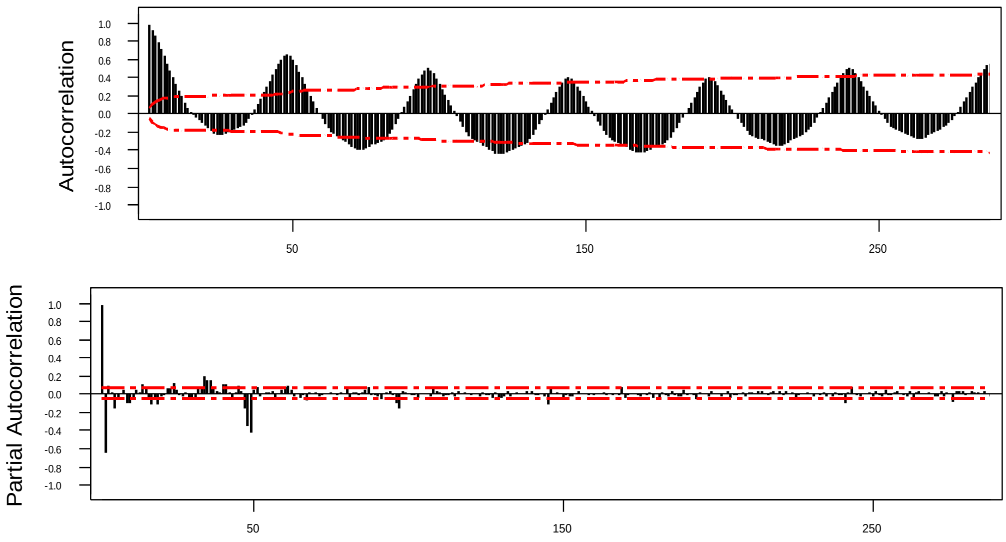

Figure 4. ACF and PACF original data

Based on ACF and PACF plots from the original data, and the data, it can be seen that the dissolution of the stationary is based on the second lag, with the value of 0.98 fakp dropped to -0.65. It appears to be seasonal differences over a periode of hours (lag 12). After seasonal differences over a periode of hours, ACF and PACF plots. ACF plot shows that the regular lag decreases very slowly and shows that the sequence differences can be necessary. Furthermore, the differences in regular and daily seasonal sequence data have ACF and PACF plots. The results show that the ACF plot at 168 and 336 lag significantly and tend to stop down very slowly. Therefore, a sequence of weekly seasonal difference is required to use (lag 24), (lag 28) and (lag56).

ACF and PACF plot the data first difference stationary, the data has been differentiated with a lag of 1 and 24. Based on ACF and PACF plots, there are several ARIMA models and seasonal ARIMA double while to be proposed, the ARIMA (1,0,1), ARIMA (1,0,2), SARIMA ((1,1,1) $(2, 1.1)_{24}$), SARIMA ((1,1,1) $(2,1,1)_{28}$) and SARIMA ((1,1,1) $(2,1,1)_{56}$). Further results on tests of parameter significance and diagnostic checks for the five proficiency level models show that the rest is white noise. Moreover, the test results on the normalization of the Kolmogorov-Smirnov test rest with the rest to show that the two models do not meet the normal distribution.

The results of the comparison between the values of forecasting accuracy Elman-RNN models can be seen in table 2. Based on the criteria of MAPE and MSE on the set of data out sample.

Table 2. The comparisons of data accuracy value out sample

Models	Criteria Out Samples	
	MAPE	MSE
ARIMA (1,0,1)	5,92	3936
ARIMA (1,0,2)	6,35	4442
SARIMA ((1,1,1) $(1,1,2)_{24}$)	4,44	3732
SARIMA ((1,1,1) $(2,1,1)_{28}$)	11,45	7745
SARIMA ((1,1,1) $(2,1,1)_{56}$).	29,32	9978
Elman-RNN (24,2,1)	1,80	2892

Based on the analysis, table 2 shows that among the other models, the ARIMA (1,0,1) has a MAPE value better than ARIMA (1,0,2), whereas SARIMA model that has the smallest MAPE values is a SARIMA model ((1,1,1) (1,1,2) 24) with the amount of MSE is 3732. In this case, this model is the best model available on the ARIMA model. While the MAPE values for models Elman-RNN (24,2,1) is worth 1.8 and MSE 2892. This means that RNN model has the smallest value of MAPE and MSE when compared with the five existing ARIMA models. This means that the RNN model is the best model because it has the smallest value of MAPE and MSE.

5. Conclusion

5.1. Conclusion

Based on the discussion above, it can be concluded that the application of fractional Model of integrated recurrent neural networks (FIRNN) for forecasting time series data in national electricity load consists of several steps such as determining the identification of long memory models using the rescale range, and conducting nonlinearity and stationary test using Terasvirta test and data analysis using ARIMA and Elman-RNN. The results of the integrated fractional forecasting recurrent neural networks (FIRNN) for forecasting of time series data in the National Electricity Load show that Elman-RNN model is better than the models generated by the ARIMA model considering the value of Mean Absolute Percentage Error (MAPE) and root of Mean Square Error (RMSE), which is smaller than the other models.

5.2. Suggestion

In this research, the genetic algorithm method which is popularly used in the optimization has not been used so it is necessary to compare it with the existing method in this study in terms of forecasting. Large genetic algorithm methods can be probably used to study forecasting for similar cases. Therefore, it makes wider opportunities for the development of related forecasting time series study using methods that have not been developed in the research.

References

[1] G.E.P. Box, G.M. Jenkins, and G.C. Reinsel, *Time Series Analysis, Forecasting and Control*, 3rd edition, Prentice Hall, Englewood Cliffs, 1994.

[2] P.J. Brockwell, and R.A. Davis, *Time Series Theory and Methods*, 2nd edition, Springer Verlag, New York, 1991.
http://dx.doi.org/10.1007/978-1-4419-0320-4

[3] J. Chen, and N. Chaudhari, Protein Secondary Structure Prediction with bidirectional LSTM networks, *In: International Joint Conference on Neural Networks,* Post conference Workshop on Computational Intelligence Approaches for the Analysis of Bio-data (CI-BIO), 2005.

[4] J.F. Chen, W.M. Wang, and C.M. Huang, Analysis of an adaftive time-series autoregressive moving-average (ARMA) model for short-term load forecasting, *Electric Power System Research*, **34** (1995), 187-196.
http://dx.doi.org/10.1016/0378-7796(95)00977-1

[5] C. Chatfield, *Time Series Forecasting*, Chapman & Hall, London, 2000.
http://dx.doi.org/10.1201/9781420036206

[6] J. Faraway, and C. Chatfield, Time series forecasting with neural network: a comparative study using the airline data, *Applied Statistics*, **47** (1998), 231-250. http://dx.doi.org/10.1111/1467-9876.00109

[7] C.W.J. Granger and R. Joyeuk, An Introduction to Long Memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis,* **1** (1980), 15-29. http://dx.doi.org/10.1111/j.1467-9892.1980.tb00297.x

[8] F.A. Gers, J. Schmidhuber, LSTM Recurrent Networks Learn Simple Context Free and Context Sensitive Languages, *IEEE Transactions on Neural Network*, **12** (2001), no. 6, 1333-1340. http://dx.doi.org/10.1109/72.963769

[9] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Springer Berlin Heidelberg, SCI 385, pp. 1-3, 2012. http://dx.doi.org/10.1007/978-3-642-24797-2

[10] A. Graves, J. Schmidhuber, Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures, *Neural Networks*, **18** (2005), no. 5-6, 602-610. http://dx.doi.org/10.1016/j.neunet.2005.06.042

[11] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, *Gradient flow in Recurrent Nets: The Difficulty of Learning Long-Term Dependencies*, In: Kremer, S.C., (Kolen, J.F. (eds) A Filed Guide to Dynamical Recurrent Neural Networks, IEEE Press, 2001. http://dx.doi.org/10.1109/9780470544037.ch14

[12] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural Computation*, **9** (1997), no. 8, 1735-1780. http://dx.doi.org/10.1162/neco.1997.9.8.1735

[13] W. Husen, *Forecasting of Maximum Short Term Electricity Consumption by Using Elman-Recurrent Neural Network*, Unpublished Final Project, Department of Physics Engineering, State Institute of Technology, Indonesia, 2001.

[14] T.H. Lee, H. White, and C.W.J. Granger, Testing for Neglected Nonlinearity in Time Series Models: A comparison of Neural Network methods and alternative test, *Journal of Econometrics*, **56** (1993), 269-290. http://dx.doi.org/10.1016/0304-4076(93)90122-l

[15] M.T. Leung, A.S. Chen, and H. Daouk, Forecasting exchange rates using general regression neural networks, *Computers and Operations Research*, **27** (2000), 1093-1110. http://dx.doi.org/10.1016/s0305-0548(99)00144-6

[16] F. Lisi, and R.A. Schiavo, A comparison between neural network and chaotic models for exchange rate prediction, *Computational Statistics & Data Analysis*, **30** (1999), 87-102. http://dx.doi.org/10.1016/s0167-9473(98)00067-x

[17] L. Motiwalla, and M. Wahab, Predictable variation and profitable trading of US equities: A trading simulation using neural networks, *Computers & Operations Research*, **27** (2000), 1111-1129. http://dx.doi.org/10.1016/s0305-0548(99)00148-3

[18] Y. Ristiana, *Autoregressive Neural Network Model (ARNN) for Forecasting Short Term Electricity Consumption at PT. PLN Gresik*, Unpublished Final

Project, Department of Physics Engineering, Institut Teknologi Sepuluh Nopember, Indonesia, (2008).

[19] N.R. Swanson, and H. White, A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks, *Journal of Business and Economic Statistics*, **13** (1995), 265-275. http://dx.doi.org/10.2307/1392186

[20] Z. Tang, C. Almeida, and P.A. Fishwick, Time series forecasting using neural networks vs. Box-Jenkins methodology, *Simulation*, **57** (1991), no. 5, 303-310. http://dx.doi.org/10.1177/003754979105700508

[21] A.S. Weigend, B.A. Hubermann, and D.E. Rumelhart, Predicting the future: A connectionist approach, *International Journal of Neural Systems*, **1** (1990), 193-209. http://dx.doi.org/10.1142/s0129065790000102

[22] A.S. Weigend, and N.A. Gershenfeld, *Time Series Prediction: Forecasting the Future and Understanding the Past*, Reading, MA, Addison- Wesley, 1993.

[23] J.W. Taylor, Short-term electricity demand forecasting using double seasonal exponential smoothing, *Journal of the Operational Research Society*, **54** (2003), 799-805. http://dx.doi.org/10.1057/palgrave.jors.2601589

[24] J.W. Taylor, L.M. Menezes, & P.E. McSharry, A comparison of univariate methods for forecasting electricity demand up to a day ahead, *International Journal of Forecasting*, **22** (2006), 1-16. http://dx.doi.org/10.1016/j.ijforecast.2005.06.006

[25] T. Thireou, M. Reczko, Bidirectional Long Short-Term Memory Networks for Predicting the Subcellular Localization of Eukaryotic Proteins, *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, **4** (2007), no. 3, 441-446. http://dx.doi.org/10.1109/tcbb.2007.1015

[26] Walid, Subanar, Dedi. R., Suhartono, *Recurrent Neural Network (RNN) on the Electrical Load in Semarang*, National Seminar Proceedings Mathematics Semarang State University, 2012.

[27] W.W.S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley Publishing Co., USA, 1990.

[28] H. White, Learning in Artificial Neural Networks: A statistical Perspective, *Neural Computation*, **1** (1989), 425-464. http://dx.doi.org/10.1162/neco.1989.1.4.425

[29] J. Yao, and C.L. Tan, A case study on using neural networks to perform technical forecasting of forex, *Neurocomputing*, **34** (2000), 79-98. http://dx.doi.org/10.1016/s0925-2312(00)00300-3

[30] Soni Sunaryo et al., "Double Seasonal Recurrent Neural Networks for Forecasting Short Term Electricity Load Demand in Indonesia", Chapter in Book: Recurrent Neural Networks for Temporal Data Processing, InTech, India, 2011, 1-14.

Received: October 21, 2015; Published: November 28, 2015