SHORT-TERM LOAD FORECASTING IN KHATULISTIWA POWER SYSTEM USING LONG SHORT-TERM MEMORY METHOD

Irfanda Husni Sahid[1], Sarjiya,S.T., M.T., Ph.D[2], Dr.Eng.Igi Ardiyanto,S.T.,M.Eng[3]

Short term load forecasting plays vital role in electrical system operation. It has function to keep the equilibrium between power demand consumption and supply. This research presents the implementation of Long Short-Term Memory (LSTM) method for short term load forecasting in the West Kalimantan System. LSTM performance will be compared to Load Coefficient as the existing method used for weekly operation by PLN. There is no significant advantage found between LSTM and Load Coefficient. LSTM has average RMSE of 15.046 MW, while Load Coefficient has average RMSE of 15.213 MW. Besides that, this research also compared between LSTM and Recurrent Neural Network (RNN), Artificial Neural Network (ANN), and Autoregressive Integrated Moving Average (ARIMA). LSTM has significant advantage against ANN and ARIMA, but LSTM has no significant advantage against RNN.

Keywords: Short term load forecasting, LSTM, Load Coefficient, ANN, RNN, ARIMA

I. INTRODUCTION

Indonesia's electric energy consumption has rose in 2018 at 5.1% with total amount of 232.296 TWh. Kalimantan island's electric energy consumption contributed 9.814 TWh from Indonesia's electric energy total consumption. This value has average growth of 8.5% per year calculated from 2011 to 2017. The largest consumption is coming from household sector, followed by Industry and Business sector. Specifically, West Kalimantan province has total electric energy consumption of 2.374 GWh with average growth of 8.33%.

Khatulistiwa power system (Power system in West Kalimantan) consisted of one 150 kV interconnected system and several isolated systems. In 2019, this system has estimated load projection at 2174 GWh with estimated average sales growth at 20.7 %, and peak load at 440 MW. However, when peak load is continually growing, the system's reliability is relatively low because many diesel engines has aged, and electric generation reserve is inadequate. Furthermore, Khatulistiwa power system is still depending on electric supply from Malaysia [1].

Consider of low reliability, operation plan, especially weekly operation plays vital role to keep electric supply and demand balance. When performing weekly operation, PT. PLN (National Electric Company) follow the weekly operation plan. The weekly operation plan contains the schedule of generating units and power balance calculation in range of one week. The fundamental of weekly operation plan is short term load forecasting. The short-term load forecasting has vital role especially in operation of power system. Accurate load forecast can give benefit to electric provider and consumer.

PT. PLN use load coefficient method to construct load forecast that used in weekly plan operation. This method relies on the ratio of datetime load and datetime peak load, and the ratio of load growth. The drawback from this method is that the result calculated needs to continuously correcting with the result of empirical load. This caused the electric provider needs to adjust the load value empirically. Therefore, accurate and systematic load forecast method is needed.

In this research, load forecasting models are developed with LSTM, RNN, ANN, and ARIMA method. These models will be compared to load coefficient model which acts as existing method.

II. LSTM SHORT-TERM LOAD FORECASTING

A. Recurrent Neural Network (RNN)

Artificial Neural network (ANN) is a model that inspired from human brain neural system. This model is a set of nonlinear adaptive method that driven by data. This model has shortcoming in learning the dependencies of the data, as ANN treat every input data independently [6]. The dependencies become problem because the time series data tends to be affected by past or historical data. This can cause ANN fails to model the data. To overcome the problem, Recurrent Neural Network was implemented. RNN constructed to process sequence data. Basically, RNN is an ANN that has looping structure in their architecture. Figure 1 on the left showed the architecture of RNN. X is the input data, S is the state of model, O is the output or the result from the model. In the other side, u,v,w are the weight of input line, state line, and output line respectively. Figure 1 in the right side shows how RNN model operating in sequence of time. General equation of RNN is stated in Equation (1).

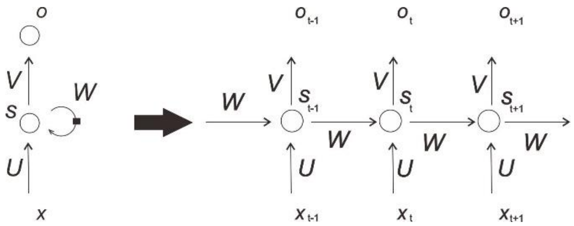

Figure 1. RNN Model Architecture

$$o_t = f(ux_t + ws_{t-1}) \quad (1)$$

Abilities of RNN to model the sequence data is caused by the lag input data. Although the dependencies problem has solved, RNN model still have another problem to model the long sequence data, or in the other terms it is called vanishing gradient. This vanishing gradient problem can decrease the RNN performance. Therefore, the LSTM model was implemented to solve the vanishing gradient problem.

B. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a variation from RNN model that have special components to overcome vanishing gradient problem. The special components of LSTM are input gate, output gate, and cell state. Illustration of LSTM architecture showed in Figure 2.

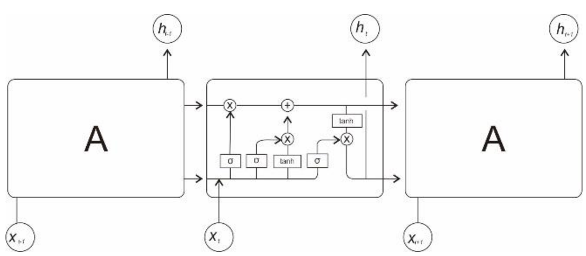

Figure 2 Long Short-Term Memory Architecture

Forget gate is used to decide the information that will be stored or forgotten from cell state. The equation of forget gate is shown in Equation (2).

$$f_t = \sigma\,(w_f[o_{t-1}, x_t] + b_f \quad (2)$$

Input gate is used to decide what information to update to cell state. The input gate will pass the sigmoid function to the input data ($h_{(t-1)}$) and (x_t) for further multiplied with element named candidate. The multiplied value is that will be used to update the cell state. The equations of input gate and candidate are stated in equation 3 and 4.

$$i_t = \sigma\,(w_i[h_{t-1}, x_t] + b_i \quad (3)$$

$$\bar{c}_t = \tanh\,(w_c[h_{t-1}, x_t] + b_c \quad (4)$$

The cell state is the storage or memory component of LSTM. It stores the information provided by input and forget gate and used to compute the output gate. The equation of cell state is stated in the Equation (5).

$$c_t = (f_t * c_{t-1}) + (i_t * \bar{c}_t) \quad (5)$$

Output gate is LSTM component used to process output result in the time -*t*. When processing the output resilt, output gate uses the previous time input that is called hidden state ($h_{(t-1)}$) and current time input (x_t). This will produce the output result and current hidden state. The equations of output gate and hidden state are stated in the Equation (6) and (7).

$$o_t = \sigma\,(w_o[h_{t-1}, x_t] + b_o \quad (6)$$

$$h_t = o_t * \tanh(c_t) \quad (7)$$

III. METHODOLOGY

This research will construct five models to forecast the electrical load consumption. These five models are LSTM, ANN, RNN, ARIMA, and load coefficient. All models will forecast the test data with total number of 30 dataset. Every dataset contains 1-hour load for 1 week. One of electric load consumption graph is shown in Figure 3.

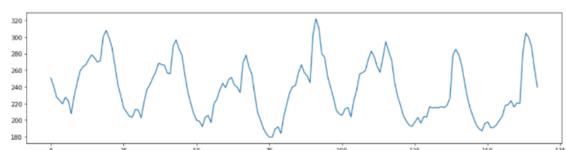

Figure 3 Electric Load Consumption Graph

The first step to construct the forecast models is separate between the train and test data. In this research the test data consist of 5040 hour of electric load consumption data or 30 weeks, every week count as one dataset, thus gives 30 dataset for test data. Train data data consist of 17640 hour of electric load consumption data or 105 weeks. Particularly for ANN, RNN, and LSTM models, the train data is divided further to model's train data, model's validation data, and model's testing data. The model's train data consist of 10080 data or 60 weeks, the model's validation data consist of 2520 data or 15 weeks, and the model's testing data consist of 5040 data or 30 weeks.

In the second step, the model's parameter will be tuned for every model. Tuning process need to be done to get the best out of parameter configuration. For

LSTM, RNN, and ANN, there are three parameter that will be tuned. These parameters are features, hidden neuron, and hidden layer. Parameter tuning process will be done using walk forward method. In a brief, walk forward method is to change one parameter and keep the others constant, and then the model will forecast the test data. This is done continuously until the parameter reach the limit selected by researcher. Then the forecast performance that measured by Root Mean Squared Error (RMSE) will be compared. The RMSE equation stated in the Equation (8). On the other side, ARIMA models will be tuned using box - jenkins method that is tuning the p, d, and q parameter. The best ARIMA model then determined by selecting the model that has minimal Bayesian Information Criterion (BIC) value. Lastly, the load coefficient model does not need any parameter tuning as the model is just constructed by the data.

$$\mathrm{RMSE} = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(Y_{actual} - Y_{prediction}\right)^2} \tag{8}$$

In the next step, all models will forecast the 30 dataset of test data. Then each model's average RMSE value will be compared. The Model's significance performance will also be justified with Diebold Mariano (DM) test. The DM test equation is stated in Equation (9).

$$\mathrm{DM} = \frac{\bar{d}}{\sqrt{[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k]/n}} \tag{9}$$

IV. RESULTS

A. Constructing LSTM Model

The initial LSTM model's parameter have the features value of 168, hidden neurons value of 168, and hidden layer value of 1. The initialized model is shown in Figure 5.

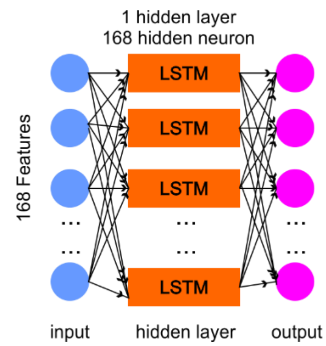

Figure 4 LSTM Initial Model

Firstly, features parameter tested with feature configuration values of 168, 336, 504, and 672 as shown in Table 1. Performance from each configuration is shown at Figure 5. The result indicates that lowest RMSE value was found at configuration 1 or 168 features. In order to find if this configuration is significantly difference than the others, the DM test was conducted. The DM test result can be seen at Table 2.

Table 1. LSTMs Features Parameter Configuration

Method	Configuration	Features	Hidden Neuron	Hidden Layer
LSTM	1	168	168	1
	2	336	168	1
	3	504	168	1
	4	672	168	1

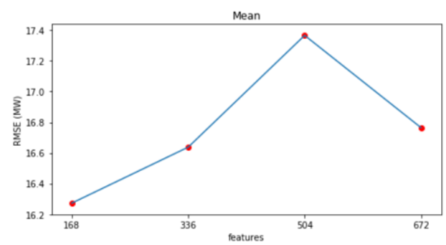

Figure 5. Average RMSE for each configuration in Features Parameter Test

Table 2. Result of Diebold Mariano Test in Features Parameter Test

Features Pair	DM Value	P Value
168 – 336	-2.347	0.018
168 – 504	-4.389	1.15e-05
168 – 673	-2.886	0.003

The DM test result found that there was significant difference in configuration 1 or 168 features as all P value falls below 0.05. Predictively, LSTM with 168 features produce better forecast compared to other features.

After the features parameter test was done. The hidden neuron parameter test was conducted. Configuration of hidden neuron parameter test is shown in Table 3.

Table 3. LSTMs Hidden Neuron Parameter Configuration

Method	Configuration	Features	Hidden Neuron	Hidden Layer
LSTM	1	168	168	1
	2		336	1
	3		504	1
	4		672	1

The average RMSE and DM test from hidden neuron configuration test is shown in Figure 6 and Table 4.

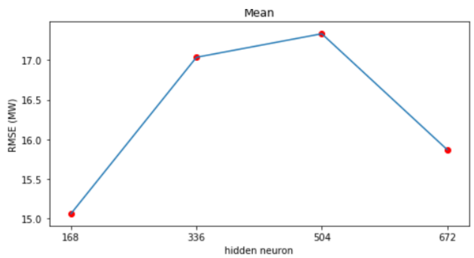

Figure 6 Average RMSE for each Configuration in Hidden Neuron Parameter Test

Table 4. Result of Diebold Mariano Test in Hidden Neuron Parameter Test

Hidden Neurons Pair	DM Value	P Value
168 – 336	-4.498	6.996e-06
168 – 504	-5.022	5.286e-07
168 – 673	-2.196	0.0281

From DM test, significant difference was found in configuration 1 or 168 hidden neuron since all P value falls below 0.05. The hidden neuron value chosen in this test was 168.

Last parameter tuned in LSTM model was hidden layer parameter. The configuration of this parameter is shown in Table 5.

Table 5. LSTMs Hidden Layer Parameter Configuration

Method	Configuration	Features	Hidden Neuron	Hidden Layer
LSTM	1	168	168	1
	2			2
	3			3
	4			4

The result of hidden layer test is shown in Figure 7. The lowest average RMSE for each hidden layer tested was found in configuration 1 or hidden layer with the value of 1. On the other side, the DM test was stated that the configuration 1 has significant difference since all P value falls below 0.05. Predictively, the LSTM with hidden layer 1 produce better forecast.

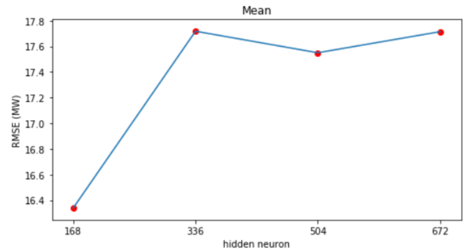

Figure 7 Average RMSE for each Configuration in Hidden Layer Parameter Test

Table 6. Result of Diebold Mariano Test in Hidden Layer Parameter Test

Hidden Layers Pair	DM Value	P Value
1 – 2	-5.072	4.075e-07
1 – 3	-6.324	2.758e-10
1 – 4	-4.863	1.185e-06

Finally, after all LSTM parameter tuning was done, LSTM model was constructed. The features of LSTM model is 168, the hidden neuron value of LSTM model is 168, and the hidden layer value of LSTM model is 1. The model's architecture is shown in Figure 8.

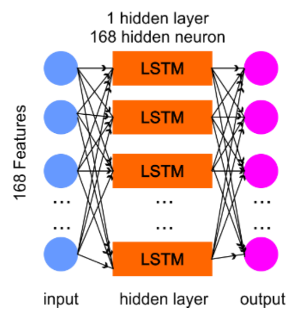

Figure 8 Final LSTM Model

B. Constructing ANN and RNN Model

The initial model of ANN and RNN have the same value as LSTM initial model. The ANN and RNN model were also constructed with the same method as LSTM model. After every parameter test was done. The final ANN model has features parameter of 672, hidden neuron value of 336, and hidden layer value of 3. The final RNN model has features parameter of 504, hidden neuron value of 336, and hidden layer value of 1. Final model of ANN and RNN can be seen in Figure 9.

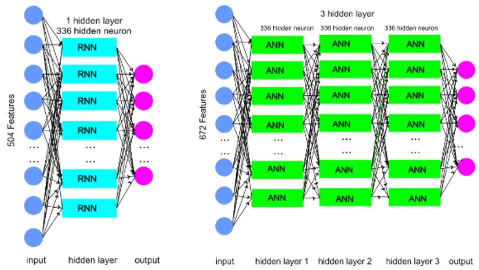

Figure 9 Model Akhir RNN dna ANN

C. Constructing ARIMA Model

First step to construct ARIMA model is observe the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) lag of train data to find every parameter need to be tested. The observation result is shown in Figure 10.

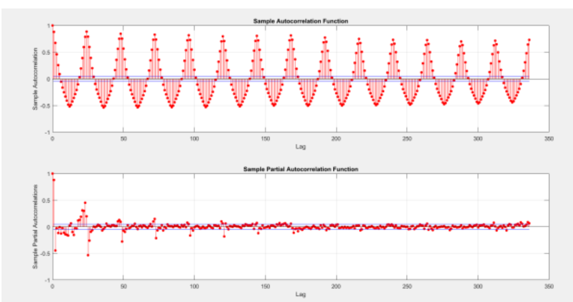

Figure 10 ACF and PACF Lag of train data

The observation above indicate that the train data are not in stationary condition and has seasonal pattern. Since seasonal pattern exists in the observation above, the data needs to be differenced seasonally and non-seasonally every 168 data and 1 data. The result of differenced is shown in Figure 11. From Figure 11, the train data is in stationary condition.

The next step is determine the parameters from ACF and PACF pattern. The parameters obtained from the observation are listed in Table 7.

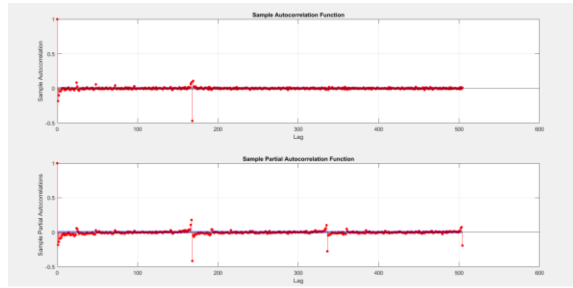

Figure 11 Differenced ACF and PACF Lag

Table 7. ARIMA's Candidate Parameter Configurations

Configuration	p	d	q	P	D	Q	M
1	0	1	0	0	1	1	168
2	0	1	1	0	1	1	168
3	0	1	2	0	1	1	168
4	1	1	0	0	1	1	168
5	1	1	1	0	1	1	168
6	1	1	2	0	1	1`	168
7	2	1	0	0	1	1	168
8	2	1	1	0	1	1	168
9	2	1	2	0	1	1	168

BIC value then estimated for every configuration in Table 7. A model with lowest BIC value will be the chosen model. The result of estimated BIC is listed in Table 8.

Table 8. Result of estimated BIC value

Configuration	BIC
1	115215.7016
2	114449.6955
3	113871.8773
4	114694.8571
5	113488.6514
6	114059.5444
7	114300.7926
8	114157.7897
9	114949.550

Configuration 5 was found to be the best ARIMA configuration with BIC value of 113488.6514.

D. Test Data Forecast

Test data that consisted of 30 datasets was tested in all forecast model. The average and standard deviation RMSE of forecast result is listed in Table 9 and Figure 12.

Table 9. Average and Standard Deviation RMSE of Test Data Forecast Result

Model	Average (RMSE)	Standard Deviation (RMSE)
LSTM	15.046 MW	6.003 MW
RNN	15.387 MW	4.725 MW
ANN	17.596 MW	6.843 MW
ARIMA	49.993 MW	9.471 MW
Load Coefficient	15.213 MW	5.529 MW

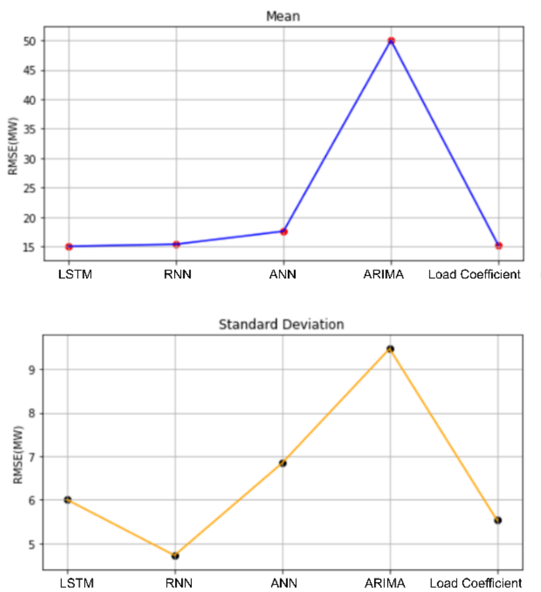

Figure 12 Average dan Standard Deviation RMSE of Test Data Forecast Result

The figure and the table showed that the lowest average RMSE found in LSTM model compared to the other models. To find the significance of the result, DM test was conducted which the result listed in Table 10.

Table 10. Result of Diebold Mariano Test

Models Pair	DM value	P value
LSTM-Load Coefficient	1.297	0.196
LSTM-ANN	8.095	7.064e-16
LSTM-RNN	-0.164	0.869
LSTM-ARIMA	23.152	8.723e-113

Table 10 has shown that LSTM model was significantly difference than the ANN and ARIMA models. However, the significant difference did not found among LSTM with RNN or Load Coefficient. The P value of LSTM-ANN and LSTM-ARIMA pairs falls below 0.05. On the contrary, the P value of LSTM-RNN and LSTM-Load Coefficient falls above 0.05. One of the test data forecast results is shown in Figure 13.

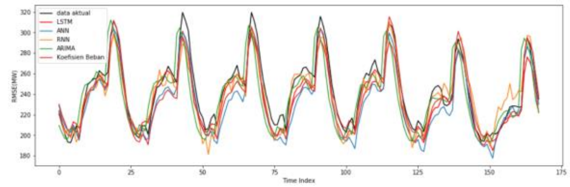

Figure 13 . Sample of Data Forecast Results

V. CONCLUSION

According to the analysis and discussion, there were several conclusions. First, the LSTM model that has been through the parameter tuned process did not have significant difference against Load Coefficient model that acts as existing forecast model. The average RMSE of LSTM model was slightly lower than Load Coefficient model with the value of 15.046 MW and 15.213 MW respectively. However, the Diebold Mariano test result showed that there was no significant difference found. The P value of DM test between LSTM model and Load Coefficient model was 0.196. On the other side, LSTM model has significant difference against ANN and ARIMA model, but has no significant difference against RNN.

REFERENCE

[1] PT.PLN (PERSERO), Rencana Usaha Penyediaan Tenaga Listrik 2019-2028, 2018.

[2] S. Alam, "Recurrent Neural Networks in Electricity Load Forecasting," KTH Royal Institute of Technology, Stockholm, Sweden, 2018.

[3] S. Bouktif and A. FIaz, "Optimal Deep Learning LSTM Model for Electric Forecasting using Feature Selection and Genetic Algorithm," MDPI, p. 20, 2018.

[4] M. S. Harsha and R. Pradyoth, "A Smart Adaptive LSTM Technique for Electrical Load Forecasting at Source," IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT), p. 5, 2017.

[5] S. Masum and Y. Liu, "Multi-step Time Series Forecasting of Electric Load using Machine Learning Model," International Conference on Artificial Intelligence and Soft Computing, vol. 10841, pp. 148 - 159, 2018.

[6] J. Zheng and C. Xu, "Electric load forecasting in smart grids using Long-Short-Term-Memory based Recurrent Neural Network," 51st Annual Conference on Information Sciences and Systems (CISS), vol. MD, pp. 1 - 6, 2017.

[7] P. P. Hyo, "DEEP LEARNING FOR SHORT-TERM ELECTRICITY," SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY, Thailand, 2017.

[8] Y.-S. B. Kyung-Bin Song, "Short-term load forecasting for the holidays using fuzzy linear regression method," IEEE Transactions on Power Systems, vol. 20, no. 1, pp. 96 - 101, 2005.

[9] N. Kim, M. Kim and J. K. Choi, "LSTM Based Short-term Electricity Consumption Forecast with Daily Load Profile Sequences," 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara,, no. 10.1109/GCCE.2018.8574484, pp. 136-137, 2018.

[10] Q. Jiang, "Electricity Power Load Forecast via Long Short-Term Memory Recurrent Neural," in 2018 4th Annual International Conference on Network and Information Systems for Computers, 2018.

[11] S. S. P. Saatwong, "Short-term electricity load forecasting for Building Energy Management System," International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 1 - 6, 2016.

[12] S. G. H. Dimitrios Asteriou, "ARIMA Models and the Box-Jenkins Methodology," in Applied Econometrics, New York, PALGRAVE MACMILLAN, 2006, p. 229.

[13] R. J. Hyndman, Forecasting: Principles and Practice, Melbourne,, Australia: OTexts, 2018.

[14] R. H. M. Ronald E Walpole, Probabilty & Statistics for Engineers & Scientist, Boston: Pearson, 2011.

[15] M. B. Nariman Mahdavi, "Short-Term Load Forecasting for Special Days Using Bayesian Neural Networks," in IEEE PES Power Systems Conference and Exposition, Atlanta, GA, USA, 2006.