Short Term Electrical Energy Consumption Forecasting using RNN-LSTM

Erliza Yuniarti
Faculty of Engineering
Universitas Muhammadiyah Palembang
Palembang,Indonesia
erlizay@yahoo.com

Nurmaini
Intelligent System Research Group
Sriwijaya University
Palembang, Indonesia
siti_nurmaini@unsri.ac.id

Bhakti Yudho Suprapto
Faculty of Engineering
Sriwijaya University
Palembang, Indonesia
bhakti@ft.unsri.ac.id

Muhammad Naufal Rachmatullah
Intelligent System Research Group
Sriwijaya University
Palembang, Indonesia
naufalrachmatullah@gmail.com

***Abstract*—The estimation of a-day forward short-term electrical consumption in this research is using time series daily data from PT Perusahaan Listrik Negara (PLN) in 5-year periods. This research proposed Recurrent Neural Network and Long Shot Term Memory (LSTM) as model to forecast the electrical load using various hidden layers consist of one, two, and three layers. In this research, secondary data pre-processing is going to be done with add empty data, remove double data, and remove time with random interval. The electrical load data for 5 years is divided into 2 types of datasets, namely training data and test data. The raining data using data consist of 4 years electrical load from 2013-2016, while the test data uses data in 2017. LSTM Model then compared with Random Forest (RF) and Support Vector Machine (SVM). From the experimental result, the Root Mean Square Error (RSME) for LSTM model with 2 has the lowest compare to SVM and RF.**

Keywords— Electrical Consumption, Forecasting, LSTM

I. INTRODUCTION

Electric energy consumption describes the characteristics of the use or use of electrical energy in a location. Profile of electricity consumption is useful for load forecasting, optimization of electricity generation and determination of consumer electricity rates. Temperature magnitude [1], [2], working days, utilization hours [3], economic growth and regional development are important factors that influence electricity consumption.

Based on time intervals, load forecasting can be separated into three categories namely short-term load forecasting (STLF) which is usually from one hour to one-week, medium-term load forecasting (MTLF) which is usually from one week to one year and term load forecasting length (LTLF) which is more than one year. Short-term load forecasting is pretty much done specifically to optimize and control generating units, especially those with fossil fuels, and the determination of electricity prices. Medium and long-term forecasting studies are rarely carried out even though this study is important for planning new generating units, expansion of electricity networks, and investment.

Many research has been done on the profile of electrical energy and load forecasting, including [2] ANOVA statistical methods, [4] using Artificial Neural Network (ANN) method, [5] comparing ANN and Multiple Linear Regression (MLP) methods, [1], [6] using Recurrent Artificial Neural Networks, [7] uses the classic Nearest Neighbor and Classic Bayesian classifications, based on Hidden Markov Models (HMM), [8] comparing ANN algorithms with Support Vector Machine (SVM) and Random Forest (RF,) [9] using Convolutional Neural Network (CNN), Long Short Term Memory (LSTM) and hybrid CNN LSTM (EPNet), [10] using an LSTM-RNN-based algorithm.

ANOVA and ARIMA statistical methods are used to determine electricity prices for one day in Turkey. Factorial ANOVA is used as pre-processing before forecasting, with independent variables consisting of prices, hours, weeks, days of the week, months of the year, working days and holidays. The output of factorial ANOVA is the basis for the formation of a model of the estimated burden or seasonal electricity prices of ARIMA.

Daily short-term electricity consumption profiles and research have been discussed by [1], [4], [7], [8], [11], [12]. Perez-Chacon et al (2016) Discusses electricity consumption data in the implementation of distributed computing parallel versions of K-Means ++ by using MLlib and Scala, where the application of electricity energy usage data grouping with this algorithm allows identification of daily electricity consumption behavior patterns related to workdays. Another study of the introduction of the [7] electricity consumption pattern was carried out at a coherent group level taking into account the features of holidays, seasons, types of houses and temperatures. This research uses the classification of Nearest Neighbor (NN) and classical Bayesian, where this probabilistic approach is used to characterize sequential data, and is applied to grouping load profiles

Other literature [8], which discusses electrical energy consumption, makes predictions of short-term electricity loads for the next 24 hours, which are based on the peak load classification model. Models are trained to identify load levels in each distribution system. The value of peak load is grouped for each week by considering days and hours. The end result compared between machine learning algorithms, the ANN model gets the greatest stability with the lowest error, compared with the SVM and RF algorithms

Load forecasting uses AI-RNN in the [1] literature to consider the profile of past demand data using an internal feedback connection, because the previous load data influences future load predictions. As input, the consumption data is used in the past and the temperature variable. AI-RNN gets the estimated load more accurately than Back propagation for a period of 24 hours a day using Matlab.

RNN using the Feed-forward neural network method, Elman and Feed-forward neural network Jordan uses annual time series data to be used for one day forecasting on the Lithuanian North Pool market [12]. The structure of Elman and Jordan RNN in this literature is designed to have several input layers, hidden layers with 1 output neuron.

The study [4] discusses wavelet transformation and Neural Network (Feed-Forward Back Propagation - FFBP) methods for forecasting medium-term expenses, namely 3 months to the next 3 years. Wavelet transformation and neural network method which is a very significant technique for forecasting the load in the interest of planning fuel reserves in the generating system. This literature uses four wavelet levels in preprocessing to decompose the original signal into detail and approximate components. After that, the estimated value component will correlate with the factors to find the value, and choose the input feature for the neural network to predict the load

Zhu, Da, Liu, Wang, & Lu (2018) has used the structure of RNN and Long Short Term Memory (LSTM) to estimate the forecast of short-term electricity prices for the next 1 day. LSTM network is very suitable to be used in learning from past data to identify and predict time series when there is a very long lag time with an unknown size. The forecasting performance of the learning method is below the input with different feature lengths and forecasting horizons. Testing the approximate model on the dataset is compared with the DT forecasting method, SVM. LSTM-based deep learning successfully estimates the power market price with satisfactory accuracy under different input lengths, and variations in data size.

The latest research [10] for the development of smart grid electricity distribution networks, forecasting accurate electrical loads is becoming increasingly important because it can help electricity companies in better load scheduling and reduce excessive electricity production. The use of the RNN-LSTM machine learning and neural network approaches with various configurations aims to build a short to medium term forecasting model. The selection of the optimal lag and number of layers to optimize the prediction performance of the model using the genetic algorithm (GA). The results obtained indicate that the LSTM-based model has shown high accuracy.

In this preliminary study discussed the profile of consumer electrical energy consumption from log sheet PT PLN (Persero) Keramasan Substation Palembang. The purpose of this study is to find out the forecast of electricity consumption one day ahead, using secondary data on daily electricity consumption of PT PLN (Persero). The RNN-LSTM learning machine method is used as a model, and the accuracy results are compared with the accuracy of the SVM and DT algorithms.

II. PREVIEWS WORK

This section discusses the approach to obtaining a profile of electrical energy consumption and short-term forecasts for the next day. An approach to evaluating the effectiveness of combining preprocessing strategies with learning methods.

The pre-processing data stage is done first to ensure that the data to be processed is good data. To get the performance and accuracy of machine learning as well as a model based on good deep learning, the data must be consistent, not lost, and not contain noisy. Preprocessing in the first datasheet is done by reducing the data on voltage, apparent power, transformer temperature and phase current, which are not used as exogenous factors of electrical energy consumption. Continuing to set inconsistent datasheet ranges, namely data with an interval of 30 minutes, so that the datasheet obtained is in the form of daily electricity consumption data in hours. Data preprocessing for time series data is shown in fig 1.

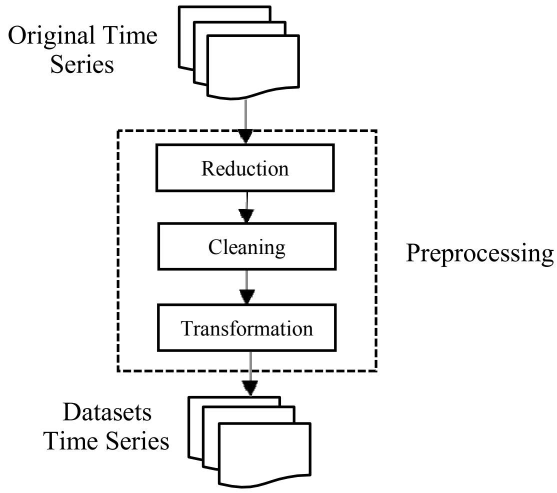


Fig. 1. Preprocessing Data Time Series

The data cleaning process, by filling in data that is empty or incomplete with the same data on the previous day. Outliers data are identified and eliminated to avoid data analysis being biased or not reflecting the actual phenomenon.

Machine learning and the LSTM-RNN model are sensitive to the input scale. The electric energy consumption datasheet is normalized as a scaling process of attribute values from the datasheet using the softmax method. Softmax is the development of linear transformation, with the output range being 0-1.

Electricity consumption data for 5 years 2013-2017 is divided into 2, with a ratio of 80/20 where 2013-2016 is used as training data and 2017 as testing data. Training data is used to train forecasting models, and the remainder is to evaluate the accuracy and validation of the proposed forecasting model.

III. METHOD AND DESIGN

A. RNN

RNN is the development of ANN by utilizing sequential information, such as time series because it combines contextual information from past inputs. RNN is quite flexible for temporal dynamics because of its repetitive nature. Information is fed to the network one by one and the nodes on the network store their status at one time step and use it to inform the next step. RNN can save memory because the current output depends on previous calculations. The disadvantage of RNN is that it only repeats in a few steps due to the loss of the gradient. RNN models the input sequence {x1, x2, ..., xn} using looping:

$$h_t = f(h_{t-1}, x_t) \tag{1}$$

B. LTSM

Short-term memory (LSTM) was initially developed [14] LSTM is a variation of Deep Recurrent Neural Networks (RNN) to enable continuity of weights and propagated through layers. LSTM for time series forecasting studies complex non-linear patterns and automatic feature extraction capabilities. LSTM is designed to avoid long-term dependency problems with increasing time on the RNN, gradient removal and exploration. LSTM is specifically designed to overcome this problem by introducing a new gate that allows better control over the gradient flow and allows better maintenance of distance. Fig. 2 illustrate LSTM architecture.

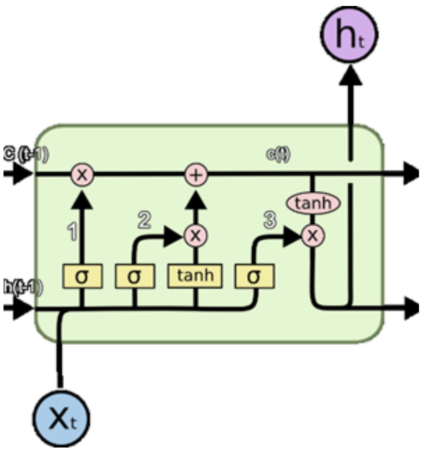


Fig. 2. LSTM Architecture

LSTM has three types of gates: a forget f_t gate, an input gate i_t and an o_t output gate, to control t-cell steps. LSTM cell state at the input gate and input gate, forget and output gate:

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{2}$$

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{3}$$

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_0) \tag{4}$$

Current cell status C_t and value of $\hat{C}$ calculates as follows:

$$\hat{C} = tanh(W_C.[h_{t-1}, x_t] + b_C)\Theta \tag{5}$$

$$C_t = f_t \odot C_{t-1}. + i_t \odot \hat{C}_t \tag{6}$$

Output from t step is calculated:

$$h_t = o_t \odot tanh(C_t) \tag{7}$$

Where *W, h, and b* are weight, output cell, and input bias respectively.

C. Dataset

The dataset used in this research comes from the daily log sheet PT. PLN (Persero), Keramasan Substation of the Simpang Tiga Department from 2013-2017, which is the operator data of a 150kV high voltage air channel transmission system. Transmission operator log sheets are

recorded on a daily basis every hour, arranged monthly and annually. Log sheet consists of data: phase current, voltage, active power, apparent power, and transformer temperature on the transmission side. Active power data every hour is the power of electricity consumption, where measurement data is carried out 26 times in 24 hours with a span of 1 hour 24 times and a range of 30 minutes 2 times during peak load hours of 18.00-20.00. Fig. 3 and 4 show the pattern of electricity usage per day and per month.

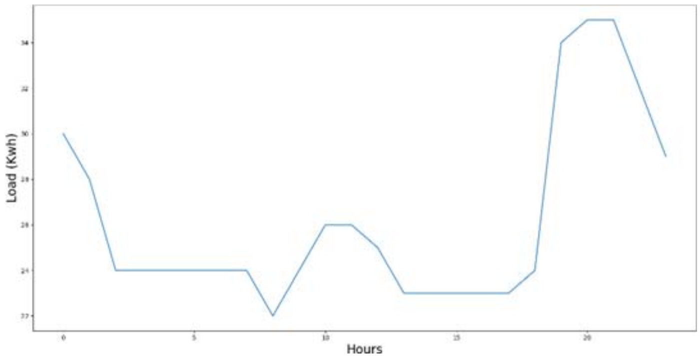


Fig. 3. Pattern of electricity usage per day (01/01/2013).

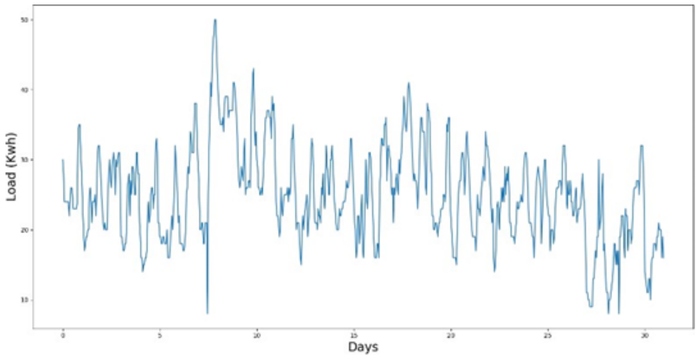


Fig. 4. Electricity Usage Pattern Per Month (January 2013).

The total data in each year consists of 8671 data, but there are missing values in certain hours. To overcome this problem, we carried out an imputation process [15] using data from one day before at the same hour.

D. Data Preprocessing

For the LSTM model, data is stationary using detrending techniques. Detrending is a technique for removing trends from time series data. A trend usually refers to changes in averages over time. By using detrending techniques, aspects that can cause distortion can be removed from the data.

Next, the data is detrending, then rescale into the range [-1, 1]. This helps the convergence process faster. In addition, this process also aims for features that have a relatively large amplitude that do not provide large weight during training so that they can produce models with good generalization.

Load data for 5 years is divided into 2 types of datasets, namely training data and test data. For training data using data for 4 years (2014-2016), while the test data uses data in 2017. Every training and testing data is set so that it can predict electricity load on day n + 1 using input from day n on the same time.

IV. Experimental Result

A. Hyperparameter

The number of hidden layers in the LSTM Model will be tested to produce a good LSTN model. In addition, this test is also conducted to determine the effect of the number of hidden layers on the errors generated. The number of layers tested is 1, 2 and 3 hidden layers. Table 1 shows the results of testing the use of a number of hidden layers in the LSTM model developed. From the test results, getting LSTM using 2 layers has the smallest MSE and RMSE values.

TABLE I. COMPARISON OF NUMBER OF HIDDEN LAYER

Num of Layer	MSE	RMSE
1 Layer	142.6812477	11.94492561
2 Layers	130.6428458	11.42991014
3 Layers	140.9930953	11.87405135

B. Comparison with othre methods

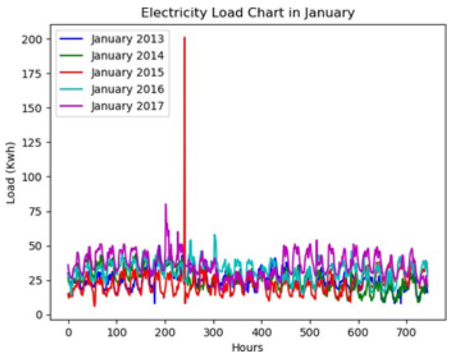
The LSTM model developed compared to other methods, namely Random Forest and Support Vector Regression. The results of the comparison of the three methods are shown in table 2.

TABLE II. COMPARISON OF LSTM WITH RANDOM FOREST AND SUPPORT VECTOR REGRESSION

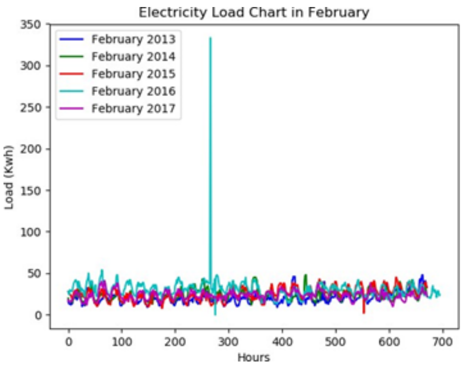
Model	MSE	RMSE
LSTM 2 Layers	130.6428458	11.42991014
Random Forest	144.0220964	12.00092065
SVM (Linear Kernel)	132.9182855	11.52901928
SVM (RBF Kernel)	142.1859158	11.92417359

V. Discusion

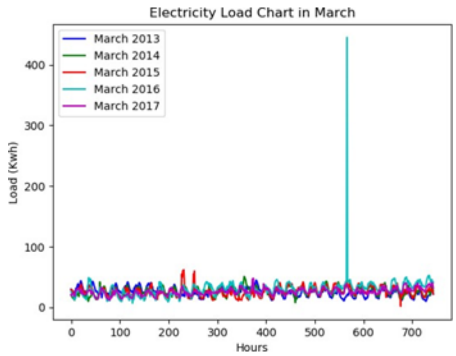
From the test results it is known that the RMSE value for all methods has a relatively high value (above 10). After analyzing the data used, it is known that there are anomalies of data in certain months. Fig. 5a-5l shows the data distribution plot every month from 2013-2017. The use of electric power in January 2017 (testing data) has a slightly different pattern from the data for 2013-2016 (training data). There were also some anomalies in training data, i.e. February, June, August, November and December. In addition, in July and September 2017 there were several testing data anomalies that could cause errors in the predicted results. We still include the data anomaly because we want to keep the purity of the data. These anomalies often occurred in real time data, so that in the future we want to construct a LSTM-based classifier that is robust with data anomalies.



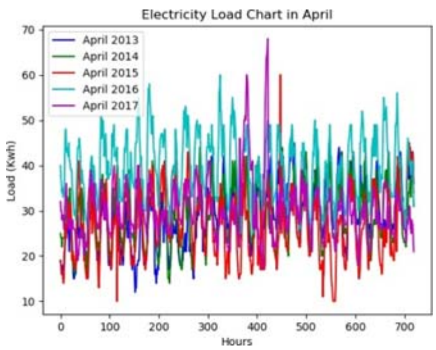
(a)



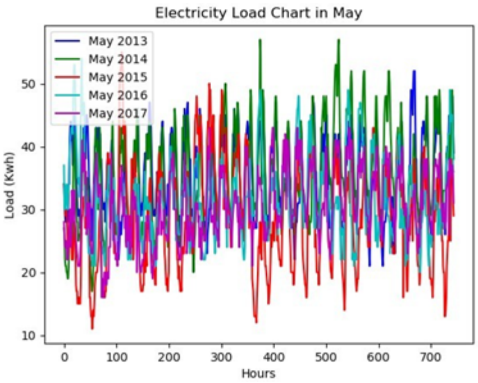
(b)



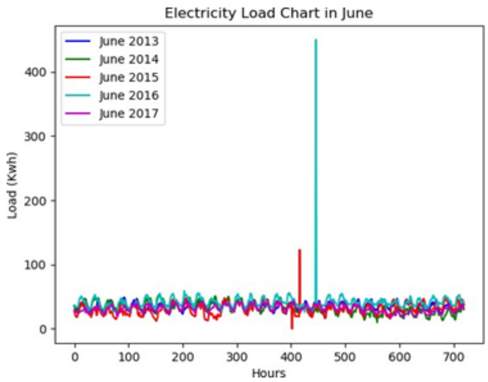
(c)



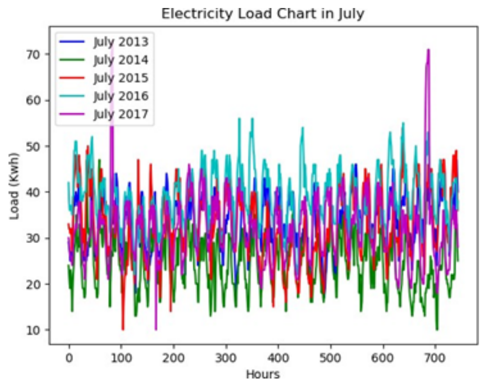
(d)



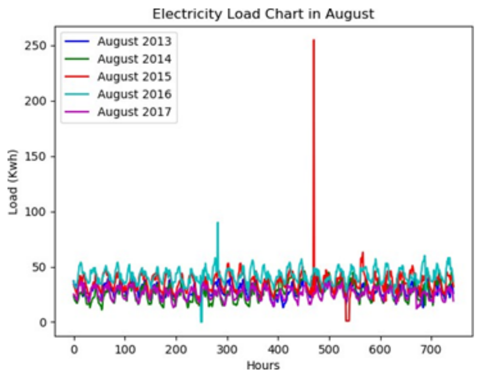
(e)



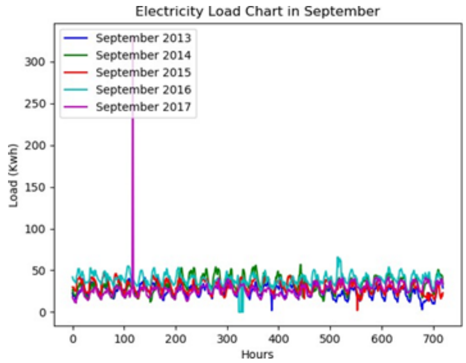
(f)



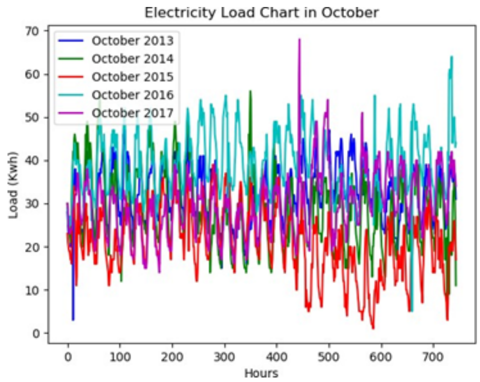
(g)



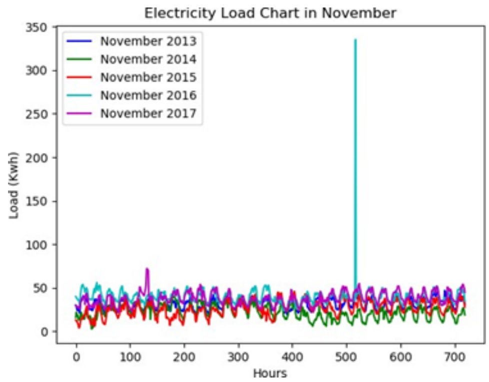
(h)



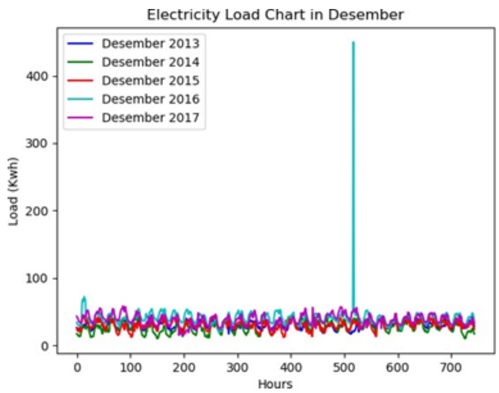
(i)



(j)



(k)



(l)

Fig. 5. Distribution Plot for each Months from 2013-2017

VI. CONCLUSIONS AND FUTURE WORKS

In this paper we aim to process the forecasting of electrical loads using the LSTM method. From the test results, the LSTM model with 2 layers has the smallest RSME value of 11.4299. However, it is also known that the RSME value for all models is relatively high (above 10). This is because there are differences in the pattern between training and testing data and there are data anomaly in several months. In the next study we will try to overcome the anomalies found in the dataset.

REFERENCES

[1] V. Mansouri and M. E. Akbari, "Efficient Short- Term Electricity Load Forecasting Using," vol. 3, no. 9, 2014.

[2] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, pp. 1–23, 2018.

[3] U. Ali, C. Buccella, and C. Cecati, "Households electricity consumption analysis with data mining techniques," *IECON Proc. (Industrial Electron. Conf.*, no. December 2018, pp. 3966–3971, 2016.

[4] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," *Energy Procedia*, vol. 14, pp. 438–444, 2012.

[5] K. Panklib, C. Prakasvudhisarn, and D. Khummongkol, "Electricity Consumption Forecasting in Thailand Using an Artificial Neural Network and Multiple Linear Regression," *Energy Sources, Part B Econ. Plan. Policy*, 2015.

[6] C. E. M. Luis and A. M. G. Marrero, "Real object mapping technologies applied to marine engineering learning process within a CBL methodology," in *Procedia Computer Science*, 2013, vol. 25, pp. 406– 410.

[7] M. Bicego, A. Farinelli, E. Grosso, D. Paolini, and S. D. Ramchurn, "On the distinctiveness of the electricity load profile," *Pattern Recognit.*, vol. 74, pp. 317–325, 2018.

[8] K. Gajowniczek and T. Zabkowski, "Two-stage electricity demand modeling using machine learning algorithms," *Energies*, vol. 10, no. 10, 2017.

[9] P. H. Kuo and C. J. Huang, "An electricity price forecasting model by hybrid structured deep neural networks," *Sustain.*, vol. 10, no. 4, pp. 1–17, 2018.

[10] S. Bouktif, A. Fiaz, A. Ouni, and M. A. Serhani, "Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches," *Energies*, vol. 11, no. 7, 2018.

[11] R. Perez-Chacon, R. L. Talavera-Llames, F. Martinez-Alvarez, and A. Troncoso, "Finding electric energy consumption patterns in big time series data," *Adv. Intell. Syst. Comput.*, vol. 474, pp. 231–238, 2016.

[12] R. Beigaite and T. Krilavičius, "Electricity price forecasting for Nord Pool data," *CEUR Workshop Proc.*, vol. 1856, no. November, pp. 37–42, 2017.

[13] Y. Zhu, R. Da, G. Liu, Z. Wang, and S. Lu, "Power market price forecasting via deep learning," *Proc. IECON 2018 - 44th Annu. Conf. IEEE Ind. Electron. Soc.*, pp. 4935–4939, 2018.

[14] S. Hochreiter, "The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions," *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.*, vol. 06, no. 02, pp. 107–116, 2003.

[15] W. L. Junger and A. Ponce de Leon, "Imputation of missing data in time series for air pollutants,"*Atmos. Environ.*, 2015.