2018 15th International Conference on
Control, Automation, Robotics and Vision (ICARCV)
Singapore, November 18-21, 2018

Day-Ahead Price Forecasting for Electricity Market using Long-Short Term Memory Recurrent Neural Network

LianLian Jiang[1] and Guoqiang Hu[2]

[1]Institute for Infocomm Research, A*STAR, Singapore, 138632
[2]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 639798
ljiang2@e.ntu.edu.sg, GQHu@ntu.edu.sg

***Abstract*—In this paper, an efficient method for the day-ahead electricity price forecasting (EPF) is proposed based on a long-short term memory (LSTM) recurrent neural network model. LSTM network has been widely used in various applications such as natural language processing and time series analysis. It is capable of learning features and long term dependencies of the historical information on the current predictions for sequential data. We propose to use LSTM model to forecast the day-ahead electricity price for Australian market at Victoria (VIC) region and Singapore market. Instead of using only historical prices as inputs to the model, we also consider exogenous variables, such as holidays, day of the week, hour of the day, weather conditions, oil prices and historical price/demand, etc. The output is the electricity price for the next hour. The future 24 hours of prices are forecasted in a recursive manner. The mean absolute percentage error (MAPE) of four weeks for each season in VIC and Singapore markets are examined. The effectiveness of the proposed method is verified using real market data from both markets. The result shows that the LSTM network outperforms four popular forecasting methods and provides up to 47.3% improvement in the average daily MAPE for the VIC market.**

***Index Terms*—Electricity price forecasting (EPF), long-short term memory neural network, multiple steps, energy market.**

I. Introduction

The changes from monopoly towards liberalized market regimes and the increasing complexity brought by policy targets for renewable energy and emissions, make electricity price forecasting (EPF) an important topic. Accurate price forecast can improve its bidding strategy and production or consumption schedule in order to reduce the risk or maximize the profits in day-ahead trading. The main factors affecting the price include electricity costs, demand and supply of the electricity in the market, and human factors (e.g. bidding strategies and market manipulation by energy monopolies). Due to these factors electricity prices generally exhibit seasonality, large fluctuations and spikes [1] and makes it quite challenging to accurately predict MCPs. In literature, existing EPF methods can be classified into seven main categories [2]: simulation models [3], multi-agent models [4, 5], statistical models [6–12], computational intelligence models [13–15], deep learning methods [16] and hybrid intelligent models [17, 18]. However, simulation models requires detailed system operation parameters and has high computational complexity to build the model. Multi-agent models like game theory are able to model the strategies of the market participants, however, they focus more on qualitative issues rather than quantitative results. Statistical models include similar-day methods [17], exponential smoothing methods (ESM) [6], general additive models (GAM) [7], regime-switching models (RSM) [19], jump-diffusion models (JDM) [20], etc. The similar-day method is simple but has low accuracy when large variations exist in price time series. ESMs perform better than the normal moving average method but it is difficult to determine the parameter of smoothing factors. GAM is a flexible nonlinear model and provides better estimation accuracy than conventional linear models. However, it works best for trends that are steady and systematic. RSMs and JDMs have the ability to take into account the asymmetry of the time series and even can simulate large spikes. However, they require accurate model parameters which are usually difficult to get. Time series based models have been widely used to predict the electricity price, such as autoregressive integrated moving average (ARIMA) [11], generalized autoregressive conditional heteroskedastic (GARCH) [21], etc., but they can be problematic when there are rapid variations and high frequency changes in the price.

Computational intelligence techniques are used in price forecasting due to its strong nonlinear modelling capabilities. They include artificial neural networks (ANNs), support vector machine (SVM), fuzzy logic (FL), etc. When directly applying ANN in forecasts, the main idea is to use ANN to conduct a nonlinear functional mapping from past observations (y_{t-1}, y_{t-2},..., y_{t-p}) to the future value, y_t. Many different kinds of ANNs have been applied in EPF, such as feed-forward neural network (NN) [22], recurrent NN [23], fuzzy NN [24], cascaded NN [25], weighted nearest neighbors (WNN) [26], adaptive wavelet NNs (AWNNs) [14], etc. Instead of using sole forecasting model, many hybrid intelligent system (HIS) methods, which combine two or more existing EPF methods, have been proposed to form an effective method to exhibit advantages of the individual technique, such as similar days method + ANN[17], fuzzy + ARIMA + ANN [27], ARFIMA + ANN [18], wavelet transform +ANFIS [28], ARIMA + GARCH [10, 11], ARIMA + wavelet transform [12], singular spectrum analysis (SSA) + modified wavelet NN (WNN) [29], etc. Among all the computational intelligence models, recurrent neural networks (RNN) have been widely used in time series predictions. To solve the problem of gradient exploding and vanishing during the learning process, the long short term memory (LSTM) network is proposed [30]. LSTM network is capable of learning long sequences with long time lags. To further improve the forecast accuracy, in this paper, we propose to use LSTM model for 24 hours ahead price forecasts due to the strong ability of the LSTM to memorize the previous price trend during training. Many exogenous factors are also considered as inputs to the network, including the forecasted system demand, historical prices, hour of the day, day of the week, week of the year, holidays information, etc. The effectiveness of the LSTM based prediction model is verified using both datasets from Australia market at Victoria region and Singapore market.

II. Methodology

For the electricity market, on day d-2 the market price at day d is forecasted after obtaining the bidding price for day d-1 at a particular hour h_c. The objective of the presented methodology is to predict the day ahead electricity price, given historical price data and exogenous variables.

A. Preprocessing

Fig.1 shows the high volatility of the electricity price in Australian market at the VIC region and the Singapore market. From Fig.1, we can see that both negative and extremely high prices appear in two markets. The negative prices can be caused by transmission constraints on the system [31]. The extremely high prices can be caused by shortages of power supply in the system. However, those negative and extreme values of the price occurs infrequently. For instance, for the VIC market during 5 years from *Jan* 1, 2010 to *Dec* 31, 2014, the possibility of price getting negative values is only 0.048$\%$ and the price higher than 1000 AUD/MWh is only 0.087$\%$, while for the Singapore market from *Jan* 1, 2013 to *Dec* 31, 2015, the negative prices only consist of 0.019$\%$ and the prices higher than 1000SGD/MWh only consist of 0.059$\%$. Therefore, to reduce the effect of abnormal events on the performance of the prediction we refine the negative and extreme prices into specific values. The negative prices are refined as 5 AUD/MWh for the VIC market and 5 SGD/MWh for the Singapore market, respectively. The prices higher than 1000 AUD/MWh in Australia market and the prices higher than 1000 SGD/MWh in Singapore market are interpolated by its neighbor prices, respectively. Due to the infrequentness of the price spikes, some authors use either raw price data or limits the spike amplitude to the mean plus/minus three times the standard deviation [16]. After redefining the prices, we use the natural logarithm of the data to do the prediction and it is defined as:

$$L_t^d = ln(P_t^d) \tag{1}$$

where, P_t^d is the electricity price of the day d at time step t.

B. LSTM neural network

In RNN, the backpropagation (BP) is commonly used by the gradient descent optimization algorithm to adjust weights between network layers during training. However,

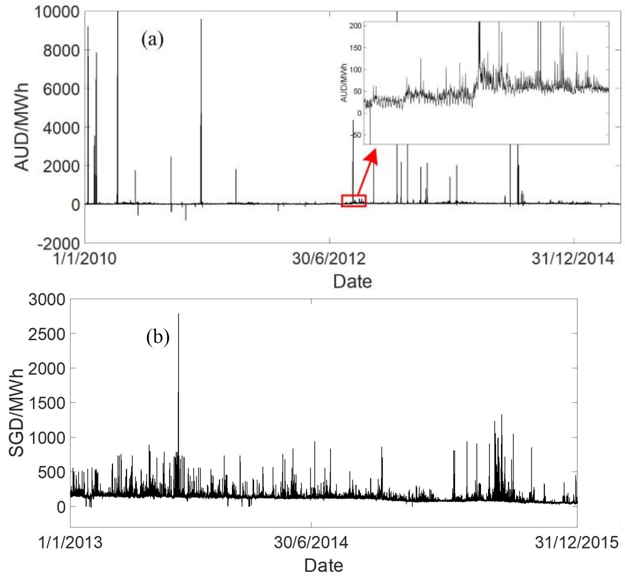

Fig. 1. The volatility of the electricity price of a) Australian market in VIC region and b) Singapore market.

the weights update scheme may stop the neural network from further training. This is because the gradient with a value within a range of (0, 1) (larger than 1) becomes extremely small (large) after a long chain, and thus causing a gradient vanishing(gradient exploding) problem. To solve this problem, Hochreiter and Schmidhuber proposed a LSTM network in 1997 [30]. The key idea behind the LSTM is to regulate the cell states using different types of gates, namely input, forget, and output gates. As shown in Fig.2, the state of each cell (c_{t-1}) passes through the LSTM cell to generate a state for the next step (c_t). Along the state flow line, the state information is controlled by adding or element-wise multiplying the results of inputs and the hidden state generated in the previous step after going through *sigmoid* or *tanh* neural network layers.

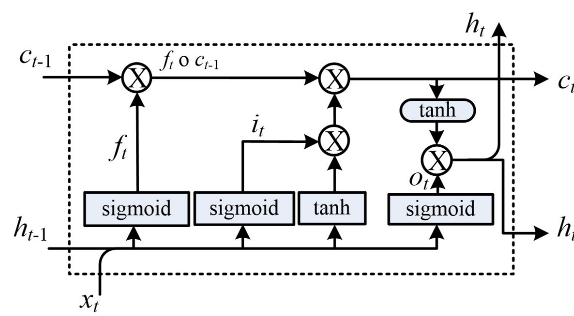

Fig. 2. The detailed structure within a LSTM cell.

The mathematical functions of three gates are defined as:

$$i_t = sigmoid(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{2}$$

$$f_t = sigmoid(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{3}$$

$$o_t = sigmoid(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{5}$$

$$h_t = o_t \odot tanh(c_t) \tag{6}$$

where i_t is the input gate which controls how much information of input (x_t) and previous hidden state (h_{t-1}) is allowed to pass into the memory cell, f_t is the forget gate which controls how much information is forgotten before passing though the cell, o_t is the output gate which controls how much information from the current memory cell can be output to the hidden state, c_t represents the cell state generated as an additional variable for the cell, W is the weight matrix and b is the biases to each layer. The symbol '$\odot$' represents the operation of element-wise multiplication. Each gate can be considered as a neural network layer. The information only comes from the current input x_t and the hidden state in the previous step h_{t-1}. The gate does not provide additional information to the cell but only has a function of limiting the amount of information passing through the memory cell. Compared to the original RNN with a hidden state update equation $h_t = tanh(W_{xh}x_t + W_{hh}h_{t-1} + b)$, the hidden state of the LSTM with an update equation (5) is generated based on the intermediate cell state. LSTM cells can also be stacked to form a multi-layer structured network.

The training algorithm of LSTM is called backpropagation (BP) through time (BPTT) [30] and it has similar scheme as the basic BP algorithm. Once the LSTM netowrk is trained, the weights and biases of the model can be obtained so that the model can be used for forecasting the electricity price.

III. ELECTRICITY PRICE FORECASTING USING LSTM

Due to the advantages stated above, we propose to apply stacked LSTM with multiple layers to predict the electricity price. The performance of the model can be influenced by the number of LSTM layers, input time steps, the structure of the forecasting manner, and input variables. In this section, we describe details of these factors in the LSTM model. The actual price values at day d are denoted as:

$$P^d = \left\{p_1^d, p_2^d, ..., p_t^d, ..., p_T^d\right\}. \tag{7}$$

The predicted price values at day d are represented as:

$$\hat{P}^d = \left\{\hat{p}_1^d, \hat{p}_2^d, ..., \hat{p}_t^d, ..., \hat{p}_T^d\right\} \tag{8}$$

where $\hat{p}_t^d$ is the predicted price at time step t. T can be 24 for a hourly market and 48 for a half hourly market.

Inputs and output: Various variables can be used as inputs, such as historical prices or loads, weather conditions, holidays, status of the day, oil prices, etc. After the correlation analysis, a vector with multiple steps, formed by multiple variables such as historical prices at previous steps, hour of the day, day of the week, holidays, the maximum temperature, the system load, etc., are taken as inputs to the LSTM. The input vectors for one day d is defined as:

$$I^d = \left\{i_1^d, i_2^d, ..., i_t^d, ..., i_T^d\right\} \tag{9}$$

where I^d is formed by the input vectors (i_t^d) for each step. The features to form the input vector i_t^d is represented by:

$$i_t^d = \left[hd_t^d, dw_t^d, wd_t^d, ll_t^d, mt_t^d, op_t^d, al_t^d, ap_t^d, lp_t^d\right] \tag{10}$$

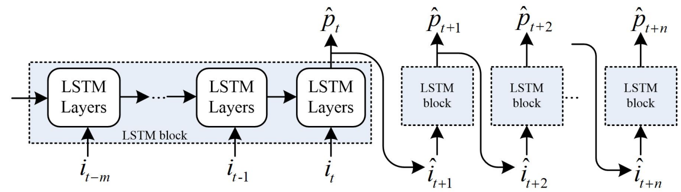

Fig. 3. The structure of the LSTM network based model in a recursive manner for multiple steps forecasting.

where hd_t^d, dw_t^d, wd_t^d, ll_t^d, mt_t^d, op_t^d, al_t^d, and lp_t^d, are the hour of the day, day of the week, flag of the working day, lagged system loads, maximum temperature, oil prices, average load of last 24 hours, average price of last 24 hours, and lagged historical prices, respectively. i_t^d is the input vector at the tth time step. In this model, m time steps are used for predicting the price in the next step. A price series with m time steps means that the data information in previous m steps is memorized during the training process of LSTM, and the trained LSTM model will be used for forecasting the price value for the next step. The input to the LSTM cell with a time step of m is defined as:

$$i_{t(m)}^d = \left\{i_{t-m}^d, i_{t-m+1}^d, ..., i_{t-2}^d, i_{t-1}^d, i_t^d\right\}. \tag{11}$$

In the input vector, the system demand up to one hour ahead the forecast hour are assumed to be known. This is reasonable because in recent years the short-term load forecasting has reached a good accuracy with about 0.621% of MAPE [32]. Therefore, the forecasted demand values for the next day can be used.

In the input vector, the previous price information within a certain number of steps are required as represented in (10). However, the actual prices for the whole target day (to be forecasted) is not available until the bidding is completed in the next day. Therefore, for multiple steps ahead prediction, we use recursive forecast method, in which the price at the (t+1)th step is predicted using the price forecasted in the last step. This process is repeated until the predefined length of prices is reached. Fig.3 shows the structure of using LSTM with multiple layers to forecast the electricity prices with n steps ahead recursively. As shown in Fig.3, the price output at tth step, p_t, is used to form the input vector at next step, i_{t+1}. Similarly, after the LSTM model outputs the price at the (t+1)th step, p_{t+1} is then continued to be used for forcasting the price value at the (t+2)th step. In our case, m and n are set to 20 and 24, respectively. This model can memorize series information within m previous steps. The Adam algorithm is selected as the optimizer due to its higher accuracy and faster convergence than SGD [33]. The weights and biases are obtained by minimizing the objective function, which is expressed as:

$$Loss = \sum_{t=1}^{T}(P_t - \hat{P}_t)^2. \tag{12}$$

The main procedure of the steps are described as follows.

Step 1: (*Data collection*) Collect datasets (including historical electricity prices, system demands, maximum temperature, oil prices, etc.) and prepare the corresponding number of the day/week values for each dataset.

Step 2: (*Initialization*) Initialize parameters for LSTM such as the time steps (m), number of hidden layers/neurons, activation functions, the optimizer, the batch size and epochs for training, etc; Split datasets into two parts for training and validation.

Step 3: (*Training LSTM*) Train LSTM network using three months historical data.

Step 4: (*One step ahead forecasting*) Forecast electricity price of one step ahead.

Step 5: (*Update inputs*) Update input variables based on the predicted price value in the last step; Check the hours of forecasts, if 24 hours of forecasted value are obtained, then go to Step 6. Otherwise, update the number of hour and go back to Step 4 to forecast the price value for the next step.

Step 6: (*Forecast prices for the next day*) Calculate MAPE values for each day, and check whether the whole week of forecast is completed. If it is completed, then go to Step 7, otherwise update the number of day and go back to Step 3 to continue forecasting the price value for the next day.

Step 7: (*Output results*) Calculate the average MAPE for the whole week and output forecast resuts.

IV. Case Study

In this section, we examine the effectiveness of the model for both Australia VIC and Singapore electricity markets. The methodology is implemented in Python 3.5.4.

A. Australian market in VIC

Data preparation: Datasets of VIC market used in the experiments are downloaded from [34]. The price data and exogenous variables from three months (90 days) prior to the first day of each test week are used as training datasets, and 20% of the training data is used for validation. Four one-week periods representing four different seasons in 2013 (*Oct* 21 to *Oct* 27, *Feb* 4 to *Feb* 10, *Apr* 22 to *Apr* 28 and *Jul* 22 to *Jul* 28) are selected from VIC market for testing. Since the price series shows strong periodical character weekly, monthly and seasonally, the electricity prices at the present day are highly correlated with the historical prices. In our model, we use lagged price set $[p_{t-1}, p_{t-2}, p_{t-3}, p_{t-24}, p_{t-25}, p_{t-48}, p_{t-49}, p_{t-72}, p_{t-73}, p_{t-96}, p_{t-97}, p_{t-120}, p_{t-121}, p_{t-144}, p_{t-145}, p_{t-168}, p_{t-169}, p_{t-192}, p_{t-193}]$ as part of input features for the VIC electricity market in Australia. Besides, the exogenous variables, including the system demand with the same lagged steps as prices (19 dimensions (D)), the hour of day (1D), the day of week (1D), the flag of the working day (1D), the maximum temperature in the region (1D), the average system load of last 24 hours (1D), and the average electricity price of last 24 hours (1D), are also taken into account. Therefore, the total dimension of the input features is 44 (19+19+1+1+1+1+1+1). The logarithm values of both prices and loads are used for both markets. With above training datasets, the dimension of the input data to the LSTM is set to (2160, 20, 44), which represents (number of samples, number of time steps, number of input features).

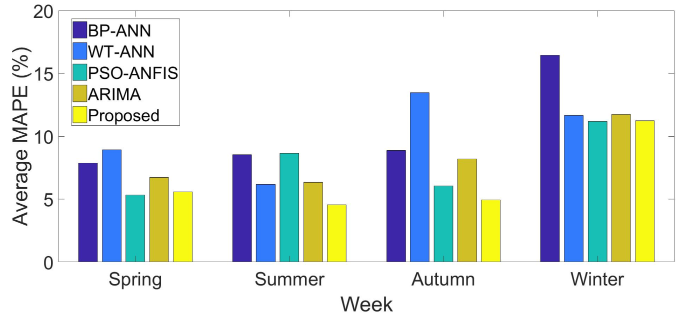

Fig. 4. The average MAPE of each week in four different seasons produced by the proposed method and other four popular forecasting methods.

The output of the model is the price value for one step ahead and has a dimension of (2160, 1), which means (number of samples, number of output features).

Building prediction model: The time step of the input variable (m) is set as 20 based on the trial and error method. A structure with three stacked layers of LSTM cells and a fully connected output layer is applied. The activation function for the hidden layer is set as *sigmoid*. The number of hidden neurons for each LSTM is given as 30, and thus the LSTM has a structure of [30, 30, 30, 1]. The batch size for the training is set to 20. A rolling price window with a size of 193 is defined during the recursive prediction process. When the new price value is available, the lagged price set is updated by discarding the original lagged price at 193th step and adding the new price value, predicted by the forecaster in the previous step, to the first step of rolling price window. Simultaneously, input variables for predicting the next price value are also updated. The network is trained daily using available datasets.

Results and discussion for VIC market: In the experiment, we use the mean absolute percentage error (MAPE) to evaluate the accuracy of the forecasting. It is defined as:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{p}_i - p_i|}{|p_i|} \quad (13)$$

where $\hat{p}_i$ is the predicted electricity price, and p_i is the actual electricity price. To calculate the MAPE of the prediction for one week, the LSTM based prediction model for each day is trained separately. For each prediction day, the model is trained, which means that for one week, the training is executed for 7 times. The model is trained 10 times and the best model with the minimum forecasting error for each day is used for calculating the MAPE for the whole week.

The results is compared with four popular models, namely conventional backpropergation (BP) based multilayer feed-forward network (BP-ANN), adaptive neuro fuzzy inference system optimized with particle swarm optimization algorithm (PSO-ANFIS), ANN with the wavelet transformation (WT-ANN), and seasonal ARIMA (SARIMA).

After applying the proposed method and existing methods stated above to the day-ahead electricity price forecasting for the VIC market, we observe that the proposed LSTM based method performs much better than other four popular

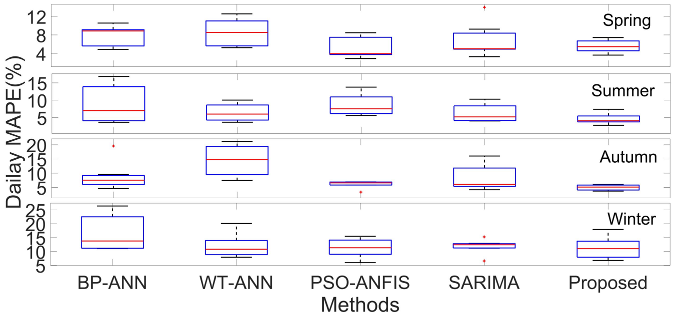

Fig. 5. The boxplots of daily MAPE for four season weeks in the Australia VIC market using five different methods.

methods in most seasons. Table I shows the comparison of the average MAPE for four weeks at different seasons between the compared methods and the proposed LSTM based method. From Table I, we can see that the proposed method has better performance than the other four methods during all these four seasons. Even though the PSO-ANFIS provides slightly lower average MAPEs during spring and winter weeks, it exhibits much higher MAPEs for other two season weeks, where it is 47.3% and 17.5% higher than that by the proposed method during the summer and autumn week, respectively. We can also see that for the other four methods, there is no such a method which provides the best performance for all the cases. The bar plot of the average MAPE for each week in four different seasons is shown in Fig. 4, which shows that all the methods provide higher forecasting error during the winter week than the rest weeks. This is due to the large fluctuations of the electricity demands in winter. The boxplots of the daily MAPE for each season are presented in Fig. 5, from which we can see that in the spring week, although the PSO-ANFIS method can provide slightly lower medium value of daily MAPE marked with the red line, the proposed method shows much smaller distribution distance of the error (length of the box). In the summer week, the proposed method provides not only a smaller medium daily MAPE but also a smaller distribution distance. In the autumn week, both a lower medium MAPE and a smaller error distribution distance are obtained. In the winter week, the proposed method shows a similar performance as the PSO-ANFIS model. The real and predicted prices by all the methods for four season weeks are presented in Fig. 6.

B. Singapore market

Building prediction model: Singapore is located near the equator. It has a typically tropical climate with high and uniform temperatures, abundant rainfall and high humidity all year round. Thus, the electricity utilization patterns do not change much around the year. Therefore, we randomly choose one week, from *Dec* 29^{th}, 2014 to *Jan* 4^{th}, 2015 to illustrate the effectiveness of the LSTM network based forecasting model. It has been found that lower fuel oil prices tend to shift energy offer prices into the lower price bands [35]. Therefore, we include the average oil price. The logarithm values for prices, demands, the maximum

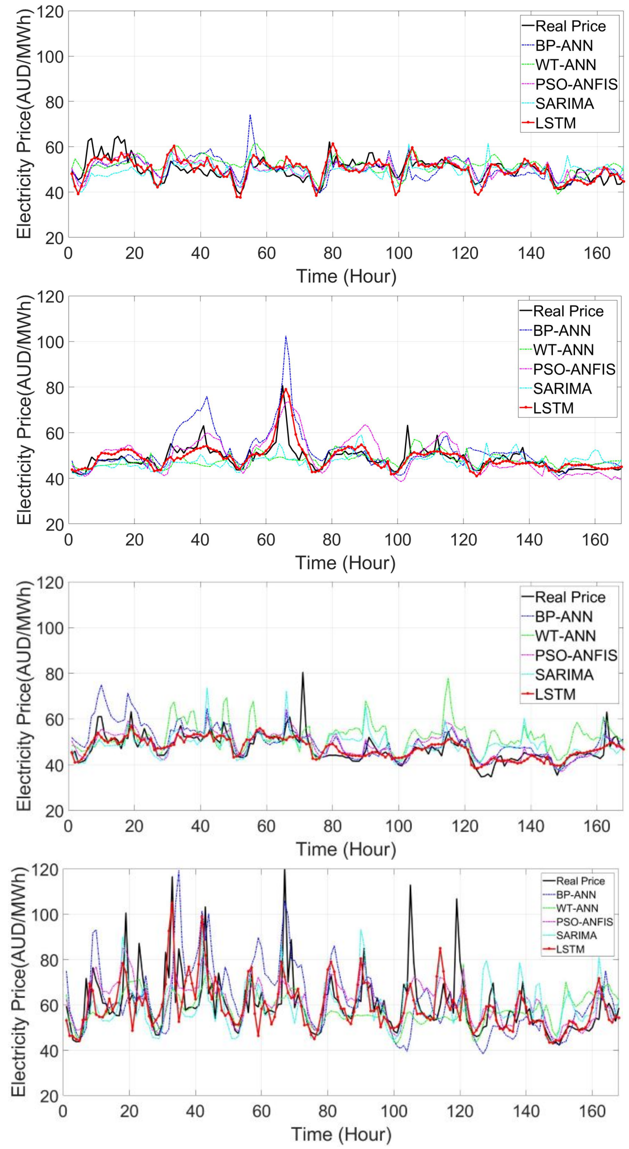

Fig. 6. The comparison between the actual electricity prices and forecasted prices by five methods for the (a) spring week, (b) summer week, (c) autumn week and (d) winter week in VIC market.

temperature, and the average value of oil prices are applied. The dimension of the input data to the LSTM is (2160, 20, 45). The output of the LSTM model is the recursive price values for the next step and its dimension is (2160, 1). The structure of the LSTM based model is set as [50, 50, 50, 1].

Results and discussion for Singapore market: For the comparison, all the parameters for other compared methods are adjusted to the best combination values so that it can generate the best accuracy. After applying all the exisiting methods and the proposed methods, we get the average MAPE of the tested week as shown in Table II, from which we can see that the proposed method show lower average MAPE than other four methods. The boxplot of the daily MAPE of the tested week is shown in Fig. 7. It illustrates that both the distribution of the error and the medium value is lower than other four methods. The comparisons of real and the predicted prices by different methods for the test week

TABLE I
THE COMPARISON OF AVERAGE MAPE FOR FOUR WEEKS DURING DIFFERENT SEASONS BY THE PROPOSED METHOD AND FIVE EXISTING METHODS FOR AUSTRALIA VIC MARKET IN 2013.

Seasons	**Test dataset**	**Multilayer ANN (%)**	**WT-ANN (%)**	**PSO-ANFIS (%)**	**SARIMA (%)**	**Proposed (%)**
Spring (Sep-Nov)	Oct21 - Oct27	7.86	8.93	5.34	6.73	5.58
Summer (Dec-Feb)	Feb4 - Feb10	8.55	6.17	8.65	6.34	4.56
Autumn (Mar-May)	Apr22- Apr28	8.86	13.48	6.05	8.19	4.95
Winter (Jun-Aug)	Jul22 - Jul28	16.46	11.67	11.19	11.74	11.25

TABLE II
THE COMPARISON OF THE AVERAGE MAPE BY FOUR EXISTING METHODS AND THE PROPOSED METHOD FOR THE SINGAPORE MARKET.

Test dataset	**Multilayer ANN(%)**	**WT-ANN(%)**	**PSO-ANFIS(%)**	**SARIMA(%)**	**Proposed(%)**
Dec 29, 2014 - Jan 4, 2015	12.40	16.07	13.49	10.53	10.23

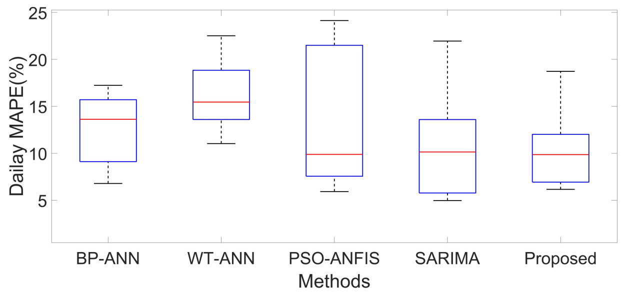

Fig. 7. The boxplots of the daily MAPE by different methods for the test week in the Singapore market.

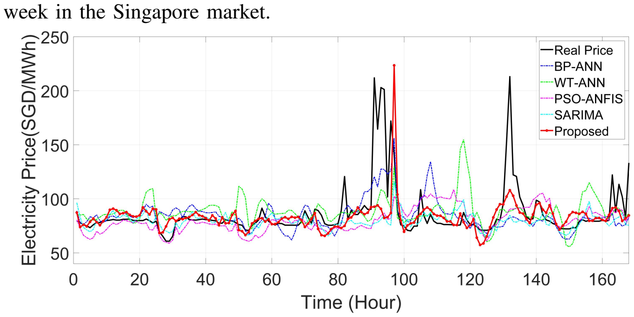

Fig. 8. The comparison of results by four existing methods and the proposed method for one week for Singapore market.

is shown in Fig 8.

V. CONCLUSION

In this work, we proposed a multilayer LSTM based model for forecasting the day-ahead electricity prices due to its ability to bridge long time lags of inputs and remembering the historical trend information in time series. The effectiveness of the proposed method is verified with four weeks of price datasets selected from four seasons in VIC Australia market and one week of dataset from Singapore market. For the VIC market, 44 features are considered as the inputs of the model. For the Singapore market, the average oil price is also taken into account as well. The 24 hours of future price values are forecasted in a recursive manner. The preprocessing data are implemented and the logarithm values of variables are utilized. The performance of the proposed method is compared with other four popular methods used in the market, namely BP-ANN, WT-ANN, PSO-ANFIS, and SARIMA. The results show that for both markets, the LSTM model outperforms other compared methods and the improvement in the average daily MAPE reaches up to 47.3%.

VI. ACKNOLEDGEMENT

This research was supported by Singapore economic development board under EIRP grant S14-1172-NRF EIRPIHL.

REFERENCES

[1] l. Cartea and M. G. Figueroa, "Pricing in electricity markets: A mean reverting jump diffusion model with seasonality," *Applied Mathematical Finance*, vol. 12, no. 4, pp. 313–335, 2005.
[2] L. L. Jiang and G. Hu, "A review on short-term electricity price forecasting techniques for energy markets," *Accepted by The 15 th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, 2018.
[3] C. Batlle and J. Barquin, "A strategic production costing model for electricity market price analysis," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 67–74, 2005.
[4] B. E. Hobbs, "Linear complementarity models of nash-cournot competition in bilateral and poolco power markets," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 194–202, 2001.
[5] S. Haili, L. Chen-Ching, and J. Lawarree, "Nash equilibrium bidding strategies in a bilateral electricity market," *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 73–79, 2002.
[6] T. Jonsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Forecasting electricity spot prices accounting for wind power predictions," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 210–218, 2013.
[7] P. Gaillard, Y. Goude, and R. Nedellec, "Additive models and robust aggregation for gefcom2014 probabilistic electric load and electricity price forecasting," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1038–1050, 2016.
[8] K. Maciejowska, J. Nowotarski, and R. Weron, "Probabilistic forecasting of electricity spot prices using factor quantile regression averaging," *International Journal of Forecasting*, vol. 32, no. 3, pp. 957–965, 2016.
[9] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055–2062, 2012.
[10] S. J. Koopman, M. Ooms, and M. A. Carnero, "Periodic seasonal reg-arfimagarch models for daily electricity spot prices," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 16–27, 2007.
[11] Z. Tan, J. Zhang, J. Wang, and J. Xu, "Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models," *Applied Energy*, vol. 87, no. 11, pp. 3606–3610, 2010.
[12] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and arima models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, 2005.
[13] N. A. Shrivastava, A. Khosravi, and B. K. Panigrahi, "Prediction interval estimation of electricity prices using pso-tuned support vector machines," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 322–331, 2015.
[14] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network-based energy price forecasting in electricity markets," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1423–1432, 2008.
[15] N. Amjady and M. Hemmati, "Day-ahead price forecasting of electricity markets by a hybrid intelligent system," *International Transactions on Electrical Energy Systems*, vol. 19, no. 1, pp. 89–102, 2008.
[16] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, 2018.
[17] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, and A. K. Srivastava, "A novel approach to forecast electricity price for pjm using neural network and similar days method," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2058–2065, 2007.
[18] N. Chabane, "A hybrid arfima and neural network model for electricity price prediction," *International Journal of Electrical Power and Energy Systems*, vol. 55, pp. 187–194, 2014.
[19] R. Ethier and T. Mount, "Estimating the volatility of spot prices in restructured electricity markets and the implications for option values," *PSerc Working Paper 98-31*, 1998.
[20] R. Weron, M. Bierbrauer, and S. Trck, "Modeling electricity prices: jump diffusion and regime switching," *Physica A: Statistical Mechanics and its Applications*, vol. 336, no. 1, pp. 39–48, 2004.
[21] R. C. Garcia, J. Contreras, M. v. Akkeren, and J. B. C. Garcia, "A garch forecasting model to predict day-ahead electricity prices," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 867–874, 2005.
[22] J. P. S. Catalo, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Electric Power Systems Research*, vol. 77, no. 10, pp. 1297–1304, 2007.
[23] S. Anbazhagan and N. Kumarappan, "Day-ahead deregulated electricity market price forecasting using recurrent neural network," *IEEE Systems Journal*, vol. 7, no. 4, pp. 866–872, 2013.
[24] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 887–896, 2006.
[25] Z. Li, P. B. Luh, and K. Kasiviswanathan, "Energy clearing price prediction and confidence interval estimation with cascaded neural networks," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 99–105, 2003.
[26] A. T. Lora, J. M. R. Santos, A. G. Exposito, J. L. M. Ramos, and J. C. R. Santos, "Electricity market price forecasting based on weighted nearest neighbors techniques," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1294–1301, 2007.
[27] O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, and M. Pasadas, "Hybridization of intelligent techniques and arima models for time series prediction," *Fuzzy Sets and Systems*, vol. 159, no. 7, pp. 821–845, 2008.
[28] J. P. S. Catalao, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-pso-anfis approach for short-term electricity prices forecasting," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 137–144, 2011.
[29] L. Xiao, W. Shao, M. Yu, J. Ma, and C. Jin, "Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting," *Applied Energy*, vol. 198, pp. 203–222, 2017.
[30] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[31] "Introduction to the national electricity market of singapore," *Energy Market Authority of Singapore*, 2009.
[32] M. K. M. A. A. P. Vu, D. H. and A. Bouzerdoum, "Short-term electricity demand forecasting using autoregressive based time varying model incorporating representative data adjustment," *Applied Energy*, vol. 205, pp. 790–801, 2017.
[33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[34] "http://www.aemo.com.au/electricity/national-electricity-market-nem/data-dashboard." Accessed in 2018.
[35] "Market report," *National Electricity Market of Singapore*, 2015.