2023 International Conference on System Science and Engineering (ICSSE) | 979-8-3503-2294-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSSE58758.2023.10227203

Short-Term Load Forecasting in Power System Using Recurrent Neural Network

Truong Hoang Bao Huy[1], Dieu Ngoc Vo[2,4], Hung Duc Nguyen[3,4], Hoa Phuoc Truong[3,4], Khanh Tuan Dang[2,4], Khoa Hoang Truong[3,4,*]

[1]*Department of Future Convergence Technology, Soonchunhyang University, Asan-si, Chuncheongnam-do 31538, South Korea*

[2]*Department of Power Systems, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam*

[3]*Department of Power Delivery, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam*

[4]*Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc City, Ho Chi Minh City, Vietnam*

*Corresponding author: trhkhoa@hcmut.edu.vn

***Abstract*—As energy demand increases rapidly, short-term load forecasting is becoming progressively vital in power system dispatch and demand response. This study proposes a short-term load forecasting approach for the power system in Vietnam. In this regard, a gated recurrent unit-based deep learning model is applied to use the historical load sequences to forecast the single-step and multi-step ahead values of the load consumption. The hourly load consumption dataset is provided by Ho Chi Minh City Power Corporation (EVNHCMC). Simulation results prove the effectiveness of the developed prediction algorithm for short-term load forecasting, especially for multi-step forecasting.**

***Keywords*—Short-term load forecasting, deep learning, gated recurrent unit, long short-term memory**

I. Introduction

Short-term load forecasting is an indispensable aspect of energy management, which involves predicting the future load demand for a given period, typically ranging from a few minutes to a few weeks ahead. For power grids to remain stable and reliable, it is crucial to have accurate short-term load forecasting that helps utilities to make informed decisions regarding the generation, transmission, and distribution of electricity. Short-term load forecasting has also become more challenging due to the increasing penetration of renewable energy sources, which are highly variable and difficult to predict.

Short-term load forecasting has been a topic of interest for several decades, and various techniques have been developed to address this problem. Conventionally, simple statistical methods, such as regression analysis [1], autoregressive integrated moving normal (ARIMA) [2], and exponential smoothing [3], were commonly used for load forecasting. However, with the increasing complexity of the power system and the availability of large amounts of data, machine learning algorithms have been developed, such as artificial neural network (ANN) [4]–[6], support vector machine (SVR) [7], and fuzzy logic [8]. However, it has been recognized that deep neural networks (DNNs) are more efficient for time series predictions compared to shallow ANNs [9]. In Ref. [10], the authors applied echo state networks (ESN) with a recursive multi-step approach for short-term forecasting. Forecasting methods were proposed in [11] and [12] using long short-term memory (LSTM) neural network to predict the short-term aggregated load. Chen et al. [13] utilized a hybrid of DNN and shallow ANN to predict a day-ahead load profile. Sun et al. [14] applied a forecasting model based on a wavelet neural network (WNN) that takes historical load data and weather forecast variables as input features. A multistage ANN-based short-term load forecaster was suggested based on temperature forecast data in [15].

Motivated by the literature review, this study proposes short-term load forecasting models based on the gated recurrent unit (GRU) architecture. Single-step and multi-step forecasting strategies are both considered in this study. The model incorporates a rolling time-index series as input features, consisting of the hour index, weekday/weekend index, and month index. This inclusion of features is demonstrated to result in a significant improvement in forecasting accuracy. Moreover, the study suggests that utilizing a specific sequence length could lead to higher prediction accuracy. The study also compares the proposed GRU model with other models using error metrics, which shows that the GRU model has higher accuracy and better performance than compared models.

II. Gated Recurrent Unit Network

A gated recurrent unit (GRU) neural network (NN) is a variant of well-known long short-term memory (LSTM) architecture, which was introduced by Cho et al. in 2014 [16]. GRU has a reduced number of parameters, which leads to a shorter training time when compared to traditional LSTM. GRUs are made up of units, which are essentially simplified versions of LSTM cells. Each unit contains two gates: a reset gate and an update gate. Fig. 1 depicts the architecture of the GRU unit. The reset gate of the GRU network adapts to incorporating new input information with the previous state memory. Meanwhile, the update gate regulates the retention of the previous state memory. The mathematical operation of the GRU network is formulated as follows:

$$r_t = \sigma(\alpha_r x_t + \beta_r h_{t-1} + b_r) \tag{1}$$

$$z_t = \sigma(\alpha_z x_t + \beta_z h_{t-1} + b_z) \tag{2}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{3}$$

$$\tilde{h}_t = \tanh(\alpha_h x_t + \beta_h (r_t \odot h_{t-1}) + b_h) \tag{4}$$

$$y_t = g(h_t) \tag{5}$$

where x_t is the input vector; r_t and z_t are the reset gate and update gate, respectively; α and β are network parameters; b is the bias of network, $\tilde{h}_t$ and h_t are the candidate hidden state and the current hidden state, respectively; σ and $\tanh$ are the non-linear activation function; $\odot$ represents an element-wise multiplication, and y_t represents the output of GRU.

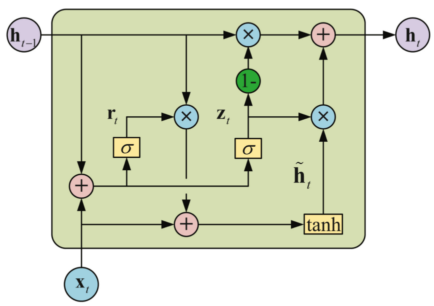

Fig. 1. The architecture of GRU.

III. GRU-Based Prediction Model

The proposed load forecasting method can be briefly described as follows:

A. Data collection and preprocessing

Historical data related to load consumption is gathered from the electric power company, followed by the examination of null values to verify the accuracy and reliability of the data.

B. Feature generation

New features are generated including the hour index $\{H = 1, 2, ..., 24\}$, weekday index $\{W = 0, 1, 2\}$, and month index $\{M = 1, 2, ..., 12\}$. The weekday index distinguishes weekdays and weekends, as well as makes a distinction between Saturdays and Sundays. In particular, for a given hour, the value of 'weekend' is equal to '0', if the hour is on a weekday, equal to '1' if the hour is on a Saturday, and equal to '2' if the hour is on a Sunday.

C. Normalization

The min-max method is used to normalize features as follows:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

D. Constructing Time Series

The input sequence length T is defined. Subsequently, the normalized input feature sequences are structured as follows: the hourly electricity load of past T timesteps is $L = \{l_t, l_{t-1}, ..., l_{t-T}\} \in R^T$; the hour index of past T timesteps is $H = \{h_t, h_{t-1}, ..., h_{t-T}\} \in R^H$; the weekday index of past T timesteps is $W = \{w_t, w_{t-1}, ..., w_{t-T}\} \in R^W$; the month index of past T timesteps is $M = \{m_t, m_{t-1}, ..., m_{t-T}\} \in R^M$. The normalized input features are transformed in the shape of a matrix: [samples, timesteps, features].

E. Building Forecasting Model

The DNN is proposed by stacking two GRU layers for short-term load forecasting, each GRU layer contains 100 neutrons with an activation function of rectified linear unit (ReLU). The GRU-based DNN model is trained with Adam optimizer to minimize the loss function of mean square error (MSE).

- For the single-step forecasting model, the output layer generates a single output value $Y_t = \{l_{t+1}\}$, which represents the forecasted load for the next time step.
- For the multi-step forecasting model, the output layer generates the forecasted output sequence $Y_t = \{l_{t+1}, l_{t+2}, ..., l_{t+N}\} \in R^T$, which represents the forecasted load for the next N time steps.

IV. Simulation Results

A. Simulation Setup

The proposed short-term load forecasting is applied to the power system in Vietnam. In this study, the historical time series data of the hourly load consumption of Ho Chi Minh City for 2011–2012 is provided by Ho Chi Minh City Power Corporation (EVNHCMC). Fig. 2 depicts the electricity consumption profile under study, whereas the maximum electricity load demand is 2876.83 MW. The dataset has 17544 samples with one-hour sampling. For the training process of the model, data from the first 16 months (January 2011 – April 2012) are used, while data from the next 4 months (May 2012 – August 2012) are used for validation. Furthermore, the performance of the proposed forecasting model is tested on data from the last four months of 2012 (September 2012 – December 2012). Simulation results and relevant analysis are presented in the next subsections.

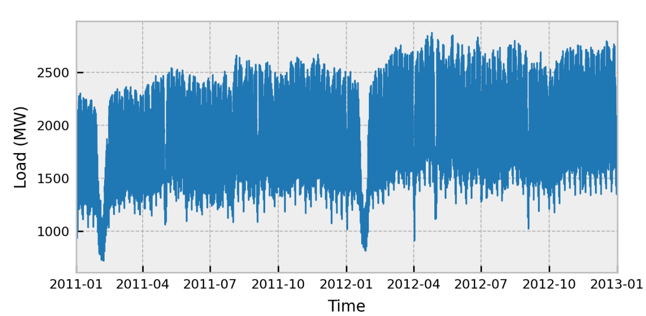

Fig. 2. Hourly consumption data of Ho Chi Minh City for 2011–2012.

B. Performance Evaluation Metrics

In this study, forecasting results are evaluated using the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{7}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{8}$$

$$\mathrm{RMSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{9}$$

where y_n is the actual value, $\hat{y}_n$ is the predicted value, and N is the size of the data samples.

C. Single-step forecasting results

The effectiveness of the forecasting model can be impacted by the length of the input sequence, which refers to the number of time steps in the look-back time window. Hence, the forecasting model is tested using various input sequence lengths, as given in Table I. Table I shows that the smallest error is obtained by using 72-step sequences. Therefore, the proposed GRU NN is trained and tested using look-back windows of 72 steps. The accuracy of the predictions over different months is shown in Table II. According to Table II, the smallest forecasting error occurs in November.

TABLE I
SINGLE-STEP PREDICTION RESULTS FOR DIFFERENT INPUT SEQUENCE LENGTHS

Timestep	MAE (MW)	MAPE (%)	RMSE (MW)
2	66.1248	3.2091	94.7110
4	33.7533	1.6866	49.2690
8	30.6031	1.5283	44.5426
12	31.6399	1.5785	46.7834
16	31.5746	1.5767	46.1332
20	30.5898	1.5502	44.8362
24	29.6143	1.4821	44.1199
48	33.6938	1.6587	48.0898
72	29.3431	1.4647	43.3596
120	30.7914	1.5335	45.2419
144	30.2482	1.5264	44.4299
168	30.4350	1.5033	44.9432

Fig. 3 illustrates the single-step forecasting results for the last four months of 2012. In addition, single-step forecasting results for typical days of these months are shown in Fig. 4. The results from Table II, Fig. 3, and Fig. 4 show that the predicted load closely aligns with the actual data, demonstrating the effectiveness of the proposed method in providing accurate single-step load forecasting.

D. Multi-step forecasting results

For the multi-step forecasting model, the proposed GRU NN is used to predict the future 24-hour horizon. Table III presents a comparison of the performance of the multi-step forecasting model based on varying input sequence lengths. As

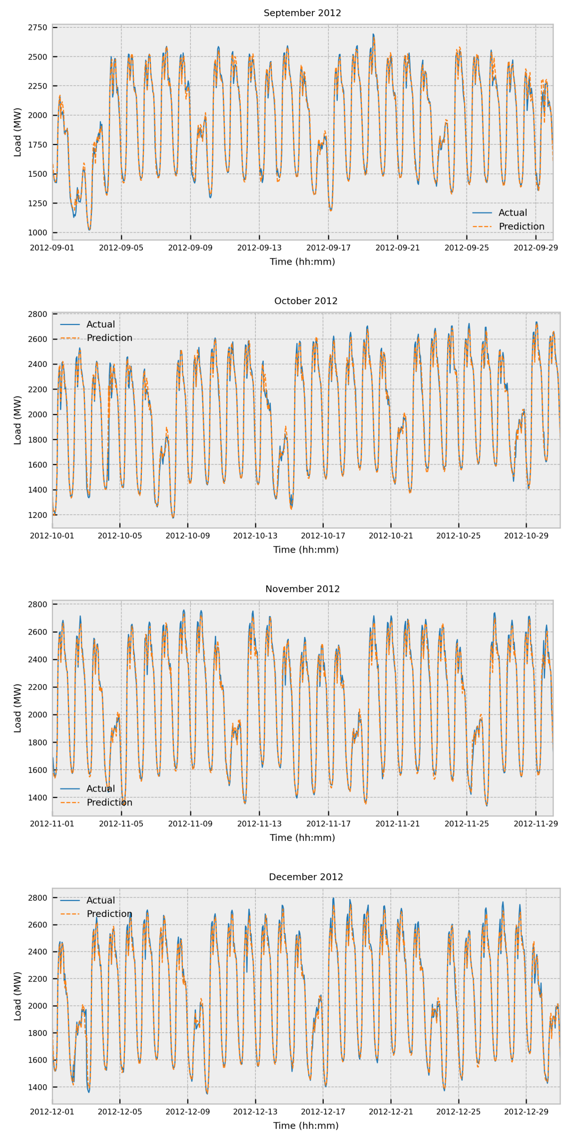

Fig. 3. Single-step prediction results for the last four months of 2012.

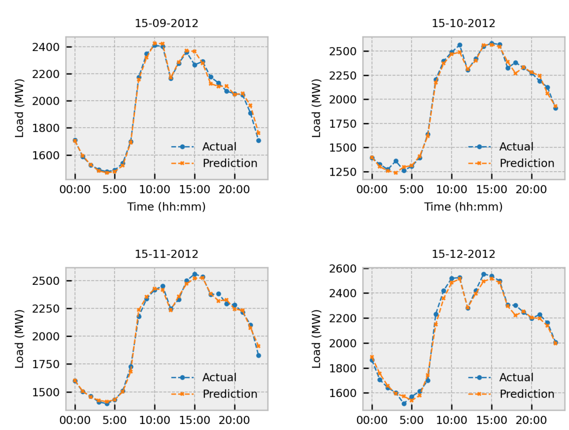

Fig. 4. Single-step prediction results for typical days of the last four months of 2012.

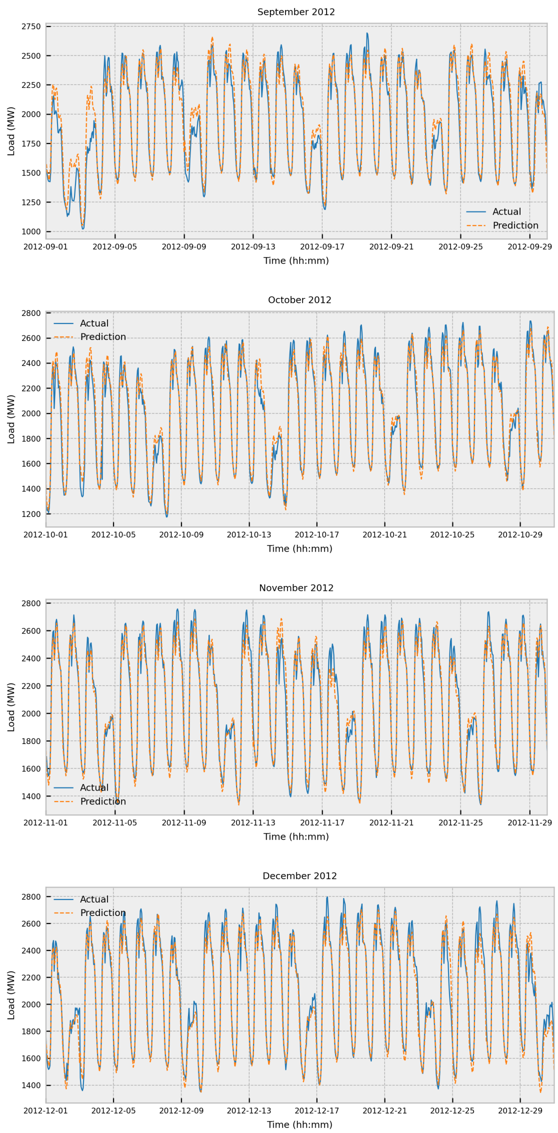

Fig. 5. Multi-step prediction results for the last four months of 2012.

TABLE II
SINGLE-STEP PREDICTION RESULTS FOR THE LAST FOUR MONTHS OF 2012

Timestep	MAE (MW)	MAPE (%)	RMSE (MW)
September	30.9769	1.6517	43.1155
October	29.4401	1.4867	49.4115
November	28.3603	1.3364	39.0328
December	28.6191	1.3863	41.0400

TABLE III
MULTI-STEP PREDICTION RESULTS FOR DIFFERENT INPUT SEQUENCE LENGTHS

Timestep	MAE (MW)	MAPE (%)	RMSE (MW)
24	67.5189	3.4688	96.6806
48	64.7490	3.2270	90.3958
72	67.2280	3.3117	91.5923
120	62.3680	3.1541	87.7012
144	70.9385	3.5786	95.9915
168	73.6020	3.6921	96.7054

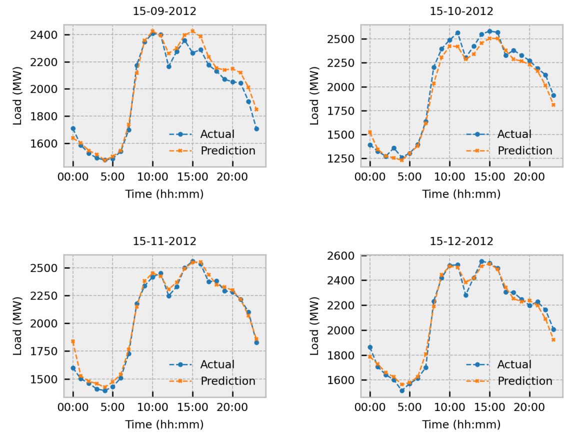

Fig. 6. Multi-step prediction results for typical days of the last four months of 2012.

TABLE IV
MULTI-STEP PREDICTION RESULTS FOR THE LAST FOUR MONTHS OF 2012

Timestep	MAE (MW)	MAPE (%)	RMSE (MW)
September	79.2788	4.3428	107.9166
October	54.8205	2.7612	73.4464
November	53.3109	2.5176	70.3057
December	62.3378	3.0144	93.9439

shown in Table III, multi-step forecasting attains the greatest precision when using 120-hour input sequences to predict 24-hour horizons. Table IV presents the accuracy of predictions for 24-hour horizons for the last four months of 2012. Similar to the forecast results from the single-step ahead model, the highest forecasting accuracy is achieved in November 2012. Fig. 5 illustrates 24-hour ahead predictions for the last four months of 2012. Additionally, Fig. 6 displays 24-hour ahead predictions for typical days of these months. From the obtained results, the predicted values are relatively consistent with the actual data. Thus, the proposed method is also effective for multi-step load forecasting.

E. Performance comparisons

To benchmark the GRU NN-based forecasting models, the forecasted results of GRU NN are compared with those of LSTM and CNN based on MAE, MAPE, and RMSE metrics. A look-back window of 120 steps is utilized for implementing the multi-step ahead model. The error metrics obtained from the GRU, LSTM, and CNN are presented in Table V. Observing the error metric values indicates that the GRU NN yields better results than LSTM and CNN. It can be concluded that the proposed GRU NN surpasses LSTM and CNN in predicting accuracy, which confirms the effectiveness of the proposed methodology for short-term load forecasting.

V. CONCLUSION

This study proposed and assessed an effective short-term load forecasting method using GRU NN with single-step and multi-step forecasting models. The proposed method was

TABLE V
FORECASTING PERFORMANCE COMPARISON

Methods	**Single-step forecasting**			**Multi-step forecasting**		
	MAE (MW)	***MAPE (%)***	***RMSE (MW)***	***MAE (MW)***	***MAPE (%)***	***RMSE (MW)***
GRU	29.34	1.47	43.36	62.37	3.15	87.70
LSTM	30.86	1.56	45.66	63.77	3.26	91.19
CNN	31.05	1.58	46.06	64.55	3.24	90.31

applied to the historical load data of the power system in Vietnam. From the forecasting results, GRU NN successfully predicts the load with low error values for both a future one-hour period (single-step forecasting) and a future 24-hour period (multi-step forecasting). The study also examined the influence of the input sequence length on the precision of the forecasting model. The analysis indicated that a specific sequence length could yield the greatest accuracy. Further, the forecasting accuracy of the proposed method was compared with LSTM and CNN using error metrics. Based on the analysis of the results, the GRU NN performed well compared to LSTM and CNN on the given dataset. Thus, the developed GRU NN is an effective approach for short-term load forecasting with high accuracy in power systems.

ACKNOWLEDGMENT

We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study.

REFERENCES

[1] W. Charytoniuk, M. Chen, and P. V. Olinda, “Nonparametric regression based short-term load forecasting,” *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 725–730, 1998.

[2] C.-M. Lee and C.-N. Ko, “Short-term load forecasting using lifting scheme and arima models,” *Expert Systems with Applications*, vol. 38, no. 5, pp. 5902–5911, 2011.

[3] W. Christiaanse, “Short-term load forecasting using general exponential smoothing,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, no. 2, pp. 900–911, 1971.

[4] D. Liu, L. Zeng, C. Li, K. Ma, Y. Chen, and Y. Cao, “A distributed short-term load forecasting method based on local weather information,” *IEEE Systems Journal*, vol. 12, no. 1, pp. 208–215, 2018.

[5] J. Xie, Y. Chen, T. Hong, and T. D. Laing, “Relative humidity for load forecasting models,” *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 191–198, 2018.

[6] Y. Wang, Q. Chen, M. Sun, C. Kang, and Q. Xia, “An ensemble forecasting method for the aggregated load with subprofiles,” *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3906–3908, 2018.

[7] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang, “Short-term electrical load forecasting using the support vector regression (svr) model to calculate the demand response baseline for office buildings,” *Applied Energy*, vol. 195, pp. 659–670, 2017.

[8] H. H. Çevik and M. Çunkaş, “Short-term load forecasting using fuzzy logic and anfis,” *Neural Computing and Applications*, vol. 26, no. 6, pp. 1355–1367, 2015.

[9] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[10] B. A. Hoverstad, A. Tidemann, H. Langseth, and P. Ozturk, “Short-term load forecasting with seasonal decomposition using evolution for parameter tuning,” *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1904–1913, 2015.

[11] K. Park, S. Yoon, and E. Hwang, “Hybrid load forecasting for mixed-use complex based on the characteristic load decomposition by pilot signals,” *IEEE Access*, vol. 7, pp. 12 297–12 306, 2019.

[12] M. S. Hossain and H. Mahmood, “Short-term load forecasting using an lstm neural network,” in *2020 IEEE Power and Energy Conference at Illinois (PECI)*, Feb 2020, pp. 1–6.

[13] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, and J. He, “Short-term load forecasting with deep residual networks,” *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3943–3952, 2019.

[14] X. Sun, P. B. Luh, K. W. Cheung, W. Guan, L. D. Michel, S. S. Venkata, and M. T. Miller, “An efficient approach to short-term load forecasting at the distribution level,” *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2526–2537, 2016.

[15] K. Methaprayoon, W. J. Lee, S. Rasmiddatta, J. R. Liao, and R. J. Ross, “Multistage artificial neural network short-term load forecasting engine with front-end weather forecast,” *IEEE Transactions on Industry Applications*, vol. 43, no. 6, pp. 1410–1416, 2007.

[16] K. Cho, B. V. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *CoRR*, vol. abs/1406.1078, 2014. [Online]. Available: http://arxiv.org/abs/1406.1078