Customer Baseline Load Estimation for Incentive-Based Demand Response Using Long Short-Term Memory Recurrent Neural Network

James Oyedokun, *IEEE Student Member**, Shengrong Bu, *IEEE Member*†, Zhu Han, *IEEE Fellow*‡, and Xue Liu, *IEEE Senior Member*§

*Systems, Power and Energy Division, University of Glasgow, Email: j.oyedokun.1@research.gla.ac.uk

†Systems, Power and Energy Division, University of Glasgow, Email: shengrong.bu@gla.ac.uk

‡Department of Electrical and Computer Engineering, University of Houston, Email: zhan2@uh.edu

§School of Computer Science, McGill University, Email: xueliu@cs.mcgill.ca

***Abstract*—The transition to an intelligent, reliable and efficient smart grid with a high penetration of renewable energy drives the need to maximise the utilization of customers demand response potential. The availability of smart meter data means this potential can be more accurately estimated and suitable demand response (DR) programs can be targeted to customers for load shifting, clipping and reduction. In this paper, we focus on estimating customer demand baseline for incentive-based DR. We propose a long short-term memory recurrent neural network framework for customer baseline estimation using previous like days data during DR events period. We test the proposed methodology on the publicly available Irish smart meter data and results shows a significant increase in baseline estimation accuracy when compared to traditional baseline estimation methods.**

***Index Terms*—Smart meter data, incentive-based demand response, customer baseline estimation, long short-term memory.**

I. Introduction

The need to reduce the carbon footprint all through the electricity supply chain and the ever increasing integration of renewable energy sources makes demand response (DR) key to ensuring a stable, reliable and high quality power supply. The non-dispatchable characteristics of renewable energy makes DR important as flexibility i n c ustomers d emand c an be exploited to follow available renewable energy resources. DR refers to a customer side effort to reduce demand in order to follow available supply which could be limited as a result of energy resources or transmission and distribution capacity constraints. There are two main categories of DR programs which are price-based DR and incentive-based DR [1] [2]. Price-based DR involves using time-varying electricity prices to change customers demand characteristics while incentive-based DR involves the utility offering customers incentives to reduce demand during a DR event. Depending on a customers demand characteristics, various DR programs can achieve different amount of demand reduction during DR events.

With the rising popularity of smart meters, smart meter data showing customers consumption can be utilized to derive intelligence for the optimal deployment of DR programs to customers. Some recent research on DR has focused on deriving intelligence from customer smart meter data for enhancing DR program implementation in smart grids [3] [4] [5]. Considering the two main load classes which are industrial and residential loads, price based DR has seen more application for residential customers while incentive-based DR is much popular with industrial customers. One of the main reasons for this is the relative ease of estimating an industrial customers demand baseline compared to the highly random nature of residential customers demand characteristics. The availability of smart meters at residential customers' location means the data can be explored to estimate a more representative baseline estimate of their demand during DR events, thereby enhancing their suitability for incentive-based DR. The main importance of baseline estimation is that customers demand reduction during DR events can be accurately estimated especially for accurate incentive payment and accurate knowledge of the aggregate DR potential of users i.e., showing the indication of how much DR resource is for a particular event.

Baseline estimation methods proposed in literature can be classified under averaging, regression, statistical and machine learning approaches. Examples of averaging approaches include direct average of X previous days; average of the highest X of Y days (HighXofY); average of the middle X of Y days (MidXofY); and average of the lowest X of Y days [6]. These methods are widely used in the industry with Mid4of6 and High4of5 used in PJM interconnection [7]. A support vector regression (SVR) method was proposed in [8] for baseline estimation. A clustering approach was implemented in [9] with self organizing map and k-means methodology proposed for estimating customer baseline. In [10], a density based clustering method together with k-means partitional clustering method was proposed to find representative baseline for customers. A statistical approach was proposed in [11] using a customer control group selection algorithm for estimating customer demand baseline. In [12], a probabilistic baseline estimation method was proposed for estimating customers consumption baseline using Gaussian process.

In this paper, we propose a machine learning approach using the long short-term memory recurrent neural network (LSTM RNN) methodology for customer baseline estimation during DR event period. We use demand data of previous like days within the DR event time span to estimate users baseline for the DR event period. The motivation for our proposed methodology is the improvement in forecasting results obtained using the LSTM RNN methodology for demand forecasting [13] [14] as well as battery state of charge estimation [15] compared to traditional methods.

The contribution of this paper are as follows:

- We propose and implement the LSTM RNN methodology for event based customer baseline estimation.
- We justify the use of mean percentage error (MPE) as a means to measure baseline estimation accuracy compared to mean absolute percentage error (MAPE) and root mean square error (RMSE).
- We compare our proposed customer baseline estimation results with the Low4of5, Mid4of6 and High4of5 methods.

The rest of the paper is organized as follows: in Section II, we present the LSTM RNN methodology for customer baseline estimation. In Section III, we present the application of our proposed model to real customer smart meter data and compare the results of our baseline estimation methodology with traditional methods. Finally, we present the conclusion of the paper in Section IV.

II. Methodology

An event based customer baseline estimation method is proposed for incentive-based DR. We propose the LSTM RNN technique for estimating the baseline demand during event period. Fig. 1 shows the framework of our proposed methodology. We characterize users demand for the event period of y previous non DR like-days[1] by normalising it using the MinMaxScaler presented in 1 below where d_t^* is the normalised data, d_t is the demand at time t, d_{min} and d_{max} is the minimum and maximum demand for the event period, respectively, and min and max represents the range of the feature. The derived feature is then fed into the LSTM network as input matrix X. We propose an input feature range min and max as -1 and 1, respectively, as it enhances an effective measure for backpropagation as proposed in [16].

The theory of RNN and LSTM is presented in subsections A and B respectively and the proposed DR event estimation model is presented in subsection C.

$$d_t^* = \frac{d_t - d_{min}}{d_{max} - d_{min}} * (max - min) + min \tag{1}$$

[1]like-days mean previous week days or weekend days

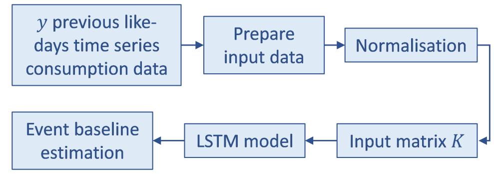

Fig. 1. Framework for customer DR baseline estimation.

A. Recurrent Neural Network

Among the main classes of artificial neural networks, RNNs has proven to be best suited for sequence learning such as time series forecasting problems [17]. Compared to feedforward neural networks (FNN), where signals travel in one direction from the input to the output via the hidden layer, RNNs allow signals to go back and forth hence allowing information in past data to be exploited for future data estimation.

Given a time series input $x = \{x_1, x_2, \cdots, x_t\}$, the hidden state and output sequence is derived using 2 and 3 respectively,

$$h_t = f(W_{hx}x_t + W_{hh}h_{t-1} + b_h), \tag{2}$$

$$y_t = g(W_{yh}h_t + b_y), \tag{3}$$

where W_{hx}, W_{hh} and W_{yx} are the input-hidden, hidden-hidden and hidden-output weight matrix, respectively. b_h and b_y represent the bias of the hidden and output layer, respectively. The activation layer of the input and output layer is represented as $f(\cdot)$ and $g(\cdot)$.

B. Long Short-Term Memory

One of the main challenges with RNN is the exploding and vanishing gradient problem [18]. This challenge occurs as a result of the loss function decaying exponentially with time. To address this challenge, LSTMs include a memory cell and gates at the hidden layers. Fig. 2 shows the architecture of an LSTM RNN.

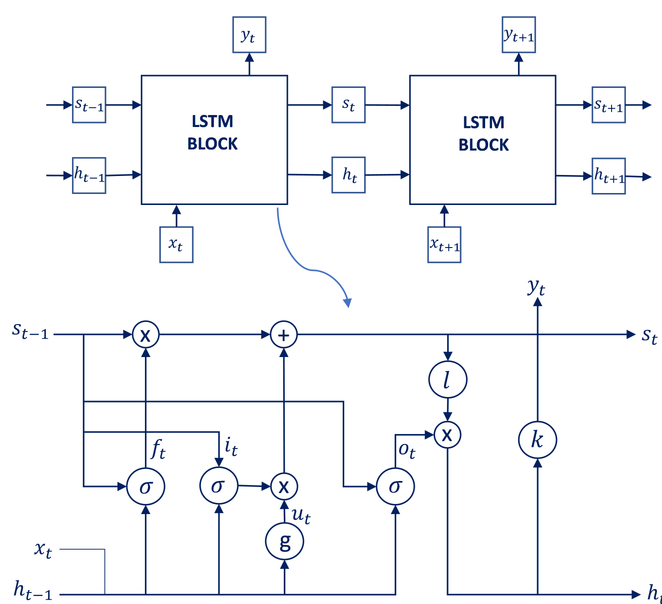

Fig. 2. LSTM RNN architecture.

The LSTM block consists of an input gate (i), output gate (o), forget gate (f) and attached memory cells (s) [13]. The input x_t and previous output h_{t-1} to the LSTM block determine the decisions of the input, output and forget gates whether to switch on or off [14]. The input gate controls what to keep in the internal state s_t, while the output gate decides the part of the internal state s_t to pass to the output h_t. The forget gate however decides what part of the previous state s_{t-1} needs to be forgotten. The equations (4), (5), (6), (7), (8), (9) and (10) below presents the operation of the gates as well as the states and output formulation.

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + W_{fs}S_{t-1} + b_f), \tag{4}$$
$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + W_{is}S_{t-1} + b_i), \tag{5}$$
$$u_t = g(W_{sx}x_t + W_{sh}h_{t-1} + b_s), \tag{6}$$
$$s_t = u_t i_t + s_{t-1}f_t, \tag{7}$$
$$o_t = \sigma(W_{ox_t} + W_{oh_{t-1}} + W_{os}s_{t-1} + b_o), \tag{8}$$
$$h_t = o_t \ell(s_t), \tag{9}$$
$$y_t = k(W_{yh}h_t + b_y), \tag{10}$$

where W_{fx}, W_{ix}, W_{fx}, W_{sx}, W_{ox}; W_{fh}, W_{ih}, W_{sh}, W_{oh}; W_{yh}; W_{fs}, W_{ic}, W_{os}; b_f, b_i, b_s, b_o, b_y are the input weight matrices; recurrent weight matrices; hidden output weight matrix; weight matrices of peephole connections and bias vectors, respectively.

C. DR Event Estimation Model

Firstly, we prepare the input data, normalise and feed the processed data into the LSTM model for baseline estimation during event period. Let C represent a set of k customers $C = \{c_1, c_2, c_3, \cdots, c_k\}$ with each enrolled for incentive-based DR. The aim is to estimate the demand baseline for a customer c_n during a DR event period between t_s and t_e. Let $D_{a,b} = \{d_{0,1}, d_{0,2}, \cdots, d_{0,48}, \cdots, d_{y,1}, d_{y,2}, \cdots, d_{y,48}\}$ represent the demand series for customer c_n where a denotes the days from event day 0 to previous like-day y, that is, $a = \{0, 1, 2, \cdots, y\}$ and $b = \{0, 1, 2, \cdots, 48\}$ daily half-hourly demand data. Given a DR day with reduction event between t_s and t_e where t_s represents the event start time and t_e represents the event end time for a day's demand. The objective is to estimate the customer baseline demand between times t_s and t_e where $t_s, t_e = \{1, 2, 3, \cdots, 48\}$ and $t_e > t_s$. The input data to the model is given as $X = D^*_{a,b} = \{d_{y,t_s}, d_{y,t_{s+1}}, \cdots, d_{y,t_e}, d_{y-1,t_s}, d_{y-1,t_{s+1}}, \cdots, d_{y-1,t_e}, \cdots, d_{2,t_s}, d_{2,t_{s+1}}, \cdots, d_{2,t_e}, \cdots, d_{1,t_s}, d_{1,t_{s+1}}, \cdots, d_{1,t_e}\}$. We set the output, i.e., the estimated baseline as $Q(t) = \{q_{t_s}, q_{t_{s+1}}, q_{t_{s+2}}, \cdots, q_{t_e}\}$.

D. Performance Metrics

In order to measure the performance of our proposed methodology for customer baseline estimation, we propose using the mean percentage error (MPE), mean average percentage error (MAPE) and root mean square error (RMSE) to measure the closeness of the estimate to the true customer demand for an event period. The MPE is defined as the average of the percentage error between the baseline estimate and the true customer demand during the event period as shown in 11, where d_t and q_t is the true baseline and estimate baseline, respectively.

$$MPE = \frac{100}{y}\sum_{t=t_s}^{t_e}\frac{d_t - q_t}{d_t}. \tag{11}$$

MAPE however measures performance based on the absolute value of the error difference between the estimated and true baseline. MAPE is presented in 12.

$$MAPE = \frac{100}{y}\sum_{t=t_s}^{t_e}\left|\frac{d_t - q_t}{d_t}\right|. \tag{12}$$

RMSE is the mean of the sum of the square error where the error is the difference between the estimate and the true baseline. RMSE is presented in 13 below.

$$RMSE = \sqrt{\frac{\sum_{t=t_s}^{t_e}(d_t - q_t)^2}{y}}. \tag{13}$$

III. Result and Discussion

For our analysis, we use customer demand data from the publicly available Irish Commission for Energy Regulation (CER) smart meter dataset [19]. Fig. 3 shows the demand profile of one of the customers which we use as a case study.

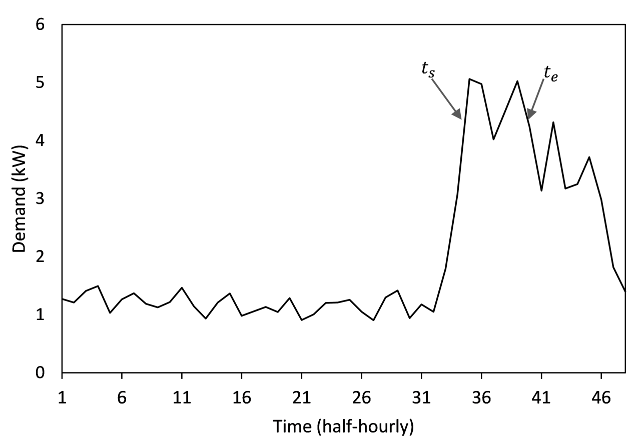

Fig. 3. Customer demand profile.

Given a DR event period between time period 34 and 40, we estimate the demand baseline for this period. We implement the proposed baseline estimation methodology using 10 previous like-days i.e., $y = 10$. Firstly, we run experiments on our model by varying the LSTM hyper-parameters which are the number of neurons and epochs. Table I presents the MPE, MAPE and RMSE of the LSTM model with varying number of epochs and neuron numbers fixed at 500.

TABLE I
DR ESTIMATION MODEL ERROR SUMMARY WITH VARYING EPOCHS

Epoch numbers	MPE(%)	MAPE(%)	RMSE
300	28.95	45.44	2.08
600	22.07	33.59	1.54
900	13.93	41.44	1.95
1200	7.36	34.80	1.78
1500	5.32	28.83	1.44

TABLE II
DR ESTIMATION MODEL ERROR SUMMARY WITH VARYING NEURON NUMBERS

Neuron numbers	MPE(%)	MAPE(%)	RMSE
100	28.22	41.10	1.93
200	21.06	46.17	2.09
300	27.54	40.10	1.92
400	7.97	26.11	1.26
500	5.32	28.83	1.44

Increasing number of epochs translated to decreasing MPE value indicating the closeness of the aggregate demand of the LSTM estimate to the true baseline. However, MAPE and RMSE did not follow the trend when epoch number was increased from 600 to 900. MAPE and RMSE use absolute values in its error calculation and this does not take into consideration the bias between both the estimated demand and true demand per unit time. In baseline estimation, the important indication for accuracy is the closeness of the aggregate demand during event period to the aggregate true demand given a non DR event. This situation makes MPE the ideal factor to measure the performance of a customer baseline estimation for incentive-based DR. We also vary the neuron numbers with epoch number fixed at 1500. Table II presents the MPE, MAPE and RMSE of the LSTM model with varying number of neurons.

The error results presented shows a decreasing MPE for increasing number of neurons. This trend shows a closer cumulative demand of the baseline estimate to the true baseline as the number of neuron increases. Also, the consequence of using more neurons is an increase in both computational time and resources. The MAPE and RMSE also in this case, does not follow a decreasing trend with increasing number of neurons as bias is not considered in the error calculation. Fig. 4 shows the baseline estimation result of the LSTM model using different neuron numbers during event time compared to the true baseline.

Based on the results of the varying hyper-parameter for the proposed LSTM model, we select neuron number as 500 and epoch number as 1500 and compare the estimation results with traditional baseline methodology i.e., Low4of5, Mid4of6 and High4of5. Fig. 5 presents the result of the LSTM estimation with the baseline estimate in red and input data to model in blue.

The proposed LSTM baseline estimation model result compared to traditional methods and the true baseline is presented in Fig. 6.

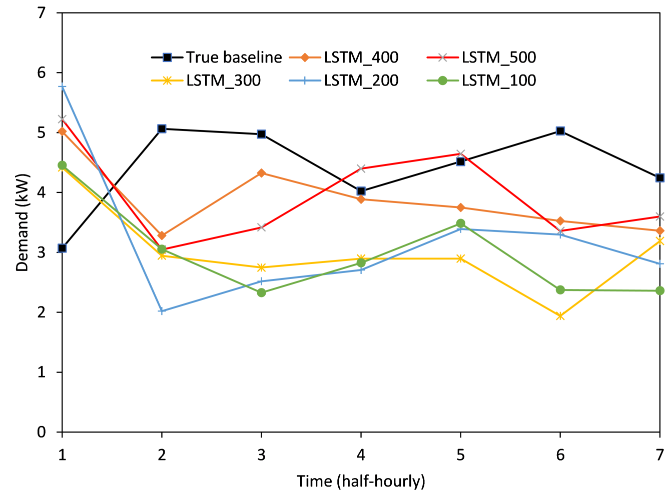

Fig. 4. Baseline estimation with varying neuron numbers.

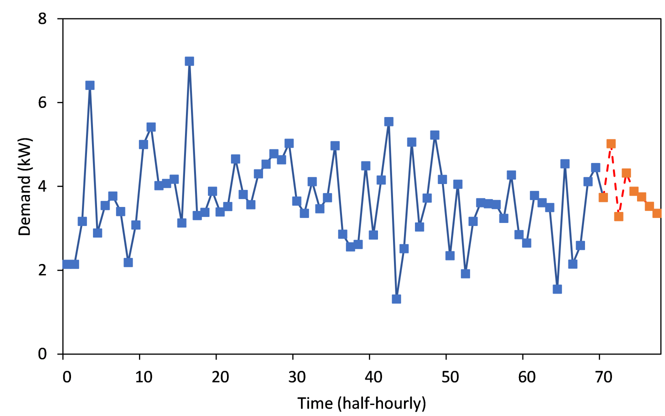

Fig. 5. Baseline estimation showing input data and output of the LSTM model.

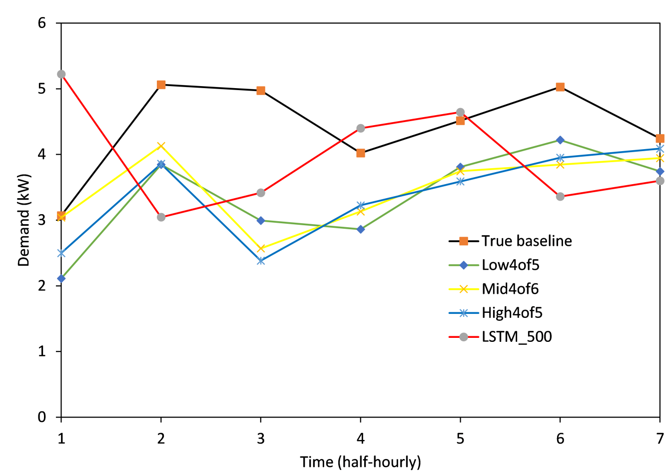

Fig. 6. Baseline estimation using LSTM compared to traditional methods.

Table III presents the MPE, MAPE and RMSE of the LSTM methodology compared to the Low4of5, Mid4of6 and

TABLE III

DR ESTIMATION MODEL ERROR SUMMARY

Estimation model	MPE(%)	MAPE(%)	RMSE
LSTM Model	5.32	28.83	1.44
Low4of5	23.91	23.91	1.14
Mid4of6	19.63	19.63	1.17
High4of5	22.84	22.84	1.26

High4of5 baseline estimation methods. Both the MAPE and RMSE shows a higher error value for the LSTM method compared to the Low4of5, Mid4of6 and High4of5.

These error values is however different from the MPE which shows the LSTM method having a significantly much lower value i.e., 5.32% compared to 23.91%, 19.63% and 22.84% for Low4of5, Mid4of6 and High4of5, respectively. The MPE gives an indication of how close the estimate of the cumulative demand during DR event is to the true baseline as can be shown in Fig. 7.

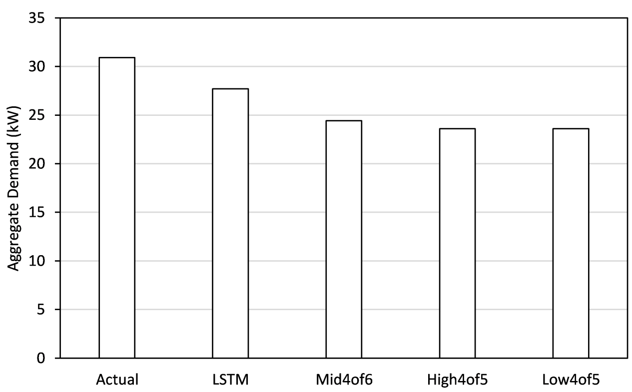

Fig. 7. Comparison of aggregate demand estimates using LSTM to traditional methods.

IV. Conclusion

In this paper, a novel customer baseline estimation methodology was proposed. We proposed a long short-term memory recurrent neural network method for estimating customer demand during DR event. To minimise the computational expense, we focus our estimation on the event period compared to the whole daily period as used in various methods proposed in literature. Our method proved to be more accurate as the cumulative demand of our estimation was much closer to the true baseline compared to traditional methods.

Acknowledgment

The authors would like to thank the Petroleum Technology Development Fund (PTDF), Nigeria and the National Universities Commission (NUC), Nigeria for their valuable financial support to this work.

References

[1] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, Jun. 2015.

[2] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, Firstquarter 2015.

[3] J. Oyedokun, S. Bu, Y. Xiao, and Z. Han, "Smart meter data characterization and clustering for peak demand targeting in smart grids," in *Proc. of IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, Sarajevo, Bosnia and Herzegovina, Oct. 2018.

[4] P. U. Herath, V. Fusco, M. N. Cáceres, G. K. Venayagamoorthy, S. Squartini, F. Piazza, and J. M. Corchado, "Computational intelligence-based demand response management in a microgrid," *IEEE Transactions on Industry Applications*, vol. 55, no. 1, pp. 732–740, Jan. 2019.

[5] J. Kwac, J. I. Kim, and R. Rajagopal, "Efficient customer selection process for various DR objectives," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1501–1508, Mar. 2019.

[6] T. K. Wijaya, M. Vasirani, and K. Aberer, "When bias matters: An economic assessment of demand response baselines for residential customers," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1755–1763, Jul. 2014.

[7] "PJM empirical analysis of demand response baseline methods," KEMA, Inc, Tech. Rep., 2011.

[8] 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, Mar. 2017.

[9] S. Park, S. Ryu, Y. Choi, J. Kim, and H. Kim, "Data-driven baseline estimation of residential buildings for demand response," *Energies*, vol. 8, no. 9, pp. 10 239–10 259, Sep. 2015.

[10] K. Li, B. Wang, Z. Wang, F. Wang, Z. Mi, and Z. Zhen, "A baseline load estimation approach for residential customer based on load pattern clustering," *Energy Procedia*, vol. 142, pp. 2042–2049, Aug. 2017.

[11] L. Hatton, P. Charpentier, and E. Matzner-Løber, "Statistical estimation of the residential baseline," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 1752–1759, May 2016.

[12] Y. Weng and R. Rajagopal, "Probabilistic baseline estimation via gaussian process," in *Proc. of IEEE Power & Energy Society General Meeting*. Denver, CO: IEEE, Jul. 2015.

[13] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network," in *Proc. of 51st Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2017.

[14] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019.

[15] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of li-ion batteries," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6730–6739, Aug. 2018.

[16] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backprop," in *Neural networks: Tricks of the trade*. Springer, 2012, pp. 9–48.

[17] Z. C. Lipton, J. Berkowitz, and C. Elkan, "A critical review of recurrent neural networks for sequence learning," *arXiv preprint arXiv:1506.00019*, 2015.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[19] Irish Social Science Data Archive. (2012), "CER smart metering project," [Online]. Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/.