LSTM Forecasts for Smart Home Electricity Usage

Rosemary E. Alden[1], Huangjie Gong[1], Cristinel Ababei[2], and Dan M. Ionel[1]
[1]Electrical and Computer Engineering, University of Kentucky
[2]Electrical and Computer Engineering, Marquette University
Email: {rosemary.alden, huangjie.gong, dan.ionel}@uky.edu, cristinel.ababei@marquette.edu

***Abstract*—With increasing of distributed energy resources deployment behind-the-meter and of the power system levels, more attention is being placed on electric load and generation forecasting or prediction for individual residences. While prediction with machine learning based approaches of aggregated power load, at the substation or community levels, has been relatively successful, the problem of prediction of power of individual houses remains a largely open problem. This problem is harder due to the increased variability and uncertainty in user consumption behavior, which make individual residence power traces be more erratic and less predictable. In this paper, we present an investigation of the effectiveness of long short-term memory (LSTM) models to predict individual house power. The investigation looks at hourly (24 h, 6 h, 1 h) and daily (7 days, 1 day) prediction horizons for four different recent datasets. We find that while LSTM models can potentially offer good prediction accuracy for 7 and 1 days ahead for some data sets, these models fail to provide satisfactory prediction accuracies for individual 24 h, 6 h, 1 h horizons.**

***Index Terms*—residential power load model; power load prediction; machine learning; LSTM**

I. Introduction and Literature Review

We are interested in the problem of estimating the energy usage in individual residences, either by simulations such as those described by [1], or through machine learning approaches specifically for smart homes [2]. Despite a lot of research done on the topic, electricity usage forecasting remains a difficult problem.

Examples of recent studies that investigated ML models include [3]–[11]. The work in [9] proposes a long short-term memory (LSTM) recurrent neural network (RNN) based framework for individual residential load forecasting. They report better prediction accuracies for aggregated residential loads compared to individual residences. The study in [12] proposed a novel pooling-based deep recurrent neural network for household load forecasting and reported that it performed ARIMA by 19.5%, SVR by 13.1% and classical deep RNN by 6.5% in terms of RMSE. The study in [13] reported that LSTM based RNN performed better compared with simple RNN and gated recurrent unit (GRU) RNN to forecast consumer electricity usage for a single user with 1-minute resolution based on one year of historical data sets.

Because the energy usage of a residential house has a higher correlation to residents' behavior, the studies in [9], [14] proposed to use an LSTM model to learn residents' life patterns in order to achieve better meter-level forecasting. The authors reported that the LSTM models achieved better forecasting compared to approaches based on feed-forward neural network (FFNN) and the k-nearest neighbor (KNN). The study in [15] uses an LSTM model to predict the energy usage of a house in order to then predict the electricity bill. They report prediction accuracy better than 80%. The work in [16] employs particle swarm optimization (PSO) to automatically determine hyperparameters for an CNN-LSTM model used to predict house global power. The study in [17] investigated several LSTM based models and reported good prediction results using standard LSTM models on the same dataset used by [15], [16]. The work in [13] reported that LSTM models performed better than deep neural networks (DNN) models for prediction a single residential energy usage.

Previous work indicates that the most promising ML models for prediction are convolutional neural networks (CNNs) [3] and RNNs [6]–[8]. More specifically, the LSTM model, as an enhanced version of an RNN model, was shown to perform at least as well or better than any other machine learning models. A summary of studies in this category is presented in Table I. That is why in this paper, we propose to develop and investigate LSTM models to predict the power in residential buildings for horizons of 24 h, 6 h, 1h, 7 days, and 1 day. This investigation is conducted on four different datasets.

II. Long Short-Term Memory (LSTM) Models

Due to lack of space and because LSTM models are not new, providing a detailed description of them is outside the scope of this paper. For details on LSTM models, we refer the reader to previous literature, including [9], [18], [19]. Here, we only present a general description.

An RNN is a type of neural network designed for sequence problems. It includes cyclic connections between different layers, which gives it the ability to "remember" past information. Outputs of units or modules from downstream layers may feedback as input to upstream layers or to the network together with the next input vector. These recurrent connections effectively add state (i.e., memory) to the model and enable it to learn and exploit the ordering of observations used as input sequences [19]. The difficulty with the standard RNN model is that it is hard to train in the case of problems that require learning long-term temporal dependencies. This is because the loss function's gradient decays exponentially with time, which is known as the vanishing gradient problem. The introduction of the LSTM model [18] attempted to address the vanishing gradient problem. The LSTM model is a more complex RNN network, which is enhanced with special units. The special units include memory cells that can store information for long periods of time. They also include three types of gates (input, forget and output) that are used to control the flow

Authors' manuscript version. The final published version is copyrighted by IEEE and is available as: R. E. Alden, H. Gong, C. Ababei and D. M. Ionel, "LSTM Forecasts for Smart Home Electricity Usage," 2020 9th International Conference on Renewable Energy Research and Application (ICRERA), 2020, pp. 434-438, doi: 10.1109/ICRERA49962.2020.9242804. ©2020 IEEE Copyright Notice. "Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works."

Table I
PREVIOUS ML MODELS FOR BUILDING ENERGY USAGE FORECASTING.

Study	Model	Model Size	Input Size	Resolution of Data	Prediction Horizon	Accuracy
[14]	LSTM	Layers 4, Units 512,512,512,1	2,3,6,12 steps	30-minute	30-minute	MAPE 22%
[9]	LSTM RNN	Layers 2, Units 20	2, 6, 12 steps	30-minute	30-minute	MAPE 44%
[12]	Pooling RNN	Layers 2,3,4,5, Units 5,10,20,30,50,100	48,96,336 steps	30-minute	30-minute	RMSE 0.45 kWh
[15]	LSTM	Layers 4, Units 75,100,100,1	10-minute	1-minute	5-minute	R^2 score 0.835
[16]	CNN-LSTM	NA	NA	Minutely	1 min, 1 h, 1 day, 1 week	MSE 0.35
[17]	LSTM, ConvLST	Layers 4, Units 100,50,50,7	7 days	Daily	7 days	RMSE 368 kW
[13]	LSTM, DNN	Layers 3, Units 50,100,1	30,40,50	Minutely	1 min	MAPE 35,24,29%

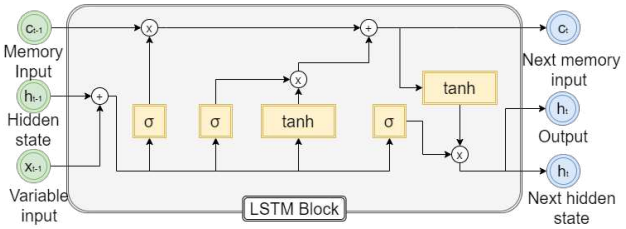


Figure 1. LSTM block diagram with inputs and outputs.

of information. These gates are of crucial importance because they decide whether to let new inputs in, erase the present cell state, and let the state impact the output at a given time step. Overall, the LSTM model can be understood as a chain of connected blocks (Fig. 1) each with three inputs, specifically implementing memory from previous blocks, and two outputs selected by weights are determined during training.

III. PUBLIC DATABASES FOR SMART HOMES

Important questions remain for the adoption of ML approaches for electricity usage forecasting, such as the right number of different variables that are measured and recorded, the total length of time continuously recorded, or time-resolution of the data. The results reported in previous literature depend also on the particular dataset that was used for investigations. In Table II, we compare four public datasets to highlight differences in their characteristics. Ideally, one would conduct model evaluation on as many such datasets as possible. Often though, results are reported only for a single dataset, and the model performance (as prediction accuracy) is quantified via different metrics (e.g., RMSE, MAE, MAPE, etc.) or just as relative comparisons between different approaches studied. For example, the study in [13] used the AMPds2 dataset while the studies in [15], [16] used the UCI dataset; and their models performance can vary quite significantly.

In this paper, we investigate all four datasets listed in Table II because we are interested in understanding the prediction capability of LSTM models for a diverse group of datasets. Specifically, we will use: 1) Honda Smart Home (HSM) dataset recorded from an advanced technological pilot home or "living laboratory" located in Davis, CA [20]. Fig. 2 presents a snap-shot of this dataset to visually highlight differences among power traces collected at different resolutions (in this case for the HVAC portion of the total net power of the house), which make the prediction problem even more complex. 2) AMPds2 dataset recorded at a house in Vancouver, Canada [21]. It includes electricity, water, and natural gas measurements with minutely resolution for two years. 3) SHINES field

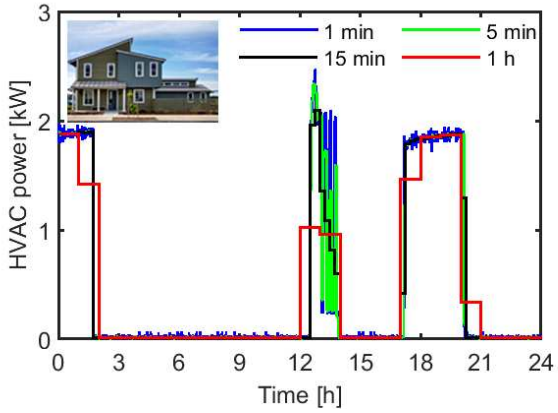


Figure 2. HVAC power traces (Dec. 15, 2018) plotted at different resolutions, from Honda Smart Home dataset.

demonstration was designed to be more representative of the typical residential smart home's construction, appliances, and monitoring capabilities that mirror a generic utility installed smart meter with a 15 minute resolution [22]. This dataset is collected by the Electric Power Research Institute and includes measurements from two houses, located in Pensacola, FL. 4) UCI dataset includes measurements from a house near Paris, France, collected with minutely resolution [23].

In our simulations, we pre-process the datasets to: 1) remove the first and the last days so that data starts on a Sunday and ends on a Saturday - this helps with the programming when predictions are done for 7 days horizons, 2) replace missing data values with the values at the same time from the previous available day, and 3) convert all power values into kWh in cases where the initial data is provided in kW.

IV. SIMULATION RESULTS FOR INDIVIDUAL RESIDENCE POWER PREDICTION

A. Simulation Results

The prediction problem was formulated in five different ways as indicated in Table III. We first looked at prediction of the *daily* net total power for seven (7) days or one (1) day ahead, when the input model was provided with the previous most recent fourteen (14) daily net total power values. In this case, each of the investigated four datasets (described in Table II) was down-sampled from the minutely original observations of *net total power* to daily totals. Then, we looked at prediction of the *hourly* net total power for twenty-four (24), six (6), and one (1) hours ahead. In each of these cases, the input features supplied to the LSTM model included the previous most recent forty-eight (48) hourly power values (i.e., a two day history).

Table II
CHARACTERISTICS OF DATASETS FOR RESIDENTIAL NET TOTAL POWER FORECASTING.

Dataset, Location	Houses	Length	Resol. (min)	Measured Variables	Instrum., Avail.	Previous Work
Honda Smart Home (HSH) Davis, CA	1	Jan. 2018 - Dec. 2018	1	HVAC (kW), WH (BTU), PV supply (kW DC), HVAC (kW), lighting (kW), appliances (kW or BTU), temperature (F), irradiance, floorplan, etc.	230+ sensors [20]	NA
AMPds2 Vancouver, Canada	1	Apr. 2012 - March 2014 24 months	1	HVAC furnance fan, mains, heat pump, clothes washer/dryer, plug loads (V, A), temperature, humidity, wind speed, etc.	Sub-meters [21]	[13], [14]
SHINES Pensacola, FL	2	Aug. 2017 - July 2020 36 months	15	Net power flow, HVAC, pool pump, battery, solar generation (kW), temperature, humidity, wind speed	Sub-meters [22]	NA
UCI Paris, France	1	Dec. 2006 - Nov. 2010 47 months	1	Mains (kW, V, A), laundry room, kitchen appliances, water heater + HVAC (Wh)	Sub-meters [23]	[15]–[17]

In this case, the original observations were down-sampled to hourly totals. When we predict multiple values into the future, the prediction problem is framed as a multi-step time series forecasting with uni-variate input data.

Based on the information reported by other authors, including for example, the information discussed in [17], we investigated several LSTM models, including multilayer LSTM, encoder-decoder LSTM, CNN-LSTM encoder-decoder, and ConvLSTM encoder-decoder models. Our current results indicate no significant differences in the performance of these models. As a result, we used a standard LSTM model with three (3) layers in the remaining of this paper.

In all our simulations, the LSTM model has a number of 56, 28, and 14 units on each of the three layers. These values were found empirically, via trial and error, and provided better performance for each specific case. A number of three layers is in alignment with most of the previous studies discussed in Section I (see Table I). We developed our LSTM model using Keras and coded in Python, drawing from and adapting the approach in [17]. All our software source-codes with the pre-processed datasets will be made available on github in order to enable complete future replication of our results for comparison purposes.

The size of the input features and the prediction horizons are listed in the first two columns of Table III. In each case, the LSTM model is trained on the first 70% portion and tested on the remaining 30% portion of each of the datasets. To report the performance of the LSTM model in each case, we use the mean absolute percentage error (MAPE) because it offers an easier to understand measure of how far are the predictions from the true values, in comparison for example with other measures such as RMSE or MAE. Each of the MAPE values reported in Table III is the best out of five (5) different runs of the same simulation experiment.

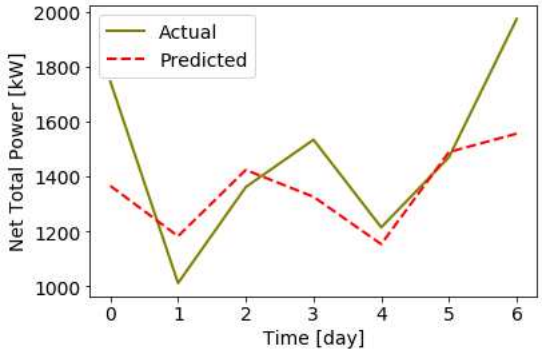
B. Discussion

Based on the insights we gained during our simulations, we make the following observations:

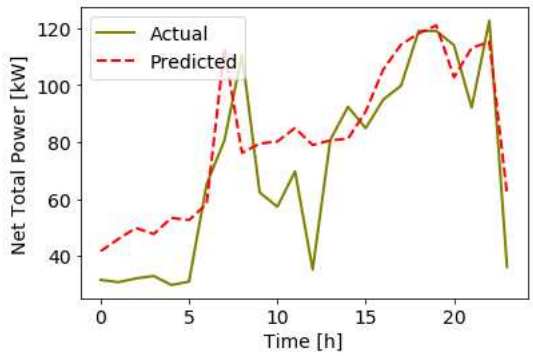
- Results in Table III indicate that the LSTM model has significantly better performance on some datasets. For example, the LSTM model has the best performance on the AMPds2 and UCI datasets. However, it does rather poorly on the HSH and SHINES datasets. Two examples of 7 days and 24 h ahead prediction for a random week and a random day from the testing portions of the UCI and AMPds2 datasets are shown in Fig. 3.
- While daily prediction for 7 or 1 days ahead can be done relatively accurately (at least for the AMPds2 and UCI datasets), the LSTM model does not achieve satisfactory prediction accuracy for hourly, 24, 6, and 1 h ahead. Hourly prediction for individual residences remains an open problem, which we plan to address in our future work.
- An important factor that affects the prediction quality is the normalization of data, which helps reduce MAPE significantly in most cases. In all our simulations, datasets are normalized to the standard interval $(0, 1)$.
- Increasing the number of model layers beyond four (4) did not improve the model performance. This is due to model over-fitting, which was also noticed in previous studies [12].
- Increasing the input feature size (for example, in the case of prediction of 24 h ahead, increasing the number of input values from 24 to 48, 96, or 168, i.e., the equivalent of 7 days) does not improve the model performance too much.
- Increasing the number of *epochs* (one of the hyperparameters that one must decide during model development) that is used during the process of model fitting has a huge impact on the MAPE value when the trained model is tested on the *training* portion of the data itself, and less on the MAPE value achieved on the *testing* portion of the data. We observed that usually a number of epochs > 200 suffices to bring the $MAPE_{train}$ to lower than 15% (which means that the model fits really well the training portion of the data). However, that does not necessarily mean that $MAPE_{test}$ is low too, especially if that particular dataset exhibits widely varying (i.e., erratic, not repeatable) patterns.
- When prediction is made for multiple steps ahead (e.g. 24 h, 6 h or 7 days), the prediction accuracy varies across future predicted steps. The MAPE variation across 7 days values during the testing of the LSTM model on the UCI dataset and across 24 h values during the testing of the LSTM model on the AMPds2 dataset is shown in Fig. 4. In other words, some hours in the day are more predictable than others; particularly, consumption during peak hours is harder to predict due to its increased variability.

Table III
SIMULATIONS RESULTS DURING PREDICTION OF NET TOTAL POWER, DAILY OR HOURLY.

		$MAPE_{test}$ (%)			
Input Size	Prediction Horizon	HSM Dataset	AMPds2 Dataset	SHINES Dataset	UCI Dataset
14 days (14 daily net total power values)	7 days (7 separate values)	42%	14%	58%	23%
14 days	1 day (1 value)	39%	13%	51%	22%
48 h (48 hourly net total power values)	24 h (24 separate values)	182%	35%	105%	64%
48 h	6 h (6 separate values)	167%	31%	107%	64%
48 h	1 h (1 value)	107%	31%	83%	46%



(a)



(b)

Figure 3. Examples of total net power predictions for 7 days (UCI dataset) and 24 h (AMPds2 dataset).

- We noticed in some cases that while the prediction of the *Net Total* power is rather poor, the prediction of individual components (such as HVAC or Lighting power) of the same dataset can be done with much better accuracy. For example, Table IV reports the MAPE values for the Lighting power component of the HSM dataset achieved with exactly the same LSTM model. This component is easier to predict because it has variation patterns that are less erratic compared to the patterns exhibited by the net total power - as it can be seen in the examples in Fig. 5. Another indicator of the difficulty of prediction is the "heat-map". For example, Fig. 6 shows such a heat-map for the SHINES dataset, which illustrates the increased randomness of the power values and the lack of repeatable patterns. Similar challenges with variable load predictions were reported in [24].

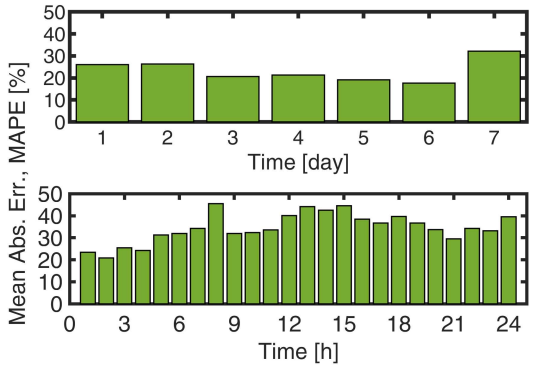


Figure 4. MAPE variation across 7 days (UCI dataset) and 24 h (AMPds2 dataset) prediction.

Table IV
MAPE VALUES FOR PREDICTION OF NET TOTAL POWER VS. PREDICTION OF ONLY LIGHTING POWER FOR HSM DATASET.

What is Predicted	MAPE 7 days	MAPE 1 day	MAPE 24 h	MAPE 6 h	MAPE 1 h
Net Total Power	42%	39%	182%	167%	107%
Lighting Power	25%	28%	109%	112%	76%

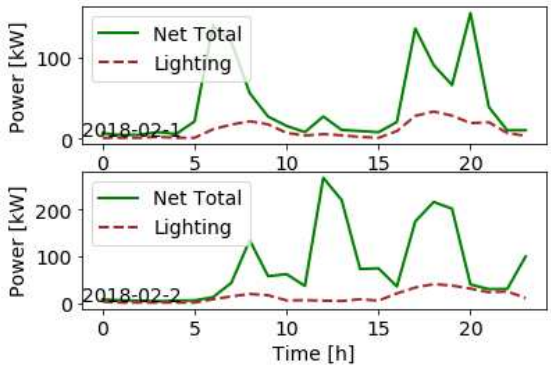
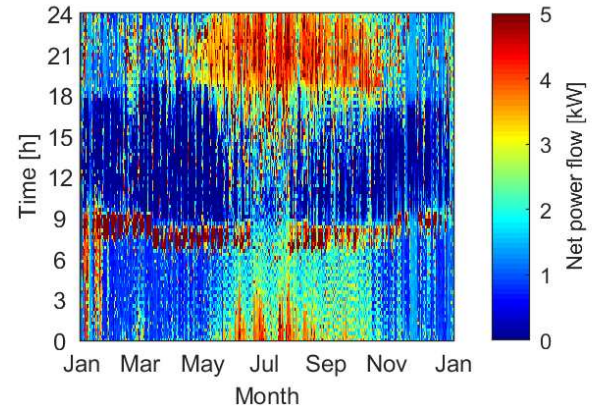


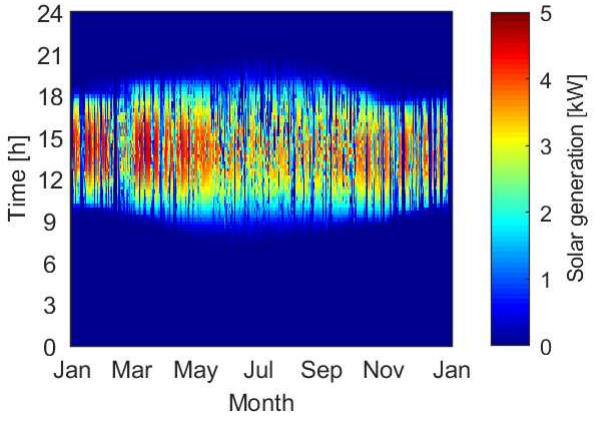
Figure 5. Power traces showing Net Total and Lighting power for HSM dataset, on the first two days of Feb. 2018.

V. CONCLUSION

We investigated the effectiveness of standard LSTM models to predict power of individual residences. The investigation looked at hourly (24 h, 6 h, 1 h) and daily (7 days, 1 day) prediction horizons for four different datasets. Extensive simulations using a three layer LSTM model indicated that while such models can potentially offer good prediction accuracy



(a)



(b)

Figure 6. Heat-maps of net total power for SHINES dataset. Solar generation explains low values during the afternoon.

for 7 and 1 days ahead for some datasets, these models do not achieve satisfactory prediction accuracies for 24 h, 6 h, 1 h prediction horizons. Future work will focus on framing the prediction problem as a multi-variate multi-step prediction problem, where additional variable measurements, such as temperatures inside and outside, solar irradiance, etc., are included into the input features.

ACKNOWLEDGMENT

This work was supported by the NSF, grant ECCF 1936494. Any findings and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

[1] S. Duerr, C. Ababei, and D. M. Ionel, "Smartbuilds: An energy and power simulation framework for buildings and districts," *IEEE Transactions on Industry Applications*, vol. 53, no. 1, pp. 402–410, 2017.

[2] R. E. Alden, P. Han, and D. M. Ionel, "Smart plug and circuit breaker technologies for residential buildings in the us," in *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*, 2019, pp. 1018–1021.

[3] K. Amarasinghe, D.L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," *IEEE Int. Symposium on Industrial Electronics (ISIE)*, 2017.

[4] P. Chen, W. Li, Y. Chen, K. Guo, and Y. Niu, "A parallel evolutionary extreme learning machine scheme for electrical load prediction," *IEEE Computing Conference*, 2017.

[5] D.C. Mocanu, E. Mocanu, P.H. Nguyen, M. Gibescu, and A. Liotta, "Big IoT data mining for real-time energy disaggregation in buildings," *IEEE Int. Conference on Systems, Man, and Cybernetics*, 2016.

[6] J. Rubio-Herrero, V. Chandan, C. Siegel, A. Vishnu, and D. Vrabie, "A learning framework for control-oriented modeling of buildings," *IEEE Int. Conference on Machine Learning and Applications (ICMLA)*, 2017.

[7] D.L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," *Annual Conference of the IEEE Industrial Electronics Society (IECON)*, 2016.

[8] A. Rahman, V. Srikumar, and A.D. Smith, "Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks," *Applied Energy*, vol. 212, pp. 372-385, 2018.

[9] 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.

[10] S. Seyedzadeh, F. Pour Rahimian, I. Glesk, and M. Roper, "Machine learning for estimation of building energy consumption and performance: a review," *Visualization in Engineering*, vol. 6:5, 2018.

[11] A. A. Mamun, M. Hoq, E. Hossain, and R. Bayindir, "A hybrid deep learning model with evolutionary algorithm for short-term load forecasting," in *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*, 2019, pp. 886–891.

[12] H. Shi, M. Xu and R. Li, "Deep learning for household load forecasting - a novel pooling deep RNN," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5271-5280, Sep. 2018.

[13] T. Hossen, A.S. Nair, R.A. Chinnathambi, and P. Ranganathan, "Residential Load Forecasting Using Deep Neural Networks (DNN)," *North American Power Symposium (NAPS)*, 2018.

[14] W. Kong, Z.Y. Dong, D.J. Hill, F. Luo, and Y. Xu, "Short-Term Residential Load Forecasting based on Resident Behaviour Learning," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1087-1088, Jan. 2018.

[15] R.A. Rashid, L. Chin, M.A. Sarijari, R. Sudirman, and T. Ide, "Machine Learning for Smart Energy Monitoring of Home Appliances Using IoT," *IEEE Int. Conf. on Ubiquitous and Future Networks (ICUFN)*, 2019.

[16] T.-Y. Kim and S.-B. Cho, "Particle Swarm Optimization-based CNN-LSTM Networks for Forecasting Energy Consumption," *IEEE Congress on Evolutionary Computation (CEC)*, 2019.

[17] J. Brownlee, "How to Develop Multi-Step LSTM Time Series Forecasting Models for Power Usage," [Online]. Available: https://machinelearningmastery.com/how-to-develop-lstm-models-for-multi-step-time-series-forecasting-of-household-power-consumption, 2018.

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

[19] Jason Brownlee, "Long Short-Term Memory Networks With Python - Develop Sequence Prediction Models With Deep Learning," Machine Learning Mastery Ebook, Ed. v1.6, 2019.

[20] A. H. M. Co. (2020) Honda Smart Home US. [Online]. Available: https://www.hondasmarthome.com/

[21] S. Makonin, "AMPds2: The Almanac of Minutely Power dataset (Version 2)," Harvard Dataverse, [Online]. Available: https://doi.org/10.7910/DVN/FIE0S4, 2016.

[22] E. P. R. Institute. (2020) DOE SHINES Residential Demonstration. [Online]. Available: https://dashboards.epri.com/shines-residential/dashboard

[23] D. Dua and C. Graff, "UCI machine learning repository," 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[24] S. Balluff, J. Bendfeld, and S. Krauter, "Short term wind and energy prediction for offshore wind farms using neural networks," in *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*, 2015, pp. 379–382.