Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Enhanced Home Energy Forecasting With Parallel LSTM Networks

Amit Kumar Upadhyay[1], Dr. Manoj Varshney[2], Dr. Pankaj Varshney[3]

1. *Department of Computer Engineering and Applications, Mangalayatan University, Aligarh.*
2. *Associate Professor, Department of Computer Engineering and Applications, Mangalayatan University, Aligarh, U.P.*
3. *Associate Professor, IITM Janakpuri, GGSIP, University, New Delhi.*

The primary obstacles in addressing the energy consumption forecasting challenge revolve around ensuring reliability, stability, efficiency, and accuracy in forecasting methodologies. The current forecasting models face difficulties due to the unpredictable nature of energy consumption data volatility. There is a need for artificial intelligence models that can anticipate abrupt irregular changes and effectively capture long-term dependencies within the data. Within this study, a pioneering AI-boosted forecasting model is presented, combining Extreme Gradient Boosting (XGBoost) with parallel long short-term memory (PLSTM) neural networks. The integration of XGBoost with PLSTM neural networks contributes to the improved performance of the overall PLSTM network. The access the suggested model using the Mean Absolute Percentage Error (MAPE).

Keywords: Long Short Term Memory, Energy Consumption, Time Series Data Analysis, Forecasting, Extreme Gradient Boosting.

1. Introduction

Energy consumption forecasting (ECF) stands as a crucial application of artificial intelligence (AI) essential for supporting the development of smart grids and smart cities [1]. The enhancement of ECF's reliability, efficiency, and accuracy contributes to increased transmission efficiency within smart grids, ensures secure energy market trading, and minimizes energy wastage [2–4]. Leveraging Internet of Things (IoT) technology introduces a pattern recognition process, utilizing sensor data collected from individual households for ECF [5,6]. This AI-enhanced data-driven approach to ECF offers valuable insights to governments, power plants, and residences, guiding them towards sustainable energy usage [7].

As part of the advanced metering infrastructure (AMI) program integral to smart grid development, the focus on ECF for individual household energy consumption data has garnered considerable attention. This heightened interest is primarily due to the challenges posed by the highly volatile time series data influenced by human behaviors [8,9]. Traditional methods, such as those based on physical models, often struggle to make accurate predictions in such dynamic scenarios [10]. Conversely, with the rapid evolution of AI, deep learning technologies, including long short-term memory neural networks, have found widespread application in addressing the ECF challenges associated with individual households.

A combining Extreme Gradient Boosting (XGBoost) with parallel long short-term memory (PLSTM) model is devised in this study for ECF. XGBoost is extracting essential features from the raw data. Then these data are preprocessed and subsets of these data are input into a Parallel Long Short-Term Memory Neural Network (PLSTM), comprising several LSTM neural networks. All LSTM neural networks operate in parallel, generating forecasting results. The final ECF outcome is derived by combining the forecasted results from all LSTM neural networks.

In summary, this study introduces a hybrid data-driven ECF model that integrates XGBoost with a PLSTM structure. The novelty lies in the design of a parallel neural network architecture, and the proposed method's performance has been rigorously verified. The primary findings of the present investigation can be summarized as follows:

1. The current study introduces a neural network structure, termed PLSTM, that integrates multiple parallel LSTM neural networks. This PLSTM configuration comprises several peer LSTM neural networks designed to operate concurrently, enhancing efficiency. Each LSTM is dedicated to training a specific subset of the data generated by the XGBoost. This integrated approach allows for accurate predictions representing distinct features.
2. A novel AI-empowered forecasting framework is proposed, which combines XGBoost and PLSTM neural networks. XGBoost is extracting essential features from the raw data. The XGBoost outputs are then input into a set of Long Short-Term Memory (LSTM) neural networks in a parallel fashion.

2. Related works

Methods for predicting time series data are generally classified into two categories: model-based and data-driven (AI-based) approaches [13,14]. In [15], it is highlighted that data-driven methods, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) neural networks, are particularly well-suited for energy consumption forecasting [16,17]. Data-driven models can be broadly categorized into two types: singular models and hybrid models. Singular models encompass decision tree [21], random forest [22], support vector regression [23], multilayer perceptron [24], convolutional neural network [25], recurrent neural network [26], and long and short-term memory neural network [27].

The author asserts that Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) [28] models outperformed the selected empirical models when estimating daily global solar radiation (H) in humid subtropical climates, both with complete and incomplete temperature and precipitation data. The XGBoost model exhibited comparable prediction accuracy to the SVM model but demonstrated greater stability and computational efficiency. Specifically, in humid subtropical regions of China, the author strongly recommends the use of the XGBoost model for estimating daily solar radiation (H) based on temperature and precipitation data, emphasizing its outstanding performance in terms of accuracy, stability, and computational efficiency.

The author developed a model that combined Short-Term Load Forecasting (STLF) models [29] using a sliding window-based Principal Component Regression (PCR) approach. The model, specifically utilizing PC 1 and a seven-day window size, demonstrated excellent prediction performance. The author referred to this model as the Combination of Short-Term Load Forecasting Models Using a Stacking Ensemble Approach (COSMOS) model. However, it should be noted that the proposed model did not satisfactorily predict building electric energy consumption on weekends, and its application was limited to a single building. Therefore, additional validation is necessary to assess the broader applicability of this forecasting model.

A novel spatial and temporal ensemble electric consumption model [19] incorporates clustering analysis to address short-term electric consumption forecasting. This innovative approach involves the examination of electric consumption profiles at the apartment level through cluster analysis, employing the k-means algorithm. The ensemble forecasting model comprises two deep learning models, namely the Long Short-Term Memory Unit (LSTM) and Gated Recurrent Unit (GRU). By grouping LSTM and GRU together, an ensemble is formed with the aim of enhancing prediction accuracy and minimizing generalization errors. The model demonstrates improved prediction accuracy when aggregating consumption at spatial scales, such as the building and floor levels. However, it is observed that the forecasting error tends to increase when extending the prediction horizon from hourly to weekly scales. Notably, the model's limitation is identified in terms of clustering the consumption profiles at the apartment level, as this further reduces the forecasting error compared to scenarios without clustering.

A model combines the strengths of the eXtreme Gradient Boosting Machine (XGB), Light Gradient Boosting Machine (LGBM), and Multi-Layer Perceptron (MLP) [20]. The Stacked XGB-LGBM-MLP model operates by generating meta-data from the XGB and LGBM models, which is then utilized to compute final predictions using an MLP network. This approach is particularly sensitive to two key factors: the forecasting horizon and the size of the data. However, it's worth noting that the study identifies a limitation in the performance of the Stacked XGB-LGBM-MLP model when forecasting 48 hours ahead. The effectiveness of the model experiences a decrease for this specific forecasting horizon.

A model is presented for Short-Term Load Forecasting utilizing XGBoost [11]. This model, designed for forecasting electrical load, involves the transformation of daily load data into weekly load data. The approach aims to enhance the set of features available for forecasting by considering the load of a lag variable. XGBoost is employed for feature selection from the converted data and subsequently trains the model for load prediction. The results indicate that the XGBoost load forecast generally aligns well with the actual load, yielding accurate predictions most of the time. However, it's worth noting that the accuracy of XGBoost diminishes when dealing with larger loads.

In a literature survey, it is generally observed that hybrid models tend to achieve higher prediction accuracy compared to singular models. In [28], the author employs a combination of Singular Spectrum Analysis (SSA) and least square Support Vector Machine (SVM) for Electric Consumption Forecasting (ECF). Yan et al. [29] argue that the Long Short-Term Memory (LSTM) neural network outperforms Support Vector Machine (SVM) in capturing dependencies among data samples. Wei et al. [30] introduce a hybrid model that combines Improved Singular Spectrum Analysis (ISSA) and LSTM for predicting daily natural gas consumption. However, the focus of the study is on the model's superiority in different climate zones rather than different time spans. It is acknowledged that data decomposition plays a crucial role in improving forecasting results, as demonstrated by Sun et al. [31] who decomposed economic factors for energy consumption forecasting.

The proposed system introduces a hybrid model that combines XGBoost with multiple LSTM neural networks. In the data pre-processing phase, XGBoost is utilized for feature extraction, and the original data is partitioned into subsets. For the model training stage, parallel LSTM networks are employed, matching the number of subsequences generated after decomposition. LSTM neural networks excel in handling non-linear and non-stationary time series data. Given the significant fluctuations in the original dataset, LSTM is particularly effective in capturing the long-term dependencies inherent in the original Energy Consumption Forecasting (ECF) time series. The anticipated outcome is that the final prediction results of this hybrid method will surpass those of current state-of-the-art methods in terms of accuracy.

3. Methodology

This section offers a thorough overview of the hybrid forecasting model that we have put forth. Initially, it outlines the pivotal step of utilizing the eXtreme Gradient Boosting Machine (XGBoost) algorithm for data processing. Subsequently, the internal structure and workflow of the Long Short-Term Memory (LSTM) neural network are presented, providing insights into why it is effective and readily available. Finally, we delve into the collaboration between XGBoost and Parallel LSTM (PLSTM), presenting the overall strategy in this study.

3.1 eXtreme Gradient Boosting Machine (XGBoost)

The Extreme Gradient Boosting (XGBoost) algorithm [12] is a novel implementation method for Gradient Boosting Machine and in particular K Classification and Regression Trees. The algorithm is based on the idea of “boosting”, which combined all the predictions of a set of “weak” learners for developing a “strong” learner through additive training strategies. XGBoost aims to prevent over-fitting but also optimize the computation resources. This is obtained by simplifying the objective functions that allow combining predictive and regularization terms, but maintaining an optimal computational speed. Also, parallel calculations are automatically executed for the functions in XGBoost during training phase. The processes of additive learning in XGBoost are explained below. The first learner is firstly fitted to the whole space of input data, and a second model is then fitted to these residuals for tackling the drawbacks of a weak learner. This fitting process is repeated for a few times until the stopping criterion is met. The ultimate prediction of the model is obtained by the sum of the prediction of each learner.

The eXtreme Gradient Boosting Machine (XGBoost) algorithm is a supervised learning algorithm that falls under the category of ensemble learning, specifically gradient boosting frameworks. Developed by Tianqi Chen, XGBoost has gained widespread popularity due to its efficiency, scalability, and high predictive performance. It is widely used for both regression and classification tasks in machine learning.

Here are key components and features of the XGBoost algorithm:
1**. Objective Function**: XGBoost minimizes a regularized objective function, which is a combination of a loss function that measures the model's prediction error and a regularization term that penalizes complex models, helping to prevent overfitting.
2. **Gradient Boosting**: XGBoost builds an ensemble of weak learners, typically decision trees, sequentially. Each tree corrects the errors of the previous ones by focusing on the residuals (the differences between actual and predicted values).
3. **Regularization**: XGBoost incorporates L1 (Lasso) and L2 (Ridge) regularization terms into the objective function. This helps control the complexity of the model and avoid overfitting.
4. **Tree Pruning**: During the boosting process, XGBoost employs a technique called tree pruning to control the depth of the trees, preventing them from growing too deep and overfitting the training data.
5. **Parallel and Distributed Computing**: XGBoost is designed for efficiency and supports parallel and distributed computing. This allows it to handle large datasets and expedite the training process.
6. **Handling Missing Values**: XGBoost has a built-in mechanism for handling missing values, reducing the need for extensive preprocessing.
7. **Cross-Validation**: The algorithm supports built-in cross-validation, allowing for the assessment of model performance during the training process.
8. **Feature Importance**: XGBoost provides a feature importance score, which helps users understand the contribution of each feature to the model's predictions.

3.2 Long short term memory (LSTM)

The intricate workings of the Long Short-Term Memory (LSTM) are depicted in Fig. 1[18]:. In the current network, the inputs and outputs are denoted as x_t and h_t, respectively. C'_t represents the state learned at the current moment. Notably, the LSTM structure comprises three gates: the input gate, the output gate, and the forget gate.

The input gate, denoted as x_t, governs the extent to which current input data is allowed to transfer to the cell state. By regulating the input gate, numerous irrelevant elements from the current input are prevented from entering the memory. The forget gate, represented by f_t, determines the amount of cell state from the previous time that is retained in the current state, preserving information from a considerable time ago. The decision to retain or discard the current state is under the control of the output gate, O_t, ultimately shaping the output cell state, C_t.

During each processing step, the LSTM repeats the operation wherein the current input and the previous cell state are fed into the current state, generating the output for the subsequent cell state. The formulas for cell states and gating units are expressed as follows, where W denotes the corresponding weight matrix, and b signifies the related bias items. It is essential to highlight that σ and tanh serve as activation functions, playing a crucial role in regulating the flow of information through the gates. The formulas for these elements are as follows:

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

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

$$\widetilde{C}_t = \tanh\,(W_C \,.\, [h_{t-1}, x_t] + b_C \quad (3)$$

$$C_t = f_t \,.\, C_{t-1} + i_t \,.\, \widetilde{C}_t \quad (4)$$

$$o_t = \sigma\,(W_o \,.\, [h_{t-1}, x_t] + b_o \quad (5)$$

$$h_t = o_t \,.\, \tanh\,(C_t) \quad (6)$$

where Ct, C_{t-1} and $\tilde{C}_t$ represent current cell status value, last time frame cell status value and the update for the current cell status value, respectively. The notations f_t, i_t and o_t represent forget gate, input gate and output gate, respectively. With proper parameter settings, the output value h_t is calculated based on $\tilde{C}_t$ and C_{t-1} values according to Eqs. (4) and (6). All weights, including:W_f , W_i , W_C and W_o, are updated based on the difference between the output value and the actual value following back-propagation through time (BPTT) algorithm [32].

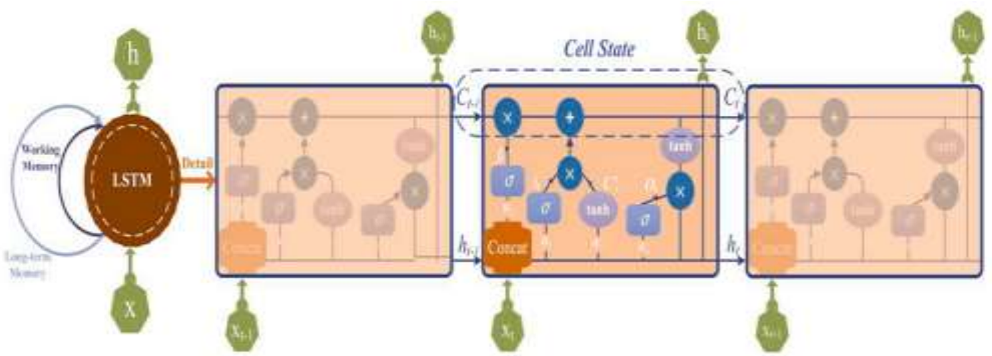

Fig-1: Detained Structure of LSTM Network

3.3. Hybrid forecasting model

The detail of the AI model algorithm is presented in Algorithm 1.

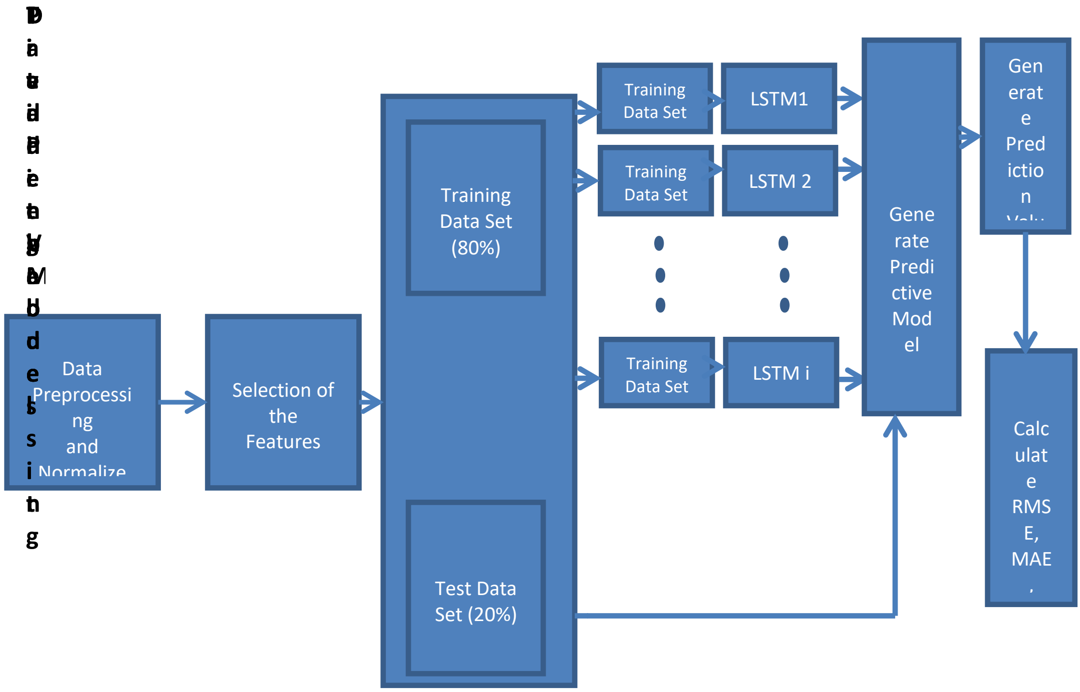

Figure 2. The schematic diagram of the proposed method.

Algorithm 1. Proposed Scheme Algorithm for Load Prediction.

1: Load daily records data
2: Do preprocessing the missing value and normalization of data
3: for k←1 to size (features) do
4: Calculate feature importance for feature
5: end for
6: Select features with importance value greater than threshold
7: Divide data into training and testing data
8: Divide the training set into i subset
9: Train parallelly each subset of training data on different LSTM model
10: Combine the output of different LSTM model
11: Predict load using trained model over testing data

The experiments have demonstrated that the hybrid model exhibits superior accuracy and stronger generalization ability. The entire process primarily involves three key steps:

Step 1: Employ the preprocessing on the actual input data by processing the missing values, data normalization and embedding. XGBoost algorithm is used for the extraction of important features from the preproceed data.

Step 2: The data obtained from Step 1 are then partitioned into multiple training sets and testing sets. These training set serve as the input for the Parallel Long Short-Term Memory (PLSTM) neural network. The network is optimized by learning the fitting process between the training dataset and the raw data of the five ECF. Combining the outcomes of each set of dataset enables the acquisition of final outcomes. After the training phase, the ECF testing dataset is utilized to obtain the ultimate result.

Step 3: The final step involves using the optimized hybrid model to predict and obtain results for the Energy Consumption Forecasting (ECF) testing dataset.

The schematic diagram of the proposed AI framework is depicted in Fig. 2. The potential advancements of the proposed method revolve around two key aspects.

Firstly, the XGBoost is utilized for feature extraction in the training dataset. Secondly, the structure of the parallel Long Short-Term Memory (PLSTM) neural network holds the potential to enhance forecasting results when compared to a centralized singular LSTM neural network. This improvement is attributed to the parallel training on the extracted features from preprocessed data. Moreover, the parallelized structure significantly reduces the training time, offering a more efficient and robust learning framework for addressing the Energy Consumption Forecasting (ECF) problem.

4. Experimental process and results

4.1. Data description

In the experiment, we get the data from the kaggle. This Electricity consumption dataset and the electric load dataset from 2016, provided by kaggle, were utilized. The experiment involved the implementation and testing of the proposed hybrid model. The results of load forecasting using this methods are depicted in Fig 3.

For the experiment, the training data encompassed a two-month period, while the test data covered one month. To assess the performance of the model, the dataset was divided into 10 partitions. In the first partition, the training data comprised energy load data collected in January and February 2016, while the test data included data collected in March 2016. The subsequent partitions followed the same analogy, with each partition representing different training and test data periods.

In Figure 3, the red curves represent the forecasting results of the respective models, while the blue curves depict the ground truth. The vertical axes indicate the energy load (MWh), and the horizontal axes signify the time (hour). The forecasting models received the energy load from the past (24 × 7) hours as input, and the predicted energy load for the next (24 × 3) hours was the output of the forecasting model. The correct information is represented by the blue curves, and the disparities between the red and blue curves signify the performance differences of the corresponding models.

To ensure fairness in comparison, testing data were not used during the training process of the models. According to the results presented in Figure 3, the proposed hybrid model demonstrates the prediction performance.

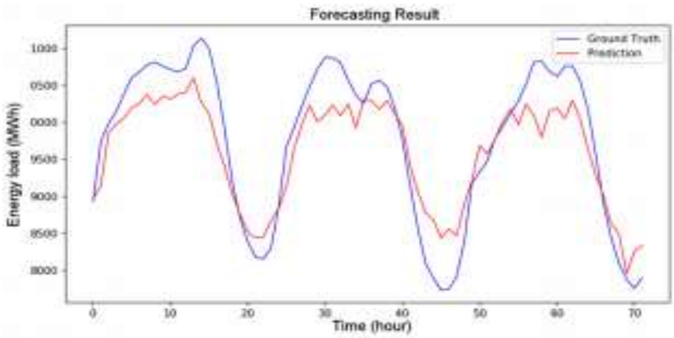

Fig 3. The forecasting result of the proposed hybrid model

To assess the forecasting models more accurately, the Mean Absolute Percentage Error (MAPE) were employed. The definitions for MAPE are provided by Equations (7) and (8), respectively, where yn represents the measured value, ˆ yn is the estimated value, and N denotes the sample size.

Table 1. The experimental results in terms of Mean Absolute Percentage Error (MAPE) given in percentages.

Test	Hybrid Model
#1	10.40804813
#2	9.970662683
#3	14.85568499
#4	12.83487893
#5	5.479091542
#6	11.7681534
#7	15.6574951
#8	7.583802292
#9	16.31443679
#10	8.390061493
Average	11.32623153

7.3 Evaluation Metrices

To assess the forecasting models more accurately, the Mean Absolute Percentage Error (MAPE) were employed. The definitions for MAPE are provided by Equations (7). The MAPE may be a live of prediction accuracy of a forecasting methodology for constructing fitted statistic values in statistics, specifically in trend estimation.it always expresses accuracy as a proportion of the error. as a result of this range may be a percentage, it may be easier to know than the opposite statistics.

$$\mathrm{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - P_t}{A_t} \right| \qquad (7)$$

where,
n = number of non-missing data points
A_t = Actual observations for tth data
P_t = Predicted Value for tth data

6. Conclusion

This paper introduces a hybrid data-driven ECF model that integrates XGBoost with a PLSTM structure, designed. The proposed model is validated through experiments using load data from the preceding seven days. In the experiment, data from the kaggle were employed, focusing on historical electricity demand from consumers. The experimental results demonstrate that hybrid model can predict energy load for the next three days. Also calculate the Mean Absolute Percentage Error (MAPE). The proposed method has the potential to reduce monitoring expenses, initial costs of hardware components, and long-term maintenance costs in future smart grids. In future, we compare the result of this model with the other existing method and use the other evaluation metices for assessing the accurate prediction of energy forecasting.

References

[1] Q. Xiao, Y. Si, Time series prediction using graph model. 2017 3rd IEEE international conference on computer and communications (ICCC), IEEE, 2017.
[2] Ke Yan, et al., Multi-step short-term power consumption forecasting with a hybrid deep learning strategy, Energies 11 (2018).
[3] B. Zhang, Y. Du, E.G. Lim, L. Jiang, K. Yan, Design and simulation of Peer-to-Peer energy trading framework with dynamic electricity price, in: 2019 29th Australasian Universities Power Engineering Conference (AUPEC), IEEE., 2019, pp. 1–6.
[4] V. Garg, N.K. Bansal, Smart occupancy sensors to reduce energy consumption, Energy Build. 32 (1) (2000) 81–87.
[5] X. Zhou, X. Xu, W. Liang, Z. Zeng, Z. Yan, Deep learning enhanced multi-target detection for end-edge-cloud surveillance in smart IoT, IEEE Internet Things J. (2021).
[6] X. Zhou, X. Yang, J. Ma, I. Kevin, K. Wang, Energy efficient smart routing based on link correlation mining for wireless edge computing in IoT, IEEE Internet Things J. (2021).
[7] P. Anand, C. Deb, K.e. Yan, J. Yang, D. Cheong, C. Sekhar, Occupancy-based energy consumption modelling using machine learning algorithms for institutional buildings, Energy Build. 252 (2021) 111478, https://doi.org/10.1016/j. enbuild.2021.111478.
[8] T. Dietz, P.C. Stern, E.U. Weber, Reducing carbon-based energy consumption through changes in household behavior, Daedalus 142 (1) (2013) 78–89.

[9] V. Masson, J. Hidalgo, A. Amoss´e, F. Belaid, E. Bocher, M. Bonhomme, et al., Urban Climate, Human behavior & Energy consumption: from LCZ mapping to simulation and urban planning (the MapUCE project). 9th International Conference on Urban Climate, 2015.
[10] X. Lü, T. Lu, C.J. Kibert, M. Viljanen, Modeling and forecasting energy consumption for heterogeneous buildings using a physical–statistical approach, Appl. Energy 144 (2015) 261–27.
[11] Raza Abid Abbasi1, Nadeem Javaid, Muhammad Nauman Javid Ghuman, Zahoor Ali Khan, Shujat Ur Rehman and Amanullah, “Short Term Load Forecasting Using XGBoost”, WAINA 2019, AISC 927, pp. 1120–1131, 2019, https://doi.org/10.1007/978-3-030-15035-8_108.
[12] Chen T, Guestrin C., “XGBoost: a scalable tree boosting system”, Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2016, p. 785–94.
[13] Y. Wang, J. Wang, G.e. Zhao, Y. Dong, Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China, Energy Policy 48 (2012) 284–294.
[14] V.G. Tran, V. Debusschere, S. Bacha, One week hourly electricity load forecasting using Neuro-Fuzzy and Seasonal ARIMA models, IFAC Proc. Vol. 45 (21) (2012) 97–102.
[15] N. Wei, C. Li, X. Peng, F. Zeng, X. Lu, Conventional models and artificial intelligence-based models for energy consumption forecasting: A review, J. Petrol. Sci. Eng. 181 (2019) 106187, https://doi.org/10.1016/j.petrol.2019.106187.
[16] H. Hu, L. Wang, L.u. Peng, Y.-R. Zeng, Effective energy consumption forecasting using enhanced bagged echo state network, Energy 193 (2020) 116778, https:// doi.org/10.1016/j.energy.2019.116778.
[17] H. Zamani Sabzi, S. Abudu, R. Alizadeh, L. Soltanisehat, N. Dilekli, J.P. King, Integration of time series forecasting in a dynamic decision support system for multiple reservoir management to conserve water sources, Energy Sources Part A 40 (11) (2018) 1398–1416.
[18] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of recurrent network architectures”, in Proc. Int. Conf. Mach. Learn., 2015, pp. 2342-2350.
[19] Khan, A.N., Iqbal, N., Rizwan, A., Ahmad, R., Kim, D.H., “An Ensemble Energy Consumption Forecasting Model Based on Spatial-Temporal Clustering Analysis in Residential Buildings”, Energies 2021, 14, 3020. https://doi.org/10.3390/en14113020.
[20] Mohamed Massaoudia, Shady S. Refaat, Ines Chihi, Mohamed Trabelsi, Fakhreddine S. Oueslati and Haitham Abu-Rub, “A novel stacked generalization ensemble-based hybrid LGBM-XGBMLP model for Short-Term Load Forecasting”, Energy 214 (2021) 118874, https://doi.org/10.1016/j.energy.2020.118874.
[21] Y.-Y. Song, L.U. Ying, Decision tree methods: applications for classification and prediction, Shanghai Arch. Psychiatry 27 (2) (2015) 130.
[22] A.J. Hill, G.R. Herman, R.S. Schumacher, Forecasting severe weather with random forests, Mon. Weather Rev. 148 (5) (2020) 2135–2161.
[23] K. Kavaklioglu, Modeling and prediction of Turkey’s electricity consumption using Support Vector Regression, Appl. Energy 88 (1) (2011) 368–375.
[24] Timo Koskela, et al., Time series prediction with multilayer perceptron, FIR and Elman neural networks, in: Proceedings of the World Congress on Neural Networks. INNS Press San Diego, USA, 1996.
[25] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, Y. Wang, Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction, Sensors 17 (4) (2017) 818, https://doi.org/10.3390/s17040818.
[26] Samy Bengio, et al., Scheduled sampling for sequence prediction with recurrent neural networks, arXiv preprint arXiv:1506.03099, 2015.
[27] F. Altch´e, A. de La Fortelle, An LSTM network for highway trajectory prediction. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2017.

[28] Junliang Fan, Xiukang Wang, Lifeng Wu, Hanmi Zhou, Fucang Zhang, Xiang Yu, Xianghui Lu and Youzhen Xiang, “Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China”, Energy Conversion and Management 164 (2018) 102–111, https://doi.org/10.1016/j.enconman.2018.02.087.
[29] Jihoon Moona, Seungwon Jung, Jehyeok Rew, Seungmin Rho and Eenjun Hwang, “Combination of short-term load forecasting models based on a stacking ensemble approach”, Energy & Buildings 216 (2020) 109921, https://doi.org/10.1016/j.enbuild.2020.109921.
[30] N. Wei, C. Li, X. Peng, Y. Li, F. Zeng, “Daily natural gas consumption forecasting via the application of a novel hybrid model”, Appl. Energy 250 (2019) 358–368.
[31] T. Sun, T. Zhang, Y. Teng, Z. Chen, J. Fang, “Monthly electricity consumption forecasting method based on X12 and STL decomposition model in an integrated energy system”, Mathe. Probl. Eng. 2019 (2019) 1–16.
[32] Jihoon Moona, Seungwon Jung, Jehyeok Rew, Seungmin Rho and Eenjun Hwang, “Combination of short-term load forecasting models based on a stacking ensemble approach”, Energy & Buildings 216 (2020) 109921, https://doi.org/10.1016/j.enbuild.2020.109921.