12-2020

Predicting Residential Energy Consumption Using Wavelet Decomposition With Deep Neural Network

Dagimawi Eneyew
Western University, deneyew@uwo.ca

Miriam A M Capretz
Western University, mcapretz@uwo.ca

Girma Bitsuamlak
Western University, gbitsuam@uwo.ca

London Hydro

Predicting Residential Energy Consumption Using Wavelet Decomposition With Deep Neural Network

Dagimawi D. Eneyew[*], Miriam A. M. Capretz[†], Girma T. Bitsuamlak[‡], Syed Mir[§]

[*][†]*Department of Electrical and Computer Engineering,* [‡]*Department of Civil and Environmental Engineering*
Western University, London, Ontario, Canada, N6A 5B9
Email: [*]deneyew@uwo.ca, [†]mcapretz@uwo.ca, [‡]gbitsuam@uwo.ca
[§]*London Hydro*, London, Ontario, Canada, N6A 4H6
Email: mirs@londonhydro.com

Abstract—Electricity consumption is accelerating due to economic and population growth. Hence, energy consumption prediction is becoming vital for overall consumption management and infrastructure planning. Recent advances in smart electric meter technology are making high-resolution energy consumption data available. However, many parameters influencing energy consumption are not typically monitored for residential buildings. Therefore, this study's main objective is to develop a data-driven energy consumption forecasting model (next-hour consumption) for residential houses solely based on analyzing electricity consumption data. This research proposes a deep neural network architecture that combines stationary wavelet transform features and convolutional neural networks. The proposed approach utilizes automatically extracted features from smart-meter readings by applying wavelet decomposition, convolution, and pooling operations. This study's findings have demonstrated the advantage of integrating wavelet features with convolutional neural networks to improve forecasting accuracy while automating feature extraction.

Index Terms—convolutional neural network, deep neural network, energy consumption prediction, wavelet decomposition

I. INTRODUCTION

Electricity is one of the primary energy sources in today's world, and its rate of consumption continues to accelerate due to economic and population growth. The residential building sector consumes 27% of global energy generated [12]. Therefore, energy consumption forecasting is crucial for residential energy demand management, electricity price market design, energy efficiency, and maintenance scheduling of large-scale complex smart power grids [12]. In addition to forecasting, anomaly detection models provide vital information to improve energy efficiency [1] [2]. Recently, large amounts of high-resolution consumption data are becoming available from smart electric meters. The availability of this information has motivated (data-driven) studies on forecasting energy consumption for residential and commercial buildings.

Residential energy consumption forecasting is a multivariate time-series prediction problem where weather, occupancy type and behavior, type of appliances, season, time of day, price, building age, type, and size contribute to electricity consumption. However, many influencing parameters are not typically monitored in residential buildings. Often the only available high-resolution information is the energy consumption meter reading. Prediction based on the consumption reading without

supporting features makes the problem univariate. Residential buildings have different consumption levels and energy consumption patterns. For example, each family may have a set of appliances selected to suit their needs based on their lifestyle. Each appliance contributes to overall energy consumption, depending on how and when it is used. Weather variables like temperature and humidity also have an impact on overall consumption. Building form and construction type affect energy consumption. In addition, calendar features affect the usage pattern. For instance, consumption during weekends, during morning hours, and late at night have significantly different patterns. The impact of all the factors mentioned above is embodied in the smart-meter data.

This research's main objective is to develop a data-driven electricity consumption forecasting model (next-hour consumption) for residential houses, solely based on analyzing electricity consumption data. Therefore, an end-to-end forecasting framework that combines the stationary wavelet transform and deep neural networks is proposed. The stationary wavelet transform operation is employed to decompose the single-source (smart-meter) readings into low frequency and high-frequency components. After the wavelet decomposition, a single layer of convolution followed by pooling operation is used to extract abstract features. Furthermore, the proposed approach can generate the final prediction from a single model by aggregating wavelet features through convolution and pooling operations. Thus, unlike existing wavelet-based methods, modeling each wavelet sub-sequence and inverse wavelet transform is not required.

The remaining sections of the paper are organized as follows. In Section II, related studies in the area of residential electricity consumption forecasting are summarized. Section III provides a brief overview of common algorithms used in energy forecasting and evaluation metrics to gauge performance. Section IV describes the proposed forecasting architecture. Section V presents the evaluation process, including a brief description of the data, experiments, and results. Section VI highlights conclusions and future work.

II. RELATED WORK

Extensive research on forecasting energy consumption for residential and commercial buildings is accessible in the

literature. Recently, numerous researchers have focused on improving electricity consumption forecasting efficiency and accuracy using deep learning architectures [13] [16]. K-pri-janovska et al. [13] used a deep residual neural network to forecast household energy consumption by combining multiple sources of information. Deep reinforcement learning has also been investigated for building energy consumption forecasting [16] and compared with conventional supervised models. The results showed that deep reinforcement learning models improved a mean absolute error for single-step and multi-step predictions.

Hybrid models combining the advantage of different algorithms are also recommended [7] [11] [12] [17] [20] [21] [26]. Hafeez et al. [7] introduced a hybrid framework by combining feature engineering, a modified enhanced differential evolution (mEDE) based optimizer and a support vector machine. The approach was evaluated using hourly load data and resulted in improvements relative to other benchmark frameworks. Related work by Shao et al. [20] modified a conventional stacked CNN-LSTM architecture for electric consumption forecasting. The proposed hybrid model achieved robust results for short and long-term forecasting scenarios.

Machine learning-based electricity consumption models have also been proposed [3] [8] [25] [28]. Tran et al . [25] introduced the Evolutionary Neural Machine Inference Model (ENMIM) for estimating residential energy consumption. Comparative results revealed that the ENMIM model surpassed various benchmark models. Related work by Zhang et al. [28] applied the Support Vector Regression (SVR) algorithm to hourly and daily aggregated residential electricity consumption data. According to their findings, daily data granularity achieved better prediction than hourly data for residential buildings.

Some researchers have used statistical methods to forecast electricity consumption [5] [23] [24]. Tang et al., [24] offered a novel probability forecasting model for long-term electricity consumption. Their case study showed that the proposed methodology has higher accuracy and adaptability than other tested methods.

Some studies have used wavelet decomposition techniques for time series forecasting [10] [22] [27]. A maximal overlap discrete wavelet transform with LSTM has been used to predict energy consumption in various sectors [10]. The prediction models were built for each wavelet decomposed series and aggregated at the final stage. The approach was evaluated using monthly prediction data from residential, commercial, and industrial sectors. This method outperformed the recurrent neural network (RNN) models. Stolojescu et al. [22] also recommended a wavelet-based forecasting framework. This approach achieved a better result than a neural network, linear regression, and a random walk.

Despite significant research, individual house-level consumption forecasting solely from energy consumption data warrant further investigations to handle challenges originating from irregular and intricate usage patterns. The studies reviewed in this paper used either multiple indicator features or only the electricity consumption variable as input. Both approaches have some downsides. The most common features used in electricity consumption forecasting are weather variables. However, the correlation between weather variables and consumption patterns is not always guaranteed or may be very weak. For example, airport weather measurements typically used for analysis may not represent the monitored building's micro-climate. Also, using the temperature reading as an indicator might not be relevant when the residential building uses other heating and cooling energy sources. On the other hand, using only the electricity consumption reading to sequence models has limitations when a complicated temporal relationship exists between consecutive readings. Wavelet-based forecasting methods proposed in previous papers also have limitations. The most widely used approach fits multiple models for each decomposed sub-sequence. Multiple model fitting will become computationally expensive for a very long input sequence with multi-level decomposition.

In this study, we propose an end-to-end forecasting framework that combines the stationary wavelet transform and convolutional neural network. The framework uses automatically extracted features from smart-meter readings by applying wavelet decomposition, convolution, and pooling operations. Unlike existing wavelet-based methods, the proposed approach can generate the final prediction without modeling each wavelet sub-sequence and inverse wavelet transform.

III. BACKGROUND

A. Neural Network

Feed Forward Neural Networks (FFNN) consists of neurons grouped into layers (input, hidden, and output) [9]. Deep feedforward neural networks provide a multi-hidden-layer approach to learn data representations [18]. The goal of the feedforward network is to approximate some function $y = f(x)$. It defines a mapping function y and learns the network parameters that result in the best function approximation. The term deep neural network generally implies multiple layers between the input and output layers [18].

B. Long Short Term Memory Recurrent Neural Network (LSTM-RNN)

Recurrent neural networks (RNN) can model sequential information by taking what has been calculated and using it in the current output computation. Due to their suitability for sequential data, RNNs are commonly used for time-series data. However, regular recurrent neural networks are usually hard to train and do not contain memory units to remember context [19]. Thus, when long-range contextual information is needed, RNNs become ineffective. Long short-term memory networks (LSTM) are a special kind of RNN designed to handle long-term dependencies, where more extended information is needed [6].

C. Convolutional Neural Network

Convolutional neural networks (CNN) are a specialized type of neural network for working with two-dimensional data.

CNN's ensure shift and distortion invariance through local receptive fields, shared weights, and spatial and temporal sub sampling [15]. One of CNN's main advantages is their ability to learn useful features automatically from high-dimensional data [14]. Recently, convolutional neural networks have been widely used in time-series forecasting due to their ability to learn filters representing repeated patterns in the time-series variable. A CNN includes two main types of layers: the convolutional layer and the pooling layer. The convolutional layer convolves the input with a pre-specified size called a filter to compute a feature map. A max-pooling layer typically follows the convolutional layer. This layer extracts the maximum value within a given pooling size [14].

D. Stationary Wavelet Transform

A wavelet transform decomposes signals into different sub-signals with different frequencies using a series of functions called wavelets. Wavelets are localized in both time and frequency. The wavelet transform has continuous and discrete versions. The stationary wavelet transform (SWT) is an extension of the discrete wavelet transform (DWT). The original sequence data are generally decomposed into an approximate (low-frequency) sub-sequence and a detailed (high-frequency) sub-sequence. Both the high-frequency and low-frequency components have the same length as the original input. Reconstruction of the original signal is feasible using an inverse wavelet transform operation [4].

E. Performance Metric

1) Mean absolute percentage error (MAPE): The mean absolute percentage error (MAPE) is the mean or average of the absolute percentage errors of forecasts. This measure is easy to understand because it provides the error in terms of percentages. It is calculated using Equation (1).

$$MAPE = \frac{100\%}{n} \sum \left| \frac{y_t - y_p}{y_t} \right| \qquad (1)$$

IV. PROPOSED METHOD

Fig. 1 shows the proposed architecture for residential electricity consumption forecasting. The architecture consists of preprocessing, feature extraction, and model fitting components. The following sections explain each component in detail.

A. Data Preprocessing

The input data for the proposed model consist of univariate electricity consumption readings. Before this data is used for model fitting, it must pass through three main preprocessing phases.

1) Daily Input Grouping: This stage groups the readings using 24-hour-long non-overlapping windows to capture the daily usage characteristics of users. Full day (24-hour) usage patterns show associated peak hours and idle hours within the time frame. The level of consumption might vary from day to day, depending on different contributing factors. However,

daily usage patterns contain enough information to localize the feature extraction in the time domain.

2) Stationary Wavelet Transform: After daily input grouping, each 24-hour reading from the previous step passes independently through a multi-level stationary wavelet decomposition to generate stationary sub-sequences. The decomposition is also separate for training and testing data. Haar's mother wavelet is used as a basis function because of its advantage in capturing a signal's local aspects. A three-level decomposition is used, which is the maximum possible for input with a length of 24. The decomposition phase generates six coefficients with the same length as the original input. However, the last-level approximate coefficient and detailed coefficients at each level are kept because they are sufficient to reconstruct the input. Finally, each day's wavelet features are stacked into a single dataset while maintaining the time order. The resulting data have four wavelet features for each consumption reading.

3) Time Series to Supervised Learning: This step aims to shape the data into a supervised learning problem. This study used 24-hour window wavelet features as input and the next-hour consumption reading as output. The final output of this phase is usable with supervised learning algorithms.

B. Feature Extraction

This component uses one-dimensional convolution and max-pooling operations to extract complex features from the input data.

1) 1D-Convolution Layer: The convolution layer convolves the input with a weight matrix called a filter to compute a feature map. Feature maps contain useful data features; however, a further operation follows this step to select the most important ones.

2) Max Pooling Layer: The max-pooling layer takes feature maps as input and sub-samples it by selecting maximum values within a given pooling size. It provides a compressed form of representation for the feature map while reducing model parameters.

C. Model Fitting

The model fitting component consists of a dense or LSTM layer, followed by an output layer. This component fits a nonlinear function between the extracted features and the output variable using hidden layers.

1) Flattening Layer: This layer takes the output of the max-pooling layer and converts it into a linear array. This step is important because the following layers expect a linear input to fit the model function.

2) Dense Layer/LSTM Layer: The dense/LSTM layers take the linear input from the flattening layer and fit a forecasting model. The neurons in dense layers use a rectified linear unit(ReLU) as an activation function, and tan-hyperbolic activation is used for LSTM layers.

3) Output Layer: The output layer contains a single neuron with a linear activation function. The output from this layer is the predicted consumption value.
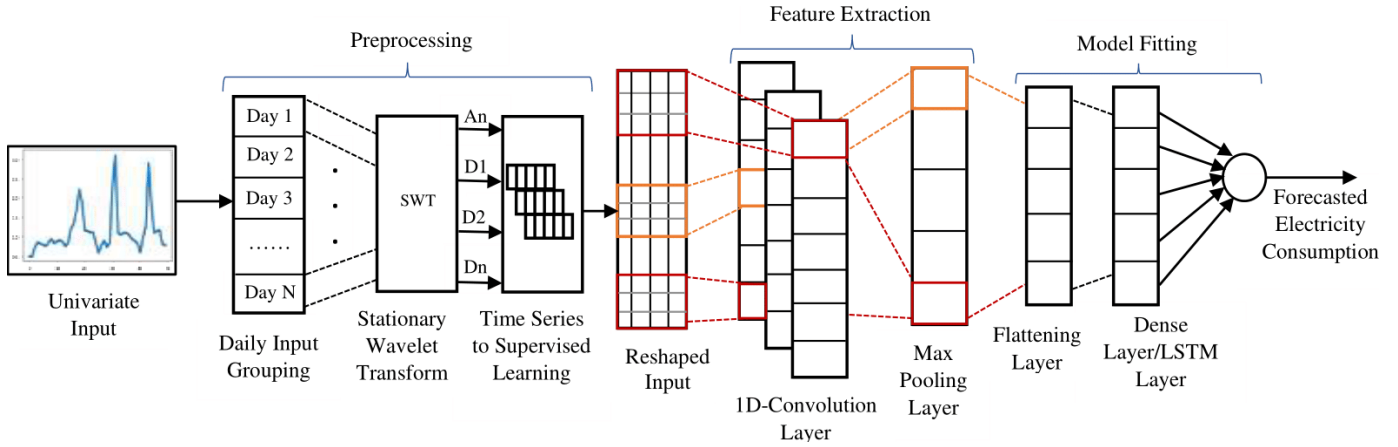
Fig. 1. Proposed architecture for electricity consumption prediction.

V. EVALUATION

A. Dataset

The proposed method was evaluated using residential electricity consumption from London Hydro (the electricity service provider in London, Ontario). The data consist of the hourly smart-meter readings in kWh for ten houses from January 2014 to January 2020. A total of 52,752 hours of readings were prepared for each house. In the experiments, a 70%-10%-20% train-validation-test split ratio was used to train and evaluate the models. Fig. 2 summarizes six years' hourly consumption for all houses as a box plot. Based on the plot, all houses had different overall consumption levels. In addition, the consumption readings for all houses contained outlier values. The lowest reading for all houses was zero, which might have been related to a temporary power outage or no appliances usage.
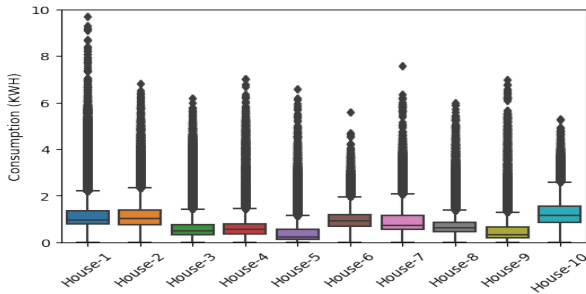

Fig. 2. Box Plot For Six Years Electricity Consumption

B. Training and Hyper-Parameter Tuning

Model training was conducted using the training and validation process. A mean square objective function was optimized during training using an Adam optimizer. Critical hyper-parameters were tuned for the convolutional layer and the model fitting layer. Hyperparameter's search space is shown in Table-I. A random search tuning approach was used to generate fifty unique random combinations for each model

due to a large number of possible combinations of these parameters. Finally, the best hyper-parameters were selected based on their validation score.

Hyper Parameter	Values Tested
Filter Size	16, 32, 64, 128
Kernel Size	2, 4, 6, 8
Pooling Size	2, 3, 4, 5
Number of Neurons	20, 30, 50, 60

C. Experiments and Results

Three experimental scenarios were considered for this study. The first experiment set the baseline results. The second and third experiments evaluated the performance of the proposed approach. The performance of the models was evaluated using the mean absolute percentage error (MAPE) metric. The following sections describe each experiment in detail.

1) Experiment-1: (1D-Convolution)-Dense Model: This experiment builds a forecasting model for the ten houses by taking the consumption data sequence as a univariate input. The input time-series data was first transformed into supervised learning data by applying a sliding window method with a twenty-four window size. The models for all houses had one convolutional layer, a max-pooling layer, a flattening layer, and three fully connected dense layers. Rectified linear unit (ReLU) activation was used for all neurons within the fully connected layers. A mean squared error objective function with an Adam optimizer was used to train the models. The training and tuning process followed the procedure described in the training and hyper-parameter tuning subsection. This experiment's results set the baseline results to compare with the proposed method (Experiment-2 and Experiment-3).

2) Experiment-2: SWT-(1D-Convolution)-Dense Model: The input for this experiment was also only the consumption reading for each house. Unlike the baseline experiment, the input was not directly used. This experiment implemented

the suggested approach using dense layers in the model fitting component. After the preprocessing step, the final input had detailed coefficients at each decomposition level and an approximate coefficient at the last decomposition level. The remaining network configurations are the same as the baseline models.

3) Experiment-3: SWT-(1D-Convolution)-LSTM Model: This experiment aimed to improve the result from Experiment-2 by replacing the fully connected layer with a three-layer LSTM. All LSTM cells within the network had a tan-hyperbolic activation function. The three-dimensional tensor from the sliding window output was further subdivided into sub-sequences to fulfill the input shape requirements of LSTM. The remaining network configurations and tuning process of this experiment were kept similar to Experiment-2.

TABLE II
EXPERIMENT-1, EXPERIMENT-2 AND EXPERIMENT-3 EVALUATION RESULTS

House	Experiment-1 MAPE	Experiment-2 MAPE	Experiment-3 MAPE
House-1	36.276	11.906	11.305
House-2	26.925	10.302	7.169
House-3	36.803	5.541	5.244
House-4	30.161	10.741	9.244
House-5	36.168	16.249	18.238
House-6	7.330	2.211	1.972
House-7	36.222	9.434	8.310
House-8	65.273	18.315	10.909
House-9	37.467	15.324	13.640
House-10	23.328	5.332	3.884

The results from the baseline experiment (Experiment-1) are shown in Table II. The best performance achieved was 7.330% for House-6, and the worst percentage error was 65.273% for House-8. For the remaining houses, the percent error was greater than 30% except for House-2 and House-10.

The results from the proposed approach (Experiment-2, Experiment-3) are also shown in Table II. The SWT-(1D-Convolution)-Dense (Experiment-2) model significantly improved the results compared to the baseline model. The lowest MAPE was 2.211% for House-6, and the highest MAPE was 18.315% for House-8. The models for House-3, House-7, and House-10 had a percentage error of less than 10%. The remaining houses achieved less than 20% MAPE. Similarly, the SWT-(1D-Convolution)-LSTM(Experiment-3) model improved the baseline results and the results of the SWT-(1D-Convolution)-Dense(Experiment-2) model. The best model attained 1.972% MAPE for House-6, and the lowest performance was 18.238% MAPE for the House-5 model. In general, the proposed method improved forecasting accuracy without using manually designed features or multiple models. The actual vs. prediction plots of the best and worst-performing SWT-(1D-Convolution)-LSTM and SWT-(1D-Convolution)-Dense models are shown in Figs. 3, 4, 5, and 6. The plots show only the first two hundred hours of prediction to display the results with better resolution.
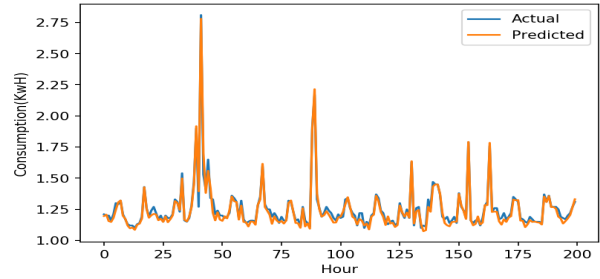

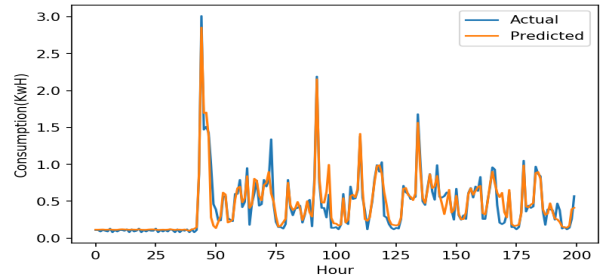Fig. 3. Actual Vs Prediction plot for best (House-6) SWT-(1D-Convolution)-LSTM Model

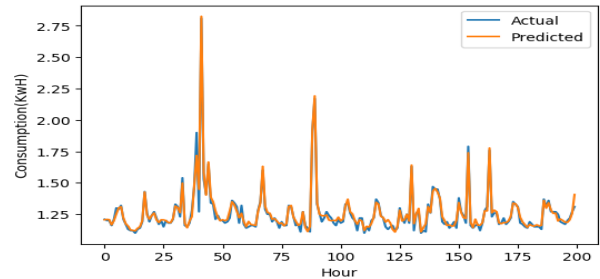Fig. 4. Actual Vs Prediction plot for worst (House-5) SWT-(1D-Convolution)-LSTM Model

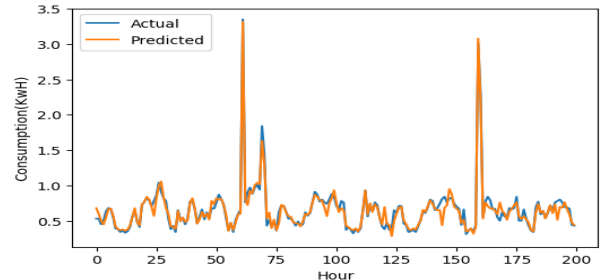Fig. 5. Actual Vs Prediction plot for best (House-6) SWT-(1D-Convolution)-Dense Model

Fig. 6. Actual Vs Prediction plot for worst (House-8) SWT-(1D-Convolution)-Dense Model

Based on Figs. 3, 4, 5, and 6, the prediction curve of the proposed model follows the actual usage values closely. In addition, the predictions of peak usage hours are very close to the actual values. For the worst-performing SWT-(1D-Convolution)-LSTM model (Fig. 4), highly fluctuating values were difficult to predict. For the best-performing models in both approaches (Fig. 3 and Fig. 5), the SWT-(1D-Convolution)-LSTM-based model resulted in a slightly better prediction than the SWT-(1D-Convolution)-Dense network.

VI. Conclusion

Advances in energy consumption data collection from smart meters motivate data-driven energy consumption forecasting approaches in residential and commercial buildings. Many studies have been conducted in residential energy consumption forecasting. This study also focused on hourly energy usage prediction for residential buildings.

This study proposed a model that combines wavelet features with convolution and pooling operations to extract robust features. Unlike existing wavelet-based methods, the proposed approach can generate the final prediction without modeling each wavelet sub-sequence and inverse wavelet transform. The research findings have confirmed that the proposed framework has significantly improved over the standard univariate input (1D-convolution)-Dense model. The results have further proved that integrating stationary wavelet features with CNN can extract relevant features to improve forecasting accuracy.

In this study, a single convolution and max-pooling layer are used for feature extraction purposes. In the future, multiple layers of convolution, max-pooling, and flexible window size should be tested for residential houses that use mixed energy sources (gas, electricity, and other local renewable energy sources).

VII. Acknowledgment

References

[1] Daniel B Araya, Katarina Grolinger, Hany F ElYamany, Miriam AM Capretz, and G Bitsuamlak. Collective contextual anomaly detection framework for smart buildings. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 511–518. IEEE, 2016.

[2] Daniel B Araya, Katarina Grolinger, Hany F ElYamany, Miriam AM Capretz, and Girma Bitsuamlak. An ensemble learning framework for anomaly detection in building energy consumption. *Energy and Buildings*, 144:191–206, 2017.

[3] Karol Bot, Antonio Ruano, and Maria da Graça Ruano. Forecasting electricity consumption in residential buildings for home energy management systems. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 313–326. Springer, 2020.

[4] Liu Chun-Lin. A tutorial of the wavelet transform. *NTUEE, Taiwan*, 2010.

[5] Cheng Fan, Yongjun Sun, Fu Xiao, Jie Ma, Dasheng Lee, Jiayuan Wang, and Yen Chieh Tseng. Statistical investigations of transfer learning-based methodology for short-term building energy predictions. *Applied Energy*, 262:114499, 2020.

[6] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. 1999.

[7] Ghulam Hafeez, Khurram Saleem Alimgeer, Abdul Baseer Qazi, Imran Khan, Muhammad Usman, Farrukh Aslam Khan, and Zahid Wadud. A hybrid approach for energy consumption forecasting with a new feature engineering and optimization framework in smart grid. *IEEE Access*, 2020.

[8] Yaoyao He, Yang Qin, Shuo Wang, Xu Wang, and Chao Wang. Electricity consumption probability density forecasting method based on lasso-quantile regression neural network. *Applied Energy*, 233:565–575, 2019.

[9] Riki Herliansyah et al. Feed forward neural networks for forecasting indonesia exchange composite index. *GSTF Journal of Mathematics, Statistics & Operations Research*, 4(1), 2017.

[10] Rabin K Jana, Indranil Ghosh, and Manas K Sanyal. A granular deep learning approach for predicting energy consumption. *Applied Soft Computing*, 89:106091, 2020.

[11] Weiheng Jiang, Xiaogang Wu, Yi Gong, Wanxin Yu, and Xinhui Zhong. Holt–winters smoothing enhanced by fruit fly optimization algorithm to forecast monthly electricity consumption. *Energy*, 193:116779, 2020.

[12] Tae-Young Kim and Sung-Bae Cho. Predicting residential energy consumption using cnn-lstm neural networks. *Energy*, 2019.

[13] Ivana Kiprijanovska, Simon Stankoski, Igor Ilievski, Slobodan Jovanovski, Matjaž Gams, and Hristijan Gjoreski. Houseec: Day-ahead household electrical energy consumption forecasting using deep learning. *Energies*, 13(10):2672, 2020.

[14] Irena Koprinska, Dengsong Wu, and Zheng Wang. Convolutional neural networks for energy time series forecasting. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[15] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

[16] Tao Liu, Zehan Tan, Chengliang Xu, Huanxin Chen, and Zhengfei Li. Study on deep reinforcement learning techniques for building energy consumption forecasting. *Energy and Buildings*, 208:109675, 2020.

[17] Jihoon Moon, Seungwon Jung, Jehyeok Rew, Seungmin Rho, and Eenjun Hwang. Combination of short-term load forecasting models based on a stacking ensemble approach. *Energy and Buildings*, page 109921, 2020.

[18] Josh Patterson and Adam Gibson. *Deep learning: A practitioner's approach*. " O'Reilly Media, Inc.", 2017.

[19] Gábor Petneházi. Recurrent neural networks for time series forecasting. *arXiv preprint arXiv:1901.00069*, 2019.

[20] Xiaorui Shao, Chang-Soo Kim, and Palash Sontakke. Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion cnn–lstm. *Energies*, 13(8):1881, 2020.

[21] Nivethitha Somu, Gauthama Raman MR, and Krithi Ramamritham. A hybrid model for building energy consumption forecasting using long short term memory networks. *Applied Energy*, 261:114131, 2020.

[22] Cristina Stolojescu, Ion Railean, Sorin Moga, Philippe Lenca, and Alexandru Isar. A wavelet based prediction method for time series. In *Proceedings of Stochastic Modeling Techniques and Data Analysis (SMTDA2010) International Conference, Chania, Greece*, pages 767–774, 2010.

[23] Souhaib Ben Taieb, James W Taylor, and Rob J Hyndman. Hierarchical probabilistic forecasting of electricity demand with smart meter data. *Journal of the American Statistical Association*, pages 1–17, 2020.

[24] Lei Tang, Xifan Wang, Xiuli Wang, Chengcheng Shao, Shiyu Liu, and Shijun Tian. Long-term electricity consumption forecasting based on expert prediction and fuzzy bayesian theory. *Energy*, 167:1144–1154, 2019.

[25] Duc-Hoc Tran, Duc-Long Luong, and Jui-Sheng Chou. Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings. *Energy*, 191:116552, 2020.

[26] Ran Wang, Shilei Lu, and Wei Feng. A novel improved model for building energy consumption prediction based on model integration. *Applied Energy*, 262:114561, 2020.

[27] Ke Yan, Wei Li, Zhiwei Ji, Meng Qi, and Yang Du. A hybrid lstm neural network for energy consumption forecasting of individual households. *IEEE Access*, 7:157633–157642, 2019.

[28] Xiaoou Monica Zhang, Katarina Grolinger, Miriam AM Capretz, and Luke Seewald. Forecasting residential energy consumption: Single household perspective. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 110–117. IEEE, 2018.