

## Customer Active Power Consumption Prediction for the Next Day Based on Historical Profile

Ahmad A. Goudah<sup>1</sup>, Mohmed El-Habrouk<sup>2</sup>, Dieter Schramm<sup>3</sup> and Yasser G. Dessouky<sup>1</sup>

<sup>1</sup>Arab Academy for Science, Technology and Maritime Transport,

<sup>2</sup>Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Egypt,

<sup>3</sup>Institute of Mechatronics University Duisburg-Essen D-47057 Duisburg, Germany,

Email: {ahmad.goudah@aast.edu, eepgmme1@yahoo.co.uk, dieter.schramm@uni-due.de, ygd@aast.edu}

Received on, 12 January 2022 - Accepted on, 07 May 2022 - Published on, 19 June 2022

### ABSTRACT

Energy consumption prediction application is one of the most important fields that is artificially controlled with Artificial Intelligence technologies to maintain accuracy for electricity market costs reduction. This work presents a way to build and apply a model to each costumer in residential buildings. This model is built by using Long Short Term Memory (LSTM) networks to address a demonstration of time-series prediction problem and Deep Learning to take into consideration the historical consumption of customers and hourly load profiles in order to predict future consumption. Using this model, the most probable sequence of a certain industrial customer's consumption levels for a coming day is predicted. In the case of residential customers, determining the particular period of the prediction in terms of either a year or a month would be helpful and more accurate due to changes in consumption according to the changes in temperature and weather conditions in general. Both of them are used together in this research work to make a wide or narrow prediction window.

A test data set for a set of customers is used. Consumption readings for any customer in the test data set applying LSTM model are varying between minimum and maximum values of active power consumption. These values are always alternating during the day according to customer consumption behavior. This consumption variation leads to leveling all readings to be determined in a finite set and deterministic values. These levels could be then used in building the prediction model. Levels of consumption's are modeling states in the transition matrix. Twenty-five readings are recorded per day on each hour and cover leap years extra ones. Emission matrix is built using twenty five values numbered from one to twenty five and represent the observations. Calculating probabilities of being in each level (node) is also covered. Logistic Regression Algorithm is used to determine the most probable nodes for the next 25 hours in case of residential or industrial customers.

*Index Terms—Smart Grids, Load Forecasting, Consumption Prediction, Long Short Term Memory (LSTM), Logistic Regression Algorithm, Load Profile, Electrical Consumption.*

## I. INTRODUCTION

SMART grids, which are robust and efficient power grids, are going to become the grids of the future. Smart grids, basically, consist of a two-way flow of electricity and information as the U.S. Department of Energy (DOE) suggests. This is used to monitor customer electrical consumption behavior on the grid [1]. At the same time, Advanced Metering Infrastructure (AMI) is collecting real time data during the day as consumption information is considered as an important and essential data logger resource. Millions of smart meters are assembled worldwide in Smart Grids. Electrical industry is deregulated globally due to the increase in demand consumption [2].

Detailed daily customer consumption individual data is obtained by AMI in order to achieve a better figure about customer behaviors. It is worthwhile to note here that customers' consumption behavior patterns are completely different even if they are from the same category (residential, commercial, industrial, etc.) [3]. Work in this paper depends on data provided by a Spain company coordinated with Oviedo university in Spain. This data set is for the hourly consumption profile (Active and reactive power) of more than 3,000,000 clients in a period of three years as shown in figure 1.

In this work, extracting customer consumption patterns from his load profiling is done by using a Hidden Markov Model to predict future consumption for that customer.

- **Dia:** The date on which the energy usage occurred, in the following format: YYYY/MM/DD
- **H1, H2, ..., H25:** The time at which the energy usage occurred. There are 25 entries due to the time change when the clocks are set back one hour on the last Sunday of October each year, which leads to a total of 25 hours on this day.
- **ACTIVA\_H1, ACTIVA\_H2, ..., ACTIVA\_H25:** Active energy consumed per hour as measured in kWh (kilowatt-hour). This is the useful energy that the customers absorb from the grid and transform into work and/or heat at home.
- **REACTIVA\_H1, REACTIVA\_H2, ..., REACTIVA\_H25:** Reactive energy consumed per hour in VArh (volt-ampere-hours). This is a supplementary usage that the customers cannot take advantage of. Currently reactive energy consumption is not billed, although it could be charged for in future as a way of improving the energy efficiency of homes.
- **DE\_MUNICIPIO:** The municipal district to which the customer belongs. Geographical reference.
- **FECHA\_ALTA\_STRO:** Date on which the customer's service was activated with the following format: QXAAA
- **TARGET\_TENENCIA\_CUPS:** Probability that the municipal district in question is already equipped with a natural gas distribution network (which does not imply that the customer had contracted natural gas service).
- **IDENTIFICADOR:** The unique, customer reference number which allows for segmentation of usage per customer.
- **CNAE:** (National Classification of Economic Activities) This value indicates whether the customer is domestic (T1) or not (T2).
- **PRODUCTO:** Tariff / electrical product that the customer has contracted; there are up to 120 products.
- **MERCADO:** This value indicates whether the customer has a regulated tariff (M1) or a free market tariff (M2)
  - Regulated Tariff: the price of the electricity is regulated periodically by the corresponding authority.
  - Free market: the price of the electricity is freely agreed upon by the provider and the customer.

Fig. 1. Electrical company data structure model

However, customer consumption prediction is useful for determining the future electrical needs for city, governorate and country now and in the coming years [1]. This information about what is predicted for the future is translated directly into profits and gradually affect companies marketing strategies [4]. Such systems are improving the overall efficiency of the electricity network and a better consumption behavior is achieved. This informative system leads to a good understanding for organizations to accurately target customer behavior modifications and act accordingly [2], [4]. Different customer behaviors are useful for electrical companies to offer some incentives to change household electrical behavior to a much proper way [2]. Electrical Demand Market now is ready for expectations released from customer consumption prediction to lead a successful energy utilization [5]. In this paper, Hidden Markov Model (HMM) is used to model customer historical data for the average of 300 days consequently recorded. Then Logistic Regression Algorithm is applied to determine the most probable consumption levels sequence during the day. The same process is applied to customers-base of different users to apply the idea and verify that it is a valid prediction model.

## A. SYSTEM BLOCK DIAGRAM

Household consumer Demand Response (DR) term concerns with three types of processes applied to customer's energy consumption. These types are categorize, predict and modify customer's energy consumption. It is an important tool for improving a utility's economic and energy efficiency, reducing emissions, and integrating renewable devices [2].

Paper [2] presents a shape-based approach and is mainly based on Dynamic Time Warping DTW. Hidden patterns of regular consumer behavior are observed and reflected by an optimal alignment between energy consumption patterns using DTW. Under DTW distance, two valuable benefits were achieved. Firstly, 50% reduction in the number of representative groups of electrical household consumers. Secondly, measured prediction accuracy is improved. Extendedly, determining which device is used in any particular hour from consumption curve analysis. Used devices and their fundamental structures are the main base to make classification for household customers' consumptions. DTW method is comparing and classifying household load curves. DTW, K- means and Gaussian based E&M algorithms are mentioned in this paper as different clustering algorithms. Three topics are discussed in [2]. First, clustering with DTW.

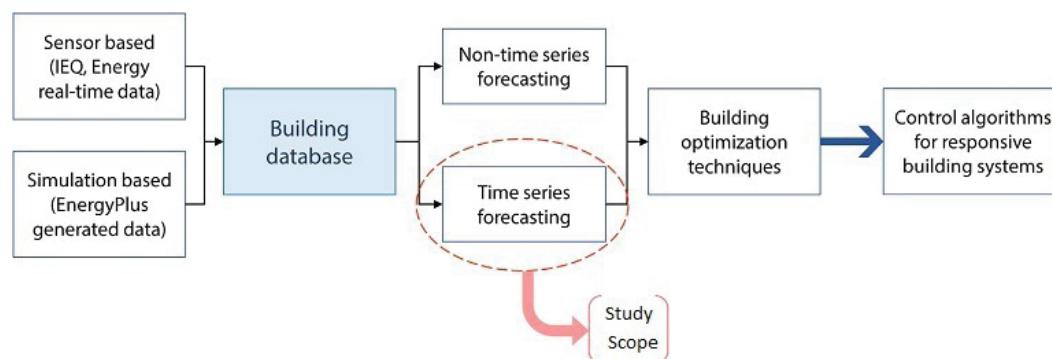


Fig. 2. Modified generalized block diagram [1]

Second, applying Markov Model based method to predict one day ahead under a shape-based measure. Finally, a new method is introduced to estimate used devices in household consumptions. Most forecasting techniques model a relationship between aggregated load demand and driver variables such as calendar effect, weather effect and lagged load demand (e.g. demand at previous hours or at the same hours of previous days). [3] Presents the definition of the clustering as “the data mining technique where similar data are placed into related or homogeneous groups without advanced knowledge of the groups’ definitions”.

Unlabeled data set of objects are distributed among groups according to the maximum similarity that has to be in the same group. Exploratory data analysis process is using clustering to form data into similar groups for summary generation as a preprocessing step. Time-series data analysis is used to do many tasks in many fields for different purposes like: subsequence matching, anomaly detection, motif discovery, indexing, clustering, classification, visualization, segmentation, identifying patterns, trend analysis, summarization, and forecasting. Time-series clustering is a special type of clustering which is working on continuous, real-valued items and is dynamic because values and observations are changed as a function of time and are considered as temporal data which is huge in size and highly dimensioned in structure. Time-series clustering is a challenging problem because of:

- data normally is larger than memory size and is stored in disks which leads to decrease in the speed of the clustering process.
- data are often high dimensional which makes handling these data difficult for many clustering algorithms.
- the similarity measures that are used to make the clusters depend on “whole sequence matching” where whole lengths of time-series are considered during distance calculation which is complicated, because time-series data are naturally noisy and include outliers and shifts which makes similarity measure select is a challenge in itself.

[4] used a big data approach to make load forecasting. The multiple linear regression model is introduced to find the optimal number of lagged hourly temperature and moving average temperature needed in a regression model that gives the best mean DTWE across 150 households in validation set. [5] presents Semi-parametric additive model to make day ahead (short term) and year ahead (middle term) load forecasting from data collected every ten minutes by ERDF at 2260 substations located at the frontier between the high voltage grid and the distribution network in France. In [6], Support Vector Machine (SVM) algorithm is used on load forecasting.

## B. MODELING AND METHODOLOGY

Time-series modeling scheme is better than Temperature that should not be considered in during the period in which the temperature does not vary much. In [7], Artificial Neural Network (ANN) hourly short-term electric load forecasting system is used as shown in I-A. It is known as ANNSTLF (Artificial Neural Network Short Term Load Forecaster) and is the most widely used ANN-based load forecaster in the USA and, possibly, the world. ANNSTLF can model the effect of two major weather variables on the load, temperature, and relative humidity.

The ANNSTLF package also includes an ANN hourly temperature forecaster and an hourly relative humidity forecaster. In [1], Time series electrical consumption forecasting models have strong contribution in optimization and planning fields in both buildings and compounds. Machine learning and statistical algorithms are used to predict and forecast future consumption based on historical readings and it is proven to be accurate and fast way than other methods. Historical data of energy consumption is analyzed often with various variables like weather, temperature, humidity, consumption season and other environmental conditions.

Combined methods of machine learning are much effective than using just one method. Reducing costs and carbon emissions in efficient buildings are commonly discussed topics related to energy, economic and environmental aspects. Energy consumption is affected by many factors like weather conditions, occupancy schedule, thermal properties of building materials, complex interactions of the energy systems like HVAC and lighting, etc. Complex relations between these factors are very difficult to simulate or using in simulation programs. Depending on data derived from consumption directly is a good scenario to study these effects of multi-dimensional problem. Algorithms used for that are depending on history readings and past patterns of electrical consumption. Time series algorithms are used as machine learning data-driven algorithms which depend on past customer electrical consumption. Electrical consumption forecasting is very important because it gives very clear boundaries for future consumption of a day, week, month and year and how this will affect economically.

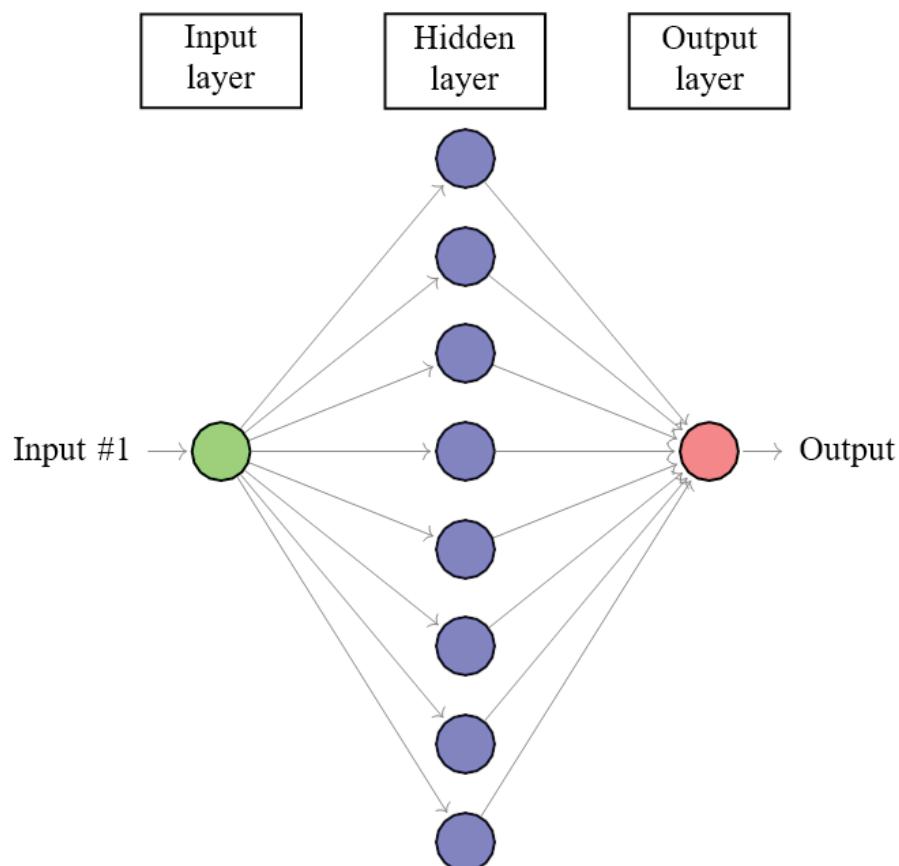


Fig. 3. Traditional shallow layer perceptron Artificial Neural Network Model with 8 neurons in hidden layer

In [8], time series forecasting model learned for customer behavior to give building managers a full view of usage patterns in spot and by comparing these consummations by recorded ones in the same conditions, a good view appears. Time series and non time series forecasting models could be used to predict occupancy and other operational factors. Forecasting and energy optimization are highly dependant. They are fully integrated because of the always need for information to calculate the best consummation configuration and future vision from electrical consumption predictors. Machine learning algorithms are used here widely to answer questions about optimal consumption behavior which takes in consideration patterns of consumption and predictions for future consumption too.

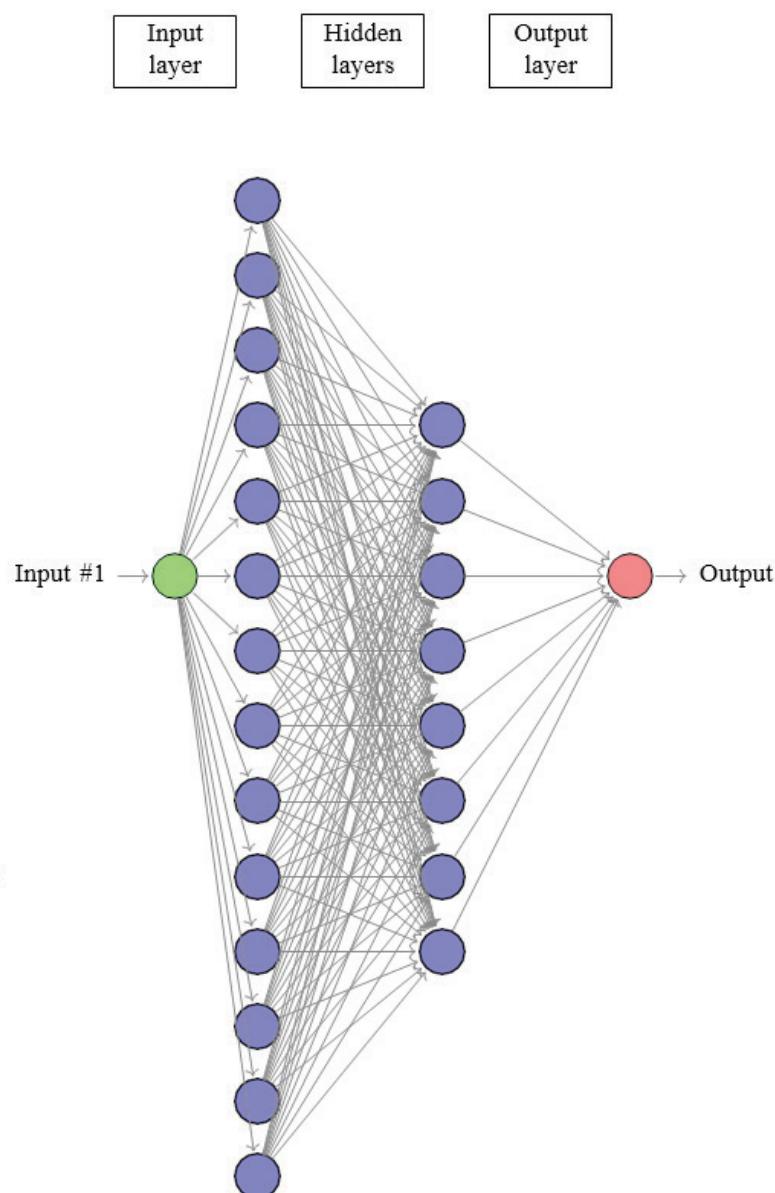


Fig. 4. Deep perceptron Artificial Neural Networks with two hidden layers with 14 neurons in first hidden layer and 8 neurons in second hidden layer

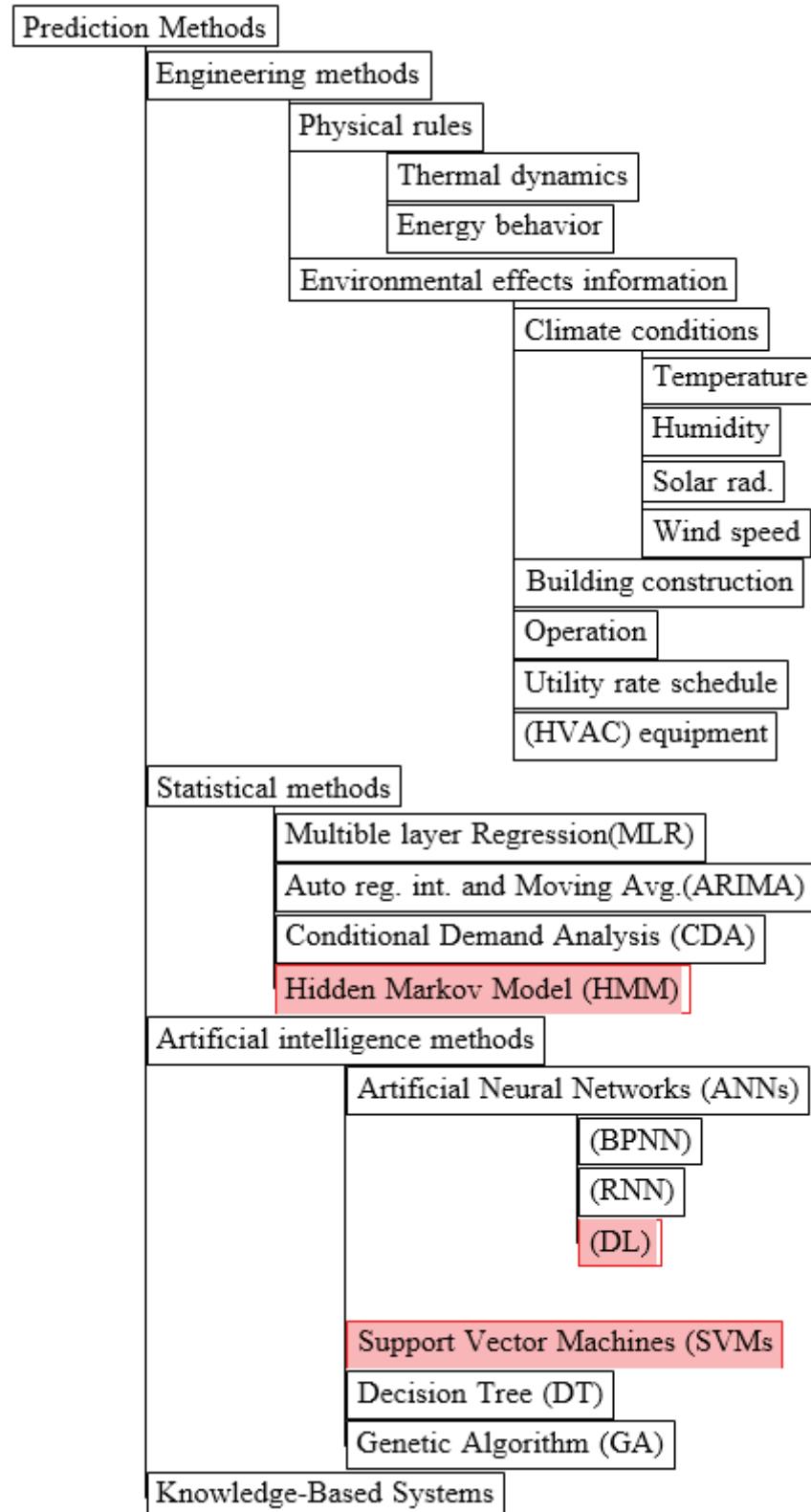


Fig. 5. Customer consumption prediction categories and methods with selected algorithms

In [9], There are many factors that highly affect electrical consumption like temperature, building construction and thermal properties of material which are used in building construction. Due to these complicated factors, prediction of household electrical consumption is very difficult.

As I-A shows, there are mainly three modeling categories or methods which are used to make customer prediction vision for building consumption. These categories are:

- **Engineering methods:** use physical rules such that thermal dynamics and energy behavior of the electrical consumption in buildings with environmental effects information,

such as:

- **External climate conditions like:**

- \* Temperature
- \*Humidity
- \* Solar radiation
- \* Wind speed
- Building construction
- Operation
- Utility rate schedule
- Heating, Ventilation and Air-Conditioning (HVAC) equipment.

- **Statistical methods like:**

- In [10], statistical regression algorithms such as:

\***Multible layer Regression (MLR)** is used and by predicting a day ahead, load forecasting is obtained in this model [11] [12]. Regression coefficients were found out with the help of method of least square estimation [13] [14]. Load in electrical power system is dependent on temperature, due point and seasons and also load has correlation to the previous load consumption (Historical data) [1] [15]. Then, the input variables are temperature, due point, load of prior day, hours, and load of prior week [16] [17]. To validate the model or check the accuracy of the model mean absolute percentage error is used. Using day ahead forecasted data weekly forecast is also obtained [11] [18]. Load forecasting mean forecasting average load in KW or total load in KWh for periods or blocks of 15 minutes, 30 minutes, 1 hour, day, week, month or a year for daily forecast, weekly forecast, monthly forecast or yearly [19] [20]. There are many factors which influence the accuracy of load forecasting like weather variables, holidays, festivals or events, tariff structures, available historical data, time of the year, day of the week and hour of the day. Weather variable includes temperature, humidity, rain and wind. Temperature and humidity has a considerable effect on power consumption because as the temperature rises people turn on air conditioners and if temperature is low air heaters will be turned on, which increases electricity demand. If there is a celebration, the electricity demand will rise due to lightnings [21] [10].

\***Auto regressive, Integrated and Moving Average (ARIMA)** is one of the most popular including time series analyses. ARIMA method and regression analysis, is one of traditional methods that based on mathematical calculations, but ARIMA is one of the most recent approaches including Kalman filtering, BoxeJenkins models, and state space models [1] [22] [20].

Hybrid model called SVRARIMA that sums up the Support Vector Regression (SVR) and ARIMA models are used to have better forecasting performance than individual model [23] [24] [25]. All of these methods can achieve electric load forecasting but cannot receive the desired prediction accuracy because of their limitations. For example, linear regression depends on historical data and cannot solve non-linear problems. Autoregressive moving average models give the result taking only into account the past and current data points while ignoring other influential elements. The grey forecasting model can only effectively solve the problem with exponential growth trends [26] [27]. Literature shows that time series analysis techniques are neither scalable to higher dimension nor are effective in highly volatile data [28] [29] [11]. For this reason time series methods such as regression models, ARIMA models, GARCH and hybrid models such as combination of ARIMA and GARCH using wavelet transform are not considered for short term forecasting [30] [31].

ARIMA processes are well suited to express the stochastic nature of the load time series [1] [14]. Modeling of multiple seasonal cycles as well as introducing exogenous variables is not a problem in ARIMA [6] [32]. The disadvantage of ARIMA models is that they are able to represent only linear relationships between variables [33] [34]. The difficulty in using ARIMA is the problem of order selection which is considered to be subjective [35] [36]. To simplify the forecasting problem, the time series is often decomposed into a trend, seasonal components and an irregular component [37] [38]. These components, showing less complexity than the original series, are modeled independently [3] [39]. The ARIMA parameters were estimated for each forecasting task (i.e. the forecast of system load at time  $t$  of the day  $d$ ) using time series fragments immediately preceding the forecasted day [40] [41]. Typical days in these fragments were replaced with the days from the previous weeks [42] [43].

Due to using short time series fragments for parameter estimation (much shorter than the annual period) and due to time series decomposition into  $n$  series [1] [44], it is not required to take into account the annual and daily seasonalities in the models [10] [45] [16]. In such case, the number of parameters is much smaller and they are easier to be estimated compared to models with triple seasonality [46] [47]. This eliminates the daily seasonality and simplifies the forecasting problem [48] [49]. To estimate parameters of ARIMA the stepwise procedures for traversing the model spaces implemented in the forecast environment for statistical computing [50] [15]. The conventional forecasting model like ARIMA works significantly worse than the best Neural Networks models [51] [52], ARIMA and ES are optimized on the time series fragments directly preceding the forecasted fragment [53] [13]. The classical statistical model like ARIMA is the simplest among tested methods. It has only one parameter to estimate. Such a model is easy to optimize and has good generalization properties. Its learning and optimization procedures are extremely fast [54] [12] [55].

**\*Conditional Demand Analysis (CDA)** is a method to model the residential end-use energy consumption at the national level [56]. There are several studies where CDA was used to model energy consumption at the regional level; however the CDA method had not been used to model residential energy consumption at the national level [9] [10]. The prediction performance and the ability to characterize

the residential end-use energy consumption of the CDA model are compared with those of a neural network (NN) and an engineering based model developed earlier. The comparison of the predictions of the models indicates that CDA is capable of accurately predicting the energy consumption in the residential sector as well as the other two models. The effects of socio-economic factors are estimated using the NN and the CDA models, where possible. Due to the limited number of variables the CDA model can accommodate, its capability to evaluate these effects is found to be lower than the NN model. [57].

**\*Hidden Markov Models (HMM)** are useful in time series data analysis, however, its application in building energy sector is not so much investigated [2]. HMM-based procedure is used for identification of individual household appliances from collective energy consumption data [3]. However, this requires HMM models to be trained for individual appliance energy consumption profile [1]. An HMM-based modeling scheme is presented in for energy level prediction in wireless sensor networks nodes. They consider node energy level as a stochastic variable having values within certain fixed range forming hidden state for their HMM model whereas nodes energy consumption is treated as observed state [18] [58] [59].

- **Artificial intelligence methods like:**

- Artificial Neural Networks (ANNs) algorithms such as:

**\*Back Propagation Neural Network (BPNN)** The most common type of ANN used in classification of remote sensing imagery is the MLP (multi-layer perceptron) networks based on the BP (back-propagation) learning algorithm [1]. This type of network is called BPNN (back propagation neural network) [13] [1] [22]. Studies on the performance of back propagation networks are still ongoing. The back propagation learning algorithm is the best algorithm among the Multi-layer perceptron algorithms [34] [56] [15]. The processing speed and classification accuracy often depend on the design and implementation of the network. In the literature, different researchers introduced different network architectures for different applications. Are there any common standards among neural network researchers to guide the design and implementation of BPNNs for remote sensing image classification? This question is still open [60] [61] [10]. Choosing the number of layers and the number of neurons in each layer is essential. The performance of the feed-forward BPNNs is affected significantly by the network layout [29] [27] [24]. The selection of this layout and its related parameters is referred to as the art of network design. Unfortunately, there is no clear answer for the layout of the neural network for a particular application; however, there are some general rules which have been collected over time and followed by researchers [20] [21] [31].

**Recurrent Neural Network (RNN) and Multilayer Perceptron's (MLPs)** algorithms are very efficient for time series forecasting processes. Figure I-A shows the Neural Network Model used in MLP algorithm used in this study. There are many advantages of using such algorithm such as:

- Robust to Noise
- In input data and in the mapping function. Support learning and prediction in the presence of missing values [21].
- Nonlinear

- Do not make strong assumptions about the mapping function.
- Readily learn linear and nonlinear relationships [10] [62].
- Multivariate Inputs
- Input features can be specified.
- Providing direct support for multivariate forecasting [33] [63].
- Multi-step Forecasts
- Output values can be specified.
- Providing multi-step and multivariate forecasting [12] [51].

### C. RESULTS AND ACCURACY

In this study, the used model depends on Feed Forward Neural Networks for Time Series Forecasting. Firstly, transforming Data for Time Series calculation is done. Then sliding window transformation technique is applied too. LSTM algorithm is applied here in time series forecasting as follows:

- \*It is defined as the number of input time steps as three via the input dim argument on the first hidden layer.
- \*Efficient Adam version of stochastic gradient descent and optimizes the mean squared error ('MSE') loss function is used as shown in figures (5) and (6).
- \* Fitted model can be used to make a prediction

Implementing these steps using Python, NumPy and Keras for deep learning is strait forward and the results are clear in figures: I-C for Results of Naive Time Series Predictions with Neural Network and I-C for Results of Window Method for Time Series Predictions with Neural Networks respectively. From these different runs, it becomes clear that frame of the prediction problem is done Error was not significantly reduced in window method compared to that of the previous section with Naïve one. The window size and the network architecture were not tuned as more runs and modifications should be used to reach tuned output of prediction system. In this work, it is utilized to use window size of ten successive days in a row. After applying Discrete Cosine Transform DCT to the results of the prediction, Validation Error reduced to be 0.9813% for all runs. This step is mandatory after using the LSTM to reduce validation error. LSTM model obtained a higher performance than the other applied models as shown in Table 1 and due to the small sample of different target variables, it is difficult to generalize if DL models may help when applied at such levels. The experiments demonstrated that the LSTM model had the best performance under each statistical indicator.

– **Support Vector Machines (SVMs)** are one of the most probability used algorithms for datasets with high number of dimensions with small and nonlinear samples. on the other hand, these algorithms are not applicable for big datasets and they have a little big time complexity of  $O(N^3)$  [57] [64] [11]. However, combining SVM with other AI models may improve forecasting accuracy [33] [65]. There is still a long way to go in order to make the methods applicable, and future work should focus on reducing the computational costs and memory while maintaining accuracy before on-line practical applications [12] [54]. SVM is a very specific technique characterized by usage of kernels, absence of local minima, sparseness of the solution and capacity control obtained by acting on the margin, or on number of support vectors [66]. The capacity of the system is controlled by parameters that do

not depend on the dimensionality of the feature space. The non-linear function is learned by linear learning machine which maps inputs into high dimensional kernel induced feature space [13] [67]. SVM is motivated to find and optimize the generalization bounds given for regression [22] [34]. They relied on defining the so called epsilon intensive loss function that ignores errors, which are situated within the certain distance of the true value [64] [14]. Using two stages forecast engine incorporating linear regression; dynamic programming; and support vector machine (SVM) [63] [2]. Fixed size least squares support vector machines (LS-SVM) using an autoregressive exogenous nonlinear autoregressive

```
# create and fit Multilayer Perceptron model
model = Sequential()
model.add(Dense(8, input_dim=look_back, activation='relu'))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(trainX, trainY, epochs=200, batch_size=2, verbose=2)
    Train Score: 49280.91 MSE (221.99 RMSE)
    Test Score: 13341.68 MSE (115.51 RMSE)

: plt.plot(dataset)
plt.plot(trainPredictPlot)
plt.plot(testPredictPlot)
plt.show()
```

Train Error: 4.19%  
 Test Error: 4.42%

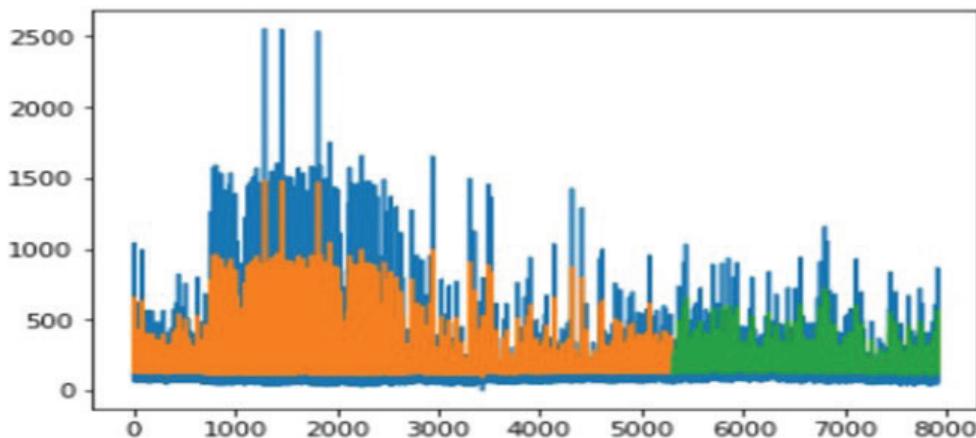


Fig. 6. Results of Naive Time Series Predictions with Neural Network

```

# create and fit Multilayer Perceptron model
model = Sequential()
model.add(Dense(14, input_dim=look_back, activation='relu'))
model.add(Dense(8, activation='relu'))
model.add(Dense(1))
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(trainX, trainY, epochs=400, batch_size=2, verbose=2)
.... plt.plot(dataset)
.... plt.plot(trainPredictPlot) Train Error: 4.077%
.... plt.plot(testPredictPlot) Test Error: 4.888%
.... plt.show()
Train Score: 46773.37 MSE (216.27 RMSE)
Test Score: 16315.15 MSE (127.73 RMSE)

```

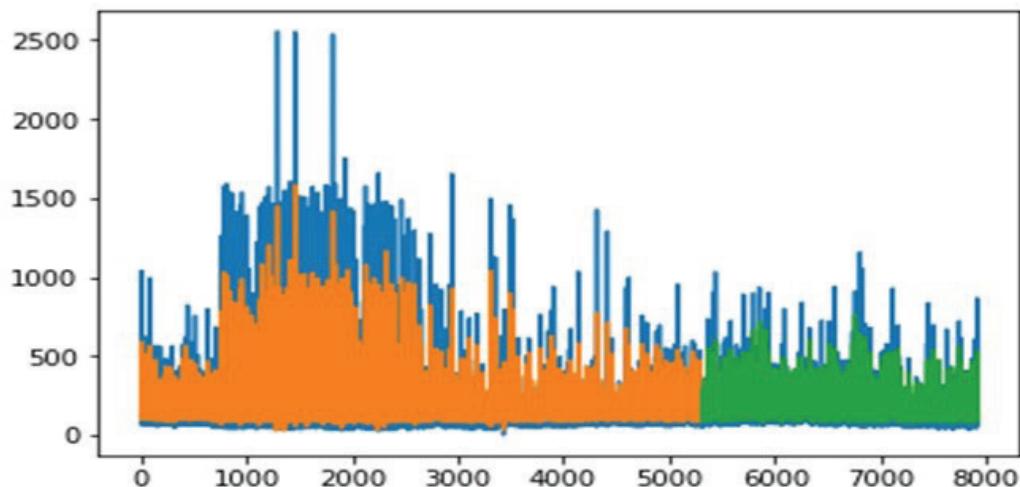


Fig. 7. Results of Window Method for Time Series Predictions with Neural Networks

exogenous structure, could also be implemented to outperform the linear model [68] [1] [56] [15]. Poor computational scalability is one of the disadvantages when using SVM's as like as using other methods such as neural networks and other computational intelligence techniques [61] [59] [16] [45]. The state of the art experimental results showed that SVM's methods are considered to be strong and have a non-linear learning capabilities for electric load forecasting which are combined with the empirical mode decomposition method and auto regression one [10] [39] [36]. The neural network easily falls into the local minimum because of the restriction on generalization ability and cannot make full use of information from selecting sample with a small sample size [28] [69].

Compared with a traditional neural network, the support vector machine (SVM) can overcome these drawbacks to improve forecasting performance. As a kernel based method, SVM employs the learning principle with structural risk minimization (SRM) to increase its generalization capability in the training process, generating better forecasts [37] [70] [60]. Because of the attractive feature and empirical performance, SVM has become one of the most promising and popular forecasting methods [11] [19] [40] [47]. Therefore, SVM is used as the forecasting method and the parameters of SVM have an important influence on the accuracy of prediction [46] [43] [71]. An alternative to solve the problem is by using heuristic optimization algorithms for parameter selection that they are

prone to be more efficient and robust than a traditional optimization algorithm, e.g., grid search algorithm [9] [50] [49] [72]. Therefore, the Cuckoo search (CS) algorithm, a heuristic optimization algorithm that has powerful ability to search for an optimal solution, is used to determine parameters of SVM [36] [11] [25] [73] [58]. In addition, Singular Spectrum Analysis (SSA), a powerful technique in time series analysis that was used for electric load forecasting, is employed to remove the high frequency components of the noisy load series in order to improve the forecasting performance of the SVM model, producing the CS- SSA-SVM model [25] [73] [62]. The combination of the data preprocessing-based technique, kernel-based method and heuristic optimization algorithm. The performance of the hybrid model is validated by forecasting the short-term electric load [24] [23].

Integrating merits of individual algorithms to enhance prediction accuracy and based on the signal reprocessing technique, the hybrid methods can robustly map input space into feature space to tackle the complex nonlinear problems [58] [21] [74]. The powerful signal processing technique is applied to decomposition and reconstruction of the original electric load data; the analysis process from embedding to diagonal averaging performs identification and extraction of different characteristics to alleviate negative effects from the noisy signal [31] [20] [53]. The powerful global search capacity of the CS algorithm is employed to serve for optimal intelligent selection of model parameters, overcoming limitations of artificially selected parameters [35] [29] [75]. Apply the hybrid model to the constructed electric load series to show its superiority compared with other benchmark models [44] [41] [6].

- **Decision Tree (DT)** have been used literally to build energy prediction system and to provide reliable promising prediction results during the past two decades [10] [72]. However, one of the major disadvantages of single prediction approach is the instability issue within each learning algorithm [10] [19]. Learning algorithms such as ANN and decision tree are unstable learners which may introduce significant variation in the output value due to some small changes made in the input data [64] [11]. This instability issue could impede these algorithms from implementation in real-time, on-the-field applications as some energy efficiency measurements rely on the reliability of the prediction [12] [76], for example, an unstable learner may lead to high false alarm ratio for building system fault detection. To overcome the limitation of instability as well as to improve the prediction accuracy, the concept of ensemble learning has been recently introduced by researchers to solve both classification and regression problems [13] [67].

To improve prediction accuracy and by using data obtained from meteorological systems and building-level occupancy and meters, the prediction performance of conventional decision tree method is improved i.e., Classification and Regression Tree (CART) by introducing bagging technique [64] [14] [34]. However, it should be noted that the feature importance results may not accurately interpret the relationship between each input feature and the output variable [77] [66]. Some of the input variables used generally, are highly correlated with each other, for example, the outdoor temperature and solar radiation variable. The decision tree may only use one of them for tree growing and put away the rest because correlated variables share the same impurity [56] [15] [60]. the feature importance of the used

variable will be high while those of the unused variables will decrease significantly. For example, the feature importance of outdoor temperature is 2.0 in module [47] [42] [78], while its correlated variable solar radiation only has feature importance of 0.5 [48] [16]. The low value of feature importance should not be taken as evidence that the variable is not strongly related to the output variable [45] [23] [21].

- **Genetic Algorithm (GA)** it is used to combine the prediction of each base model and output the final results as the prediction of the ensemble model [10] [11] [12]. Literature preliminary results showed that the proposed ensemble model provided higher prediction accuracy than the typical single model [54] [53] [14]. Recently, gene expression programming (GEP) was proposed as a new function model mining algorithm. Compared with traditional genetic algorithms (GA) and Genetic Programming (GP), GEP has advantages in terms of convergence speed and ability to solve complex problems [1] [22]. At present, research on GEP focused on symbolic regression, function finding, combinatorial optimization and prediction. In symbolic regression and function mining, an improved GEP algorithm named SGEP is proposed, which is especially suitable for dealing with symbolic regression problems [34] [79]. On the other hand, hybrid models for load forecasting algorithm included combination of genetic algorithm and ant colony optimization for feature selection and multi-layer perceptron for hourly load prediction are commonly used [79] [7].
- **Knowledge-Based Systems** Knowledge-based filtering is emerging as an important field which uses knowledge about users and products to pursue a knowledge-based approach to generating recommendations, reasoning about what products meet the user's requirements [11] [67] [14]. Gradual incorporation of different types of information (e.g., explicit ratings, social relations, user contents, locations, use trends, knowledge-based information) has forced forecasting system to use hybrid approaches. Once the memory-based, social and location-aware methods and algorithms are consolidated, the evolution of system demonstrates a clear trend toward combining existing collaborative methods [1] [34] [77]. The major reason behind using commercialized knowledge-based systems (KBS), is the lack of a strict validation step in their life cycle. There is a widespread agreement that KBS cannot be designed in a linear fashion. This is due to the typical problems they have to resolve; these will require them either to adapt traditional techniques of software development or to use new techniques relevant to artificial intelligence (AI) systems [41] [47] [46].

A significant problem in the development of Knowledge-Based Systems (KBS) is its verification step. By using Machine Learning techniques to progressively improve the quality of expert system Knowledge Bases and by coping with two major KB anomalies: incompleteness and incorrectness [9] [72] [78]. In agreement with the current tendency, KBs considered in our approach are expressed in different formalisms. Results obtained with two different learning algorithms, confirm the hypothesis that integrating machine learning techniques in the verification step of a Knowledge-Based System life cycle, is a promising approach [48] [48] [45]. Knowledge based systems (KBS) or expert systems emulate the human expert behavior in a certain knowledge area [10] [3] [32]. They constitute aid systems to take decisions in different areas such as educational strategic selection,

environmental variables control [74], neonatology fans configuration, agreement in judicial process or the attended generation of activity maps of software development projects. Knowledge based systems to aid decision taking is a particular knowledge based system [6] [31] [19].

Although, hybrid approaches are used by combining multiple methods to enhance prediction performance. In [10], for classification in data-driven models, algorithms that are usually considered are:

- **K-means Clustering Techniques**

- **Fuzzy K-means Clustering**

To improve the accuracy of short-term electric load forecasting for individual users, a short-term power load forecasting model based on K-means [3] [69]. By analyzing the users' electricity consumption features, K-means is applied to group users into two clusters. [2] [10].

- **Fuzzy C-means Clustering**

For users with strong correlation at adjacent moments, local similar data are filtered out with the help of improved Fuzzy C-Mean clustering (FCM), integrating the load value of the adjacent moments into new input features [80] [55]. For users with weak correlation at adjacent moments, the local similar daily data are utilized as features [60] [40]. Finally, the feather vectors are used as input data for BP Neural Network, which is utilized to forecast the short-term load [34] [15].

- **Self-organizing map (SOM)** Among different neural network classifiers, Self Organizing Map (SOM) is one of the most effective methods. SOM has two valuable features, namely, pattern recognition and pattern complementarity [1] [22] [2]. These two SOM properties are exploited for short-term load classification and for forecasting, respectively [70] [41]. SOM is using an iterative active learning technique based on self-organizing map (SOM) neural network and support vector machine (SVM) classifier [47] [50] [3]. The technique exploits the properties of SVM classifier and that of SOM neural network to identify uncertain and diverse samples to be included in the training set [6] [36] [29]. It selects uncertain samples from low density regions of the feature space by exploiting the topological properties of SOM [27] [69] [62].

In addition to the extensive use of classification methods for different applications in electrical engineering, classification methods particularly have been used for load management and load forecasting [21] [81] [67]. Doing forecasting for the short-term (the next day) load, while used system focuses only on the method of classifying daily electric load and the classification of loads with the same behavior for the time period selected (e.g. one week, one month, or one year, or any desired time interval from one day to several years) [10]. Although none of the outputs of the SOM classifier is directly used for the prediction purpose, the results of the SOM network classification process can be used as an effective way for sampling the appropriate training patterns for training and forecasting 24-hour active power consumption of the nationwide electricity. This can be used in a separate SOM network. According to the outcomes of the classifier, every Normal day of a week has its specific load consumption curve and holidays have distinguished ones [82][83].

- **Hierarchical clustering** Hierarchical time series prediction plays an important role in various applications, such as retail in business, electricity supply and environmental protection [11] [67]. In the electricity power supply and management, the power supply and consumption is usually organized in a hierarchical structure according to administrative divisions [2] [56] [60]. In big data applications, hierarchical time series prediction is an important element of decision-making and concerns the inherent aggregation consistency, which is maintained by reconciliation methods [47] [71]. By using hierarchical forecasting approach, the k means clustering [55] based multiple alternative clustering strategy is employed to cluster the time series with different cluster number k to obtain a large number of time series clusters [10] [3] [11]. It brings more chances to build good aggregate hierarchies as the input of our hierarchical prediction model. Furthermore, instead of dealing with the clusters constraints and the geographical constraints at two separate steps, these different aggregation constraints are integrated as a whole for optimal prediction reconciliation [18] [75]. Compared with the state-of-the-art methods, the one-step ahead forecasts of the proposed method for electricity load and solar power data are improved than ordinary multiple stages hierarchy [69] [81].

Big amount of data is generated from energy power consummations meters. Many algorithms are developed with data- driven approach area and target several types of energy applied applications like:

- Load forecasting.
- Prediction.
- Energy pattern profiling.
- Regional energy-consumption mapping.
- Bench-marking for building stocks.
- Global retrofit strategies.
- Guideline making.
- Etc.

Classification methods are developed too in same area of interest like:

- K-mean clustering.
- Self-organizing map.
- Hierarchy clustering.

Both of them (Prediction and Classification) are very important for achieving efficient energy consumption building and enhancing performance. Reducing consumption and environmental impact is also a clear target for applying these strategies. Building prediction and classification models for energy consumption with a very high accurate result requires

very huge efforts. Instead of that, it is useful to build these models by a moderate accuracy. These models have many advantages to be built:

- 1) Conserve energy based on predicted customer behaviors.
- 2) Build Demand-Side management (DSM) system according to electrical consummation scenarios.
- 3) Visualizing predicted and actual energy consumption curves.
- 4) Implementing a benchmark databases for multilevel consumption.

- 5) Build a combined model facilitated with designing, running and editing functions in the new buildings.
- 6) offering an image of energy footprint of the building under study.
- 7) introducing a good imagination for related financial aspects.
- 8) giving an important tool for decision makers such as:
  - Policymakers.
  - Building owners.
  - Investors.
  - Operators.
  - Engineers.

Dataset of historical energy consumption for a group of customers is used in learning stage to train the algorithm. Pattern of consumption is achieved for each customer to predict next days consumption and to group similar customers in consumption behavior together. Household consumption could be caused from many end-uses devices like heating, ventilation and air-conditioning (HVAC) system, domestic hot water, lighting, plug-loads, elevators, kitchen equipment, ancillary equipment and appliances.

In [11], Load Forecasting is essential in business perspective because, for example, it is important for:

- Power systems planning and operations
- Revenue projection
- Rate design
- Energy trading
- Design evaluation [15]
- Operation strategies [15]
- Enhancing demand and supply management [15]

These fields are always discussed in many sectors of various companies like:

- Electric utilities
- Regulatory commissions
- Industrial and big commercial companies
- Banks
- Trading firms
- Insurance companies

In [15], reducing CO<sub>2</sub> emissions, global warming, environmental pollution and energy consumption which are mainly generated from fossil fuel are very hot topics nowadays to save environment and resources. A big percentage of total energy consumption and total CO<sub>2</sub> emissions is from buildings. Energy consumption prediction is very important to achieve the international standards for CO<sub>2</sub> emissions levels.

## D. FIELD TRENDS AND WORK COMPARISONS

By reviewing the work in the field for the last twenty years, it could be concluded that there are two main types of household customer electrical consumption prediction.

- Resolution: which is the time step of the electrical data.

- Forecast Horizon: which is the window size of the predicted future (day, week, and etc.) Resolution and Forecast Horizon, for example could be hourly time step data predicting a horizon of 24 hours in advance. Forecast horizon can be classified as long term (greater than three years), medium (two weeks to three years), and short-term (less than two weeks) shown in [84] and illustrated in Figure 8.

Table 1: shows the Validation Error obtained by the LSTM with DCT, compared other benchmark methods when predicting the test set. It can be seen that the proposed LSTM using the Discrete Cosine Transform DCT significantly improves the validation error obtained by the other prediction models.

Table 1: Validation Error Obtained by The Proposed LSTM Compared with Other Methods [85]

Validation Error %	
<b>LR</b>	<b>7.3395</b>
<b>DT</b>	<b>2.8783</b>
<b>GBT</b>	<b>2.7190</b>
<b>RF</b>	<b>2.2005</b>
<b>DFFN</b>	<b>1.6769</b>
<b>LSTM + CVOA</b>	<b>1.5898</b>
<b>TFT</b>	<b>1.5148</b>
<b>LSTM + Random</b>	<b>1.4472</b>
<b>LSTM + DCT</b>	<b>0.9813</b>

## II. CONCLUSION

This paper presented five methods to forecast the energy consumption in a residential building over different time horizons with different time resolutions. Notably, it proposed the use of Deep Learning, more exactly Conditional Restricted Boltzmann Machines and Factored Conditional Restricted Boltzmann Machines for the prediction of energy consumption. The analysis performed showed that FCRBM is a powerful method which outperformed the state-of-the- art prediction methods such as HMM, ANNs, SVMs, RNNs

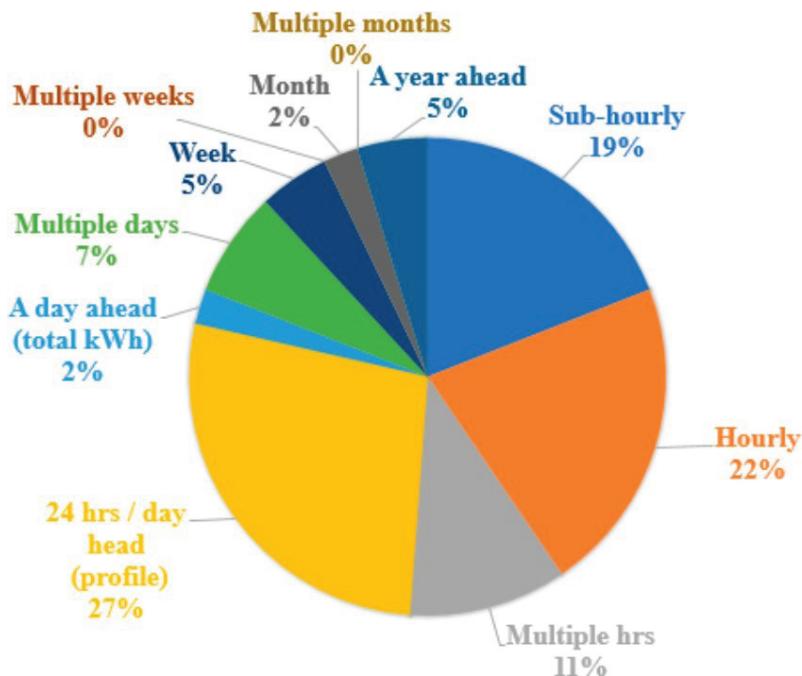


Fig. 8. Compositional breakdown of the forecast horizons [84]

and CRBMs. It is worth mentioning that as the prediction horizon is increasing, FCRBMs and CRBMs seem to be more robust and their prediction error is typically half then that of the ANN. All methods presented showed comparable prediction time, in the order of few hundred milliseconds, and are therefore suitable for near real-time exploitation in applications such as home and building automation systems. From all the experiments, it can be observed that all methods perform better when predicting the aggregated active power consumption than predicting the demand of intermittent appliances(e.g.electric water-heater) recorded with the three sub-metering. Although versatile and successful, CRBMs and FCRBMs come with their own challenges, similar to other ANNs. At this stage, the feasibility of the various methods has been proven. Furthermore, fine-tuning, such as the choice of the optimal number of hidden units or the learn in grage, might improve the performance of these models.

## REFERENCES

- [1] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902-924, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032117303155>
- [2] T. Teeraratkul, D. O'Neill, and S. Lall, "Shape-Based Approach to Household Electric Load Curve Clustering and Prediction," *IEEE Transactions on Smart Grid*, p. 1, 2017.
- [3] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah, "Time-series clustering – A decade review," *Information Systems*, vol. 53, pp. 16–38, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306437915000733>
- [4] P. Wang, B. Liu, and T. Hong, "Electric load forecasting with recency effect: A big data approach," *International Journal of Forecasting*, vol. 32, no. 3, pp. 585–597, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207015001557>

[5] Y. Goude, R. Nedellec, and N. Kong, "Local Short and Middle Term Electricity Load Forecasting With Semi-Parametric Additive Models," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 440–446, 2014.

[6] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1821–1830, 2004.

[7] A. Khotanzad, R. Afkhami-Rohani, T.-L. Lu, A. Abaye, M. Davis, and D. J. Maratukulam, "ANNSTLF-a neural-network-based electric load forecasting system," *IEEE Transactions on Neural Networks*, vol. 8, no. 4, pp. 835–846, 1997.

[8] J. Yang, M. Santamouris, S. E. Lee, and C. Deb, "Energy performance model development and occupancy number identification of institutional buildings," *Energy and Buildings*, vol. 123, pp. 192 – 204, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778815304552>

[9] H. xiang Zhao and F. Magoules, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586 – 3592, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032112001438>

[10] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, and X. Zhao, "A review of data-driven approaches for prediction and classification of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 1027–1047, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S136403211731362X>

[11] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207015001508>

[12] X. Qiu, P. N. Suganthan, and G. A. J. Amaralunga, "Ensemble incremental learning Random Vector Functional Link network for short-term electric load forecasting," *Knowledge-Based Systems*, vol. 145, pp. 182–196, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950705118300236>

[13] J.-S. Chou and D.-S. Tran, "Forecasting Energy Consumption Time Series using Machine Learning Techniques based on Usage Patterns of Residential Householders," *Energy*, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544218319145>

[14] L. Wang, R. Kubichek, and X. Zhou, "Adaptive learning based data-driven models for predicting hourly building energy use," *Energy and Buildings*, vol. 159, pp. 454–461, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778817318728>

[15] K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032117306093>

[16] P. Lusis, K. R. Khalilpour, L. Andrew, and A. Liebman, "Shortterm residential load forecasting: Impact of calendar effects and forecast granularity," *Applied Energy*, vol. 205, pp. 654–669, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917309881>

[17] J. Luo, T. Hong, and S.-C. Fang, "Benchmarking robustness of load forecasting models under data integrity attacks," *International Journal of Forecasting*, vol. 34, no. 1, pp. 89–104, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169207017300900>

[18] Q. Duan, J. Liu, and D. Zhao, "Short term electric load forecasting using an automated system of model choice," *International Journal of Electrical Power & Energy Systems*, vol. 91, pp. 92–100, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0142061516300710>

[19] Z. Wang, Y. Wang, and R. S. Srinivasan, "A novel ensemble learning approach to support building energy use prediction," *Energy and Buildings*, vol. 159, pp. 109–

122, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778817335272>

[20] M. A. M. Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1108–1118, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032116310619>

[21] Z. Wang and R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 796–808, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032116307420>

[22] J. Zhang, Y.-M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, vol. 158, pp. 774–781, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S036054421831065X>

[23] P. Vrablecová, A. B. Ezzeddine, V. Rozinajová, S. Sárik, and A. K. Sangaiah, "Smart grid load forecasting using online support vector regression," *Computers & Electrical Engineering*, vol. 65, pp. 102–117, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0045790617320645>

[24] Y. Liu, W. Wang, and N. Ghadimi, "Electricity load forecasting by an improved forecast engine for building level consumers," *Energy*, vol. 139, pp. 18–30, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544217313348>

[25] Y. T. Chae, R. Horesh, Y. Hwang, and Y. M. Lee, "Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings," *Energy and Buildings*, vol. 111, pp. 184–194, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778815304102>

[26] J. Nowotarski, B. Liu, R. Weron, and T. Hong, "Improving short term load forecast accuracy via combining sister forecasts," *Energy*, vol. 98, pp. 40–49, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544216000116>

[27] G.-F. Fan, L.-L. Peng, and W.-C. Hong, "Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model," *Applied Energy*, vol. 224, pp. 13–33, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261918306238>

[28] E. Yukseltan, A. Yucekaya, and A. H. Bilge, "Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation," *Applied Energy*, vol. 193, pp. 287–296, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917301848>

[29] G.-F. Fan, L.-L. Peng, W.-C. Hong, and F. Sun, "Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression," *Neurocomputing*, vol. 173, pp. 958–970, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231215012230>

[30] Y. Chen, H. Tan, and X. Song, "Day-ahead Forecasting of Non-stationary Electric Power Demand in Commercial Buildings: Hybrid Support Vector Regression Based," *Energy Procedia*, vol. 105, pp. 2101–2106, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1876610217306422>

[31] X. Zhang, J. Wang, and K. Zhang, "Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm," *Electric Power Systems Research*, vol. 146, pp. 270–285, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S037879617300445>

[32] S. Fan and R. J. Hyndman, "Short-Term Load Forecasting Based on a Semi-Parametric Additive Model," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, 2012.

[33] N. Ghadimi, A. Akbarimajd, H. Shayeghi, and O. Abedinia, "Two stage forecast

engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting," *Energy*, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544218313859>

[34] A. I. Saleh, A. H. Rabie, and K. M. Abo-Al-Ez, "A data mining based load forecasting strategy for smart electrical grids," *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 422–448, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1474034616301331>

[35] F. Zhang, C. Deb, S. E. Lee, J. Yang, and K. W. Shah, "Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique," *Energy and Buildings*, vol. 126, pp. 94–103, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778816303899>

[36] A. Lahouar and J. B. H. Slama, "Day-ahead load forecast using random forest and expert input selection," *Energy Conversion and Management*, vol. 103, pp. 1040–1051, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0196890415006925>

[37] Y. Chen, P. Xu, Y. Chu, W. Li, Y. Wu, L. Ni, Y. Bao, and K. Wang, "Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings," *Applied Energy*, vol. 195, pp. 659–670, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917302581>

[38] S. Mishra and V. K. Singh, "Monthly Energy Consumption Forecasting Based On Windowed Momentum Neural Network," *IFAC- PapersOnLine*, vol. 48, no. 30, pp. 433–438, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2405896315030591>

[39] K. Gajowniczek and T. Zabkowski, "Short Term Electricity Forecasting Using Individual Smart Meter Data," *Procedia Computer Science*, vol. 35, pp. 589–597, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050914011053>

[40] M. Mordjaoui, S. Haddad, A. Medoued, and A. Laouafi, "Electric load forecasting by using dynamic neural network," *International Journal of Hydrogen Energy*, vol. 42, no. 28, pp. 17 655–17 663, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360319917310455>

[41] W.-C. Hong, "Electric load forecasting by support vector model," *Applied Mathematical Modelling*, vol. 33, no. 5, pp. 2444–2454, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0307904X08001844>

[42] M. Rana and I. Koprinska, "Forecasting electricity load with advanced wavelet neural networks," *Neurocomputing*, vol. 182, pp. 118–132, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231215019141>

[43] J.-S. Chou and N.-T. Ngo, "Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns," *Applied Energy*, vol. 177, pp. 751–770, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261916306717>

[44] C. Kuster, Y. Rezgui, and M. Mourshed, "Electrical load forecasting models: A critical systematic review," *Sustainable Cities and Society*, vol. 35, pp. 257–270, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2210670717305899>

[45] S. Aman, Y. Simmhan, and V. K. Prasanna, "Holistic Measures for Evaluating Prediction Models in Smart Grids," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 2, pp. 475–488, 2015.

[46] J.-S. Chou and N.-T. Ngo, "Smart grid data analytics framework for increasing energy savings in residential buildings," *Automation in Construction*, vol. 72, pp. 247–257, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0926580516000030>

[47] M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load

demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352-1372, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032115003354>

[48] S. Ahmad, A. Lavin, S. Purdy, and Z. Agha, "Unsupervised real-time anomaly detection for streaming data," *Neurocomputing*, vol. 262, pp. 134-147, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231217309864>

[49] C. Lynch, M. J.O'Mahony, and R. A. Guinee, "Electrical Load Forecasting Using An Expanded Kalman Filter Bank Methodology," *IFAC- PapersOnLine*, vol. 49, no. 25, pp. 358-365, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2405896316327070>

[50] Z. Guo, K. Zhou, X. Zhang, and S. Yang, "A deep learning model for short-term power load and probability density forecasting," *Energy*, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544218313872>

[51] W. He, "Load Forecasting via Deep Neural Networks," *Procedia Computer Science*, vol. 122, pp. 308-314, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050917326170>

[52] N. Ding, Y. Besanger, and F. Wurtz, "Next-day MV/LV substation load forecaster using time series method," *Electric Power Systems Research*, vol. 119, pp. 345-354, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378779614003575>

[53] A. S. Ahmad, M. Y. Hassan, M. P. Abdullah, H. A. Rahman, F. Hussin, H. Abdullah, and R. Saidur, "A review on applications of ANN and SVM for building electrical energy consumption forecasting," *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 102-109, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032114000914>

[54] H. Jiang, Y. Zhang, E. Muljadi, J. J. Zhang, and D. W. Gao, "A Short-Term and High-Resolution Distribution System Load Forecasting Approach Using Support Vector Regression With Hybrid Parameters Optimization," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3341-3350, 2018.

[55] X. Fu, X.-J. Zeng, P. Feng, and X. Cai, "Clustering-based short-term load forecasting for residential electricity under the increasing-block pricing tariffs in China," *Energy*, vol. 165, pp. 76-89, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544218319285>

[56] W. Ma, S. Fang, G. Liu, and R. Zhou, "Modeling of district load forecasting for distributed energy system," *Applied Energy*, vol. 204, pp. 181-205, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917308747>

[57] H. Shiraki, S. Nakamura, S. Ashina, and K. Honjo, "Estimating the hourly electricity profile of Japanese households - Coupling of engineering and statistical methods," *Energy*, vol. 114, pp. 478-491, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544216311227>

[58] L. M. Candanedo, V. Feldheim, and D. Deramaix, "Data driven prediction models of energy use of appliances in a low-energy house," *Energy and Buildings*, vol. 140, pp. 81-97, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778816308970>

[59] S. Naji, A. Keivani, S. Shamshirband, U. J. Alengaram, M. Z. Jumaat, Z. Mansor, and M. Lee, "Estimating building energy consumption using extreme learning machine method," *Energy*, vol. 97, pp. 506-516, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S036054421501587X>

[60] C. Tu, X. He, Z. Shuai, and F. Jiang, "Big data issues in smart grid - A review," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 1099-1107, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1364032117307748>

[61] O. C. Ozerdem, E. O. Olaniyi, and O. K. Oyedotun, "Short term load forecasting using particle swarm optimization neural network," *Procedia Computer Science*, vol. 120, pp. 382-393, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050917324663>

[62] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," *Renewable and Sustainable Energy Reviews*, vol. 54, pp. 1311-1322, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S136403211501196X>

[63] C. Tong, J. Li, C. Lang, F. Kong, J. Niu, and J. J. P. C. Rodrigues, "An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders," *Journal of Parallel and Distributed Computing*, vol. 117, pp. 267-273, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S074373151730196X>

[64] Y. Yang, J. Che, Y. Li, Y. Zhao, and S. Zhu, "An incremental electric load forecasting model based on support vector regression," *Energy*, vol. 113, pp. 796-808, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544216310167>

[65] Y. Chen, Y. Yang, C. Liu, C. Li, and L. Li, "A hybrid application algorithm based on the support vector machine and artificial intelligence: An example of electric load forecasting," *Applied Mathematical Modelling*, vol. 39, no. 9, pp. 2617-2632, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0307904X14005769>

[66] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, "Machine learning approaches for estimating commercial building energy consumption," *Applied Energy*, vol. 208, pp. 889-904, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917313429>

[67] L. Hernández, C. Baladron, J. M. Aguiar, B. Carro, A. Sánchez- Esguevillas, and J. Lloret, "Artificial neural networks for short-term load forecasting in microgrids environment," *Energy*, vol. 75, pp. 252- 264, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544214008871>

[68] I. M. Coelho, V. N. Coelho, E. J. da S. Luz, L. S. Ochi, F. G. Guimarães, and E. Rios, "A GPU deep learning metaheuristic based model for time series forecasting," *Applied Energy*, vol. 201, pp. 412- 418, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917300041>

[69] T. Yang, M. Ren, and K. Zhou, "Identifying household electricity consumption patterns: A case study of Kunshan, China," *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 861-868, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S136403211830265X>

[70] A. Boustani, A. Maiti, S. Y. Jazi, M. Jadliwala, and V. Namboodiri, "Seer Grid: Privacy and Utility Implications of Two-Level Load Prediction in Smart Grids," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 2, pp. 546-557, 2017.

[71] J. Yang, C. Ning, C. Deb, F. Zhang, D. Cheong, S. E. Lee, C. Sekhar, and K. W. Tham, "k-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement," *Energy and Buildings*, vol. 146, pp. 27-37, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778817305352>

[72] M. W. Ahmad, M. Mourshed, and Y. Rezgui, "Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption," *Energy and Buildings*, vol. 147, pp. 77-89, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378778816313937>

[73] S. Deng, C. Yuan, L. Yang, and L. Zhang, "Distributed electricity load forecasting model mining based on hybrid gene expression programming and cloud computing," *Pattern Recognition Letters*, vol. 109, pp. 72-80, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167865517303665>

[74] Y. Chen, M. Kloft, Y. Yang, C. Li, and L. Li, "Mixed kernel based extreme learning machine for electric load forecasting," *Neurocomputing*, vol. 312, pp. 90-106, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231218306532>

[75] W. Yu, F. Zhuang, Q. He, and Z. Shi, "Learning deep representations via extreme learning machines," *Neurocomputing*, vol. 149, pp. 308–315, 2015. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231214011461>

[76] Z. Ding and M. Fei, "An Anomaly Detection Approach Based on Isolation Forest Algorithm for Streaming Data using Sliding Window," *IFAC Proceedings Volumes*, vol. 46, no. 20, pp. 12–17, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1474667016314999>

[77] C. Fan, F. Xiao, Y. Zhao, and J. Wang, "Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data," *Applied Energy*, vol. 211, pp. 1123–1135, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306261917317166>

[78] W. Zhang, H. Quan, and D. Srinivasan, "Parallel and reliable probabilistic load forecasting via quantile regression forest and quantile determination," *Energy*, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360544218313148>

[79] M. Shepero, D. van der Meer, J. Munkhammar, and J. Widen, "Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data," *Applied Energy*, vol. 218, pp. 159–172, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S030626191830299X>

[80] C. Lopez, W. Zhong, and M. Zheng, "Short-term Electric Load Forecasting Based on Wavelet Neural Network, Particle Swarm Optimization and Ensemble Empirical Mode Decomposition," *Energy Procedia*, vol. 105, pp. 3677–3682, 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1876610217309323>

[81] G. Chicco, R. Napoli, and F. Piglione, "Comparisons among clustering techniques for electricity customer classification," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933–940, 2006.

[82] G. Dudek, "Neural networks for pattern-based short-term load forecasting: A comparative study," *Neurocomputing*, vol. 205, pp. 64–74, 2016. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231216302764>

[83] V. Thouvenot, A. Pichavant, Y. Goude, A. Antoniadis, and J. M. Poggi, "Electricity Forecasting Using Multi-Stage Estimators of Nonlinear Additive Models," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3665–3673, 2016.

[84] J. Runge and R. Zmeureanu, "A review of deep learning techniques for forecasting energy use in buildings," *Energies* 2021, vol. 14, p. 608, 2021. [Online]. Available: <https://www.mdpi.com/1996-1073/14/3/608/pdf>

[85] F. M.-A. J. F. Torres and A. Troncoso, "A deep lstm network for the spanish electricity consumption forecasting," *Neural Computing and Applications*, 2021. [Online]. Available: <https://link.springer.com/content/pdf/10.1007/s00521-021-06773-2.pdf>