# Electricity Load/demand Forecasting in Sri Lanka using Deep Learning Techniques

Abeysingha A.A.K.U
*Dept. of Electrical & Electronic Engineering*
*University of Jaffna*
Kilinochchi, Sri Lanka
udarakasun27@gmail.com

Sritharan A.S
*Dept. of Electrical & Electronic Engineering*
*University of Jaffna*
Kilinochchi, Sri Lanka
subashinisri3@gmail.com

Valluvan R
*Dept. of Electrical & Electrical Engineering*
*University of Jaffna*
Kilinochchi, Sri Lanka
Valluvan.r@eng.jfn.ac.lk

Ahilan K
*Dept. of Electrical & Electronic Engineering*
*Univercity of Jaffna*
Kilinochchi, Sri Lanka
ahilan.eng@gmail.com

Jayasinghe D.H.G.A.E
*Dept. of Engineering Technology*
*University of Jaffna*
Kilinochchi, Sri Lanka
jakila@tech.jfn.ac.lk

***Abstract* – Most of the time electricity cannot be stored, it should be generated as soon as it is demanded. Therefore, electricity demand forecasting is a vital process in the planning of electricity industry and the operation of electric power systems. Two major scenarios should be considered when forecasting the electricity demand. They are short term and long term forecasting scenarios. The short term scenario is more critical since many features have to be considered. In this research study, deep learning techniques such as Recurrent Neural Network (RNN), Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) were considered for electricity demand forecasting of Sri Lankan demand profile. Further, the results of deep learning approaches were compared with traditional techniques such as Linear Regression, Lasso Regression, Light Gradient Boosting Model (LGBM) and Random Forest Regressor. It was found from our studies that LSTM based approach performs better than other approaches.**

***Keywords - Demand forecasting, short term forecast, long term forecast, deep learning, neural network, regression***

## I. Introduction

Electricity demand forecasting can be divided into three types based on the time frame that is used to forecast. They are; short term load forecasting (an hour to few weeks), medium term load forecasting (few weeks to a few months) and long term load forecasting (few months to a few years) [1]. Recently there were many researches about the Short Term Load Forecasting (STLF) [2]. STLF is very useful for utilities by providing accurate predictions of future demand which could be used to precise planning, estimation of suppliers and determining prices. Similarly, (Long term load forecasting) LTLF is useful to get an insight into future consumption requirements and needed generation to satisfy future demand. Hence, system engineers can plan the future power plants as well as the maintenance periods for the existing power plants accordingly. Both STLF and LTLF leads to reduction in extra generation and operational cost, increase the net profit and maintain the reliability of the power system [3]. Over the past few years of research in STLF and LTLF, there are several models that are proposed to forecast electricity demand in different countries [4]. Nowadays, most of them are focused on machine learning based techniques rather than classical techniques because of the accuracy and easiness to forecast few hours to few years ahead [5].

Generally, deep learning techniques are highly precise and are useful for time series forecasting. These algorithms predict the future demand patterns by combining previous system data as well as weather data [6]. There are several factors that influence the electricity demand forecasting such as time (hour of the day, day of the week and etc.), weather (temperature and humidity), economy and possible customers groups are the few of the major factors among them. The time factor consists time of the year, day of the week, weekend/week days and special day or season of the year [7]. Sri Lankan government offers 25 formal annual holidays and it is divided into two groups such as P.B.M (Public, Bank & Mercantile) and P.B (Public & Bank). Similarly, peak electricity demand can be observed in monthly Buddhists' holiday called "Poya day". The unique and unsteady peak demand pattern is shown because of the Sinhalese and Tamil cultural festival season called as "New Year Season" during April and Buddhists' religious activities for "Wesak Poya" in May [8]. It is necessary to clean and scale data to zero mean and unit variance under the preprocessing step, after collecting the data of highly correlated factors to the electricity demand forecasting [9]. The deep neural network (DNN) architectures are very important to grab the information from the historical demand sequence which contains rich information to predict the demand in both STLF and LTLF. The Recurrent Neural Network (RNN) contains feedback connections which can be external or internal. Therefore, RNN has the ability of learning the patterns from the past records, generalize and project the future patterns for unseen data.

In this study, the deep neural network models have been used as deep architectures and regression models as the traditional methods to forecast electricity demand profile in Sri Lanka for both STLF and LTLF and compare each model based on the performance evaluators such as Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). Then two best models from both deep neural network models and regression models were selected to forecast electricity demand. This work makes use of Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Long-Short-Term-Memory for the deep architectures. For traditional methods, the linear regression, lasso regression, light gradient boosting model and random forest regression models have been considered. The data is based on ten year of Sri Lankan demand data taken from the system control center, Ceylon Electricity Board (CEB) of Sri Lanka.

## II. BACKGROUND

### *A. Traditional methods*

The Linear Regression (MLR), LASSO Linear Regression (LASSO), Light Gradient Boosting Model (LGBM) and Random Forest Regression Model were used as traditional model to forecast electricity demand in Sri Lanka [5]. The reason of selecting these models is the popularity of the literature and usage in the industry. The brief description about each of them is provided below.

**Linear Regression (MLR)** model is good to estimate the general trends of the data and it cannot predict nonlinearities. It basically produces coefficient vector that can be multiplied with any input to provide an estimated output. The modeling is performed by considering the relationship of demand consumption and other influence factors such as weather (Temperature, Humidity), year, quarter and hour. The basic equation of a linear regression is given below. Here; $\boldsymbol{A}$ is the slope, and $\boldsymbol{B}$ is the intercept.

$$Y = Ax + B + \boldsymbol{\varepsilon}$$
$$x = (x_1, x_2, x_3 \dots \dots, x_r)$$

Where $r$ is the number of predictors, $\boldsymbol{\varepsilon}$ is the random error.

**LASSO Linear Regression** is also a type of the linear regression model. This method could produce thinly distributed solutions and when the numbers of features are less compared to the number of observations this method performs very well [10]. Also in our method, the $\boldsymbol{l1}$ regularization has been used. Because it helps to skip the data from the over fitting. This is a time-consuming process.

**Light GBM Regression** contains two procedures those are Gradient-based One-Side Sampling and Exclusive Feature Bundling to manage huge number of information occurrences and enormous number of highlights individually [11]. It is based on the decision tree algorithm.

Objective of the function = training loss + regularization

$$Obj(\theta) = L(\theta) + \Omega(\theta)$$

**Random Forest Regression** is a ductile method, it is easy to use machine learning algorithm that produces magnificent results most of the time and spent minimum time on hyper-parameter tuning. Further, it is a simple algorithm [12].

### *B. Deep Neural Networks*

Deep Neural Networks (DNN) are complex and accurate models than traditional models. Therefore, it is possible to get more reliable predictions of future demand. Different architectures were used such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long-Short-Term-Memory and brief description of each architectures are discussed below.

**Convolutional Neural Network (CNN)** is a type of deep neural network, it can be used for natural language processing and analyzing of time series data [13]. The series of convolutional layers can be seen in the hidden layers of a CNN and they are convolved with a multiplication or other dot product. The process of convolution having a filter that can be sliding over the time series. In this case, unlike the image, filters have only one dimension (time) instead of two dimensions (width and height). The filter can be introduced as a generic non-linear transformation of a time series.

The general way of applying the convolution for centered time stamp *t* is given by following equation.

$$C_t = f\left(\omega \times X_{t-l/_2} : t + {}^{l}/_{2} + b\right) | \forall t \in [1, T]$$

Where $\boldsymbol{C}$ the result of a convolution is applied on the univariate time series $\boldsymbol{X}$, of length $\boldsymbol{T}$ with a filter ω of length $\boldsymbol{l}$ and bias parameter $\boldsymbol{b}$. The $\boldsymbol{f}$ is the final non-linear function called as Rectified Linear Unit (ReLU). The output from the convolution of an input time series $\boldsymbol{X}$, can be considered as another univariate time series $\boldsymbol{C}$ and it is also undergo a filtering process. Therefore, the multivariate time series whose dimensions are equal to the number of filters used can be obtained by applying several filters on a time series. It will be used to learn multiple discriminative features useful for classifications.

**Recurrent Neural Network (RNN)** models are neural networks which consists of feedback loops allowing the persistence of the information. Keeping the memory is very important to learn about long term dependencies in a sequence. The very important part of the RNN is a layer made by memory cells. The most famous cell at that moment is the Long-Short-Term-Memory (LSTM) that is described below.

**Long-Short-Term-Memory (LSTM)** is very popular deep neural network and it contains special units called memory blocks in the recurrent hidden layer. Each memory block has an input gate and an output gate. The input gate controls the flow of input activations into the memory cell. The output gate controls the output flow of cell activations into rest of the network.

An LSTM network computes a mapping from an input sequence x = ($x_1$ …., $x_t$ ) to an output sequence y = ($y_1$ ..., $y_t$) by calculating the network unit activations using following equations iteratively from t = 1 to T,

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i)$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f)$$

$$c_t = f_t.c_{t-1} + i_t.g(W_{cx}x_t + W_{cm}m_{t-1} + b_c)$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o)$$

$$m_t = o_t.h(c_t)$$

$$y_t = \varphi(W_{ym}m_t + b_y)$$

where the $\boldsymbol{W}$ terms denote weight matrices (e.g. $\boldsymbol{W_{ix}}$ is the matrix of weights from the input gate to the input) and $\boldsymbol{W_{ic}}$ , $\boldsymbol{W_{fc}}$ , $\boldsymbol{W_{oc}}$ are diagonal weight matrices for peephole connections, the $\boldsymbol{b}$ terms denote bias vectors ($\boldsymbol{b_i}$ is the input gate bias vector), $\boldsymbol{\sigma}$ is the logistic sigmoid function and $\boldsymbol{i}$, $\boldsymbol{f}$, $\boldsymbol{o}$ and $\boldsymbol{c}$ are respectively the input gate, forget gate, output gate and cell activation vectors. All of which are the same size as the cell output activation vector $\boldsymbol{m}$, "." is the element-wise product of the vectors, $\boldsymbol{g}$ and $\boldsymbol{h}$ are the cell input and cell output activation functions. Generally, $tanh$ and $\boldsymbol{\varphi}$ is the network output activation function.

## III. Analysis

### *A. Data Description*

The dataset obtained from CEB consists of hourly samples over the ten years period from 31st January 2006 to 31st December 2015. The seven (influence factors) features were used such as temperature, humidity, weekend/weekdays, year, quarter, GDP (Gross Domestic Product) and population growth rate apart from the electricity demand data. The data set was split into two sets for training and testing with the sizes of 80% and 20% respectively. Then correlation of features with the demand data was checked and less correlated features such as temperature, humidity, and GDP and population growth rate were dropped. As Sri Lanka is a tropical region country there is no significant variation of electricity demand due to temperature or humidity. There were several missing data samples also available in the dataset. Therefore, the forward filling method has been used to fill the missing data.

### *B. Method*

Data pre-processing was the first step in the time series analysis. Because the dataset had missing data, an approach indicate in the previous section had been followed to fill the missing samles. However, it could be a reason for having outliers too. Therefore, one of the objectives of pre-processing was removing the outliers from the dataset to mitigate the unexpected variations of the data. After removing the outliers from the dataset, the normalization has been done to take the data samples between 1 and 0. Normalization was done as given in the following equation.

$$X_{sc} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

After finishing the cleaning and normalization of data samples, all the traditional techniques were cross-validated to determine the appropriate values for the hyper-parameters. Finally, the performance of the models were evaluate with metrics such as MAE, MSE and RMSE. They are defined as;

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|A_i - F_i|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - Q_i)^2$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(P_i - Q_i)^2}{N}}$$

### *C. Numerical Results*

TABLE I. PERFORMANCE OF DNN MODELS

| Models | MAE | MSE | RMSE |
|---|---|---|---|
| CNN | 0.4103 | 0.1874 | 0.4328 |
| RNN | 0.4386 | 0.2239 | 0.4732 |
| LSTM | 0.3257 | 0.1158 | 0.3404 |

Initially, the performance of each model was checked by changing the model parameters such as activation function, number of neurons and number of iterations. Therefore, the number of neurons and epochs were kept constant as it was keeping under fitting condition while changing the activation function. In addition to the features mentioned above, there were five layers used for the CNN with filter size of (1*1), four layers for the RNN and two layers for the LSTM. The number of neurons used for the CNN were 50, 32 and 2 respectively and for the RNN used similar number of neurons that was 50. The best activation function was selected as ReLU by rejecting the Sigmoid and $tanh$ based on the performance shown by the models. The best results was given by LSTM and the performance was 0.3257, 0.1158, and 0.3404 for the MAE, MSE and RMSE respectively. The Fig.

1 shows how the model performs when data was trained with number of epochs. Initially, MSE error was taken higher value and it was reduced and steady while increasing number of epochs.

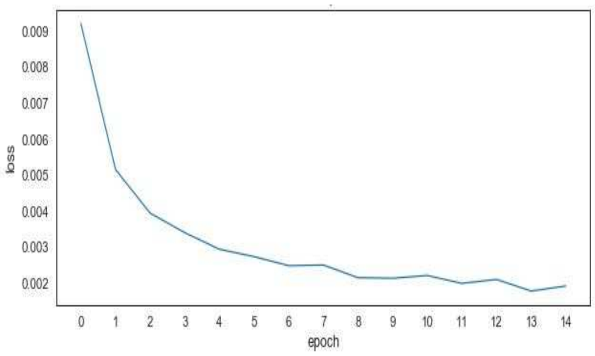


Fig. 1: Performance of train data with number of epochs

Then, the number of neurons were changed by setting the activation function as ReLU and keeping number of epochs same as previous step. Number of neurons were changed step by step such as 50, 500 and 1000 neurons. The maximum performance was observed when 50 neurons were used and best performance was given by LSTM as before. They were 0.3362, 0.1280, and 0.3577 for the MAE, MSE and RMSE. Then, the number of epochs were changed step by step such as 15, 70 and 200 by setting the activation function as ReLU and number of neurons as 50. The best performance was given at 70 iterations by LSTM model. The performance was 0.3257, 0.1158 and 0.3404 for the MAE, MSE and RMSE. Same procedure was applied for all DNNs and summary of the performance is tabulated in TABLE I. The performance of the regression model is given in TABLE II.

TABLE II. PERFORMANCE OF REGRESSION MODELS

| **Models** | **MAE** | **MSE** | **RMSE** |
|---|---|---|---|
| Linear regress | 1.0950 | 1.6889 | 1.2996 |
| Lasso regress | 1.1626 | 1.9876 | 1.4098 |
| LGBM | 0.9751 | 1.3738 | 1.1721 |
| Random forest | 0.8555 | 1.2548 | 1.1202 |

### *D. STLF and LTLF using LSTM model*

The optimized LSTM model was used to test the fit model by forecasting both short term and long term scenarios. Since LSTM gave best performance and lower error, it has performed better than other DNN techniques. Similarly, random forest regressor gave the best performance among the traditional methods and their performance is given in TABLE *II*.

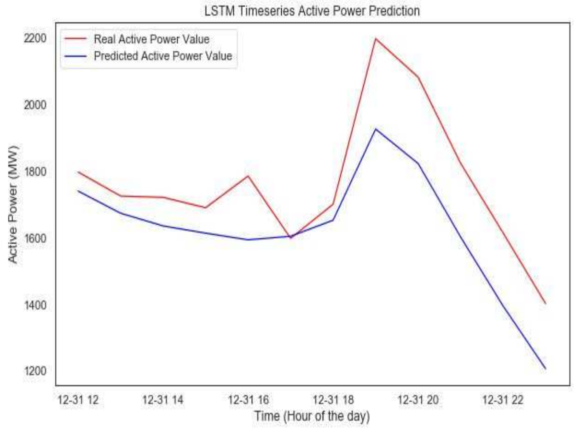


Fig. 2. STLF using LSTM model

The Fig. 2 shows test data performance of LSTM model for a very short time period such as few hours. The predicted pattern is close to real pattern and it has small deviation since features were taken as their average values. On the other hand STLF can have sudden variation of demand pattern and features may be static only for particular time frame.

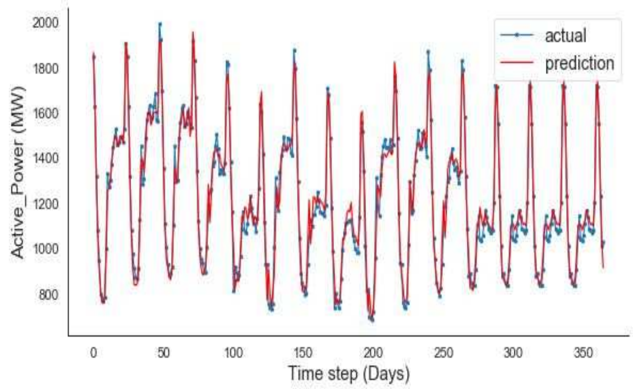


Fig. 3. LTLF using LSTM model

The Fig. 3 is shows test data performance of LSTM model for few weeks of time frame. The predicated demand pattern closely follows the real demand pattern. Therefore, the difference in the demand pattern reduces as the time frame increases.

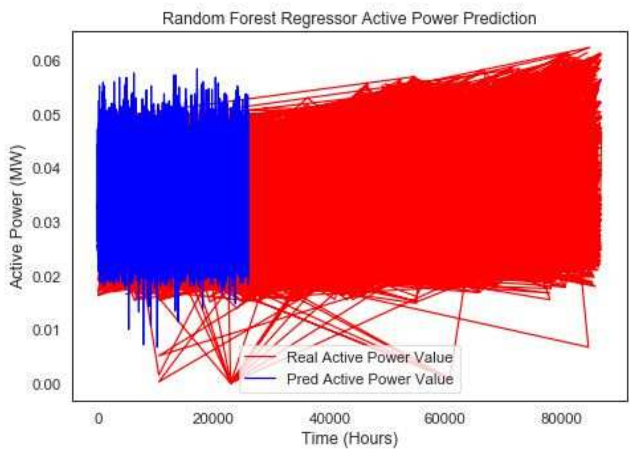


Fig. 4. LTLF using random forest regression

## E. Seasonal/Special days Pattern Analysis

The seasonal data analysis plays an important role when dealing with the time series forecasting. The term seasonality can be defined as fixed or known period. Therefore, seasonal time series can be called as periodic time series as well. New Year period electricity demand in Sri Lanka was analyzed. The celebration time always starts from the beginning of the April month. Therefore, the peak demand can be observed at the beginning of April month every year and it slightly reduces when it becomes towards the end of the month. Similarly, the peak demand in April month has been increasing year by year.

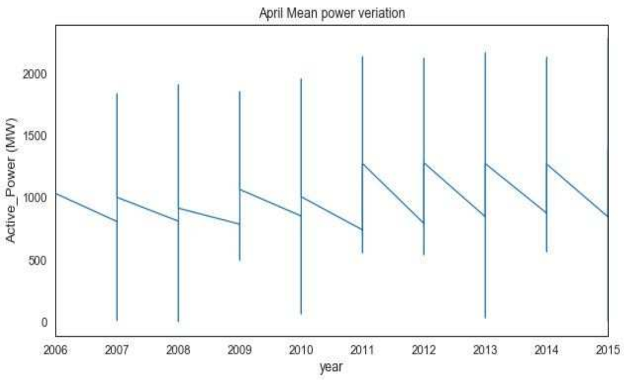


Fig. 5. The demand variation in April month

Then weekend/weekdays electricity demand variation in Sri Lanka was analyzed. The day of the week is a very important factor when demand forecasting model is developed. Therefore, attention is required on whether it is a weekday or weekend day. The weekday demand is much higher than weekend day demand according to the observation. The reason can be most of the working places stop operation during weekend and continuously work in week days. However, both weekday and weekend day demand has increased year by year.

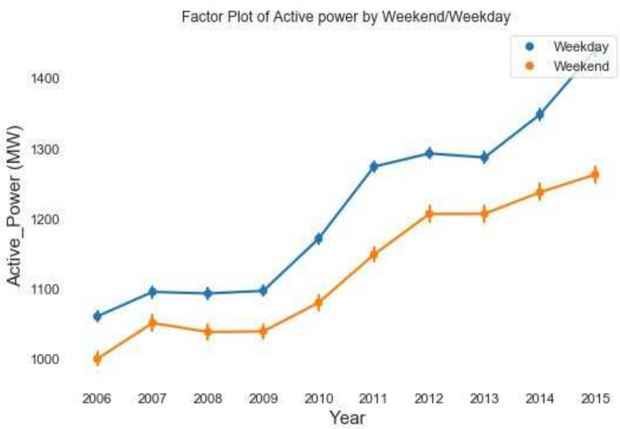


Fig. 6. Weekend/weekday demand variations

## IV. Conclusion

Most of the time electricity cannot be stored, it should be generated as soon as it is demanded. Therefore, electricity demand forecasting is a vital process in the planning of electricity industry and the operation of electric power systems. In this paper, deep learning techniques such as Recurrent Neural Network (RNN), Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) were used for electricity demand forecasting of Sri Lankan demand profile. Further, the results of deep learning approaches were also compared with Linear Regression, Lasso Regression, Light Gradient Boosting Model (LGBM) and Random Forest Regressor. It was found from our research studies that LSTM based approach perform better than RNN, CNN, Linear Regression, Lasso Regression, Light Gradient Boosting Model (LGBM) and Random Forest Regressor.

## References

[1] D. H. G. A. E. Jayasinghe, K. R. S. Lankapriya, R. Valluvan, and V. Cynthujah, “Scenario Analysis of Future Load Profile of Sri Lanka Considering Demand Side Management Initiatives,” in *International Conference on Innovative Smart Grid Technologies, ISGT Asia 2018*, 2018, pp. 1251–1256.

[2] A. Baliyan, K. Gaurav, and S. Kumar Mishra, “A review of short term load forecasting using artificial neural network models,” in *Procedia Computer Science*, 2015, vol. 48, no. C, pp. 121–125.

[3] W. He, “Load Forecasting via Deep Neural Networks,” in *Procedia Computer Science*, 2017, vol. 122, pp. 308–314.

[4] S. Hosein and P. Hosein, “Load forecasting using deep neural networks,” in *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2017.

[5] H. Sangrody, N. Zhou, S. Tutun, B. Khorramdel, M. Motalleb, and M. Sarailoo, “Long term forecasting using machine learning methods,” in *2018 IEEE Power and Energy Conference at Illinois (PECI)*, 2018.

[6] E. Almeshaiei and H. Soltan, “A methodology for Electric Power Load Forecasting,” *Alexandria Eng. J.*, vol. 50, no. 2, pp. 137–144, Jun. 2011.

[7] J. Bedi and D. Toshniwal, “Deep learning framework to forecast electricity demand,” *Appl. Energy*, vol. 238, Mar. 2019.

[8] Beenish Khan and Nadeem Javaid, “Short-Term Residential Load and Price Forecasting based on Convolutional Neural Network,” 2018.

[9] D. C. Idoniboyeobu, A. J. Ogunsakin, and B. A. Wokoma, “Forecasting of Electrical Energy Demand in Nigeria using Modified Form of Exponential Model,” *Am. J. Eng. Res. (AJER*, no. 7, pp. 122–135, 2018.

[10] P. Ray, S. Sen, and A. K. Barisal, “Hybrid methodology for short-Term load forecasting,” in *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems, PEDES 2014*, 2014.

[11] T. Anwar, B. Sharma, K. Chakraborty, and H.

Sirohia, “Introduction to Load Forecasting,” *Int. J. Pure Appl. Math.*, vol. 119, no. 15, pp. 1527–1538, 2018.

[12] M. D. Olagoke, A. A. Ayeni, and M. A. Hambali, “Short Term Electric Load Forecasting using Neural Network and Genetic Algorithm,” *Int. J. Appl. Inf. Syst.*

[13] L. Hernandez *et al.*, “A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings,” *IEEE Commun. Surv. Tutorials*, vol. 16, no. 3, pp. 1460–1495, 2014.