An SVR-based Building-level Load Forecasting Method Considering Impact of HVAC Set Points

Ashraful Haque, Manisa Pipattanasomporn and Saifur Rahman
Bradley Department of Electrical and Computer Engineering
Virginia Tech – Advanced Research Institute
Arlington, VA, USA

Sri Raghavan Kothandaraman, Ahmadreza Malekpour
Esa Aleksi Paaso and Shay Bahramirad
Commonwealth Edison Company (ComEd)
Chicago, IL, USA

***Abstract*— This paper focuses on a technique to determine the impact of HVAC set point adjustments on building-level electrical load (kW) utilizing Support Vector Regression (SVR) with the minimum possible set of input variables. The paper uses two SVR-based forecasting methods, namely single-step and recursive models. These models are used to forecast hourly electrical loads of a commercial building in Chicago area for the summer period from 8AM to 8PM. The model accuracy is observed to be higher than 95% for hour-ahead load forecasts, and higher than 93% for 12-hour ahead load forecasts. The models presented in the paper can be used to quantify the reduction in electrical load (kW) based on HVAC set point adjustments during peak hours in buildings.**

Index Terms*— *load forecasting, SVR, regression model, machine learning, demand response

I. Introduction

Power grid as a whole is a large and distributed system consisting of generation, transmission and last mile distribution lines. This vast plethora of interconnected systems need to work in unison for the power grid to function properly. Any mismatch between the supply and demand can cause the system to shut down. System-level load forecasting has long been a topic of great interest that can help proper generation scheduling, plan for reserve margin and ancillary services, thereby maintaining power system stability. Since 1980s, much work has been done in the area of system-level load forecasting. Authors in [1] summarize various load forecasting techniques for large-scale power systems. The knowledge-based approach was introduced in [2, 3].

However, with the advent of smart grid and Internet-of-Things (IoT) devices, the attention has shifted from the system-level to building-level operation. Smart grid operation allows instant, secure and reliable monitoring and control of electrical equipment over different communication infrastructures and protocols [4, 5]. This capability results in an increasing popularity of developing building-level load forecasting techniques and methods. Several methods have been used for building-level load forecasts, including linear regression [6, 7], Artificial Neural Network (ANN) [8, 9], Support Vector Regression (SVR) [10, 11] and Fuzzy Neural Network (FNN) [12].

The above mentioned work typically performs baseline building-level load forecasts using input parameters, like historical load data, time of the day and outdoor temperature profiles. Incorporating HVAC set points provides additional advantages and allows the estimation of peak demand reduction potentials in a building when set points are raised during a demand response period. However, since most of the HVAC systems in building operation do not have a log mechanism, HVAC set points are normally not considered as input parameters for building-level load forecasting.

There are a few prior studies that incorporate HVAC set points along with outdoor temperatures in load forecasting using linear regression techniques [13, 14]. However, relationship among HVAC set points, outdoor temperatures and building electrical load are often non-linear and vary from buildings to buildings. Therefore, kernels along with linear and polynomial properties of SVR are useful to understand these non-linear relationships. The purpose of this paper is to formulate an SVR-based load forecasting model with a minimum possible set of input variables that can capture the impact of HVAC set point changes on building-level hourly electrical load (kW), focusing mainly during the peak demand hours. The paper discusses single-step and recursive models for load forecasting and compares their forecasting accuracy by using the models to forecast hourly electrical load of a building in Chicago.

II. Mathematical Background

Load forecasting tasks can be accomplished using different techniques, such as time series modeling, curve fitting, ARMA model, Kalman filtering and data smoothing. Most popular modern techniques are SVR and ANN. ANN models are efficient where there exist highly complex relationships among non-linear parameters. However, for load forecasting, ANN does not have any reliable theory to determine the structure of the network. This is due to the application of ERM (Empirical Risk Minimization) principles into ANN. SVR overcomes this limitation by introducing SRM (Structural Risk Minimization). Important concept of SRM is minimization of error by introducing regularization parameters. The accuracy of SVR depends largely on the selection of input variables. The more relevant the variables are the more accurate the result is. Irrelevant data can over or under fit the result.

In this work, SVR is used for load forecasting. Support Vector Machine (SVM) algorithm was introduced in [15]. SVR is an extension of that algorithm. SVR uses the “Hyperplane” separating the data points with minimum error as the fitted

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

978-1-5386-8232-6/19/$31.00 ©2019 IEEE

model to predict future outputs. For a given data points (x_i, y_i) a linear function in (1) can be developed.

$$f(x) = \boldsymbol{w}x + b \tag{1}$$

Where, $\boldsymbol{w}$ is the weight vector and b is the bias term. This function can be trained by a linear regression. The objective of SVR is to find the flattest fit. Therefore, the objective function for SVR is:

$$\begin{aligned} \min \quad & \frac{1}{2}||\boldsymbol{w}^2|| + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ \text{s.t.} \quad & y_i - f(x_i, \boldsymbol{w}) \le \epsilon + \xi_i^* \\ & f(x_i, \boldsymbol{w}) - y_i \le \epsilon + \xi_i \\ & \xi_i, \xi_i^* \ge 0, i = 0,1,2,3\ldots n \end{aligned} \tag{2}$$

Here, C is the cost function, which represents the trade-off between points lying outside the tolerance boundary and the flatness of the fitted curve. ϵ represents the tolerance boundary and n is the number of test cases. Value of C and ϵ are user dependent. ξ_i and ξ_i^* are the errors for the points lying outside the tolerance boundary. After the "Hyperplane" generated with the minimum value of objective function is found, that can be used as the fitted model to predict value y_i for given set of inputs, x_i.

By introducing Lagrangian multipliers α, α^* and solving the quadratic optimization problem with inequality constraints, weight factor $\boldsymbol{w}$ in (1) can be obtained as:

$$w = \sum_{i=1}^{n}(\alpha^* - \alpha)\ x_i \tag{3}$$

The SVR regression equation can then be written as:

$$f(x) = \sum_{i=1}^{n}(\alpha^* - \alpha)\ k(x_i, x) + b \tag{4}$$

Where, $k(x_i, x)$ is the kernel function, which maps the data to higher dimensional spaces that enables better manipulation of non-linear relationship in dataset. Kernels used for the methods developed in this paper are: Linear, Polynomial and Gaussian.

III. METHODOLOGY

This paper focuses on using SVR-based machine learning method to develop a model that can capture the impact of HVAC set point changes on building-level hourly electrical load. This section describes the reason underlying the selection of input variables, SVR models, the preparation of training and testing datasets and model validation criteria.

A. *Reasoning for Input Variables Selection*

The accuracy of load forecasts using SVR depends largely on the proper selection of input variables and on the relevance of the variables rather than the volume of the data. The aim is to account for as minimum number of variables as possible with high percentage of forecasting accuracy.

It is intuitive that the power consumption of a building is dependent on HVAC set points. That is, lower set points result in higher electricity consumption in buildings. Authors in [16] show that there exists relationship between outdoor temperature and preferred HVAC set point.

Fig. 1 plots the simulated hourly electrical loads of the target building (kW, Y-axis) against outdoor temperatures (°F, X-axis) during May 1st – October 31st. It can be seen that the hourly electrical loads are less than 175 kW, when the outdoor temperatures are lower than the baseline set point of 72°F. However, the hourly electrical loads increase non-linearly as the outdoor temperatures increase beyond the baseline set point. As there is a relationship between outdoor temperatures and hourly building electrical load, outdoor temperatures along with HVAC set points are considered as the input parameters in the developed SVR-based load forecasting model.

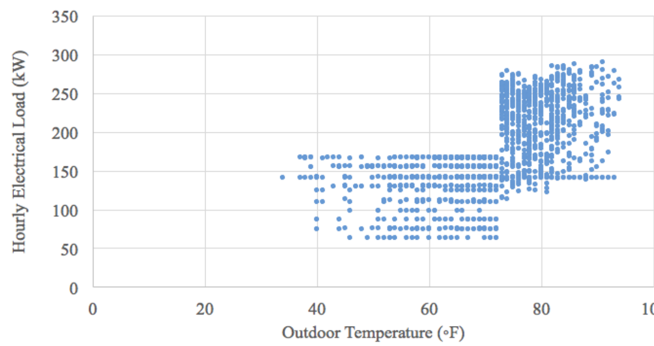

Fig. 1. Relationship between outdoor temperature (°F) and hourly electrical load (kW) of the target building.

B. *SVR Models*

Two types of SVR-based load forecasting models are used: single-step and recursive models. These models are trained separately for each hour with separate training data. The single-step model aims at forecasting building electrical loads (kW) at the beginning of each hour for the next hour. On the other hand, the recursive model performs electrical load forecasting once at the beginning of the forecasting hour, using the single-step model to predict the first hour load and then recursively uses the predicted loads to forecast future hours' electrical loads.

Input variables to the models are chosen to include:

- HVAC set point (°F) at hour *i*
- HVAC set point (°F) at hour *i+1*
- Outdoor temperature (°F) at hour *i*
- Outdoor temperature (°F) at hour *i+1*
- Building electrical load (kW) at hour *i*
  - ✓ Actual load for the single-step model
  - ✓ Predicted load for the recursive model

MATLAB function *fitrsvm* has been used as a basis for the building-level load forecasting method. "Grid-Search" has been done by setting the *OptimizeHyperparameter* function of *fitrsvm* to *all* to find out the best possible combination of the relevant parameters (cost function, kernel function, kernel scale, epsilon, polynomial order) thus minimizing the objective function of SVR. This implies the forecasting error has the lowest possible value.

C. Preparing Training and Testing Datasets

While time-series building-level hourly electrical load data are readily available via the corresponding smart meter, HVAC set points are not easily obtained. Hence, in this study, a building simulation model has been developed using eQUEST and validated against the historical hourly electrical load data of a commercial building in Chicago, which is obtained from smart meter. The validated eQUEST model was then used to generate input training and testing datasets for the SVR model described in the previous section.

The building in consideration is three-story building with total building area of around 100,000 square feet. Inputs used for eQUEST model development are weather data, building structure and envelope details, seasonality, HVAC system model, occupancy and other loads like lighting, plug and water heating. Outdoor temperature data have been obtained from Weather Underground https://www.wunderground.com/, for the Chicago area. The developed eQUEST building model containing HVAC set point information has been validated against hourly electrical load data obtained from the corresponding smart meter. The simulated hourly electrical load closely follows the actual one with the average error of around 5% during the peak demand hours. This indicates that the developed eQUEST model can represent the actual hourly building load, thereby validating the developed building simulation model.

Once validated, the developed eQUEST model has been used to prepare training and testing datasets for each hour starting from 8AM to 8PM.

- The training dataset has been prepared with different combinations of HVAC set points during different hours of each day and with the weather data of three summer periods from May 1st to October 31st in 2015, 2016 and 2017 when HVAC is operational.
- The testing dataset of the baseline scenario (Case I in Section IV) has been prepared by keeping the set point constant at 72°F for each hour of each day and by using the weather data from May 1st to October 31st in 2017 when HVAC is operational.
- The testing dataset of the case with HVAC set point adjustment (Case II in Section IV) has been prepared by increasing the HVAC set point from the baseline to 74°F between 12PM and 4PM, using the weather data from May 1st to October 31st in 2017 when HVAC is operational.

Since, for each hour of the day starting from 8AM to 8PM different fitted SVR models are developed, individual training and testing datasets have been prepared for each hour, comprising hourly electrical load at the beginning of each hour, outdoor temperatures and HVAC set points. The ratio of training to testing data is 2:1. Training and testing datasets for both single-step and recursive models are the same.

D. Validation Criteria

The Mean Absolute Percentage Error (MAPE, %) and the goodness of fit (R^2) have been selected as the validation criteria. MAPE (%) represents the numerical accuracy of the forecast and R^2 represents how closely the predicted model follows the actual load pattern. See (5) and (6).

$$\text{MAPE}\ (\%) = \frac{1}{n}\sum_{i=1}^{n} \frac{\left|y_{p,i} - y_{a,i}\right|}{y_{a,i}} \cdot 100 \tag{5}$$

$$R^2 = \frac{\sum_{i=1}^{n}(y_{p,i} - y_{m,i})^2}{\sum_{i=1}^{n}(y_{a,i} - y_{m,i})^2} \tag{6}$$

In this particular study, $y_{p,i}$ is the predicted load (kW) at hour i by the SVR model, $y_{a,i}$ is the actual load (kW) at hour i and $y_{m,i}$ is the mean value of the actual load (kW) at hour i.

IV. Load Forecasting with SVR and Errors

The developed SVR models have been used to perform baseline load forecasts (Case I) and load forecasts with HVAC set point adjustments between 12PM to 4PM (Case II) for the targeted building. The time for HVAC set point adjustments has been chosen to coincide with the demand response period to see the impact of HVAC set point change on the building's hourly electrical loads.

A. Case I: Baseline Load Forecasts

For Case I, the HVAC set point is assumed to be constant at 72°F between 8AM to 8PM.

Fig. 2 shows the actual vs predicted hourly electrical loads of this building from 12PM to 4PM using the single-step model with baseline set point of 72°F. For the single-step model, forecasting is done at the beginning of each hour. For the recursive model, on the other hand, at each hour, the forecasted hourly electrical load from the previous hour is used as an input of the current hour along with current set point and outdoor temperature. Hence, the forecasting has been done once at the beginning of the forecasting period, i.e., at 8AM. Fig. 3 shows the actual vs predicted hourly electrical loads from 12PM to 4PM for the recursive model with the baseline set point. Except at the beginning of the forecasting period, the recursive model does not require any information of actual electrical loads. Rather, it recursively uses the forecasted load from the previous hour to predict the electrical load for the next hour.

Fig. 2 and Fig. 3 show that the forecasted hourly loads from both models closely follow the actual hourly building loads, and the single-step model showing slightly better accuracy.

B. Case II: Load Forecasts with HVAC Set Point Adjustments

For Case II, the HVAC set point is assumed to be raised by +2°F from the baseline set point 72 °F during 12PM to 4PM, representing a demand response request to reduce peak demand in buildings. During other hours, HVAC remains at the baseline set point.

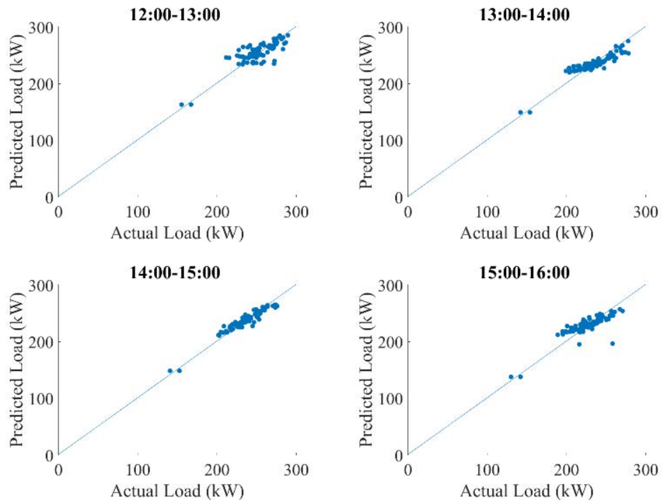

Fig. 2. Actual vs predicted hourly electrical loads from 12PM to 4PM for the single-step model with baseline set point of 72°F(Case I–single-step model).

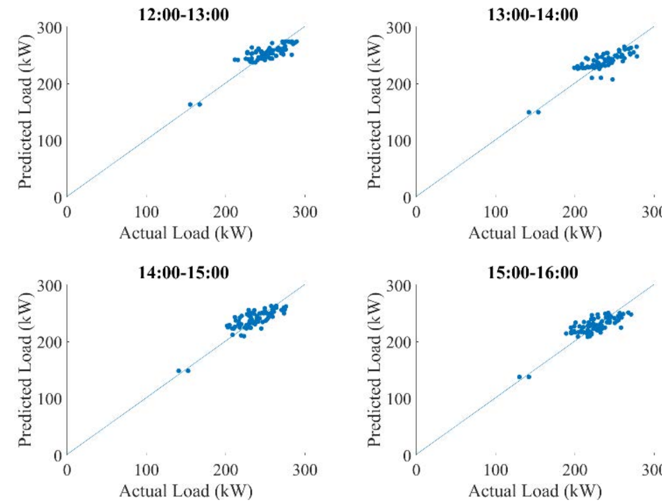

Fig. 3. Actual vs predicted hourly electrical loads from 12PM to 4PM for recursive model with baseline set point of 72 °F (Case I – recursive model).

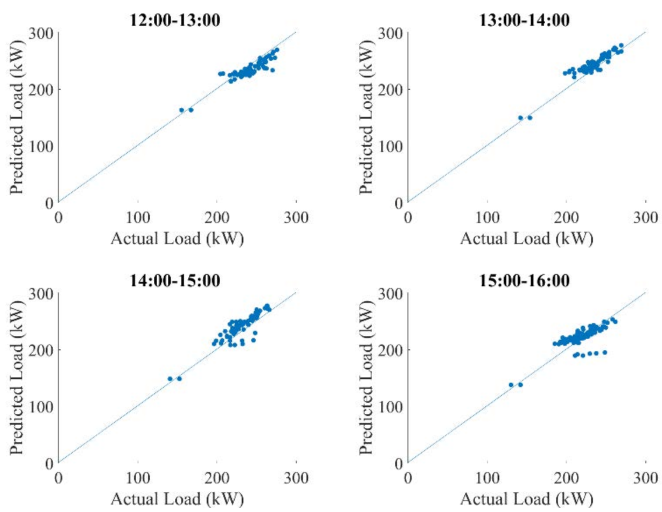

Fig. 4. Actual vs predicted hourly electrical loads from 12PM to 4PM for the single-step model with set point raised (Case II – single-step model).

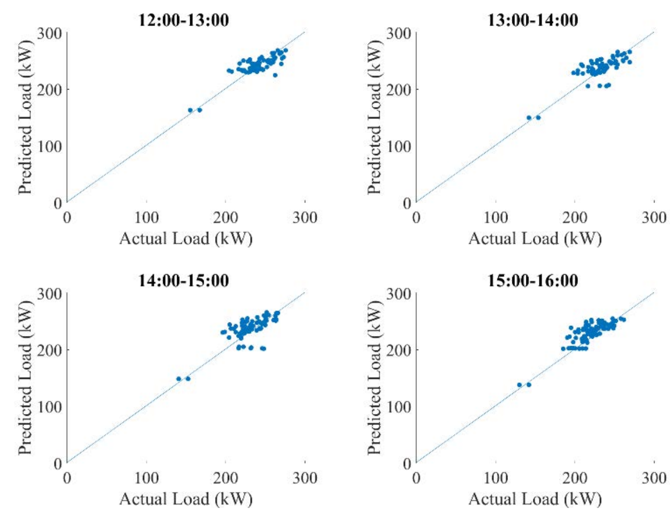

Fig. 5. Actual vs predicted hourly electrical loads from 12PM to 4PM for the recursive model with baseline (Case II – recursive model).

Fig. 4 and Fig. 5 show the actual vs predicted hourly electrical loads from 12PM to 4PM with set points raised by +2 °F from the baseline for the single-step model and recursive model, respectively. The plots show that the forecasts from both models are able to resemble the hourly electrical loads when the set point was raised by +2°F from the baseline with high accuracy, and thereby capturing the impact of set point change.

As an example to see how closely the forecasts are comparing to the actual loads, Fig. 6 plots the actual vs forecasted loads during the forecasting period, i.e., from 8AM to 8PM using the recursive model. It can be seen that the developed SVR model can perform load forecasting quite accurately for both Case I (with constant set point) and Case II (with the raised set point).

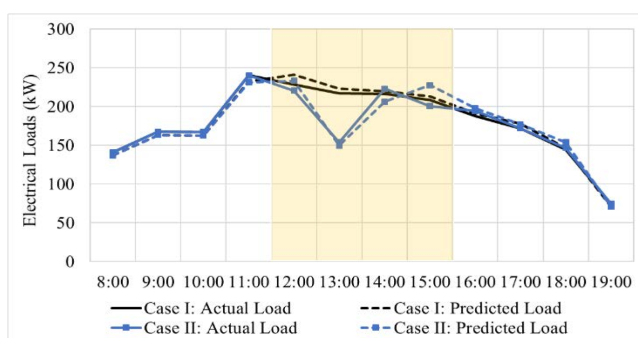

Fig. 6. Actual vs predicted hourly electrical loads (kW) for a single day with the baseline set point (Case I) and baseline +2°F (Case II).

Since the recursive model only requires the electrical load data at the beginning of the forecasting period as its input, along with weather forecasts, it is a useful method to predict hourly electrical loads of the following day if HVAC set points are adjusted during demand response period.

Table I summarizes the MAPE(%) and R^2 values for each hour load forecasting for both Cases I and II.

TABLE I
MAPE (%) AND R^2 FOR CASES I AND II

Time	Case I				Case II			
	MAPE		R^2		MAPE		R^2	
	Single-step	Recursive	Single-step	Recursive	Single-step	Recursive	Single-step	Recursive
08:00-09:00	3.81	3.81	0.97	0.97	3.81	3.81	0.97	0.97
09:00-10:00	3.32	3.25	0.96	0.97	2.65	3.52	0.98	0.95
10:00-11:00	3.95	3.86	0.94	0.95	3.99	3.78	0.94	0.95
11:00 12:00	3.45	3.61	0.97	0.95	3.66	3.83	0.95	0.94
12:00-13:00	3.95	3.77	0.94	0.94	3.39	3.89	0.95	0.94
13:00-14:00	2.97	3.98	0.98	0.94	4.19	4.34	0.94	0.93
14:00-15:00	4.01	4.94	0.93	0.89	4.31	5.08	0.95	0.91
15:00-16:00	3.25	4.22	0.97	0.94	4.45	5.41	0.91	0.89
16:00-17:00	4.47	5.26	0.95	0.92	5.8	5.51	0.89	0.92
17:00-18:00	4.46	5.71	0.96	0.92	4.47	5.92	0.94	0.91
18:00-19:00	4.39	6.46	0.98	0.93	5.35	6.29	0.96	0.93
19:00-20:00	4.81	7.39	0.98	0.93	5.01	7.61	0.96	0.92

For all hours in both the Cases I and II, MAPE (%) values for the single-step model are around 5% or lower. For the recursive model, MAPE (%) values are between 3.25% and 7.61%. R^2 values of both models are around 0.9 and above. Since the recursive model incorporates the error in previous hours' forecasting, it creates an avalanche effect, and thereby increasing MAPE (%) of later hours.

V. CONCLUSION

The models used in this paper are able to perform load forecasting in both the constant HVAC set point scenario and the raised set point scenario. Hence, the impact of HVAC set point raise on building electrical loads during a peak demand period can be quantified. Running building simulation software, like eQUEST, is resource hungry and requires lots of inputs and run time. The paper proposes that such a load simulation software to be run only once to prepare the training dataset and then use the fitted SVR model to forecast hourly electrical loads of a building at each one-hour intervals. Since the recursive model does not require any electricity consumption data as the input (except the actual load at the beginning of the forecasting period), the SVR-based models used in this paper can be utilized for day-ahead optimal operation planning, and can be beneficial to an electric utility and/or a Distribution System Operator (DSO) in discovering demand response potentials of a group of buildings. On the other hand, these models can be useful for a demand response aggregator in helping program participants to properly adjust HVAC set points during a demand response period to meet the required demand reduction target.

ACKNOWLEDGEMENT

The work presented in this paper was supported by the project "Development of Dynamic Input-Output Building Models for System-Level Energy Savings and Peak Demand Reduction Assessment", sponsored by the Commonwealth Edison Company, Chicago.

REFERENCES

[1] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," in IEEE Transactions on Power Systems, vol. 4, no. 4, pp. 1484-1491, Nov. 1989.

[2] S. Rahman and R. Bhatnagar, "An expert system based algorithm for short term load forecast," in IEEE Transactions on Power Systems, vol. 3, no. 2, pp. 392-399, May 1988.

[3] S. Rahman and O. Hazim, "A generalized knowledge-based short-term load-forecasting technique," in IEEE Transactions on Power Systems, vol. 8, no. 2, pp. 508-514, May 1993.

[4] Y. Yan, Y. Qian, H. Sharif and D. Tipper, "A Survey on Smart Grid Communication Infrastructures: Motivations, Requirements and Challenges," in IEEE Communications Surveys & Tutorials, vol. 15, no. 1, pp. 5-20, First Quarter 2013.

[5] V. C. Gungor et al., "Smart Grid Technologies: Communication Technologies and Standards," in IEEE Transactions on Industrial Informatics, vol. 7, no. 4, pp. 529-539, Nov. 2011.

[6] T. Hong, M. Gui, M. E. Baran and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," IEEE PES General Meeting, Providence, RI, 2010.

[7] Grzegorz Dudek, "Pattern-based local linear regression models for short-term load forecasting", Electric Power Systems Research,Volume 130,2016,Pages 139-147,ISSN 0378-7796.

[8] T. Senjyu, H. Takara, K. Uezato and T. Funabashi, "One-hour-ahead load forecasting using neural network," in IEEE Transactions on Power Systems, vol. 17, no. 1, pp. 113-118, Feb. 2002.

[9] Luis Hernández, Carlos Baladrón, Javier M. Aguiar, Belén Carro, Antonio Sánchez-Esguevillas, Jaime Lloret, "Artificial neural networks for short-term load forecasting in microgrids environment,Energy",Volume 75,2014,Pages 252-264,ISSN 0360-5442.

[10] Abdollah Kavousi-Fard, Haidar Samet, Fatemeh Marzbani, "A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate Short Term Load Forecasting", Expert Systems with Applications, Volume 41, Iss. 13, 2014, Pages 6047-6056.

[11] Bo-Juen Chen, Ming-Wei Chang and Chih-Jen lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," in IEEE Transactions on Power Systems, vol. 19, no. 4, pp. 1821-1830, Nov. 2004.

[12] A. Jain, E. Srinivas and R. Rauta, "Short term load forecasting using fuzzy adaptive inference and similarity," 2009 World Congress on Nature & Biologically Inspired Computing (NaBIC), Coimbatore, 2009, pp. 1743-1748.

[13] Rongxin Yin, Emre C. Kara, Yaping Li, Nicholas DeForest, Ke Wang, Taiyou Yong, Michael Stadler, "Quantifying flexibility of commercial and residential loads for demand response using setpoint changes", Applied Energy, Volume 177, 2016, Pages 149-164, Sept. 2016.

[14] J. L. Mathieu, P. N. Price, S. Kiliccote and M. A. Piette, "Quantifying Changes in Building Electricity Use, With Application to Demand Response," in IEEE Transactions on Smart Grid, vol. 2, no. 3, pp. 507-518, Sept. 2011.

[15] Cortes, C. & Vapnik, V. Mach Learn (1995), "Support-vector networks" https://doi.org/10.1007/BF00994018.

[16] Thermal comfort and building energy consumption implications – A review, Liu Yang, Haiyan Yan and Joseph C. Lam, Applied Energy, 2014, vol. 115, issue C, 164-17.