Journal
IJCSM

Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM
e-ISSN: 2788-7421 p-ISSN: 2958-0544

Enhancement Methods for Energy Consumption Prediction in Smart House based on Machine Learning

Musa A. Hameed [1], Esam T. Yassen [1], Wesam M. Jasim [1]

[1] Department of Computer Science, College of Computer Science and Information Technology University of Anbar, Anbar, IRAQ

*Corresponding Author: Musa A. Hameed.

DOI: https://doi.org/10.52866/ijcsm.2023.04.04.008
Receive July 2023; Accepted September 2023; Available online October 2023

ABSTRACT: Energy efficiency in modern homes has recently become a significant concern due to the emergence of smart-home infrastructure. Numerous public structures, such as homes, hospitals, schools, and other institutions, use more energy. To come close to meeting the actual energy demand, it is crucial that as much energy is created. The utilization of machine learning has various advantages in improving the effectiveness and efficiency of smart-home systems and appliances, including the management and the reduction of energy use. Additionally, as a key component of the smart-home idea, the potential integration of machine learning based on some algorithm methodologies should be explored as a way to improve power energy consumption system and control. The models used to identify patterns for smart-home and variations in energy consumption. This study's conclusions are acquired by analyzing case studies about energy-consumption forecast. Detection Change (of used and generation) for all appliances foresees excessive energy use and stops when a rise in usage is detected. Predict Future Energy uses meteorological data and maximizes the supply of energy to forecast future energy generation and use. Finally, five machine-learning algorithms, including the linear regression (LR), gradient boosting regression (GBoostR), decision tree regression (DTR), stochastic gradient descent regression (SGDR), and Bayesian ridge regression (BRR), measure the mean absolute error (MAE), mean squared error (MSE), root mean absolute error (RMAE), and root mean squared percentage error (RMSPE) in order to determine how well the models perform.

Keywords: Weather Information (WI), Smart Home Energy Management System (SHEMS), and Machine Learning (ML).

1. INTRODUCTION

The dependability, stability, and robustness issues facing the power grid are not adequately addressed by the traditional power system [1]. Therefore, new infrastructure is required to effectively address these concerns and lessen the strain on the environment [2]. The increase in demand for residential energy is a result of population growth, technological advancements, and heavy load usage. Additionally, the reckless attitude of users in the residential sector is to blame for this increase in energy usage [3]. The residential sector uses more than 45% of the world's energy, while careless and inefficient management practices waste millions of dollars [4]. To address the rising demand for residential energy and user comfort, majority of the nations are concentrating on creating new smart city infrastructures. In addition, the idea of smart houses offers significant economic, social, and environmental advantages [5]. Enhancing the sustainability, dependability, and power conservation of the home consumer is made possible through demand-side management of energy usage. It successfully addresses the dynamically shifting energy consumption pattern brought on by different consumer demands during the pandemic [6]. Energy management schemes have been built using a variety of optimization and programming techniques, including management of batteries, RES, and management [7]. Because of their accuracy, efficiency, and speed, "machine learning (ML) models are crucial for predictive modelling of production, consumption, and demand in EMS" [7], [8].

Furthermore, machine-learning models can be utilized to understand the operation of energy systems within complex human interactions [9]. The integration of information and communication technology (ICT) is ubiquitous in smart grid systems, spanning from the initial stages of energy production to the ultimate consumption by end users [10]. Of significant importance, as a fundamental constituent of the SG, they possess the capability to automatically balance the processes of production, consumption, and distribution, as shown in Fig. 1 [10].

 *Corresponding author: mus20c1007@uoanbar.edu.iq

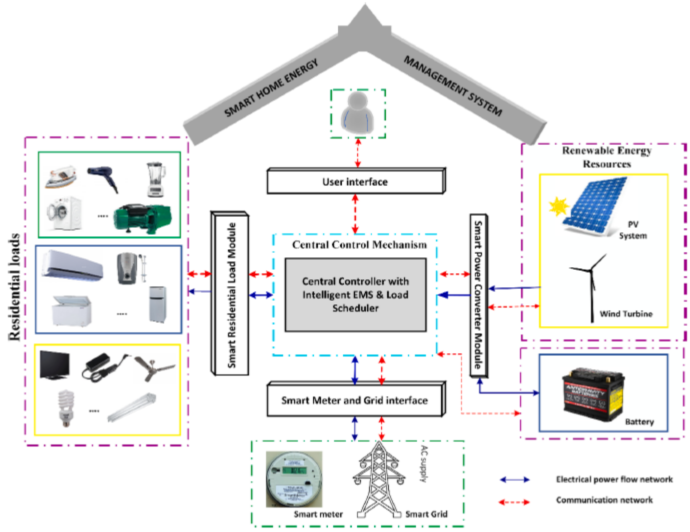

FIGURE 1. Home energy management system architecture

A smart controller with a "user interface, schedulable appliances, a smart meter, distributed energy resources (DERs), a communication network, and a hybrid energy storage system (HESS) are the main components of a smart home energy management system (SHEMS)" [11], [12]. Users now have the opportunity to more effectively and efficiently control their home energy use, thanks to machine-learning techniques in smart homes. These methods assisted us in identifying the optimal energy consumption pattern to lower electricity expenses. This paper demonstrates that home energy management (HEM) is a challenging endeavor since it requires smart appliances to be optimally predicted to energy demand. Additionally, ML models can be employed to comprehend how energy systems operate in intricate human relationships. This research applies five machine-learning algorithms to identify the most efficient pattern of energy usage in order to save power bills.

2. Problem Statement

The problems related to intelligent energy management based on machine learning are the following:

1. One of the most important problems in this field is how design the load balance management system to control the electricity consumption to reduce the costs of the energy and to distribute energy over the daytime.
2. Proposed algorithm depends on the machine learning to predict of energy consumption to ensure effective and robust to evaluated model.

3. The Objectives

The objective is related to intelligent energy management using machine learning. Load balancing, cost reduction, and energy distribution are indeed crucial aspects in this field. Designing an effective load balance management system that optimizes electricity consumption can have numerous benefits, including cost savings, reduced environmental impact, and efficient resource utilization. Machine-learning algorithms can play a significant role in addressing these challenges. This paper provides an expanded explanation of the issues and the proposed solution:

1. In summary, the main objective of this research is to design an energy management system to control and enhance the performance of the use of the electrical home energy supplied from a different supplier in a distributed network with the aim of minimizing the cost of the consumption energy based on the used daily energy from each type of energy.

The sub-objectives of the research are the following:

a. Design an energy management system based on machine-learning approach to manage and control the use of different types of energy sources to minimize the cost of the power consumed.
b. Deal with the distributed network in a smart home to improve the daily-hour use of different energy types.

4. RELATED WORK

Regression analysis serves to establish a correlation among two or more categories of power variables for the purpose of predicting power data. Electricity consumption patterns can be predicted by utilizing past anomalous observations. Several advanced techniques have been developed for the evaluation of the energy usage of smart buildings. This section presents a comprehensive overview of the latest machine-learning models utilized for the purpose of identifying anomalies. In reference [13], an alarm system for energy consumption in the country was meticulously crafted by the authors of the present study. The model under consideration was employed to predict the available direction of energy consumption by utilizing past data.

The primary focus of the author in reference [14] pertains to the growth of the smart city, covering its architecture, characteristics, and other urban features that contribute to its intelligence. Furthermore, the author explains on a variety of tools and techniques employed in the implementation of smart cities. Moreover, the authors of [2] have shown considerable interest in the urban Internet of Things (IoT) system, which has been designed to help implement the concept of a smart city. The primary objective of this system is to leverage advanced communication technology to deliver value-added services to both the city's administration and its residents. The authors proposed a method for detecting specific anomalies in energy consumption that can be applied to smart buildings, as reported in their publication [9]. The authors of reference [15] proposed an unsupervised dynamic threshold technique in detecting anomalies in time-series data related to temperature. Wang et al. [16] proposed an improved long short-term memory (LSTM) model for forecasting energy consumption.

The authors give a summary of technical solutions, protocols, architecture, and supporting developments. Numerous methods for EMS prediction based on ML are shown in table 1 above.

Table 1. Overview of techniques that used of prediction of EMSS.

Author's	Feature	Methods	Type Learning
Yassine et al. [14]	Intelligent detection of inappropriate power consumption	Knearest neighbors (KNN)	Supervisor learning
Khan et al. [11]	real versus expected consumption differences	Neural networks (NN)-ARIMA	Supervisor learning
Buzau et al.[7]	Data about historical and asynchronous power	MLP and DNN	Supervisor learning
Cui & Wang [13]	Time series for power consumption	Gaussian distribution and Polynomial regression and	Supervisor learning
Mentel et al [4]	Temperature and power usage	difference between measurements and simulations	Supervised
This Work	Prediction of electricity production and consumption based on weather data.	DTR, SGDR, BRR, LR and GBoostR	Semi-supervised

Based on the literature, Table 1 demonstrates that machine-learning methods are not well explored despite its potential in energy consumption detection. The objective of this study is to develop the most notable model to control energy consumption by using machine learning for efficient consumption detection.

5. DATASET DESECRPTION

The smart-home datasets linked to weather variables that were utilized for testing and validation are described in this paper. Information on the weather as well as the energy usage of home appliances was gathered. The smart meters are used to collect data on various weather-related attributes such as house energy consumption, dishwasher, furnace, home office, fridge, wine cellar, garage door, kitchen, barn, well, microwave, living room, and solar-power generation that are measured in kilowatts (kw).

The dataset has the following dimensions: 32 columns and 503,911 rows. Using this dataset, it is possible to investigate the relationship between appliance of time and energy consumption, identify instances of aberrant appliance use, and explain the connection between meteorological data and solar energy production. The dataset includes local meteorological information as well as readings for household appliances from a smart meter for a one-minute period. The dataset's complete features are listed in Table 2, along with a brief description of each characteristic. In order to reduce the problem of missing values, the actual weather data is often categorized into daily intervals, as the variability of weather variables tend to be relatively low over short timeframes. Table 3, on the other hand, presents a collection of data rows that were selected as representative samples for time series analysis, with the absence of any missing values. The dataset comprises 500,000 rows, each of which represents the status of a residence at a specific minute.

6. METHODOLOGY

Fig. 2 is the proposed model in predicting the energy consumption in an EMS home based on ML methods.

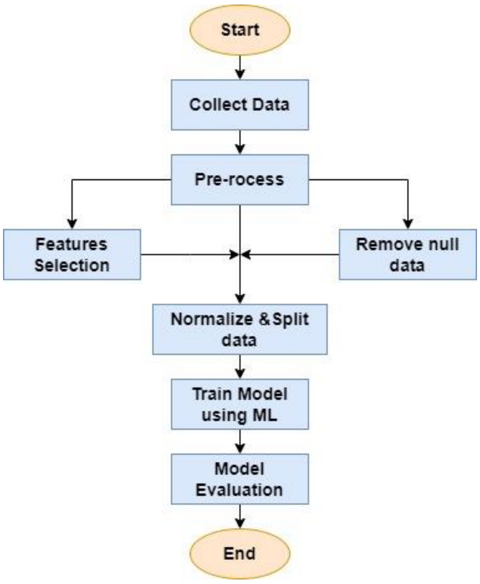

FIGURE 2. Proposed flowchart model of Machine Learning

The proposed model acquires data relating to energy consumption through sensors and smart devices. The sensors also provide environmental data such as temperature, pressure, and wind speed, while columns contain null values. The null values are identified with the zeros and blanks. The dataset must handle null values differently. Machine-learning models cannot use null values in the dataset when they appear in the dataset. The datasets are processed by removing outliers, the imputation of missing values, and the reduction of variables. After preprocessing, the datasets are ready for predictive modeling at the next level.

6.1 MACHINE LEARNING

It is extremely difficult to process and analyze data from sources like weather and energy monitors using only human labor. Methods of machine learning are employed to uncover buried data. Machine-learning methods for energy applications like the linear regression (LR) can optimize the energy utilization to improve human comfort. In this paper, ML methods were used to predict energy consumption, including LR gradient boosting regression (GBoostR), decision tree regression (DTR), stochastic gradient descent regression (SGDR), and Bayesian ridge regression (BRR).

6.2 PREPROCESSING MODEL

Data preprocessing is also carried out in this work in addition to model training and validation, as shown in the flowchart below.

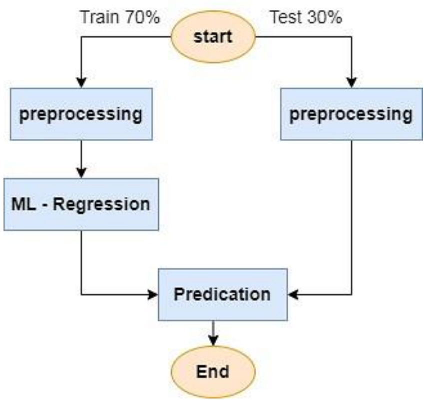

FIGURE 3. Proposed system of prediction using ML-Regression

The preprocessing phase involves the technique of attribution, which involves replacing missing data with a valid observation within the same attribute. The process to train the models that have been suggested is of significant importance as it serves to facilitate the detection of changes and enables the prediction of probable energy consumption. The effectiveness of predictive techniques and the instructional programmers is dependent on the quantity of available data. In order to minimize prediction errors, it is essential to utilize the data samples that were employed during the machine learning or training phases. The partitioning of the dataset into two distinct subsets was performed with the intention of facilitating the training and evaluation of the models. The data stated above is first employed for the purpose of training all models, thereby improving their effectiveness. The subsequent dataset is frequently denoted as the evaluation dataset in academic literature. The validation dataset is separate from the training dataset and is utilized with the objective of evaluating the efficacy of ML models that have been trained using the latter. This document presents kilowatt data that was gathered over a period of 350 days from a smart meters that were connected to a household appliance. According to Table 4, the proportion of the training set is 70% of the total data, whereas the testing set accounts for 30%. The dataset, initially measuring (503,910, 25), underwent partitioning into distinct training and testing sets, with a division of 352,737–151,173 observations correspondingly.

Table 2. The specifics of using the dataset for testing and training.

Dataset size	Training	Testing
503,910	352,737	151,173

7. RESULT AND DISCUSSION

This paper includes two case study analyses: one that looks at change detection and the other at energy usage prediction. Change Detection foresees excessive energy use and stops a rise in usage charges. Predict Future Energy use uses meteorological data and maximizes the supply of energy to forecast future energy generation and use. Tables 3, 4, 5, and 6 measure mean absolute error test (MAET), mean squared error test (MSET), root mean absolute error test (RMAET), root mean squared percentage error (RMSPE), the five algorithm of machine learning, to prove the effectivity and efficiency of the system model. .

Table 3. Analyze the results using MAET.

Method	LR	GBoostR	DTR	SGDR	BRR
MAET	0.04844	0.09612	0.09611	0.09754	0.09754

Table 4. Analyze the results using MSET.

Method	LR	GBoostR	DTR	SGDR	BRR
MSET	0.00891	0.02764	0.02764	0.02765	0.00891

Table 5. Analyze the results using RMAET.

Method	LR	GBoostR	DTR	SGDR	BRR
RMAET	0.22010	0.31002	0.31002	0.31232	0.22010

Table 6. Analyze the results using RMSPE.

Method	LR	GBoostR	DTR	SGDR	BRR
RMSPE	0.09441	0.16626	0.16626	0.16627	0.09442

The results show that BRR models outperform the other four algorithm ML models: LR, DTR, GBoostR, and SGDR. According to the BRR model, the most accurate forecast for total energy consumption was 0.04839, MSET 0.00891, RMAET 0.21999, and RMSPE 0.09442. Compared to the LR, DTR, GBoostR, and SGDR models, the Prophet model provides the best estimate of the energy consumption in the wine cellar. The four Prophets and models perform differently because they can use time and weather information as features regression. An analysis of the proposed method is performed in this work in comparison with state-of-the-art methods.

On the other hand, Fig. 1, 2, 3, and 4 depict the comparison between the proposed model and established state-of-the-art method.

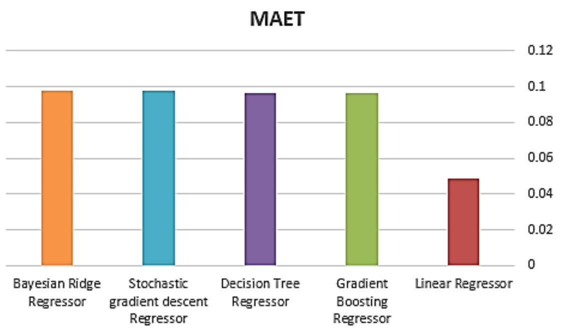

FIGURE 4. The result analysis using five ML algorithms to find MAET

Fig. 4 shows the following results: LR 0.04844 kw, DTR 0.02764 kw, GBoostR 0.09612 kw, and SGDR 0.09754 kw and BRR 0.09754 kw. According to the model prediction, the most accurate forecast for the total prediction of energy consumption was LR, that the Prophet model provides the best estimate of wine cellar energy consumption using the MAET.

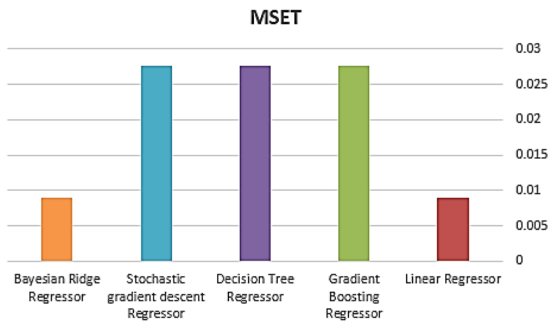

FIGURE 5. The result analysis using five ML algorithms to find MSET

Fig. 5 shows the following results: LR 0.00891 kw, DTR 0.09611 kw, GBoostR 0.02764 kw, and SGDR 0.02765 kw and BRR 0.00891 kw. According to the model prediction, the most accurate forecast for the total prediction of the energy consumption was LR and BRR, that the Prophet model provides the best estimate of the wine cellar energy consumption using MSET.

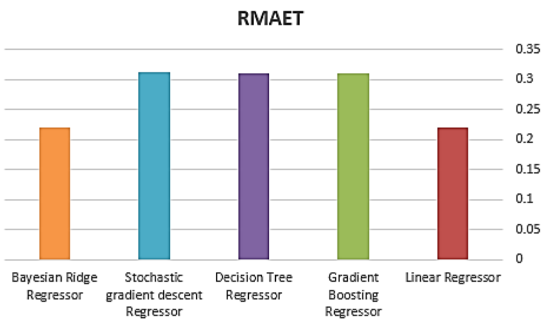

FIGURE 6. The result analysis using five ML algorithm to find RMAET

Fig. 6 shows the following results: LR 0.22010 kw, DTR 0.31002 kw, GBoostR 0.31002 kw, and SGDR 0.31232 kw and BRR 0.22010 kw. According to the model prediction, the most accurate forecast for the total prediction of energy consumption was the Prophet model as it provides the best estimate of the energy consumption in the wine cellar using RMAET.

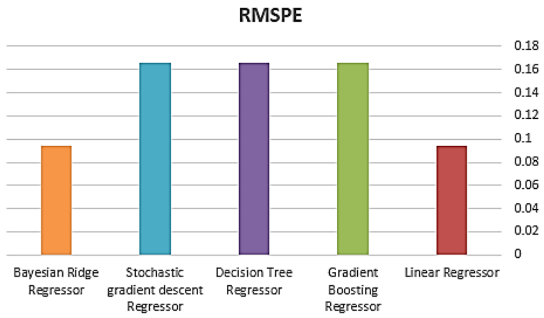

FIGURE 7. The result analysis using five ML algorithm to find RMSPE

Finally, Fig. 7 shows the following results: LR 0.09441 kw, DTR 0.16626 kw, GBoostR 0.16626 kw, and SGDR 0.16627 kw and BRR 0.09442 kw. According to this model prediction, the most accurate forecast for the total prediction of energy consumption was the Prophet model as it provides the best estimate for the energy consumption of the wine cellar using RMSPE.

A comparative analysis has been undertaken to evaluate various ML models, revealing the necessity of optimization for improved performance. The study recommends the utilization of deep-learning models such as RNN or LSTM. The evaluation of the system model's efficacy and efficiency is conducted through the utilization of various statistical measures, including the MAET, MSET, RMAET, and RMSPE. Based on the experimental results obtained from the test dataset, it can be observed that the BRR model exhibits superior performance models.

8. CONCLUSION

Smart-home systems have the potential to yield significant advantages in the context of smart city implementations, particularly in the areas of energy management and conservation. This can enable governmental entities to achieve substantial cost savings. The present investigation scrutinizes the various energy consumption methodologies employed in smart homes and their potential for prognostication. The study establishes that such prognostication can facilitate the optimization of energy utilization, thereby enhancing energy efficiency. This study has unveiled distinct patterns in the energy utilization of household appliances. The objective of this study is to display the integration of machine-learning methodologies with extensive intelligent systems to accomplish predictive maintenance of power systems through the utilization of five distinct machine-learning algorithms. Selected models are employed to detect early-stage changes in energy consumption trends. A comparison of four algorithmic ML models, LR, DTR, GBoostR, and SGDR, shows that BRR models predict future energy consumption better than LR, DTR, GBoostR, and SGDR models. BRR's model predicted total energy consumption at 0.04839, MSET at 0.00891, RMAET at 0.21999, and RMSPE at 0.9442. In comparison to LR, DTR, GBoostR, and SGDR models, the Prophet model provides the most accurate estimate of wine cellar energy consumption.

Funding

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

None

REFERENCES

[1] A. Zainab, S. S. Refaat, and O. Bouhali, "Ensemble-based spam detection in smart home IoT devices time series data using machine learning techniques," Information vol. 11, no. 7, p. 344, 2020, doi: https://doi.org/10.3390/info11070344.

[2] A. L. Bukar, C. W. Tan, L. K. Yiew, R. Ayop, W.-S. Tan, "A rule-based energy management scheme for long-term optimal capacity planning of grid-independent microgrid optimized by multi-objective grasshopper optimization algorithm," *Energy Convers Manage* vol. 221, no. 113161, 2020.

[3] A. S. Farsangi, S. Hadayeghparast, M. Mehdinejad, H. Shayanfar, "A novel stochastic energy management of a microgrid with various types of distributed energy resources in presence of demand response programs," *Energy*, vol. 160, pp. 257–74, 2018.

[4] G. Mentel, J. Nakonieczny, "Weather risk and energy consumption in Poland," in *Energy Transformation Towards Sustainability*. Elsevier, 2020, pp. 113–132. [Online]. Available: https://doi.org/10.1016/b978-0-12-817688-7.00006-9.

[5] M. Han *et al*., "A review of reinforcement learning methodologies for controlling occupant comfort in buildings," *Sustainable Cities and Society*, vol. 51, p. 101748, Nov. 2019, doi: 10.1016/j.scs.2019.101748.

[6] M. Munir, S. A. Siddiqui, M. A. Chattha, A. Dengel, S. Ahmed, FuseAD. "Unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models," *Sensors*, vol. 19, no. 11, p. 2451, 2019, doi: https://doi.org/10.3390/s19112451.

[7] M.-M. Buzau, J. Tejedor-Aguilera, P. Cruz-Romero, A. Gomez-Exposito. "Hybrid deep neural networks for detection of nontechnical losses in electricity smart meters," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1254–1263, 2020, doi: https://doi.org/10.1109/tpwrs.2019.2943115.

[8] M. E. Mondejar, *et al*., "Digitalization to achieve sustainable development goals: Steps towards a Smart Green Planet," *Sci Total Environ*, vol. 794, no. 148539, 2021.

[9] A. Elrayyah, S. Bayhan. "Multi-channel-based microgrid for reliable operation and load sharing." *Energies*, vol. 12, no. 11, p. 2070, 2019.

[10] M. A. Hameed, E. T. Yassen, W. M. Jasim. "A survey of intelligent energy management based on learning heuristic," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 1242–1246, 2023, doi: https://doi: 10.11591/ijeecs.v30.i2.pp1242-1249.

[11] A. Khan, N. Javaid, A. Ahmad, M. Akbar, Z. A. Khan, M. Ilahi. "A priority-induced demand side management system to mitigate rebound peaks using multiple knapsack." *J Ambient Intell Hum Comput*, vol. 10, no. 4, pp. 1655–78, 2019.

[12] O. A. Sianaki, M. A. Masoum, V. Potdar. "A decision support algorithm for assessing the engagement of a demand response program in the industrial sector of the smart grid." *Comput Ind Eng* vol. 115, pp. 123–37, 2018.

[13] W. Cui, H. Wang. "Anomaly detection and visualization of school electricity consumption data," in the *IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 2017, doi: https://doi.org/10.1109/icbda.2017.8078707.

[14] Y. Himeur, A. Alsalemi, F. Bensaali, A. Amira. "Smart power consumption abnormality detection in buildings using micromoments and improved k-nearest neighbors," *Int. J. Intell. Syst*., vol. 36, no. 6, pp. 2865–2894, 2021, doi: https://doi.org/10.1002/int.22404.

[15] Y. Liu, Z. Pang, M. Karlsson, S. Gong. "Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control, Build." *Environ.*, vol. 183, p. 107212, 2020, doi: https://doi.org/10.1016/j.buildenv.2020.107212.

[16] Z. Wang and T. Hong. "Reinforcement learning for building controls: The opportunities and challenges," *Applied Energy*, vol. 269, p. 115036, Jul. 2020, doi: 10.1016/j.apenergy.2020.115036.