Journal of Mines, Metals and Fuels, *73(6): 1663-1674; 2025. DOI: 10.18311/jmmf/2025/46322*

Print ISSN : 0022-2755

Journal of Mines, Metals and Fuels

Contents available at: *www.informaticsjournals.co.in/index.php/jmmf*

Hybrid Deep Learning and Internet of Things based Building Energy Management System: A Case Study on Energy Conservation in a University

Deepshikha Shrivastava[1,2]* and Prerna Goswami[2]

[1]*Pimpri Chinchwad College of Engineering and Research, Pune - 411044, Maharashtra, India; shrivastavadeepshikha@gmail.com*

[2]*Institute of Chemical Technology, Mumbai - 400019, Maharashtra, India*

Abstract

Building energy consumption monitoring has huge untapped potential, especially in a developing economy like India. Current solutions have not focused on a cost-effective, small-scale, accurate and open-source data-driven methodology. Further, several obstacles hinder research in this domain, such as security and privacy concerns, high investment and uncertainty of research outcomes. Therefore, we have implemented an automated hybrid deep learning and Internet of Things (DL-IoT) based Building Energy Management System (BEMS) for energy conservation. The DL-IoT integrates a hybrid of DL and fuzzy logic to handle the uncertainty and noise in electrical features. Our DL- IoT regression model achieves low mean absolute error and mean squared error as well as a coefficient of determination of 0.99 out of sample. Twenty-seven input electricity consumption features have been selected for extraction from the raw data for training the model. Experimental results indicated a linear relationship between the features and the energy consumption. The proposed model successfully predicted certain features that could lead to energy savings, viz., Power Factor and Power in Y Phase. The model estimated that a one-unit increase in Power in Y Phase and Power Factor will lead to a reduction in the energy consumption. The findings of the experiment indicated that the model captured the variability of the data better than other models. The results demonstrated the superiority of the proposed model over other mainstream existing models. Through the results of this paper, a more efficient energy data management and consumption plan can be established.

Keywords: *Building Energy Management System, Electricity Disaggregation, Energy Consumption Monitoring, Hybrid Deep Learning, Internet of Things*

1.0 Introduction

Buildings are responsible for 33% of final energy consumption and 40% of direct and indirect CO_2 emissions globally[1]. Hence, electricity consumption monitoring or in other words, electricity disaggregation, is an important aspect in the field of energy management and conservation[2]. Studies have been conducted to explore the use of Arduino Uno as a platform for electricity consumption monitoring[3]. These have reported a surge in interest in using data science or Machine Learning (ML) and DL tools like supervised or unsupervised learning methods[4] for tracking consumption patterns of Micro, Small, and Medium Enterprises (MSME). The advantage of electricity monitoring systems is that such systems are designed to provide real-time information about energy consumption in order to improve energy efficiency and reduce costs. The majority of implementations are using a low-cost Arduino-based electricity monitoring system that uses energy sensors to measure the current and voltage of electrical appliances[5]. The data collected by the sensors is then processed by the Arduino Uno,

*Author for correspondence

which calculates the power consumption and sends the information to a central monitoring unit[6]. However, the major hindrance is the capability of Arduino Uno in terms of shelf life, tolerance for heat, lack of wireless capabilities and low memory and processing power. All such factors hinder the acceptance at the MSME level[7].

The majority of research in this domain concentrates on preventing electricity theft, and a few notable exceptions concentrate on energy consumption optimization on Wireless Sensor Networks (WSN)[8]. A multitude of methodologies are employed for predictive purposes across various domains. These methodologies are typically appraised based on two principal criteria: the ease of implementation and the precision of their outcomes. They broadly fall into two categories: mathematical and artificial intelligence-based methods. Mathematical approaches encompass techniques such as regression analysis and exponential smoothing[9], while the domain of artificial intelligence introduces methods like ANNs, knowledge-based expert systems, and fuzzy logic[10]. One strategy to enhance the predictive accuracy is the amalgamation of multiple methodologies, thereby leveraging the strengths of one to offset the weaknesses of another[11]. Integral to the effective application of these predictive methods is the availability of historical data, which provides the necessary framework for forecasting future trends[4]. However, certain factors hinder the research in this field. These factors are the time-consuming nature of data collection at ground level, lack of access to enterprises' infrastructure due to privacy and security concerns, initial hardware investment, uncertainty of research outcomes, and continuous monitoring. Due to such factors, research groups are hesitant to participate or undertake research in the domain of electricity disaggregation.

Our work aims to overcome the drawbacks mentioned above. Initially, we have collected data from the Pimpri Chinchwad College of Engineering and Research (PCCoER), Pune, India, after surveillance of two years. The data encompasses various characteristics not previously integrated into predictive analytics at this scale[12,,13]. The data was collected continuously, 24/7 and at an interval of one minute. Several parameters related to electricity consumption were recorded in a cloud database. These parameters are identified through previous literature, *viz*., energy, frequency, power, apparent power, current, and volt-ampere hours. As far as we know, the current work will be the longest exercise of data collection performed in a university setting in India. Further, we analyzed the data to identify patterns and trends in electricity consumption throughout the day. The goal was to understand the relationship between the electricity parameters and to validate the theoretical relationship between these parameters. Further, preprocessing identified anomalies in the data collection and applied data transformations to resolve them. Subsequently, we developed a hybrid DL-Iot-based building energy management system for energy conservation to predict electricity consumption based on time of day, weather conditions, and other relevant factors. This model was needed to understand the independent variables for calculating energy consumption.

In summary, the following are the main contributions of the paper:

- Collect and preprocess electricity consumption data from a university campus over two years.
- Analyze the data to identify patterns and trends in electricity consumption throughout the day.
- Develop a Hybrid DL-Iot-based building energy management system for energy conservation
- Implement the optimized device operation strategies based on the machine learning models.
- Evaluate the effectiveness of the optimized strategies in reducing energy consumption and costs.

After an extensive literature survey, to the best of our knowledge, our work is one of the first efforts to practically study electricity consumption through an IoT setup[7,3]. The experiments and data collection have been conducted through open-source tools and low-cost hardware. Therefore, the research will enable deployment on the shop floor for those interested to carry it forward. Finally, though in Western and south- east Asian countries, there have been efforts to map energy consumption. Our work is different from such efforts as the technological landscape is ever changing, and so an updated study is needed in this domain.

The remaining sections of the paper are as follows: Section 2 explains the previous literature and an overview of the state of the art in the domain of Iot-based

automated electricity consumption monitoring. Further, Section 3 describes the data collection, the mathematical model, data preprocessing, feature description and experimental setup. Finally, the results of the experiments are discussed in Section 4, followed by concluding remarks and possibilities for furthering the current research in Section 5.

2.0 Related Works

The review sheds light on the merits and demerits of existing works in the domain of electricity consumption monitoring. The papers reviewed are from the field of non-intrusive and intrusive consumption monitoring at building, grid and cloud environment levels. Both non-intrusive and intrusive methods have their pros and cons. The intrusive load monitoring techniques rely on multiple dedicated sensors installed on each appliance to obtain its electric consumption information. Unlike intrusive load monitoring techniques, Non-Intrusive Load Monitoring (NILM) only requires measurement at the main power service entry to realize power load monitoring and identification efficiently[13]. The main advantage of intrusive systems is the accuracy of measuring the consumption at the device level. Whereas, NILM is proven to be more inexpensive to install and maintain, and more acceptable to consumers in terms of privacy and convenience.

With an aim to detect non-technical losses using meter manipulation/malfunctioning, Kim *et al.*[14] investigated the application of Arduino Uno for monitoring electricity consumption in residential buildings. They developed a smart metering system using Arduino Uno and tested it in a real residential environment. The system was able to accurately measure and monitor the electricity consumption of different appliances in the house. However, the implementation requires a heavy investment for the enterprise due to its multi-tier architecture. A similar attempt at load monitoring for prevention of electricity theft was attempted by Ibrahem *et al*[15]. The authors explored a similar approach by using Arduino Uno to monitor electricity consumption in commercial buildings. They developed a system that integrated Arduino Uno with energy meters to monitor and analyze the electricity usage of different devices in a commercial building. The authors proposed a multi-tier architecture, which raises the cost for implementation and is unsuitable for small or medium-sized industries. Both groups[15,16] have advocated the importance of data science techniques like ML, DL for the detection of trends in temporal data. Q Liu *et al.*[4].

Although unrelated to energy consumption monitoring in MSME, Biçici[16] has introduced an application named Cloud Monitor to reduce energy consumption in the cloud environment. The author argued for a linear programming optimization approach towards the task of cloud load optimization on cloud computing platforms to reduce energy consumption. The only drawback was that the results were not calculated in a real-world environment but in a simulated one. Hannan *et al.*[17] conducted a thorough review of Building Energy Management Systems (BEMS) and highlighted the challenges in implementing a BEMS to avoid suffering from inadequacies such as data loss and network problems. The authors argued that BEMS should be energy efficient, low cost, simple to install and tamper-proof. Their suggestions have been an important input in deciding the architecture of the current research.

Another research interest is to understand if BEMS impact energy consumption; researchers are of differing opinions on this topic. While previous research has shown that feedback can be useful for short-term energy saving, the work of Alahmad *et al.*[18] indicates otherwise. The authors conclude after conducting a survey in the US that feedback from BEMS did not translate significantly into energy savings. Martínez-Sánchez *et al.*[19] demonstrate how energy savings lead to profitability for an MSME and at the same time impact the Sustainable Development Goals (SDGs). With the same intention, we too selected a target institution, PCCoER, where energy monitoring was non-existent. The institute would pass on the energy consumption charges to the learners in the form of tuition fees. Thus, any energy savings would directly benefit the hundreds of learners at the institution in the form of financial savings. Our approach is data-driven, similar to the work of Kang *et al.*[20], with a key difference that while we have focused on BEMS, the authors have focused on a particular machine. The work of Zhuang *et al.*[21] was also influential as the authors have proposed a methodology of data collection, preprocessing and ML pipeline for electricity consumption forecasting. Zhang

et al.[22] designed a cloud-fog cooperation scheduling algorithm to reduce energy consumption. The authors proposed a task offloading algorithm to reduce energy consumption using energy and delay functions with queue theory. The strategy is suitable for a cloud infrastructure environment.

We investigated approaches and technology proposed in the field of energy grid monitoring, such as Atalik[23,24]. However, the methods could not apply to BEMS as these were designed for larger infrastructures. Yet, the power quality parameters mentioned in these studies were useful to us in feature engineering. Since, our work focuses on only BEMS other important aspects related to security and malware attacks have been left out and readers interested to carry forward our research by including cyber-attack proofing the system can refer to the work of Jiang *et al.*[25]. Similarly, those interested in reading case studies on electricity consumption monitoring by implementing solutions for smart grids can read the case study of Janthong *et al*[26].

These studies collectively demonstrate that the role of ML/DL has been important for data analysis and inference. The limitations of the current methods can be summarized as,

- High cost and lack of open-source techniques,
- Limited suitability for small-scale buildings,
- Lack of data-driven methodology.

2.1 Preliminaries: Basics of ML/DL Models

Load forecasting predicts future electrical demand across various time scales. It includes short-term (hourly to weekly), medium-term (weekly to monthly), and long-term (annual) forecasting. Accurate predictions are crucial for unit commitment, economic dispatch, capacity management, and strategic planning. Challenges include variable complexity and weather impacts. Predictive methods range from mathematical, like regression analysis, to artificial intelligence-based.

Although a deep review of ML/DL is not our objective in this section, a general idea of the fundamentals of the models used in this paper is mentioned to enlighten the reader. Linear regression is one of the simplest models, fitting a linear equation to observed data. It is computationally efficient but requires a linear relationship between variables. Ridge and Lasso regression are regularized versions of linear regression. Ridge regression adds a penalty on the size of coefficients, reducing overfitting. Lasso regression, in addition to shrinking coefficients, can set some coefficients to zero, performing feature selection. Polynomial regression extends linear regression by incorporating polynomial terms, allowing it to model non-linear relationships. However, it can lead to overfitting if the polynomial degree is too high. Support Vector Regression (SVR) uses support vector machine principles to predict continuous values. SVR fits the best line within a threshold value, known as the epsilon, around which the data points should lie. Decision Trees and Random Forests are non-linear models. Decision Trees partition the data into subsets to make predictions. Random Forests, an ensemble method, combines multiple decision trees to improve accuracy and reduce overfitting. Gradient Boosting Machines (GBM) and Extreme Gradient Boosting (XGBoost) enhance predictive performance by iteratively correcting previous errors. Artificial Neural Networks (ANN) and their variations, such as deep learning models, can capture complex patterns in data. They are powerful but computationally intensive.

2.2 Hybrid DL-IoT-based BEMS: Materials and Methods

2.2.1 Hybrid DL-IoT-based BEMS: Proposed Architecture

Data collection set up for electricity consumption monitoring (Figures 1-8) required the installation of a

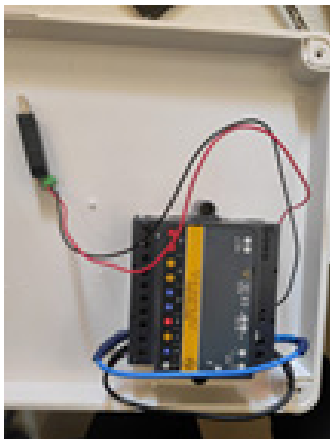


Figure 1. Multifunction meter.



Figure 2. Multifunction meter.



Figure 3. Raspberry Pi.

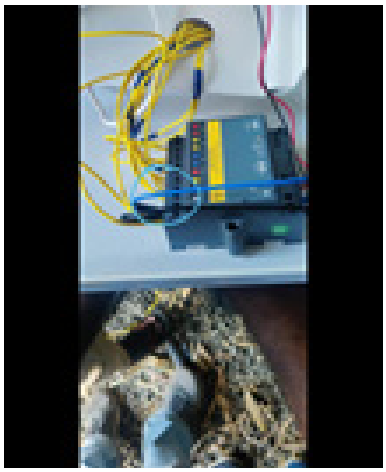


Figure 4. Multifunction meter.



Figure 5. Current transformer to the mainline.



Figure 6. USB connector and Wifi modem.

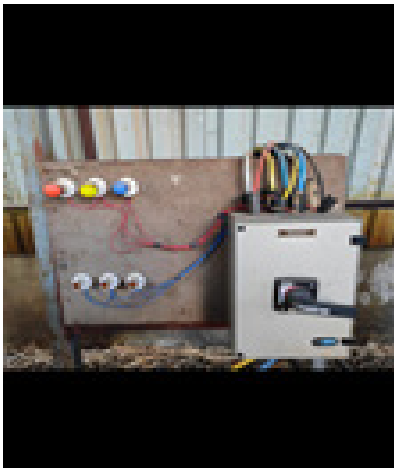


Figure 7. Power Discom main supply line.

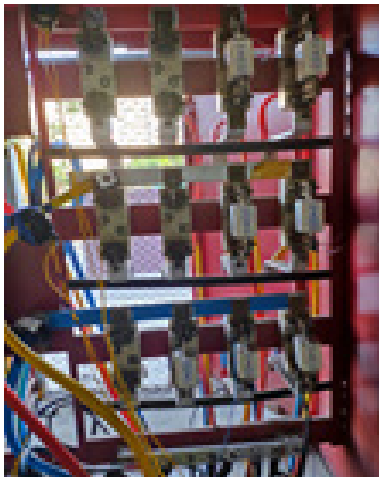


Figure 8. Current transformer.

multifunction meter, Raspberry Pi 3b, USB connector, GSM (Wifi) modem and current transformer. The USB connector (Figure 6) combines data from the meter and sends it to the Raspberry Pi through the USB.

The Raspberry Pi (Figure 3) is a microcontroller to store, fetch and upload data on Dropbox (cloud storage) through an internet connection. The Wi-Fi modem provides internet to the Raspberry Pi through a SIM card. The current transformer reduces the value of current for a specific voltage to simplify power calculation. The heat generated in the main supply line caused the battery powering the Raspberry Pi to swell and reduce its lifespan. Thrice during data collection, the battery's power was exhausted, and data collection was interrupted. A total of 50 min of data was lost due to these interruptions.

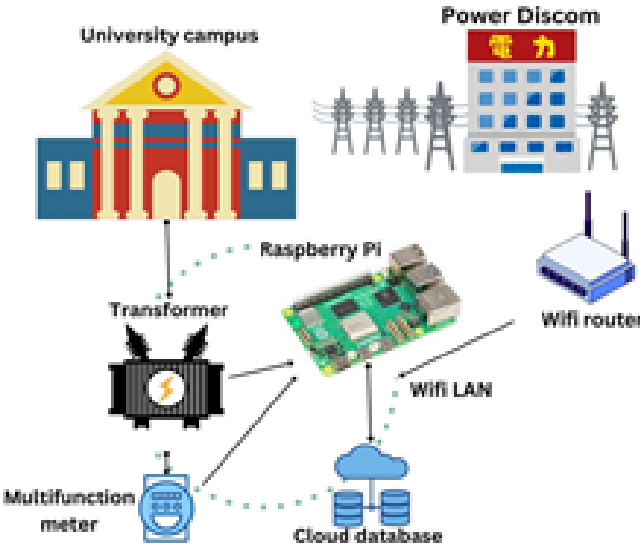


Figure 9. Proposed architecture of energy consumption monitoring.

Further, the data analysis and ML/DL modelling were performed using Visual Studio Code on a laptop having Intel(R) Core(TM) i5-1035G1 CPU @ 1.00GHz with 1.19 GHz, 64-bit operating system, x64-based processor, 8.00 GB RAM and Windows 11 Home Single Language OS. The overview of the architecture is illustrated in Figure 9.

2.2.2 *Hybrid DL-IoT: Objective Function*

We propose a novel approach for predicting energy units by integrating a deep learning neural network with a custom fuzzy logic layer. The model comprises an input layer, a fuzzy logic layer, a hidden dense layer, and an output layer (Figure 10). The model's architecture can be formulated as follows: Let $\mathbf{x} \in R^2$ be the input feature vector.

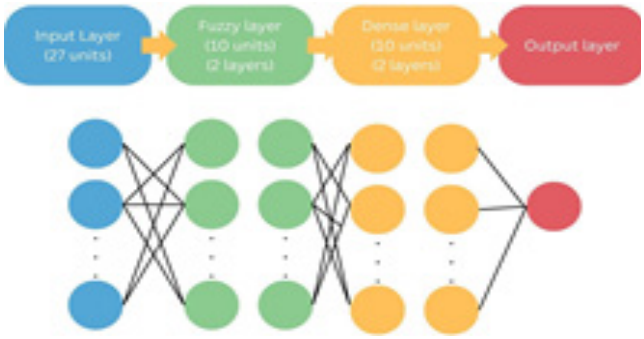


Figure 10. Hybrid DL-IoT backbone: Neuro fuzzy model.

The fuzzy logic layer transforms the input vector using a linear combination of the inputs, followed by a sigmoid activation function. This can be expressed mathematically as:

$\mathbf{z} = \sigma(\mathbf{xW})$

where:

$\mathbf{W} \in R^{2\times10}$ is the weight matrix.

σ is the sigmoid activation function defined as $\sigma(x) = 1/1+e{-}x$.

$\mathbf{z} \in R^{10}$ is the output of the fuzzy logic layer.

The hidden dense layer applies a ReLU (Rectified Linear Unit) activation function to the output of the fuzzy logic layer. It is given by:

$\mathbf{h} = \text{ReLU}(\mathbf{zV} + \mathbf{b})$

where:

$\mathbf{V} \in R^{10\times10}$ is the weight matrix.

$\mathbf{b} \in R^{10}$ is the bias vector.

ReLU(x) = max(0, x) is the ReLU activation function.

$\mathbf{h} \in R^{10}$ is the output of the hidden layer.

The final output layer predicts the energy unit value as a linear combination of the hidden layer output:

$\hat{y} = \mathbf{hu} + c$

where:

$\mathbf{u} \in R^{10}$ is the weight vector.

$c \in R$ is the bias.

$\hat{y} \in R$ is the predicted energy unit.

The model is trained using the mean squared error (MSE) as the loss function:

where:

$$\text{MSE}=\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2$$

y_i is the true value of the energy unit.

$\hat{y}_i$ is the predicted value.

n is the number of training samples.

The proposed model combines the interpretability of fuzzy logic with the robust learning capacity of deep neural networks to effectively predict energy units. This hybrid approach leverages the strengths of both techniques to achieve superior prediction performance. The code built using numpy, matplotlib, skfuzzy, and tensorflow is shared on the github repo of the project.

3.0 Experimental Work and Results

3.1 Performance Evaluation Metrics and Results

Performance of the proposed BEMS model is evaluated and compared using three common metrics: Mean Absolute Error (MAE) as calculated by Eq. (1), which is a straightforward measure of average absolute errors.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \quad (1)$$

Second is the Mean Squared Error (MSE) as calculated by Eq. (2), which penalises larger errors more strongly due to squaring.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \quad (2)$$

Coefficient of Determination (R^2) as calculated by Eq. (3), measures the proportion of variance explained by the model.

$$R^2 = 1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (3)$$

In equations (1), (2), and (3), y_i represents the true values, $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the mean of the true values. Comparing MAE, MSE, and R^2 provides a comprehensive understanding of model performance. Figure 11 illustrates the scatter plot that shows how well the predicted values (y-axis) match the actual values (x-axis). For the proposed model, the predicted points are close to the red line.

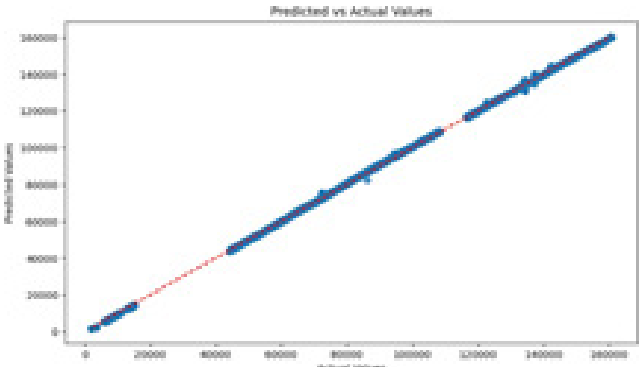


Figure 11. Hybrid DL-IoT BEMS: Predicted and actual values for the test set.

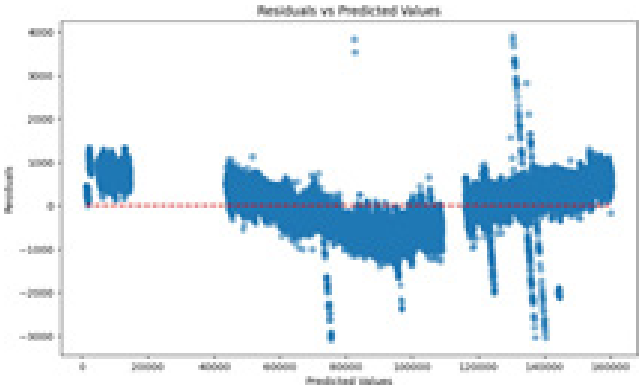


Figure 12. Hybrid DL-IoT BEMS: Residuals vs predicted values.

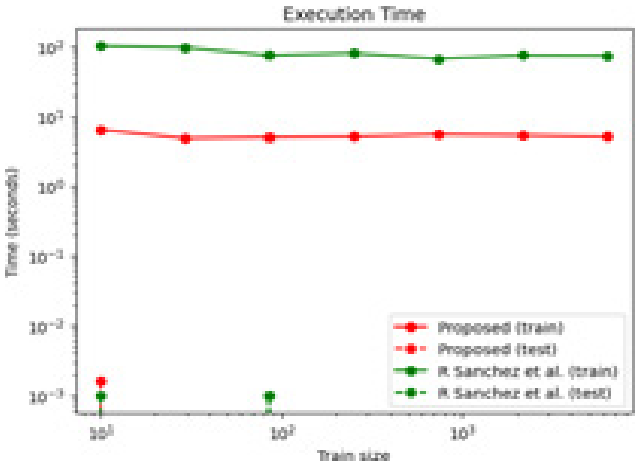


Figure 13. Hybrid DL-IoT BEMS: Time vs train size.

Figure 12 illustrates the residuals (errors) on the y-axis versus the predicted values on the x-axis. The horizontal red dashed line at $y = 0$ represents no error. For the proposed model, the residuals are randomly distributed.

Figure 13 compares the time for fitting and prediction of the proposed model and Lasso for different sizes of the training set.

Fitting R Sanchez *et al.* BEMS is faster than a proposed system for medium-sized training sets (less than a few thousand samples); however, for larger training sets proposed system's hybrid DL-Fuzzy model scales better. With regard to prediction time, the proposed model should be faster than Lasso for all sizes of the training set because of the learned sparse solution, however, this is not necessarily the case in practice because of the implementation details of the Sklearn library in which both models are implemented. Figure 14 visualises the training objective to minimise the MSE with respect to the training set size.

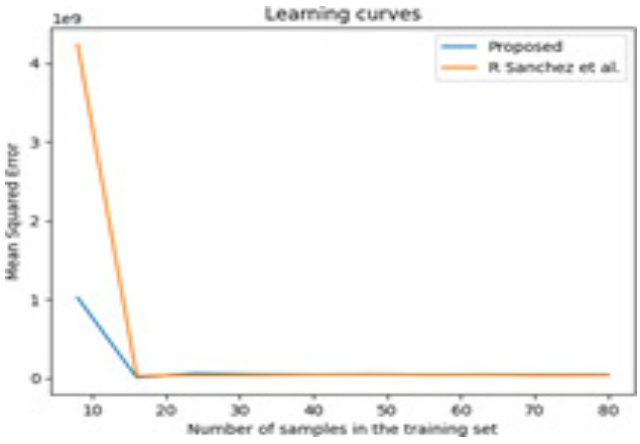


Figure 14. Hybrid DL-IoT BEMS: MSE vs samples in training set.

Both the proposed models and R Sanchez et al. are able to minimise the MSE, with the proposed model achieving its objectives with a smaller sample size of the training set. Further, as Table 1 shows, the proposed method performs well with respect to the metrics vis-à-vis other BEMS used in the literature.

Table 1. Performance metrics for different BEMS and the Hybrid DL-IoT BEMS

Model	MAE	MSE	R2
R Sanchez et al.	386.77	221053.2	0.96
J Kim et al.	401.9	234414.7	0.9
E Bicici	401.5	227856.3	0.89
Proposed method	385.7	221052.14	0.99

3.2 Feature Analysis using T-Test

Initially, the Ordinary Least Squares (OLS) model from the Python 3 statsmodels library was used to understand which feature has a statistically significant ($p \leq 0.05$) effect. Table 2 gives the results of the model.

From the above Table 3, the OLS regression model for predicting 'Energy' shows a perfect fit with an R-squared of 1.000, though this may indicate overfitting given the high F-statistic of 2.232e+08 and condition number of 1.09e+16, suggesting multi-collinearity. Most coefficients are statistically significant ($p-values \leq 0.05$), indicating strong relationships with 'Energy'. Diagnostic tests show non-normal residuals and positive autocorrelation (Durbin-Watson = 0.048). The model includes 22 predictors with 658,006 observations. While the high R^2 and low AIC/BIC suggest a good fit, further validation and addressing multicollinearity and residual issues are recommended for robustness.

3.3 Case Study at University Campus for Energy Savings

The data collection apparatus continuously monitored the following parameters of the electricity consumption

Table 2. OLS regression results

	Coefficient	Std Error	t-Statistic	p≤0.05
const	2.68e+04	369.892	72.446	0.000
Ptot	-5.223	0.754	-6.929	0.000
Pr	4.989	0.767	6.504	0.000
Py	-34.042	1.106	-30.786	0.000
Pb	23.829	2.356	10.115	0.000
PFa	-31.792	1.843	-17.250	0.000
PFr	31.988	1.976	16.192	0.000
PFy	28.112	2.203	12.762	0.000
PFb	-91.892	5.538	-16.593	0.000
St	81.890	1.646	49.751	0.000
Sr	-81.856	1.647	-49.700	0.000
Sy	-360.209	24.784	-75.301	0.000
Sb	523.955	25.157	101.606	0.000
VLa	31.158	174.358	0.179	0.858
Vry	19.863	58.120	0.342	0.733
Vyb	56.814	58.122	0.978	0.328
Vbr	32.679	58.120	0.562	0.574
Va	440.008	174.416	2.523	0.012
Vr	-208.909	58.139	-3.593	0.000
Vy	-228.410	58.140	-3.929	0.000
Vb	-274.984	58.143	-4.729	0.000
It	-19.649	0.365	-53.846	0.000
Ir	19.673	0.365	53.886	0.000
Iy	100.629	1.200	83.867	0.000
Ib	-139.951	1.209	-115.778	0.000
freq	-411.549	7.454	-55.214	0.000
VAh	0.951	1.53e-05	6.23e+04	0.000

at an interval of one minute. The starting timestamp and end timestamp for the project were 2021-12-30 17:38:00 and 2023-12-04 04:49:00, respectively. The dataset contains several key parameters that provide a comprehensive overview of electrical measurements and system performance. The Timestamp indicates the exact date and time when the data was recorded, while the Slave ID uniquely identifies the device or sensor reporting the data. Power metrics include Total Power (Pot), representing overall power consumption, and phase-specific measurements for the R (Red), Y (Yellow), and B (Blue) phases. Similarly, the Power Factor Total (PFa) and phase-specific power factors (R, Y, B) reflect the efficiency of power utilization, with an ideal value close to 1. Apparent power is captured as Total Apparent Power (St), along with phase-specific values (Sr, Sy, Sb), measured in volt-amperes. Voltage-related metrics include the Line Voltage Average (VLa), phase-to-phase voltages (Vry, Vyb, Vbr), phase-specific voltages (Vr, Vy, Vb), and the Voltage Average (Va) for individual phases. Current measurements are provided as Total Current (It) and phase-specific currents (Ir, Iy, Ib), measured in

Table 3. Metrics for evaluating OLS models prediction performance

R-squared	1.000
Adj. R-squared	1.000
F-statistic	2.232e+08
Prob (F-statistic)	0.00
Log-Likelihood	-4.9834e+06
No. Observations	658006
AIC	9.967e+06
Df Residuals	657983
BIC	9.967e+06
Df Model	22
Covariance Type	nonrobust
Omnibus	3901.131
Prob(Omnibus)	0.000
Durbin-Watson	0.048
Jarque-Bera (JB)	4190.859
Prob(JB)	0.00
Skew	-0.161

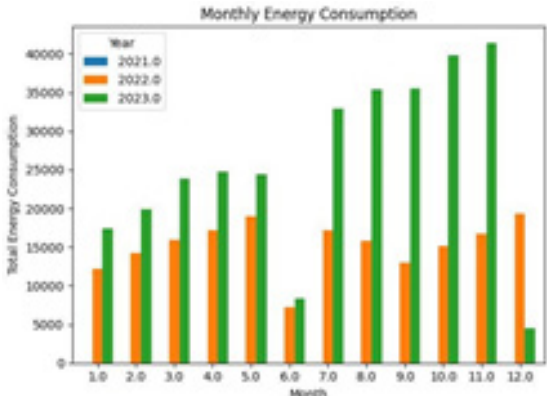


Figure 15. Year-wise monthly trends of energy consumption.

amperes. Additionally, the dataset includes the Frequency of the AC supply, typically 50 or 60 Hz, and energy consumption metrics such as Total Energy in kilowatt-hours (kWh) and Volt-Ampere-Hour (Vah), representing apparent power usage over time. These parameters collectively enable detailed analysis of the electrical system's performance and efficiency.

Figure 15 shows the consumption of electricity or the total energy consumption, generally measured in kilowatt-hours (kWh) throughout the project lifetime.

The monitoring indicates that the electricity consumption has a seasonal pattern. The consumption increases from January and tapers in May. Then again, it rises from July and tapers in November/December. This seasonality is due to the two academic semesters, *viz.*, Jul-Dec and Jan-Jun. Usually, academic activity and energy consumption increase slowly as the semester commences. Classes are conducted, and energy consumption increases as laboratories and computers are used. Usually, the electricity consumption increases after 9 am when the campus opens and tapers towards 6 pm when the campus closes.

Using the Hybrid DL-IoT BEMS, it is possible to understand the focus areas for energy savings. A partial dependence plot, as illustrated in Figure 16, is a method for the same. It can plot the relationship between input features and response, marginalizing over the values of all other input features. Results show that both Power in Y Phase (Py) and Power Factor (PFa) have an inverse relationship with energy consumption. Thus, these two parameters will have to be increased from current levels and that shall translate into energy and cost savings. From Table 1, one unit increase in Py will lead to 34-unit reduction in the energy consumption, and similarly, one unit increase in PFa will lead to 31-unit reduction in the energy consumption.

Algorithm 1: Energy Prediction using DL-IoT model

1: **Input**: Predictors X

2: **Output**: Response Y = Energy

3:

4: **procedure** PREDICTENERGY

5: **Step 1:** Split the dataset into training set (X_{train}, Y_{train}) and testing set (*Xtest, Ytest*)

6: **Step 2:** Standardise the features using Standard Scaler

7: $\mu = \text{mean}(X_{train})$

8: $\sigma = \text{std}(X_{train})$

9: Scale both X_{train} and X_{test}:

10:

11: Xtrain = $(Xtrain - \mu)/\sigma$

12: **Step 3:** Use ML model

Xtest= $(Xtest - \mu)/\sigma$

13: **Step 4:** Perform hyperparameter tuning

14: Define a hyperparameter grid

15: Use cross-validation to find the best hyperparameters

16: **Step 5:** Train the model using the best hyperparameters on X_{train}

17: **Step 6:** Evaluate the model on X_{test} to predict Y_{test}

18: **Step 7:** Compute evaluation metrics

19: **Step 8:** Output the final model and evaluation scores

20: **end procedure** if error, then return

In summary, the key learnings from the paper indicate that with features, it is possible to achieve a highly accurate model for energy consumption. The equipment needed can be regular open-source and easily available hardware. Further, open source software toolkits can be used for data-driven modelling. For training the Hybrid DL-IoT models, a large amount of data was not required since the features are intuitive and informative.

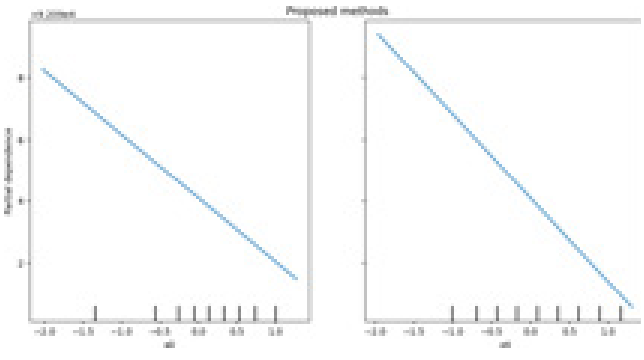


Figure 16. Hybrid DL-IoT models: Partial dependence of Py and PFa on energy consumption.

4.0 Concluding Remarks and Future Works

Our work aimed to achieve a Hybrid DL-IoT BEMS focused on a cost-effective, small-scale, accurate and open-source data-driven methodology. The Hybrid DL- IoT approach was used in place of a black-box approach seen in the literature. The reason was to achieve state-of-the-art results and focus on interpretability or explain ability. The proposed model was developed, which achieved a mean absolute error of 385.78, mean squared error of 221052.14 and coefficient of determination of 0.99 out of sample. Since, Hybrid DL-IoT model was developed, and the training features selected were intuitive and informative, problems of high data requirement were avoided. Further, device-level consumption monitoring and utilization can be the focus for achieving additional power savings.

We will also focus on implementing the developed model on the university campus and calculating the actual energy savings achieved with its recommendations. Further, we shall also build a web portal with a dashboard for tracking all the metrics. Scheduling algorithms will be used to organize the energy-intensive activities in the off-peak season and minimize activities during the peak season to achieve cost savings. Finally, we shall use clustering algorithms to understand the different groups in our data with respect to energy consumption.

5.0 Acknowledgement

The data collection and setup were installed at the premises of Pimpri Chinchwad College of Engineering and Research (PCCoER), Pune, India, for surveillance for two years.

6.0 References

1. Nabavi SA, Motlagh NH, Zaidan MA, Aslani A, Zakeri B. Deep Learning in energy modeling: Application in smart buildings with distributed energy generation. IEEE Access. 2021; 9:125439-125461. https://doi.org/10.1109/ACCESS.2021.3110960
2. Abu-Shikhah N, Elkarmi F, Aloquili OM. Medium-term electric load forecasting using multivariable linear and non-linear regression. Smart Grid Renew Energ. 2011; 02(02):126-135. https://doi.org/10.4236/sgre.2011.22015

3. Adepoju GA, Ogunjuyigbe SOA, Alawode KO. Application of neural network to load forecasting in Nigerian electrical power system. Pac J Sci Technol. 2007; 8(1):68-72.
4. Liu Q, Kamoto KM, Liu X, Sun M, Linge N. Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models. IEEE Trans Consum Electron. 2019; 65(1):28-37. https://doi.org/10.1109/TCE.2019.2891160
5. Al-Anbuky A, Bataineh S, Al-Aqtash S. Power demand prediction using fuzzy logic. Control Eng Pract. 1995; 3(9):1291-1298. https://doi.org/10.1016/0967-0661(95)00128-H
6. Moutaib M, Ahajjam T, Fattah M, Farhaoui Y, Aghoutane B, El Bekkali M. Application of Internet of Things in the health sector: Toward minimizing energy consumption. Big Data Min Anal. 2022; 5(4):302-308. https://doi.org/10.26599/BDMA.2021.9020031
7. Amjady N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. IEEE Trans Power Syst. 2001; 16(4):798-805. https://doi.org/10.1109/59.962429
8. Elashry SS, Abohamama AS, Abdul-Kader HM, Rashad MZ, Ali AF. A chaotic reptile search algorithm for energy consumption optimization in wireless sensor networks. IEEE Access. 2024; 12:38999-39015. https://doi.org/10.1109/ACCESS.2024.3374781
9. Learning artificial neural networks (ANN); 2016.
10. Ismail Z. A new approach to peak load forecasting based on EMD and ANFIS. Indian J Sci Technol. 2013; 6(12):1-7. https://doi.org/10.17485/ijst/2013/v6i12.9
11. Montgomery DC, Johnson LA, Gardiner JS. Forecasting and time series analysis. 2nd edn., New York: McGraw-Hill; 1990.
12. Yildiz B, Bilbao JI, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renew Sustain Energy Rev. 2017; 73:1104-1122. https://doi.org/10.1016/j.rser.2017.02.023
13. Yang D, Gao X, Kong L, Pang Y, Zhou B. An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance. IEEE Trans Consum Electron. 2020; 66(2):173-182. https://doi.org/10.1109/TCE.2020.2977964
14. Kim JY, Hwang YM, Sun YG, Sim I, Kim DI, Wang X. Detection for Non-Technical Loss by Smart Energy Theft With Intermediate Monitor Meter in Smart Grid. IEEE Access. 2019; 7:129043-129053. https://doi.org/10.1109/ACCESS.2019.2940443
15. Ibrahem MI, Nabil M, Fouda MM, Mahmoud MMEA, Alasmary W, Alsolami F. Efficient privacy-preserving electricity theft detection with dynamic billing and load monitoring for AMI networks. IEEE Internet Things J. 2021; 8(2):1243-1258. https://doi.org/10.1109/JIOT.2020.3026692
16. Biçici E. A cloud monitor to reduce energy consumption with constrained optimization of server loads. IEEE Access. 2024; 12:25265-25277. https://doi.org/10.1109/ACCESS.2024.3365674
17. Hannan MA, Faisal M, Ker PJ, Mun LH, Parvin K, Mahlia TMI, *et al*. A review of internet of energy based building energy management systems: Issues and recommendations. IEEE Access. 2018; 6:38997-39014. https://doi.org/10.1109/ACCESS.2018.2852811
18. Alahmad MA, Wheeler PG, Schwer A, Eiden J, Brumbaugh A. A comparative study of three feedback devices for residential real-time energy monitoring. IEEE Trans Ind Electron. 2012; 59(4):2002-2013. https://doi.org/10.1109/TIE.2011.2165456
19. Martínez-Sánchez RA, Rodríguez-Reséndiz J, Álvarez-Alvarado JM, García-Martínez JR. An approach for energy and cost savings for a seafood processing plant. IEEE Access. 2023; 11:51820-518337.
20. Kang HS, Lee JY, Lee DY. An Integrated Energy Data Analytics Approach for Machine Tools. IEEE Access. 2020; 8:56124-56140. https://doi.org/10.1109/ACCESS.2020.2981696
21. Zhuang W, Jiang W, Xia M, Liu J. Dynamic generative residual graph convolutional neural networks for electricity theft detection. IEEE Access. 2024; 12:42737-42750. https://doi.org/10.1109/ACCESS.2024.3379201
22. Zhang C, Ma L, Luo Z, Han X, Zhao T. Forecasting building plug load electricity consumption employing occupant-building interaction input features and bidirectional LSTM with improved swarm intelligent algorithms. Energy. 2024; 288:129651.https://doi.org/10.1016/j.energy.2023.129651
23. Atalik T, Čadirci I, Demirci T, Ermis M, Inan T, Kalaycioglu AS, and Salor Ö. Multipurpose platform for power system monitoring and analysis with sample grid applications. IEEE Trans Instrum Meas. 2013; 63(3):566-582. https://doi.org/10.1109/TIM.2013.2281556
24. Steven M Blair, Campbell D Booth, Gillian Williamson, Alexandros Poralis, and Victoria Turnham. Automatically detecting and correcting errors in power quality monitoring data. IEEE Transactions on Power Delivery, 32(2):1005-1013, 2016.
25. Jiang Y, Yin S, Kaynak O. Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond. IEEE Access. 2018; 6:47374-47384. https://doi.org/10.1109/ACCESS.2018.2866403
26. Janthong S, Chalermyanont K, Duangsoithong R. Unbalanced data handling techniques for classifying energy theft and defective meters in the provincial electricity authority of Thailand. IEEE Access. 2023; 11:46522-46540. https://doi.org/10.1109/ACCESS.2023.3274543