GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Journal of Computer, Signal, and System Research

Article

Energy Consumption Pattern Analysis in a Large Public Library of China

Siyuan Huang [1,*]

[1] Santa Barbara City College, Santa Barbara, California, USA
* Correspondence: Siyuan Huang, Santa Barbara City College, Santa Barbara, California, USA

Abstract: This study analyzes energy consumption in a large public library in China using regression and machine learning models. Data from 238 circuits revealed key usage patterns, including midnight surges and dual peaks, along with anomalies like unexpected energy spikes. Model comparisons identified temperature, humidity, and time of day as major influencing factors. The study recommends optimizing HVAC and mechanical circuit operations to improve efficiency. Findings contribute to reducing operational costs and carbon emissions, offering insights for energy management in public institutions. The methodologies can be applied to other buildings seeking sustainability and efficiency improvements.

Keywords: energy consumption; public library; energy efficiency

Received: 07 February 2025
Revised: 12 February 2025
Accepted: 18 February 2025
Published: 20 February 2025

1. Introduction

The feasibility of our present energy infrastructure is being called into question as the world's population expands and energy consumption rises quickly. This has caused several property owners to concentrate on lowering their energy use in an effort to save money and help the environment. The public library system is one such entity. They frequently consist of big, multifunctional buildings that use a lot of energy.

In order to uncover methods to save expenditures and energy usage, public libraries may find it helpful to do energy analyses. For instance, it may assist librarians in comprehending how energy is used across the building, identifying potential inefficient devices, and assessing the performance of energy saving programs.

Energy analysis is the process of examining energy use patterns in a building or system in order to identify opportunities for energy efficiency improvements. This involves collecting data on energy consumption, as well as other relevant variables such as weather conditions and building occupancy and using statistical techniques to identify patterns and relationships between these variables.

Energy data may be effectively analyzed using regression, and energy analytics is a crucial tool for comprehending and optimizing energy consumption in public libraries. By using regression models to particular case studies involving public libraries and assessing the outcomes to offer insights and suggestions for enhancing energy efficiency, this study can contribute to the area.

Regression analysis is a statistical technique that is commonly used in energy analysis to identify patterns and relationships between variables. This involves building a regression model that relates energy consumption to other variables, such as building size, weather conditions, and occupancy. By analyzing the coefficients of the model, researchers can identify which variables have the greatest impact on energy consumption and use this information to make recommendations for energy efficiency improvements.

1.1. Importance of Energy Analysis for Libraries

Overall, energy analysis is a critical tool for reducing costs and environmental impact in public buildings like libraries. By identifying areas of high energy consumption, determining the most effective energy conservation measures, assessing the effectiveness of those measures, and identifying opportunities for renewable energy, libraries can reduce their energy consumption and costs while also helping to protect the environment.

Overall, analyzing patterns of energy consumption can help the library save money, reduce its environmental impact, improve operations, enhance comfort for patrons and staff, and engage with the community. These benefits are particularly important for public institutions like libraries, which have a responsibility to operate efficiently and sustainably while providing valuable services to the community. Identifying patterns of energy consumption can help the library identify opportunities to reduce energy consumption and costs. By making changes to the library's energy usage patterns, equipment, or systems, the library can save on energy costs over time. By conducting an energy analysis, you can identify which parts of a building or system are consuming the most energy. This can help you target areas for energy efficiency improvements, such as by upgrading equipment or adjusting usage patterns.

1.2. Purpose of the Paper

The main purpose of this paper is described in aspects: (1) Comparison of energy consumption between different sections or floors of the library, leading to a better understanding of the energy usage patterns in the building. (2) Identification of energy-saving opportunities for the library, such as reducing energy use during non-peak hours or implementing more efficient lighting and HVAC systems. (3) Identification of the factors that influence energy consumption, such as temperature, humidity, occupancy, and outdoor weather conditions.

2. Literature Review

At the analysis stage, there are many previous studies on energy pattern analysis in public buildings, some in similar libraries. Noranai et. al. proposed saving measures to decrease the number of lamps, achieving 10% reduction from the total of saving cost [1]. Song et al. analyze how various factors influence the power consumption of buildings [2]. The impact of each factor on overall power consumption is examined using the control variables method. Simulation software such as DesignBuilder is used to show that the cooling loads can be decreased by 8.4% and 16.6% respectively, using the combination of two measures [3].

Statistics showed that peak demand for the building is proportional by air conditioning system, others electric equipment and lighting system at 80%, 10% and 7%, respectively [4], but no further optimization suggestions were provided. It is found that the window area in the south-facing reading space is beneficial to the energy efficiency of the whole library. Thus, the south-facing window area should be increased as much as possible [5]. Among the statistical models, the linear regression analysis has shown promising results because of the reasonable accuracy and relatively simple implementation when compared to other methods [6].

Table 1 below is a summary of related research, showing objectives and methods used in similar existing papers.

Table 1. A concise summary of existing research.

research	objective		method		further optimization suggestions
	energy pattern discovery	energy prediction	regression methods	other methods	
[1]	yes	no	no	yes	yes
[3]	yes	no	yes	no	no
[7]	no	yes	yes	yes	yes
[8]	no	yes	yes	no	yes
[9]	yes	yes	no	yes	no
[10]	no	yes	yes	no	no

Building energy use prediction is widely recognized as a tool for energy conservation and informed decision-making aimed at reducing energy consumption [11]. A study on a library building in China demonstrated that the hybrid GA-ANFIS model outperforms ANN in terms of prediction accuracy [12]. Liu et al. conducted accuracy analyses and compared models of machine learning used for building energy consumption prediction [7]. A new vector field-based support vector regression method has been proposed, where multi-distortions in the sample data space or the high-dimensional feature space mapped by a vector field help identify the optimal feature space, allowing the high nonlinearity between inputs and outputs to be approximated linearly [8].

Abnormal analysis is a common practice when certain prediction models are in place, along with the machine learning algorithms used for prediction and the performance metrics applied for evaluation [13]. For instance, one study found that air conditioning energy consumption was abnormal over a four-day period in September [9]. Relevant policies and suggestions are proposed based on the causal analysis. This research is expected to provide theoretical guidance and a practical data reference for building operations management. The prediction accuracy is often combined with other three methods, i.e., ensemble learning without energy consumption pattern classification [10], CNN and Bi-LSTM [14]. A noval research considered the limitation of missing certain measuring equipments, and new prediction models with the reduced secondary variables are retrained to explore the relationship between the prediction accuracy and the potential input variables [15].

3. Methodology

3.1. Framework of Library Energy Analysis and Prediction

The research project consists of five primary steps. Firstly, data cleaning is performed to ensure that the data is accurate, complete, and consistent. This step involves identifying and correcting errors, inconsistencies, and missing values in the dataset. Descriptive analysis is then conducted to summarize the main characteristics of the dataset, such as the mean, median, mode, standard deviation, and range. This step helps to identify patterns and trends in the data and aids in the selection of appropriate statistical methods for further analysis.

The second step involves the selection of the appropriate electric circuit for a specific application. This task can be facilitated by referring to a library of electric circuits that offers a range of circuits based on the application's specifications and requirements.

The third step is model selection for energy regression. This involves selecting a regression model that accurately predicts energy consumption based on various factors such as weather, time of day, and other relevant variables. The model's performance is then evaluated using metrics such as mean squared error, R-squared, and other relevant metrics.

The fourth step involves energy consumption prediction using the selected regression model. This step is crucial in energy management as it helps optimize energy usage and reduce energy costs. Various techniques such as time series analysis, regression analysis, and machine learning algorithms can be used to accurately predict energy consumption depending on the data available and the specific application.

The final step involves providing quantitative optimization suggestions to solve a given problem using a library of optimization techniques. These techniques can range from linear programming, integer programming, and dynamic programming, and can help identify the optimal solution for a specific problem. By referring to the library, researchers can quickly identify the appropriate algorithm for a specific problem and use it to find the optimal solution. Figure 1 presents the analysis flow of this paper, illustrating the sequence of these steps.

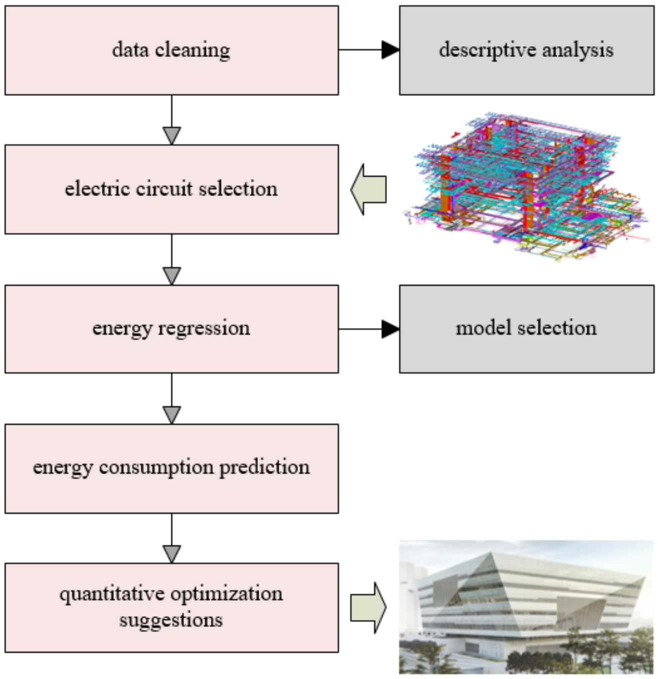


Figure 1. Analysis flow of this paper.

3.2. Data Collection and Preparation

In order to collect electric usage data, we installed sub-meters on all circuits in the library. These sub-meters were installed at the circuit level in order to provide a granular view of energy consumption throughout the building. The sub-meters were connected to a data logger that recorded electricity usage data at 15-minute intervals (shown in Figure 2). The data logger was configured to collect data for a period of one month.

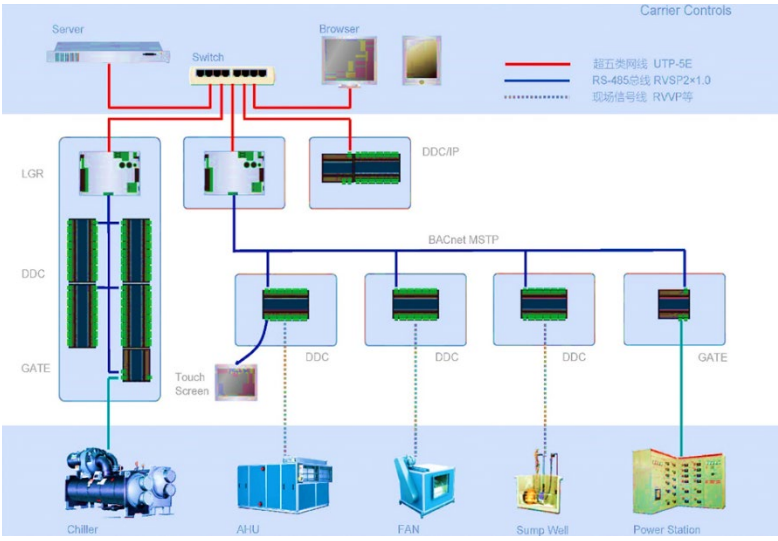


Figure 2. Installation of intelligent meters inside the library.

The data collected from the sub-meters provided a detailed view of electricity usage throughout the library. As shown in Table 2, there are 238 electric circuits.

Table 2. summary of 238 collected energy circuits.

system name	B2F	B1F	1F	2~7F	Roof	Total
light	11	12	11	70	0	104
special	11	12	0	1	0	24
powertrain	16	14	0	4	6	40
airconditioner	4	6	8	47	1	66

3.3. Description of Regression Models Used in the Analysis

3.2.1. Linear Regression

Linear regression is a statistical technique that is used to analyze the relationship between a dependent variable and one or more independent variables. It involves fitting a linear equation to a set of data points to find the best fit line that represents the relationship between the variables. The linear regression model assumes a linear relationship between the dependent variable and independent variables and provides a way to predict the value of the dependent variable based on the value of the independent variables.

Therefore, this paper first uses a linear model to find elementary patterns in energy data of the public library.

3.2.2. Multi-Layer Perceptron (MLP)

MLPs are artificial neural networks known for their ability to capture intricate data relationships. This paper employs an MLP network comprising four hidden layers to accurately fit the energy consumption pattern. By leveraging the MLP's non-linear activation functions and backpropagation algorithm, we aim to enhance our understanding of energy usage trends and enable more informed decision-making in energy management and conservation strategies. The experimental results demonstrate the strong ability of the proposed MLP model in analyzing energy consumption patterns.

3.2.3. Decision Tree and Random Forest

Random Forest (RF) is an ensemble learning method that combines multiple decision trees to make predictions. Decision tree (DT) is a supervised machine learning algorithm that is widely used for classification and regression analysis. It works by constructing a tree-like model of decisions and their possible consequences. The model is built by recursively splitting the dataset into smaller subsets, with each split based on a feature or attribute of the dataset that is most informative for separating the different classes or predicting the target variable. The goal is to create a tree that predicts the target variable with high accuracy while keeping the tree as small and simple as possible to avoid overfitting. Decision trees are easy to interpret and explain, making them a popular choice for solving complex problems.

Since DT can handle both categorical and numerical data, it is suitable to predict energy consumption with discrete attributes such as weekday or different electric circuits.

4. Energy Analysis Results

4.1. Descriptive Analysis of Energy Consumption in the Library

The data collection process involved gathering information from the 238 circuits in the library over a specific time period. These circuits are responsible for powering various aspects of the library, including lighting, heating, cooling, equipment, and other electrical devices. The dataset, comprising 585,000 data points, was recorded at regular intervals. To narrow down our analysis, we selected the top 10 circuits with the highest energy consumption. This selection was based on aggregate energy usage, and it helped us pinpoint areas of the library that may have a significant impact on overall energy costs and sustainability efforts. Figure 3 provides a clear visual representation of the energy consumption distribution in the library, highlighting the top 20 circuits/systems that account for the majority of energy usage. Among them, 10 circuits were selected for further analysis.

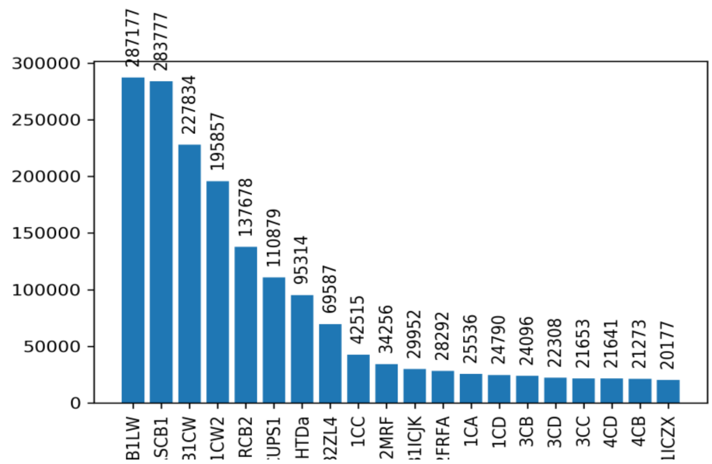


Figure 3. Top 20 highest energy consumption circuits.

In Figure 4, a graphical representation illustrates the proportions of four distinct types of subsystems over a month's period, showing the dynamics and trends in a system where different subsystems play critical roles. Figure 5 displays a series of curves that represent the typical behavior of important circuits within a system over a daily time frame. These figures are valuable tools for visualizing and analyzing complex systems and circuits, enabling better decision-making and understanding of system dynamics.

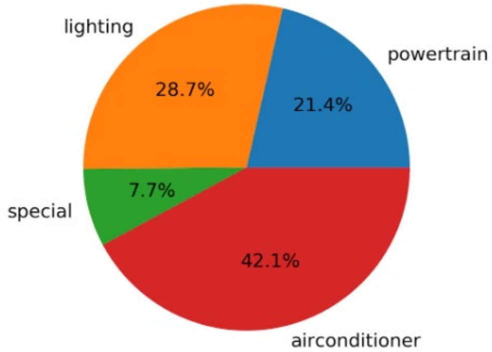


Figure 4. Proportion of four types of subsystems.

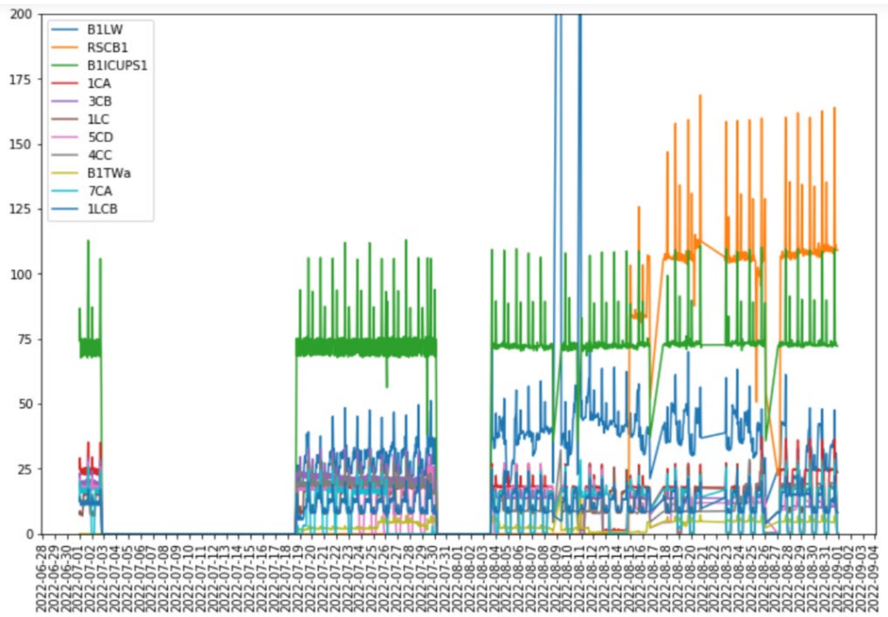


Figure 5. Typical curves of important circuits (by day).

A pair plot (Figure 6) explored relationships between multiple variables simultaneously, showing relationships between energy, circuit, weekday, hour, and electric flow. It can be concluded that not all variables were significant, and quantity relations remained complex.

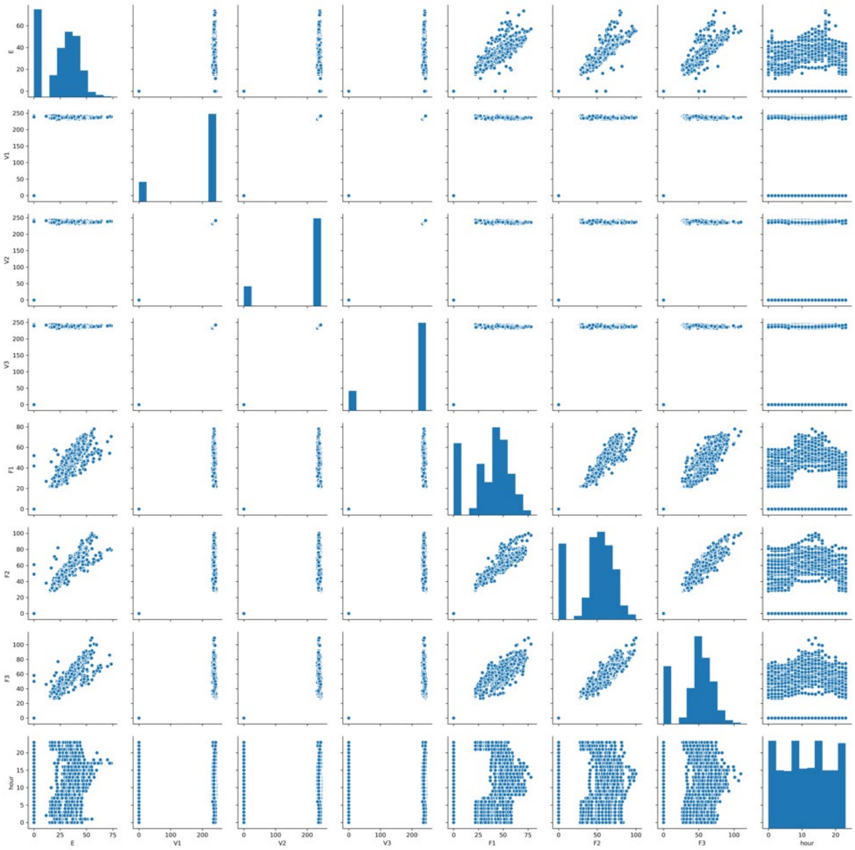


Figure 6. Pair plot of analyzed variables.

4.2. Comparison of Regression Models and Their Performance

4.2.1. Hyper Parameters

Table 3 shows hyper parameters used in three prediction models. The MLP model was tuned several times, and the best neural network structure was recorded.

Table 3. Values of hyper parameters.

Model	Model parameters
Linear model	loss: RMSE
MLP model	hiddenLayers = 4, neurons = (5, 30, 30, 5)
Random forest model	treeNum = 30

4.2.2. Model Comparison and Selection on Important Circuits

In the context of analyzing and optimizing important circuits, selecting the right predictive model is crucial. To make an informed choice, the above-mentioned three distinct models were employed and their performance was rigorously compared. To determine which model performs best for predicting outcomes related to important circuits, a thorough comparison of their accuracy was conducted as Table 4. Accuracy was assessed through Root Mean Squared Error (RMSE). For the MLP models, random initial weight parameters were used. Therefore, ten runs were executed and the model with best accuracy was recorded in the table. Different electric circuits use different models due to specific nonlinear behaviors.

Table 4. Result of model comparison and selection "×" denotes insufficient performance.

Circuit	Linear model		MLP model		Random forest model		Selected model
	Train	Predict	Train	Predict	Train	Predict	
B1LW	0.625	0.773	0.700	0.859	0.951	0.719	MLP
RSCB1	0.328	×	0.936	0.584	0.984	0.552	MLP
B1ICUPS1	×	×	×	×	0.884	0.777	RF
1CA	×	0.429	×	0.448	0.923	0.579	RF
3CB	0.183	0.589	0.165	0.645	0.945	×	MLP
1LC	0.949	0.959	0.954	0.946	0.995	0.960	RF
5CD	0.654	0.746	0.667	0.760	0.986	0.708	MLP
4CC	0.629	0.836	0.621	0.846	0.935	0.824	MLP
B1TWa	0.881	0.798	0.883	0.804	0.937	0.719	MLP
7CA	0.574	0.659	0.566	0.698	0.933	0.727	RF
1LCB	0.527	0.686	0.489	0.681	0.917	0.826	RF

To better assess the impact of variables on the model's predictive accuracy and help the following decision-making processes, the variable importances of RF models was investigated. As shown in Figure 7, the order of important variables, apart from pure electric metrics (current and voltage), is as follows: hour of day, humidity, temperature/feel temperature, wind power, and weekday.

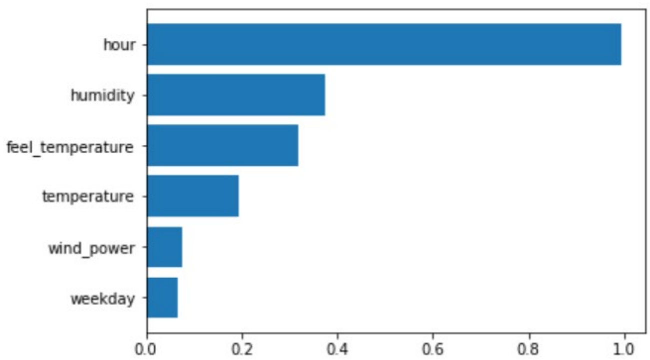


Figure 7. Variable importance derived from RF models.

The hour of the day can capture diurnal patterns and variations in energy data. It may have a strong influence on the target variable if there are specific time-related trends or events that impact the outcome. Humidity levels and outdoor temperature can significantly affect various processes and phenomena. They are fundamental environmental factors that affect many natural and human systems, making it an important predictor in the models. However, the weekday variable does not have a significant impact and holds the lowest importance. This can be attributed to the operational model of libraries, which differs from other public buildings like offices or commercial establishments that experience discernible variations in electricity demand between weekdays and weekends. Libraries operate continuously throughout the year, and their energy consumption patterns remain relatively consistent regardless which day of the week.

4.3. Pattern/Abnormal Identification and Suggestions

4.3.1. Behavior Patterns from Daily Curves

This section outlines and analyzes typical energy consumption patterns observed in the public library environment. The examination encompasses three distinctive scenarios, each shedding light on different aspects of energy utilization:

1) Midnight Surge

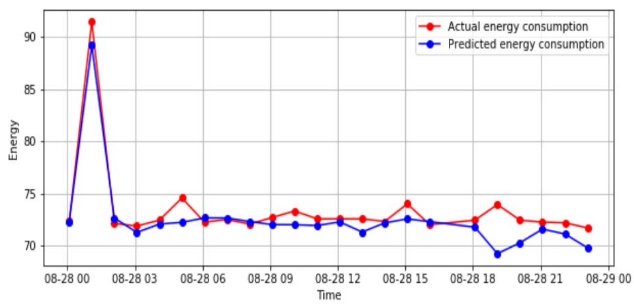
One of the prominent energy consumption patterns observed is the substantial surge in energy usage during the midnight hours (Figure 8-a). This surge predominantly corresponds to the operation of specific mechanical circuits within the library's infrastructure. The data reveals a pronounced spike in energy consumption, typically occurring between the hours of 12:00 AM and 2:00 AM. This pattern implies a distinctive operational requirement during these late hours.

2) Operating Hours Consistency

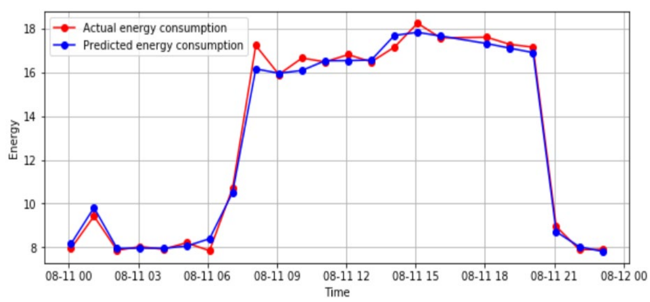
Another common pattern that emerges is the consistent energy consumption during the library's operational hours (Figure 8-b). The data reflects an expected energy demand during the usual opening hours, such as from 6:00 AM to 9:00 PM. This consistent energy usage is largely attributed to essential functions within the library, such as lighting and air conditioning systems.

3) Dual Peaks

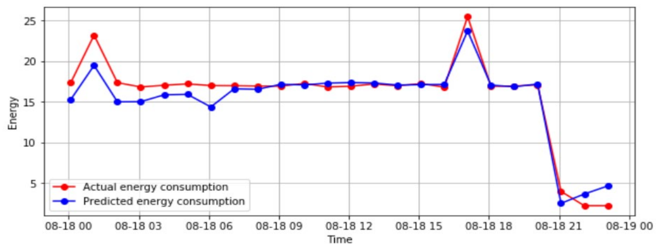
The third distinctive pattern observed presents dual peaks in energy consumption, one during the midnight hours and another in the early evening (Figure 8-c and Figure 8-d). Unlike the midnight surge mentioned earlier, these dual peaks are associated with separate mechanical circuits, possibly serving different functions within the library.



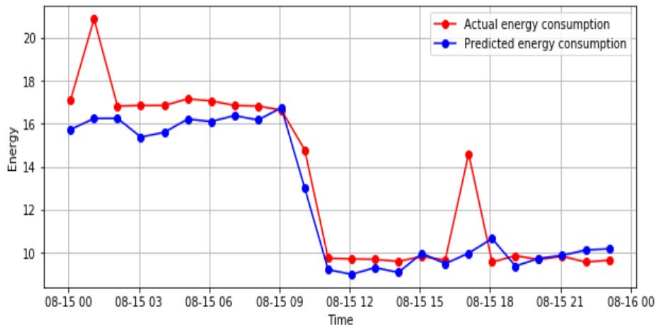
(a) Pattern 1 in Aug 28.



(b) Pattern 2 in Aug 11.



(c) Pattern 3 in Aug 18.



(d) Pattern 4 in Aug 15.

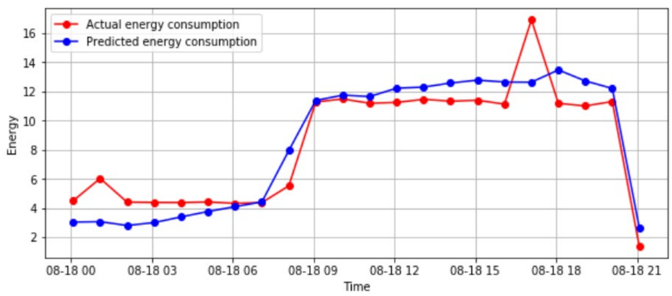
Figure 8. Typical energy consumption patterns.

4.3.2. Abnormal Behaviors

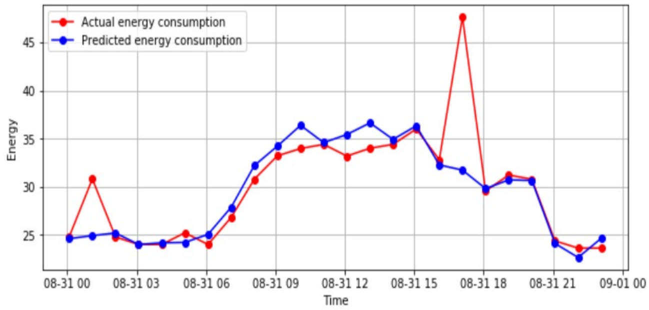
In addition to the typical energy consumption scenarios, two noteworthy abnormalities within the library's energy consumption patterns were discovered. These anomalies deviate significantly from the anticipated usage patterns and warrant closer examination:

1) 5 PM Surge (Figure 9 a ~ Figure 9 c)

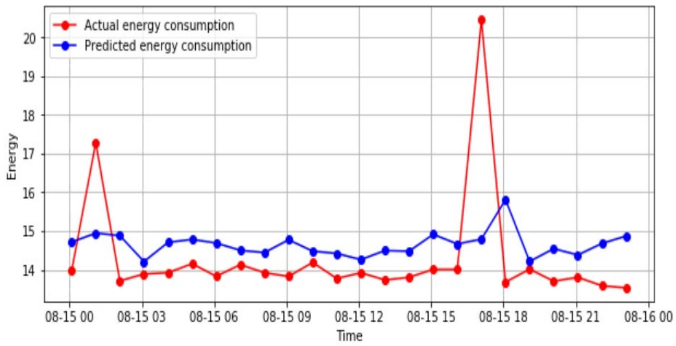
At 5 PM, an unexpected and sudden spike in energy consumption was detected. This surge was notably larger than the typical energy demand observed at this time of day. Further investigation revealed that the heightened energy usage was attributed to a specific mechanical circuit. The source and purpose of this sharp energy increase at 5 PM pose a compelling anomaly, as it indicates an uncharacteristic operational behavior within the library. It is suggested to ensure energy efficiency and to potentially identify areas for optimization.



(a) Abnormal 1 in Aug 20.



(b) Abnormal 2 in Aug 31.

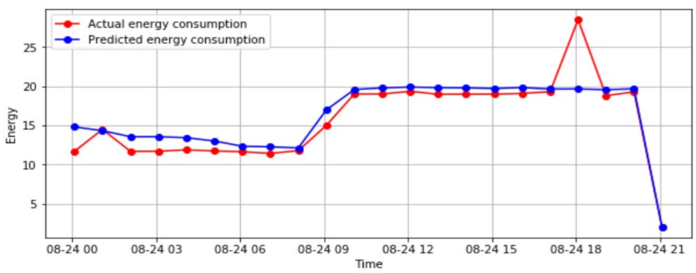


(c) Abnormal 3 in Aug 15.

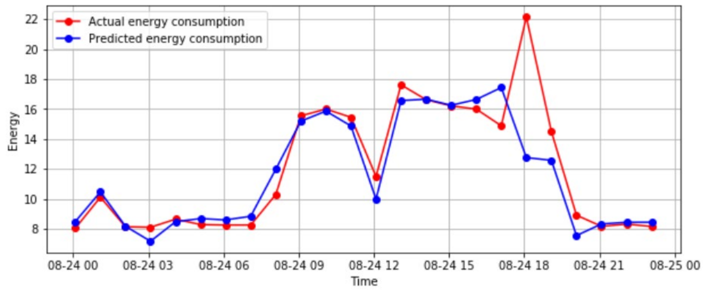
Figure 9. Abnormalities of 5 pm surge.

2) 6 PM Air Conditioning Surge (Figure 10 a and b)

Another abnormality occurred at 6 PM, with a substantial surge in energy consumption attributed to an air conditioning circuit. Unlike the steady energy usage observed during the library's operational hours, this abrupt increase in energy demand at 6 PM is inconsistent with the expected behavior. The anomaly signifies an unanticipated surge in the air conditioning system's power consumption, indicating a deviation from standard environmental control patterns. It is suggested to identify the root cause of this abnormal energy consumption for improving energy efficiency and the overall comfort of library patrons.



(a) Abnormal 4 in Aug 24.



(b) Abnormal 5 in Aug 24.

Figure 10. Abnormalities of 6 pm surge.

5. Discussion and Conclusion

This paper analyzed behavior of some electric circuits using three regression methods, and found out some typical behavior and abnormalities, then some suggestions for the library about energy usage were provided. In conclusion, this paper delved into the critical domain of energy analysis, focusing on the specific context of a public library. The study employed a comprehensive approach by scrutinizing the behavior of various electric circuits within the library infrastructure. Leveraging the power of three regression methods, the investigation unearthed valuable insights into the typical behavior and anomalies in energy consumption patterns.

The findings revealed a spectrum of electric circuit behaviors, ranging from expected fluctuations in energy usage during peak hours to unanticipated irregularities. This in-depth analysis not only enriched our understanding of the library's energy dynamics but also pinpointed certain scenarios where energy optimization measures could be implemented.

The findings and suggestions of this research can serve as a blueprint for energy analysis in other public institutions and contribute to the broader goals of reducing carbon emissions and promoting environmental sustainability. By adhering to these suggestions, the public library can not only minimize its environmental impact but also potentially reduce energy-related operational costs, thus fostering a greener, more sustainable future.

References

1. Z. Noranai and A. D. F. Azman, "Potential reduction of energy consumption in public university library," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 243, no. 1, p. 012023, Sep. 2017, doi: 10.1088/1757-899X/243/1/012023.
2. J. Song, X. Zhang, and X. Meng, "Simulation and analysis of a university library energy consumption based on EQUEST," *Procedia Eng.*, vol. 121, pp. 1382-1388, 2015, doi: 10.1016/j.proeng.2015.09.028.
3. Y. Li and J. He, "Evaluating the improvement effect of low-energy strategies on the summer indoor thermal environment and cooling energy consumption in a library building: A case study in a hot-humid and less-windy city of China," *Build. Simul.*, vol. 14, pp. 1423–1437, 2021, doi: 10.1007/s12273-020-0747-6.
4. E. Boonthum, "Study of energy consumption in library building," *Kasem Bundit Eng. J.*, 2018. [Online]. Available: https://ph01.tci-thaijo.org/index.php/kbej/article/view/123496. [Accessed: Feb. 2, 2025].
5. J. Li, Q. Guan, and H. Yang, "Winter energy consumption in reading space of green libraries in cold regions," *Int. J. Heat Technol.*, vol. 36, no. 4, pp. 1256-1261, 2018, doi: 10.18280/ijht.360413.
6. N. Fumo and M. R. Biswas, "Regression analysis for prediction of residential energy consumption," *Renew. Sustain. Energy Rev.*, vol. 47, pp. 332-343, 2015, doi: 10.1016/j.rser.2015.03.035.
7. Z. Liu, D. Wu, Y. Liu, Z. Han, L. Lun, J. Gao, G. Jin, and G. Cao, "Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction," *Energy Explor. Exploit.*, vol. 37, no. 4, pp. 1426–1451, 2019, doi: 10.1177/0144598718822400.
8. H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, "Vector field-based support vector regression for building energy consumption prediction," *Appl. Energy*, vol. 242, pp. 403-414, 2019, doi: 10.1016/j.apenergy.2019.03.078.
9. Y. Liu, H. Chen, L. Zhang, X. Wu, and X. J. Wang, "Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in China," *J. Clean. Prod.*, vol. 272, p. 122542, 2020, doi: 10.1016/j.jclepro.2020.122542.

10. Z. Dong, J. Liu, B. Liu, K. Li, and X. Li, "Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification," *Energy Build.*, vol. 241, p. 110929, 2021, doi: 10.1016/j.enbuild.2021.110929.
11. R. Olu-Ajayi, H. Alaka, I. Sulaimon, F. Sunmola, and S. Ajayi, "Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques," *J. Build. Eng.*, vol. 45, p. 103406, 2022, doi: 10.1016/j.jobe.2021.103406.
12. K. Li, H. Su, and J. Chu, "Forecasting building energy consumption using neural networks and hybrid neuro-fuzzy system: A comparative study," *Energy Build.*, vol. 43, no. 10, pp. 2893-2899, 2011, doi: 10.1016/j.enbuild.2011.07.010.
13. K. Amasyali and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192-1205, 2018, doi: 10.1016/j.rser.2017.04.095.
14. T. Le, M. T. Vo, B. Vo, E. Hwang, S. Rho, and S. W. Baik, "Improving electric energy consumption prediction using CNN and Bi-LSTM," *Appl. Sci.*, vol. 9, no. 20, p. 4237, 2019, doi: 10.3390/app9204237.
15. J. Q. Wang, Y. Du, and J. Wang, "LSTM based long-term energy consumption prediction with periodicity," *Energy*, vol. 197, p. 117197, 2020, doi: 10.1016/j.energy.2020.117197.