INTELLIGENT SYSTEMS AND SUSTAINABLE ENERGY
E-ISSN: 3154-7532
VOLUME 1, ISSUE 1, 2026, pages 1 – 8
DOI: https://doi.org/10.15282/isse.1.1.2026.13539

RESEARCH ARTICLE

Enhancing chiller energy consumption prediction using gradient boosted trees

M. N. I. Mahamad Yazid[1], M. I. Mohd Rashid[1], Z. Mustaffa[2], Y. Durachman[3], M. H. Sulaiman[1]*

[1]Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
[2]Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
[3]Faculty of Science and Technology, Universitas Islam Negeri Syarif Hidayatullah, Jakarta, Indonesia

Abstract - Chiller systems account for a considerable amount of energy consumption in commercial buildings, which highlights the need of precise energy prediction for operational efficiency and efficient energy management. The standard prediction approaches have difficulties due to the nonlinear interactions between the external variables, the building loads and the chiller operating parameters. The study aims to evaluate the efficiency of Gradient Boosting Trees (GBT) in predicting chiller energy consumption and compare its accuracy with eXtreme Gradient Boosting (XGBoost) and Feedforward Neural Networks (FFNN). The models were constructed based on operational data from a commercial building located in Singapore, with factors such as external temperature, building load and chiller operating parameters. Data pretreatment and feature engineering were performed to improve prediction performance, and grid search was used for hyperparameter optimization. The model evaluation was carried out using the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE). GBT (R^2 = 0.9617) had the best prediction accuracy and was slightly better than XGBoost (R^2 = 0.9586), whilst FFNN exhibited the lowest prediction performance. These results show the effectiveness of GBT in modeling energy consumption of chillers and the potential of tree-based ensemble learning for building energy prediction. Future study can be extended by considering more operational variables and new optimization techniques to further enhance the prediction accuracy.

Article History
Received : 26 November 2025
Revised : 19 December 2025
Accepted : 20 December 2025
Published :

Keywords
Chiller energy consumption
Machine learning
Gradient boosting trees
Extreme gradient boosting
Feedforward neural networks

1. Introduction

Chiller systems represent a substantial part of energy consumption in commercial buildings and play a significant part in providing indoor thermal comfort. Hence, enhancement of chillers efficiency is an important aspect in building energy management. Accurate prediction of chiller energy use enables facility managers to improve system efficiency, reduce energy expenditures, and manage facilities more effectively. This is a common use of traditional forecasting methods, such as regression algorithms or thermodynamic models. However, these methods typically necessitate extensive domain expertise and fail to adequately capture the complex nonlinear interactions within building operating conditions [1–4]. In recent years, the breakthroughs in artificial intelligence have opened the way for the application of machine and deep learning methods for energy forecasting. Machine learning algorithms have been found to have good performance in predicting building energy use since they are able to learn from previous data without explicit physical modeling [7–11]. Deep learning models can also reflect complicated interactions between environmental and operational variables, when huge datasets and nonlinear dynamics are involved [12–17]. These technical advances have motivated the development of data driven methodologies to improve the accuracy of chiller energy consumption prediction.

Gradient Boosting Trees (GBT) are one of the widely used machine learning approaches that have received great attention owing to their capacity to learn nonlinear relationships through sequential learning and error correction [18–21]. GBT builds an ensemble of decision trees, each new tree being trained to correct the prediction mistakes of the previous trees. This approach has been successfully applied to several prediction and forecasting settings. The eXtreme Gradient Boosting (XGBoost) is a highly optimized gradient boosting algorithm that provides outstanding predictive performance owing to its improved computing efficiency and regularization [22–27]. In addition, Feedforward Neural Networks (FFNN) are still popular deep learning architectures for simulating complex input-output interactions in energy systems [28–33]. Although previous studies have reported excellent results using a single machine learning or deep learning model, comparable studies under uniform modeling settings are rare, notably in the field of chiller energy consumption prediction. Moreover, the effect of operational and environmental variables on the predictive performance of the model is not always fully defined. Therefore, choosing an optimal forecasting model for practical application is a challenging task. Therefore, this work uses GBT to estimate chiller energy usage with operational data from a commercial building in Singapore. We compare the performance of GBT with XGBoost and FFNN on the same datasets, pre-processing techniques and evaluation measures. Model performance is tested using coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), maximal error and forecast variability. The results provide insight into

***CORRESPONDING AUTHOR | M. H. Sulaiman** | ✉ herwan@umpsa.edu.my

the relative strengths of tree-based ensemble learning and neural network approaches for modelling chiller energy usage and hence assist the selection of appropriate forecasting models for building energy management.

2. Materials and Methods

Figure 1 depicts the proposed framework. The methodology includes data pre-processing, feature engineering, model development, hyper-parameter tuning and performance evaluation. Three prediction models, Gradient Boosting Trees (GBT), eXtreme Gradient Boosting (XGBoost) and Feedforward Neural Networks (FFNN) were built and evaluated using the same dataset and modeling technique. After the model training, the predicted results were compared with the actual energy consumption data to find out the most appropriate prediction model.

Data for prediction of energy usage were collected from a commercial building in Singapore equipped with a chiller system [7]. The data ranges from January 1, 2022 to December 31, 2024 and consist of a total of 13,561 samples. Data preprocessing was performed before analysis to improve the quality of data by removing outliers, addressing missing values and normalizing numerical features between 0 and 1. Then, feature engineering was performed to improve the prediction performance. The main features are holidays, settlement periods and time-based elements such as month, week, year and quarter. These were included to account for the seasonal and operational fluctuations in energy consumption. The goal variable was the Transmission System consumption (TSD), which is the entire electricity consumption of the chiller system.

The prediction task was formulated as a supervised regression problem with operational, temporal and environmental factors as input features to predict chiller energy consumption. The dataset was randomly split into training (70%), validation (15%) and testing (15%) sets. Model parameters were estimated using the training dataset and hyperparameter tuning was performed using the validation dataset. The testing dataset was used for the final model evaluation on unseen data. Repeated random subsampling cross-validation was used throughout the model development process to increase the reliability of model evaluation.

Hyperparameter optimization was used to find the best setting for each model based on a grid search approach. The optimization was performed over 500 iterations and 10 simulations. The FFNN architecture includes an input layer for the engineering features, the optimized hidden layers, and a single output neuron representing the chiller energy usage. The optimization for GBT and XGBoost was based on features such as tree depth, learning rate, and the number of estimators.

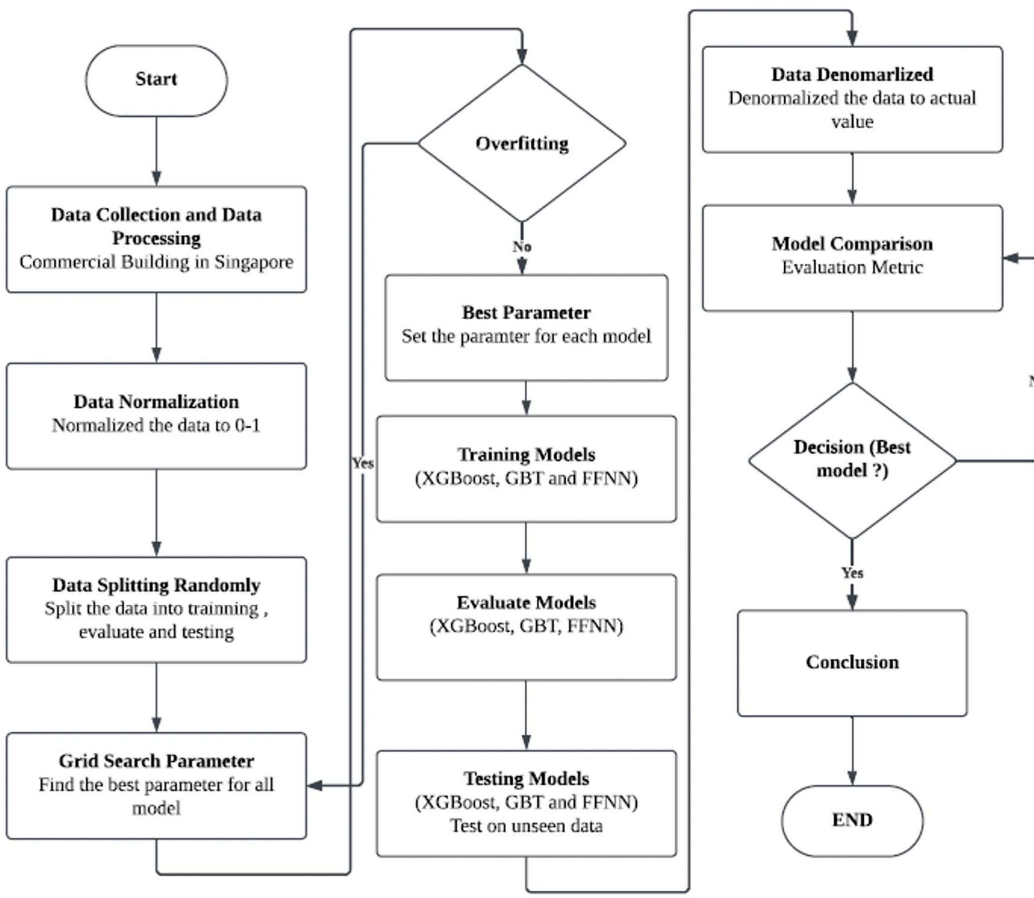

Figure 1. Flowchart of the proposed work

Additional climatic variables were incorporated in the modeling framework including wet-bulb temperature, relative humidity, and precipitation from Singapore weather records.Time-lagged variables from 0 to 6 hours were included to account for delayed thermal responses in the building and chiller system. A seven-day exponentially weighted moving average of TSD was added to capture short-term demand patterns. A binary post-COVID occupancy indicator was added to distinguish the transition period after the return-to-office phase in 2022 from the more stable occupancy trends observed in 2023–2024. All engineering variables were assessed using the Variance Inflation Factor (VIF), and only variables meeting the threshold of VIF < 5 were retained for model construction.

The evaluation of model performance utilized the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and standard deviation of prediction errors. The coefficient of determination (R^2) quantifies the fraction of variation in the observed data that is elucidated by the model. Values approaching 1 signify superior prediction performance, as delineated in Eq. (1).

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

The RMSE measures the average prediction error with the square operation giving more weight to larger errors. The lower the values, the better the prediction and the agreement between the measured values. MAE is a statistic that measures the average size of the errors in a set of predictions, regardless of their direction. The RMSE and MAE are calculated as Eq. (2) and Eq. (3) accordingly.

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

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

The standard deviation of prediction errors indicates the variation of the model mistakes around their average value. Eq. (4) . Lower standard deviation value represents more reliable model prediction .

$$\text{Standard Deviation} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y - \bar{y})^2} \tag{4}$$

where y_i represents the measured chiller energy consumption, $\hat{y}_i$ denotes the predicted value, $\bar{y}$ represents the mean measured value, and $n$is the total number of observations.

3. Results and Discussion

The same dataset and modelling process were used to compare the prediction performance of GBT, XGBoost, and FFNN. Models were compared based on prediction accuracy, error metrics, and performance on previously unidentified test data. The findings provide a foundation for evaluating each model's viability for predicting chiller energy use.

Figure 2 shows the actual and predicted chiller energy consumption calculated using the GBT model. The measured energy usage is shown by the solid blue line, and the projected numbers are shown by the dashed red line. As a rule, the projected values are closely matched to the measured consumption pattern indicating that the model was able to capture the relationship between input variables and energy demand accurately. Some discrepancies are seen during peak-load conditions in which the model tends to underestimate the highest energy consumption figures. However, the measured and anticipated statistics are in good agreement under most operational situations. The finding is in line with the performance parameters presented in Table 1, indicating that GBT had the highest coefficient of determination and lowest prediction errors of the assessed models.

Figure 3 illustrates the largest GBT model prediction error with the actual energy usage being 178.5 kW and the predicted value of 127.447 kW. This indicates that the model tends to underpredict the energy consumption for this peak-load condition. The graph shows that the error is only in one localized area of high demand and the overall trend is in agreement with the observed data. The GBT model generally was good at monitoring the energy of chiller usage under normal operating conditions but was less accurate at capturing high demand peaks. The model exhibited good predictive performance as seen from the highest R 2 value and lowest RMSE and MAE values in Table 1.

Figure 4 depicts the largest prediction error of the XGBoost model. The measured energy consumption of the chiller was 165.3 kW, while the predicted value was 122.099 kW. Similar to the GBT model, the largest prediction deviation occurred during a peak-demand event, where the model underestimated the actual energy consumption. As can be seen from the figure, the predicted values were generally consistent with the measured trend, suggesting satisfactory predictive performance of the model under most operating conditions. However, the increased deviation at high energy demand implies that the model was less successful in capturing the peaks of the energy consumption. This is consistent with the performance metrics in Table 1, where the XGBoost model achieved high overall prediction accuracy but was slightly inferior to the GBT model in terms of RMSE, MAE and R^2.

The highest prediction error of the FFNN model is shown in Figure 5. The actual energy consumption of the chiller was 183.3 kW but the predicted value was 112.387 kW. In comparison to GBT and XGBoost, FFNN had a wider deviation for peak-demand conditions, indicating lower accuracy in capturing extreme energy consumption events. The graph shows that the model was able to capture the overall trend of the energy consumption but its predictive performance deteriorated with significantly increased energy demand. This is also supported by the performance metrics shown in Table 1 where FFNN had the highest RMSE and MAE values, and the lowest coefficient of determination (R^2). The results suggest that the FFNN model was not as effective as the tree-based models in predicting chiller energy consumption for the investigated dataset.

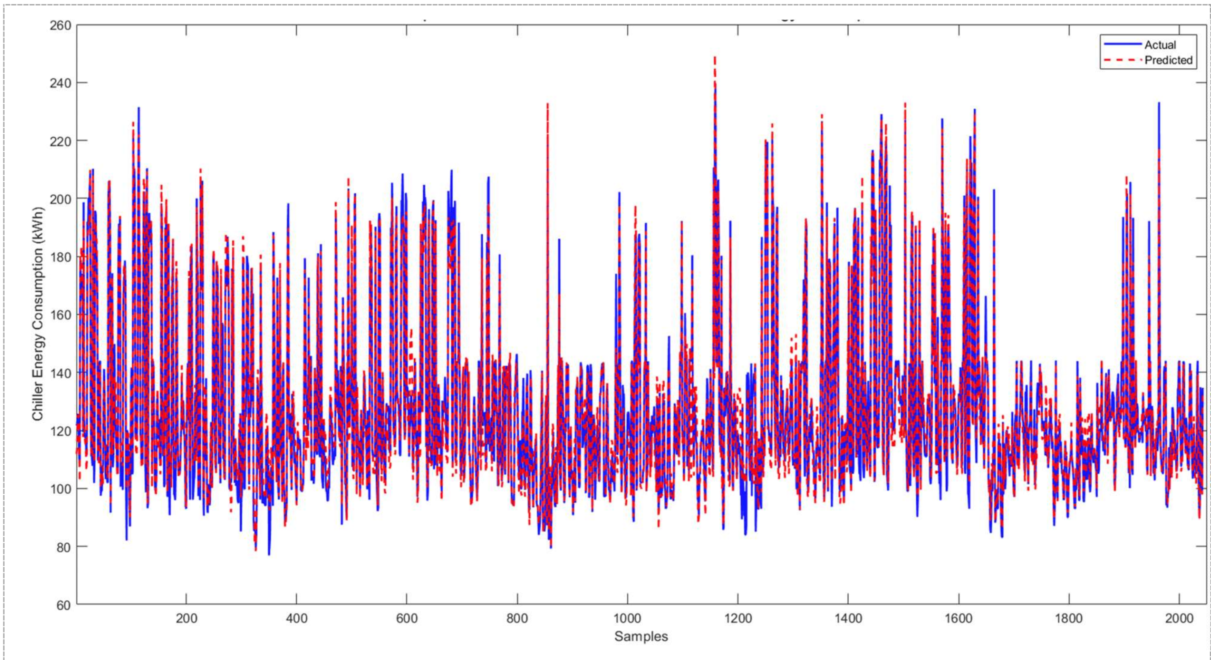

Figure 2. Actual versus predicted using the GBT model

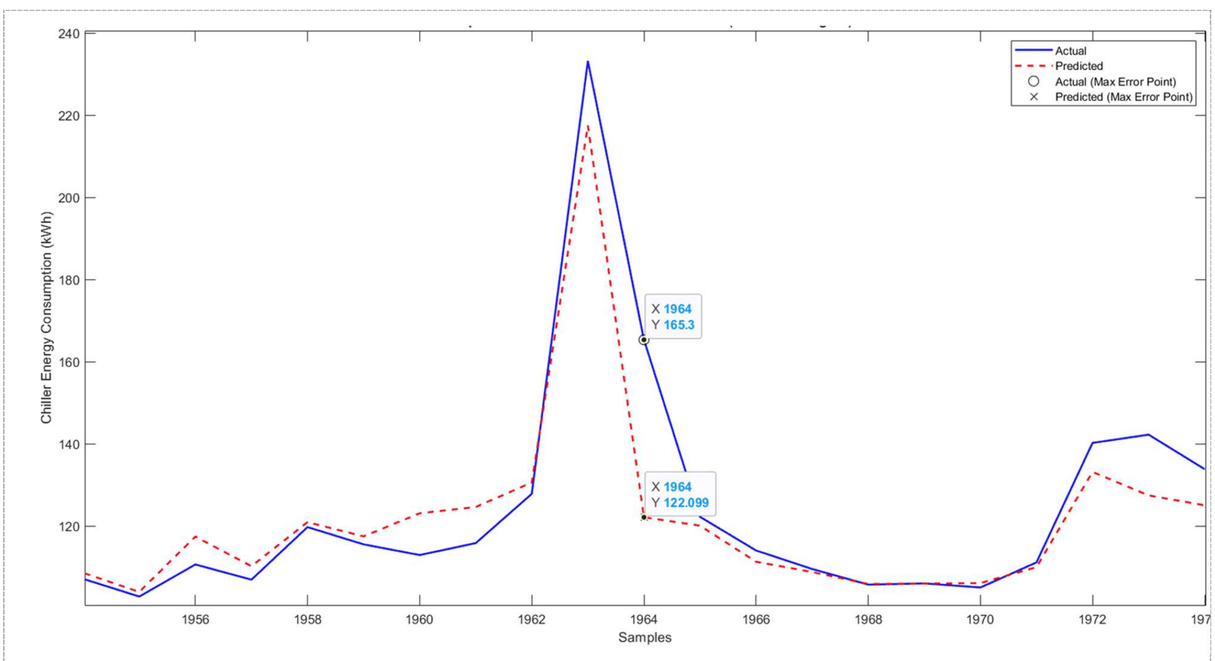

Figure 3. Max error obtained by the GBT model

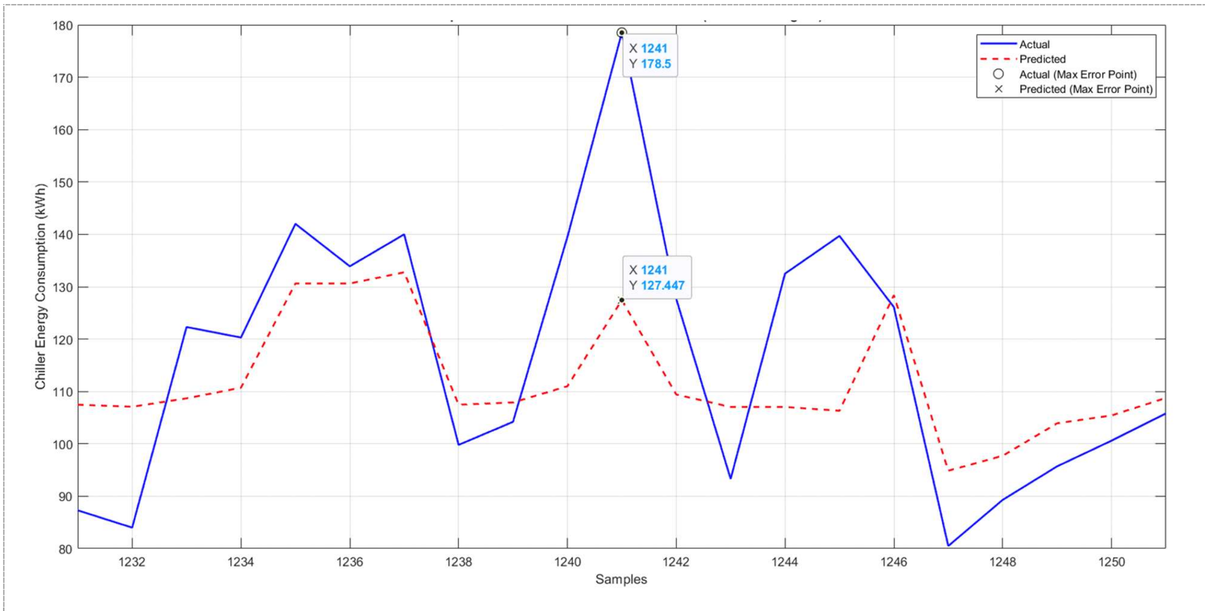

Figure 4. Max error obtained by the XGBoost model

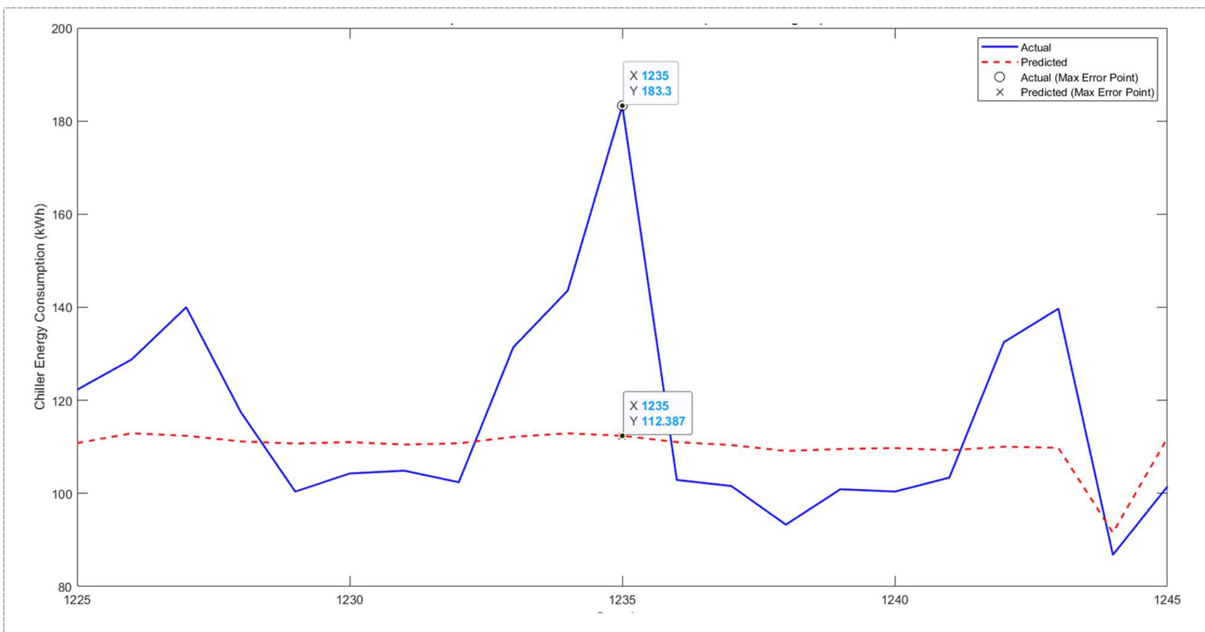

Figure 5. Max error obtained by the FFNN model

Table 1 summarizes the prediction performance of GBT, XGBoost and FFNN models for the prediction of chiller energy consumption. GBT model provides the lowest RMSE (6.0909), the lowest MAE (3.9734) and the highest R^2 (0.9617) indicating the best overall prediction accuracy. XGBoost model provides comparable results with RMSE of 6.1166, MAE of 4.2843 and R^2 of 0.9586 indicating the strong prediction capability across the dataset. FFNN model provides the highest RMSE (7.6519), the highest MAE (5.3725) and the lowest R^2 (0.9361) indicating the comparatively lower prediction accuracy. The results in Table 1 are in line with the graphical results shown in Figures 2–5. GBT and XGBoost captured the overall trend of chiller energy consumption relatively well with relatively small prediction errors, while FFNN prediction errors were relatively larger, particularly at peak-demand conditions. Although XGBoost had the lowest maximum error (43.2012), GBT had lower average prediction errors and higher coefficients of determination, demonstrating more consistent overall performance. Despite achieving competitive RMSE and R^2 values, XGBoost had a much larger standard deviation of prediction errors (29.2138), indicating more variability in prediction performance across the dataset. Moreover, the lower standard deviation of prediction errors from GBT also indicates a higher prediction stability under different operating conditions. To conclude, based on the complete assessment of RMSE, MAE, R^2 and the prediction variability, GBT achieved the most balanced performance among the models investigated, representing a feasible approach for predicting chiller energy consumption in commercial buildings.

Table 1. Performance of GBT, XGBoost, and FFNN for chiller energy consumption prediction

Model type	GBT	XGBoost	FFNN
RMSE	6.0909	6.1166	7.6519
MAE	3.9734	4.2843	5.3725
Max Error	51.0531	43.2012	70.9128
Standard Deviation	5.9010	29.2138	7.6446
R^2	0.9617	0.9586	0.9361

4. Conclusions

In conclusion, this research assessed the performance of GBT, XGBoost and FFNN in predicting the energy consumption of a commercial building chiller based on operational data. GBT outperformed the other models with the lowest RMSE (6.0909), lowest MAE (3.9734) and highest coefficient of determination (R^2 = 0.9617). The findings indicate that tree-based ensemble learning models are effective in modeling the nonlinear relationship between operational variables and chiller energy consumption. Despite the strong predictive ability of XGBoost, GBT offered the most balanced combination of accuracy, robustness and generalization capability. On the other hand, FFNN demonstrated relatively lower prediction accuracy, especially when the chiller was operating at peak-demand conditions. The overall results indicate GBT is a suitable approach to predict chiller energy consumption and can play a role in data-driven energy management of commercial buildings. The results also indicate the potential of tree-based ensemble-learning approaches to facilitating better informed decisions for building energy management and operational planning. Future research can incorporate more operational variables and advanced optimization approaches to further improve the prediction performance.

Acknowledgement

The authors gratefully acknowledge Universiti Malaysia Pahang Al-Sultan Abdullah for supporting this work.

Funding

This work was supported by Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) under Industrial Matching Grant (RDU242404).

Declaration of Competing Interest

The authors declare no conflict of interest.

CRediT Authorship Contribution Statement

M. N. I. Mahamad Yazid (Writing - original draft; Formal analysis; Visualisation; Data curation)
M. I. Mohd Rashid (Resources; Funding acquisition)
Z. Mustaffa (Conceptualization; Formal analysis)
Y. Durachman (Writing - review & editing)
M. H. Sulaiman (Supervision; Writing - review & editing)

Availability of the Data and Materials

The data used to support the findings of this study are included within the article.

Ethical Declaration

This research did not involve any human participants, animals, or sensitive personal data. Therefore, ethical approval was not required. All data used in this study were obtained from publicly available sources and used in accordance with relevant guidelines and regulations.

Generative Artificial Intelligence Declarations

During the preparation of this work, the author(s) used ChatGPT to improve the clarity of technical descriptions in the methodology section and ensure consistency in engineering terminology. The AI tool was not used for computational modelling, simulation, or analysis of experimental results. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

REFERENCES

[1] G. Hong and N. Seong, “Optimization of the ANN Model for energy consumption prediction of direct-fired absorption chillers for a short-term,” *Buildings*, vol. 13, no. 10, p. 2526, Oct. 2023, doi: 10.3390/buildings13102526.
[2] T. Sathesh and Y. C. Shih, “Optimized deep learning-based prediction model for chiller performance prediction,” *Data Knowl. Eng.*, vol. 144, p. 102120, Mar. 2023, doi: 10.1016/j.datak.2022.102120.

[3] R. Morcillo-Jimenez, J. Mesa, J. Gómez-Romero, M. A. Vila, and M. J. Martin-Bautista, “Deep learning for prediction of energy consumption: an applied use case in an office building,” *Appl. Intell.*, vol. 54, no. 7, pp. 5813–5825, Apr. 2024, doi: 10.1007/s10489-024-05451-9.
[4] P. Kumar, N. Kamalakshi, and T. Karthick, “Towards sustainable architecture: Machine learning for predicting energy use in buildings,” in *Proc. Int. Conf. Adv. Comput. Commun. Appl. Informatics (ACCAI)*, 2024, pp. 1–6, doi: 10.1109/ACCAI61061.2024.10602461.
[5] I. H. Sarker, “Machine learning: Algorithms, real-world applications and research directions,” *SN Comput. Sci.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.
[6] S. Ardabili, L. Abdolalizadeh, C. Mako, B. Torok, and A. Mosavi, “Systematic review of deep learning and machine learning for building energy,” *Front. Energy Res.*, vol. 10, p. 786027, Aug. 2022, doi: 10.3389/fenrg.2022.786027.
[7] M. H. Sulaiman, Z. Mustaffa, and A. S. Abdullah, “Chiller energy prediction in commercial building: A metaheuristic-enhanced deep learning approach,” *Energy*, vol. 297, p. 131159, Jun. 2024, doi: 10.1016/j.energy.2024.131159.
[8] F. Wahid, R. Ghazali, A. S. Shah, and M. Fayaz, “Prediction of energy consumption in the buildings using multi-layer perceptron and random forest,” *Int. J. Adv. Sci. Technol.*, vol. 101, pp. 13–22, Apr. 2017, doi: 10.14257/ijast.2017.101.02.
[9] S. M. Sojib Ahamed, H. Imran, and I. Saral, “Load prediction of HVAC systems using deep learning,” *Glob. Netw.*, vol. 12, no. 7, pp. 928–972, 2024, doi: 10.5281/zenodo.13134333.
[10] Y. Wang, “Application of deep learning model in building energy consumption prediction,” *Comput. Intell. Neurosci.*, vol. 2022, p. 4835259, Apr. 2022, doi: 10.1155/2022/4835259.
[11] A. Bhansali, N. Narasimhulu, R. Pérez de Prado, P. B. Divakarachari, and D. L. Narayan, “A review on sustainable energy sources using machine learning and deep learning models,” *Energies*, vol. 16, no. 17, p. 6236, Aug. 2023, doi: 10.3390/en16176236.
[12] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[13] L. Deng and D. Yu, “Deep learning: Methods and applications,” *Found. Trends Signal Process.*, vol. 7, no. 3–4, pp. 197–387, 2014, doi: 10.1561/2000000039.
[14] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.
[15] J. Runge and R. Zmeureanu, “Forecasting energy use in buildings using artificial neural networks: A review,” *Energies*, vol. 12, no. 17, p. 3254, Aug. 2009, doi: 10.3390/en12173254.
[16] U. Kanewala, S. Weerakoon, and R. Nawarathna, “Exploring deep learning and tree-based ensemble models for chiller energy consumption predictions,” in *Proc. IEEE 16th Int. Conf. Ind. Inf. Syst. (ICIIS)*, 2021, pp. 306–311, doi: 10.1109/ICIIS53135.2021.9660662.
[17] W. Li and K. L. E. Law, “Deep learning models for time series forecasting: A review,” *IEEE Access*, vol. 12, pp. 92306–92327, 2024, doi: 10.1109/ACCESS.2024.3422528.
[18] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001, doi: 10.1214/aos/1013203451.
[19] Y. Zhang and A. Haghani, “A gradient boosting method to improve travel time prediction,” *Transp. Res. Part C Emerg. Technol.*, vol. 58, pp. 308–324, Sep. 2015, doi: 10.1016/j.trc.2015.02.019.
[20] G. Zheng, Z. Feng, M. Jiang, L. Tan, and Z. Wang, “Predicting the energy consumption of commercial buildings based on deep forest model and its interpretability,” *Buildings*, vol. 13, no. 9, p. 2162, Aug. 2023, doi: 10.3390/buildings13092162.
[21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009. *(Note: Your original text cited a tiny 2-page book review from 2004 instead of the actual world-renowned textbook. I changed this to point directly to the textbook itself).*
[22] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[23] M. Nalluri, M. Pentela, and N. R. Eluri, “A scalable tree boosting system: XG Boost,” *Int. J. Res. Stud. Sci. Eng. Technol.*, vol. 7, no. 12, pp. 35–41, 2020, doi: 10.22259/2349-476X.0712005.
[24] F. Wan, “XGBoost based supply chain fraud detection model,” in *Proc. IEEE 2nd Int. Conf. Big Data Artif. Intell. Internet Things Eng. (ICBAIE)*, 2021, pp. 355–358, doi: 10.1109/ICBAIE52039.2021.9390041.
[25] J. Ye, J. H. Chow, J. Chen, and Z. Zheng, “Stochastic gradient boosted distributed decision trees,” in *Proc. 18th ACM Conf. Inf. Knowl. Manage. (CIKM)*, 2009, pp. 2061–2064, doi: 10.1145/1645953.1646301.
[26] Z. A. Ali, Z. H. Abduljabbar, H. A. Tahir, A. B. Sallow, and S. M. Almufti, “eXtreme gradient boosting algorithm with machine learning: A review,” *Acad. J. Nawroz Univ.*, vol. 12, no. 2, pp. 320–334, May 2023, doi: 10.25007/ajnu.v12n2a1612.
[27] S. Ramraj, N. Uzir, R. Sunil, and S. Banerjee, “Experimenting XGBoost algorithm for prediction and classification of different datasets,” *Int. J. Control Theory Appl.*, vol. 9, no. 40, pp. 651–662, 2016.
[28] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.
[29] S. S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2009.
[30] P. W. Tien, S. Wei, J. Darkwa, C. Wood, and J. K. Calautit, “Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality – A review,” *Energy AI*, vol. 10, p. 100198, Nov. 2022, doi: 10.1016/j.egyai.2022.100198.

[31] K. Katić, R. Li, J. Verhaart, and W. Zeiler, "Neural network based predictive control of personalized heating systems," *Energy Build.*, vol. 174, pp. 199–213, Sep. 2018, doi: 10.1016/j.enbuild.2018.06.033.
[32] B. Parizad, H. Ranjbarzadeh, A. Jamali, and H. Khayyam, "An intelligent hybrid machine learning model for sustainable forecasting of home energy demand and electricity price," *Sustainability*, vol. 16, no. 6, p. 2328, Mar. 2024, doi: 10.3390/su16062328.
[33] M. Sisode and M. Devare, "A review on machine learning techniques for predictive maintenance in Industry 4.0," in *Proc. Int. Conf. Appl. Mach. Intell. Data Anal. (AMIDA)*, 2023, pp. 774–833, doi: 10.2991/978-94-6463-136-4_67.
[34] M. Kuhn and K. Johnson, *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Boca Raton, FL, USA: CRC Press, 2019, doi: 10.1080/00031305.2020.1790217.