UBC Smart City, Department of Civil Engineering

Demand Forecasting for UBC Buildings using Artificial Neural Network and Multivariate Time Series Regression

Bowen Tian, Ezekiel Camacho, Peter Kim | Dr. Omar Swei

BACKGROUND

The University of British Columbia (UBC), like many large institutions, faces the challenge of balancing growth with environmental stewardship. This study presents a data-driven approach to optimizing energy usage in existing buildings, contributing to UBC's sustainability goals, and serving as a model for other universities. Analyzing a decade of energy data from 69 buildings with diverse structures and functions, we employed both time series regression and artificial neural networks to develop models for predicting future energy demand. This analysis uncovered key factors influencing energy use, such as temperature, day of the week and time of the year. Further study of the results allowed us to identify potential optimization strategies for different building types and usage patterns under changing climate environments.

THE CAMPUS VISION 2050

We hope for this research not only to support UBC in achieving their sustainability goals, specifically, Campus Vision 2050, but also provide valuable insight and methodologies for other universities seeking to expand their community while keeping the environment safe.

Our dataset is available to all UBC students.

DATASET

METHOD

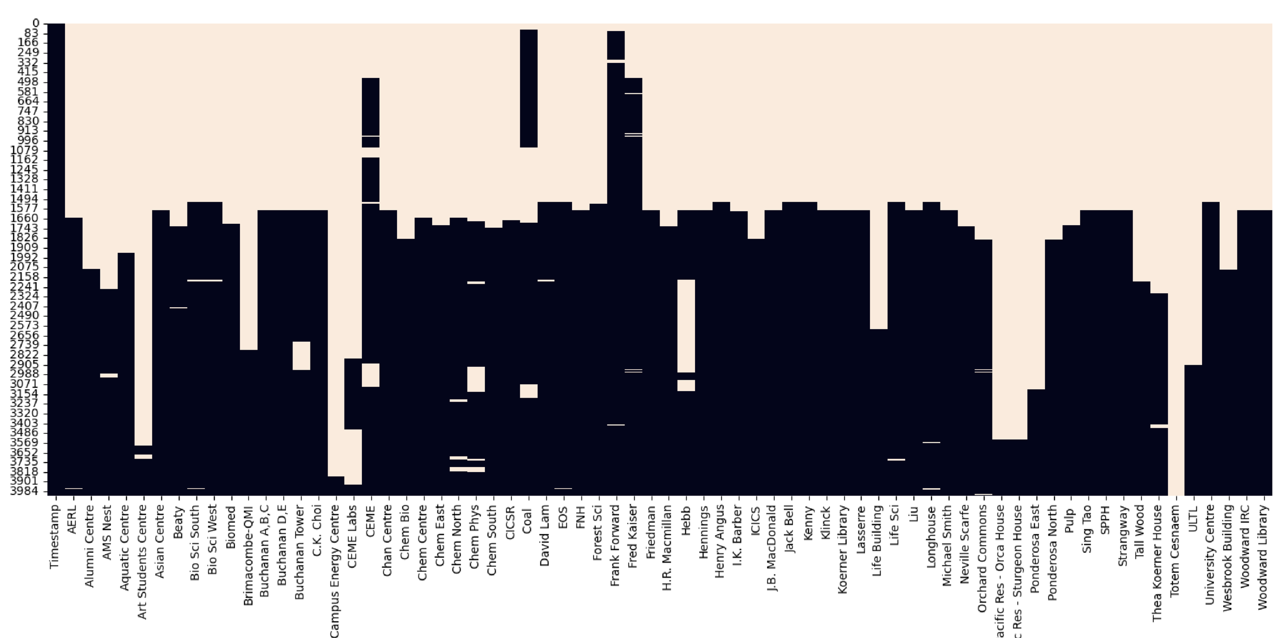


Figure 1.
Missing data visualization based on time (y-axis).

The study utilized three sources of data: UBC Skyspark, a cloud server housing building energy and feature data; UBC GeoJSON, containing detailed information about each building such as construction materials and LEED ratings; and the System Advisor Model, providing hourly weather data from 1998 to 2023. After preprocessing, the initial pool of 107 buildings was narrowed down to 69, each possessing over 10 years of data. The weather data is aggregated and standard data science procedures are followed with the combined dataset. It was found that using additional dummy variables such as day of the week vastly improves model functionality.

Multivariate time series regression and dynamic harmonic regression models were built using R, and ANN was done through Python, with energy use intensity being the response variable. Summary statistics were computed to validate model performance.

RESULTS

Initial data exploration revealed that buildings constructed after the establishment of the National Building Code in 1970 have reduced average EUI and its variability. This is attributed to an industry-wide movement to improve energy efficiencies in building operations as well as its construction. The type of the building is also a key factor. Of the 107 buildings studied, concrete was the primary construction type, with a relative frequency of 0.97. It has a higher thermal resistance compared to wood, and lower thermal resistance than steel, with an overall higher carbon footprint.

Forecast plots were generated for all 69 buildings, and for most, the predictability was high. One notable discrepancy was the pandemic period, particularly during the lockdown, where for a majority of the buildings, the energy use was overpredicted. This effect is most noticeable in large venue or operational buildings, and negligible for residential buildings. Buildings with regular energy use patterns were most predictable, with R-squared values of up to 0.9, and buildings of irregular activity spikes were poor in this regard. The most important variables were temperature and day of the week, confirming the suspicion that air conditioning and building operation were primary energy use contributors. Regression methods were better able to capture said aspects, but underperformed in predictability compared to black-box methods such as ANN.

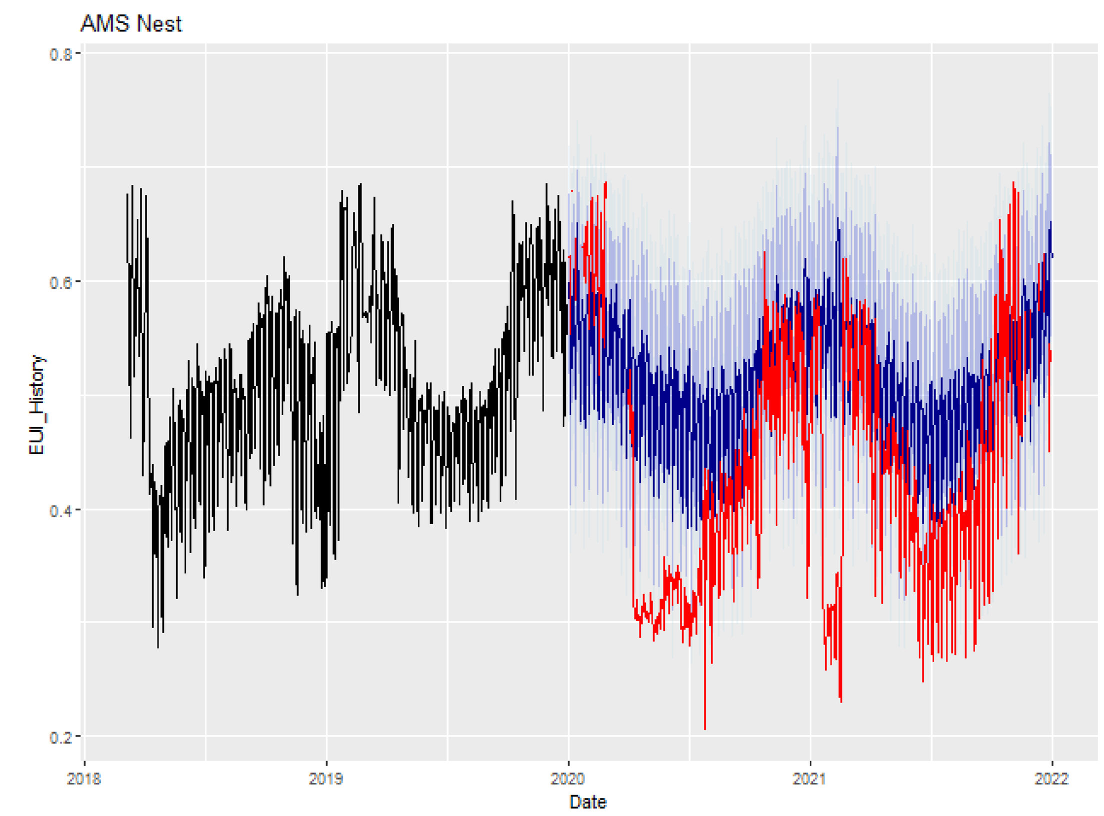


Figure 2.
AMS Nest, showing a over-prediction (in blue) over actual values (red), particularly noticeable during lock-downs. A brief under-prediction is shown in September 2021 during UBC opening.

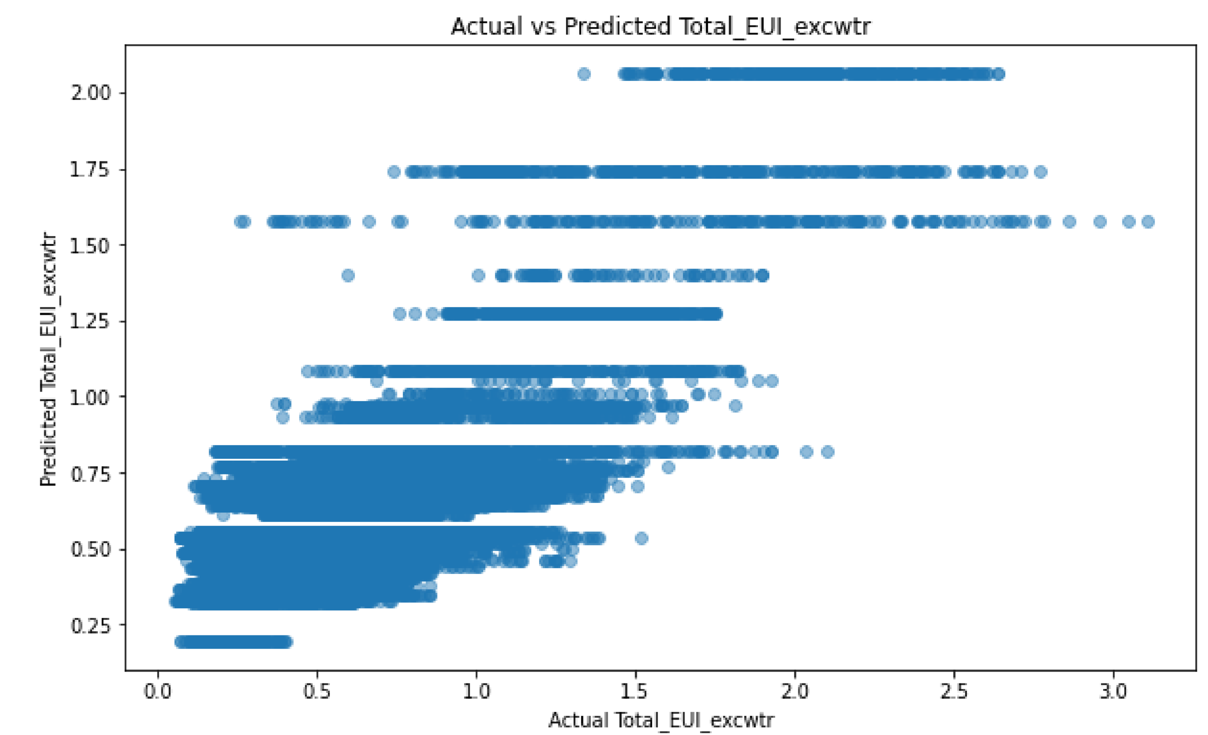


Figure 3.
Actual vs. Predicted Total EUI from ANN Model.

Artificial neural network models are applicable for building energy modelling as it allows for detection of patterns within large multivariate datasets. Input features for the ANN model was derived from the building features including height, construction type, total floors, floor area, and floor space percentage by room type. The model was configured with a rectified linear activation function (ReLU) and Adam optimizer. Construction type was manually encoded based on increasing thermal resistance of the material (wood, concrete, steel, and other mixed material in increasing order). The testing data used was 20% of the total data. The model yielded an R-squared value of 0.62 and MSE of 0.070.

DISCUSSION AND CONCLUSION

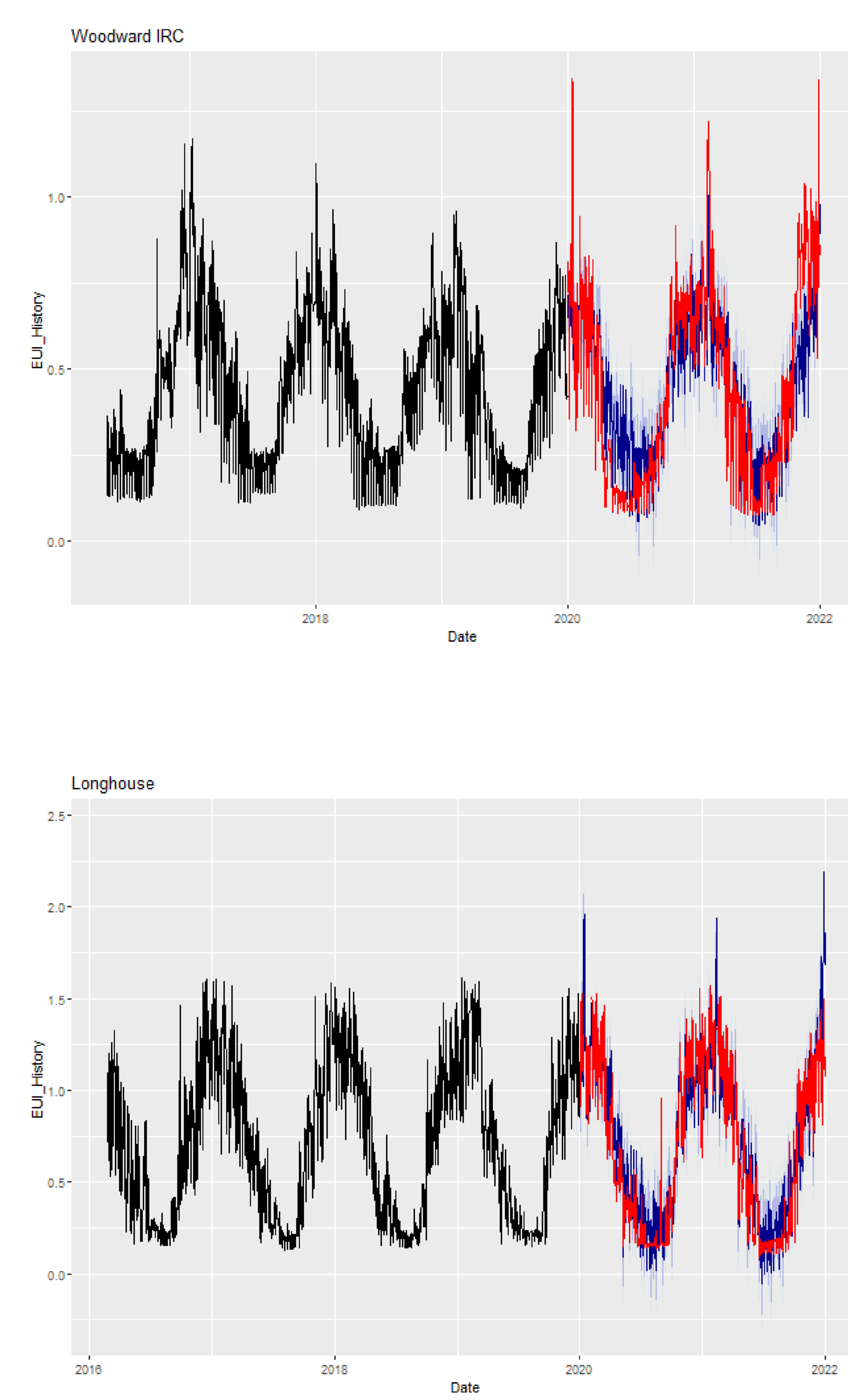


Figure 4.
Forecasted EUI for Woodward IRC (top) and Longhouse (bottom), showing regularity and high predictability.

Discussion
Residential buildings were better predicted with only weather and calendar data, and that venue buildings need additional variables which were not available, such as operational times, events and closure, in order to adequately model. Buildings with large open spaces also tend to fluctuate in energy use more based on weather conditions. Since Canada is relatively colder compared to the rest of the world, energy use is heating dominant and winter seasons consistently showed higher levels. Although the ANN model's performance is comparatively lower than other methods employed in this study, supporting research has shown its power in predictive energy modeling with sufficient inputs to accurately match the recorded EUI data.

Conclusion
Overall, the research was helpful in establishing a framework for EUI prediction and modeling. Future branches of this study may enclose in-depth analysis on the building envelope material, envelope thickness, and influence of the building geometry which may account for the errors captured in the ANN model. Weather patterns may also be explored to account for the available outside energy that can penetrate through the envelope material during different time periods. Our research could also be further developed for energy demand-supply optimization, which can be expanded upon given more time than was available to us. Weather data can also be used for renewables production such as solar and wind calculations, and an optimization model would likely be more mathematically involved, with system dynamics and agent-based modeling in effect. However, we regard this research as highly impacting as the intermittency of solar and wind power is an ongoing issue needing to be resolved.

BACKGROUND

The University of British Columbia (UBC), like many large institutions, faces the challenge of balancing growth with environmental stewardship. This study presents a data-driven approach to optimizing energy usage in existing buildings, contributing to UBC's sustainability goals, and serving as a model for other universities. Analyzing a decade of energy data from 69 buildings with diverse structures and functions, we employed both time series regression and artificial neural networks to develop models for predicting future energy demand. This analysis uncovered key factors influencing energy use, such as temperature, day of the week and time of the year. Further study of the results allowed us to identify potential optimization strategies for different building types and usage patterns under changing climate environments.

We hope for this research not only to support UBC in achieving their sustainability goals, specifically, Campus Vision 2050, but also provide valuable insight and methodologies for other universities seeking to expand their community while keeping the environment safe.

Our dataset is available to all UBC students.

THE CAMPUS VISION 2050

DATASET

METHOD

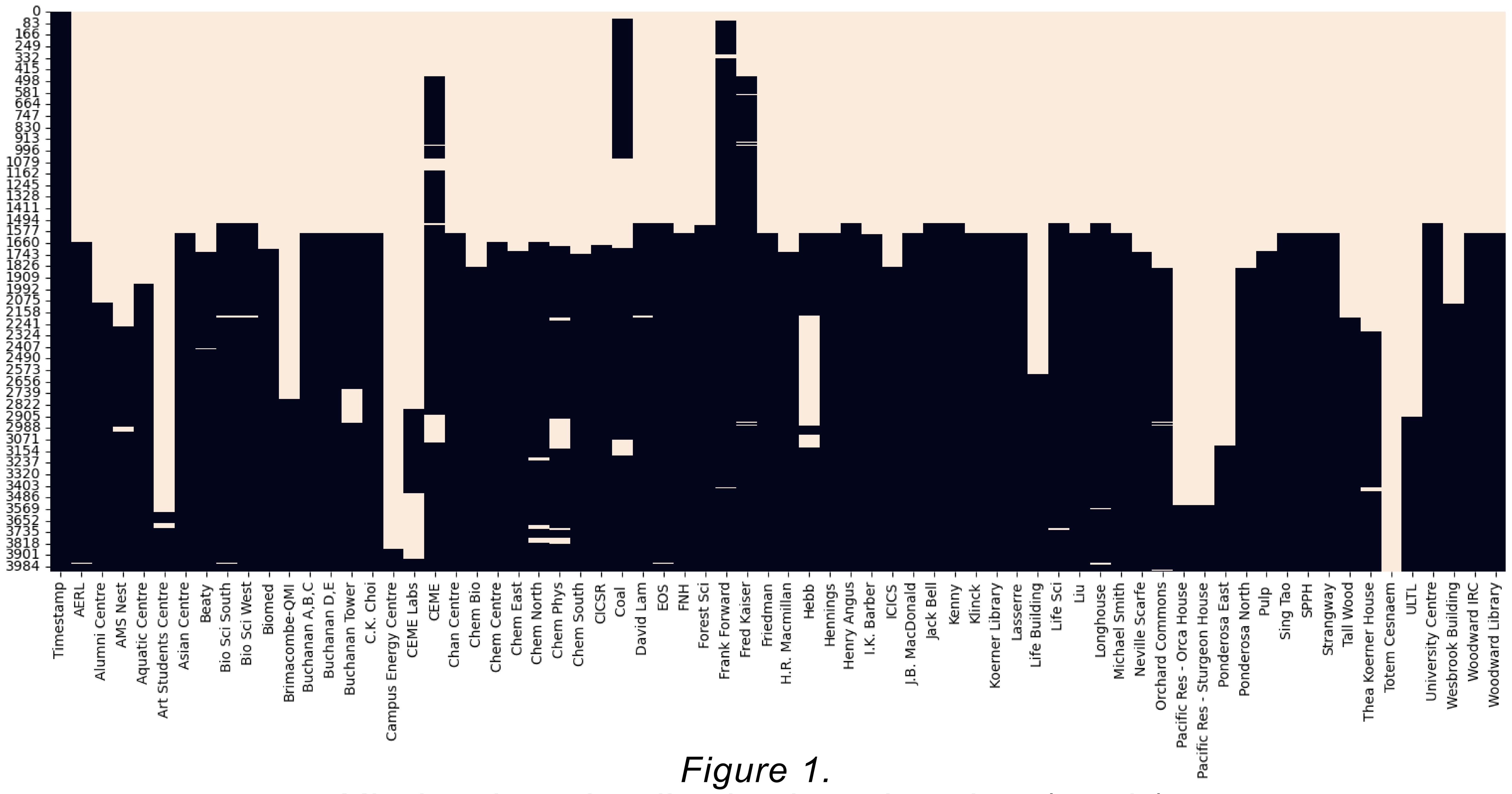


Figure 1.
Missing data visualization based on time (y-axis).

The study utilized three sources of data: UBC Skyspark, a cloud server housing building energy and feature data; UBC GeoJSON, containing detailed information about each building such as construction materials and LEED ratings; and the System Advisor Model, providing hourly weather data from 1998 to 2023. After preprocessing, the initial pool of 107 buildings was narrowed down to 69, each possessing over 10 years of data. The weather data is aggregated and standard data science procedures are followed with the combined dataset. It was found that using additional dummy variables such as day of the week vastly improves model functionality.

Multivariate time series regression and dynamic harmonic regression models were built using R, and ANN was done through Python, with energy use intensity being the response variable. Summary statistics were computed to validate model performance.

RESULTS

Initial data exploration revealed that buildings constructed after the establishment of the National Building Code in 1970 have reduced average EUI and its variability. This is attributed to an industry-wide movement to improve energy efficiencies in building operations as well as its construction. The type of the building is also a key factor. Of the 107 buildings studied, concrete was the primary construction type, with a relative frequency of 0.97. It has a higher thermal resistance compared to wood, and lower thermal resistance than steel, with an overall higher carbon footprint.

Forecast plots were generated for all 69 buildings, and for most, the predictability was high. One notable discrepancy was the pandemic period, particularly during the lockdown, where for a majority of the buildings, the energy use was overpredicted. This effect is most noticeable in large venue or operational buildings, and negligible for residential buildings. Buildings with regular energy use patterns were most predictable, with R-squared values of up to 0.9, and buildings of irregular activity spikes were poor in this regard. The most important variables were temperature and day of the week, confirming the suspicion that air conditioning and building operation were primary energy use contributors. Regression methods were better able to capture said aspects, but underperformed in predictability compared to black-box methods such as ANN.

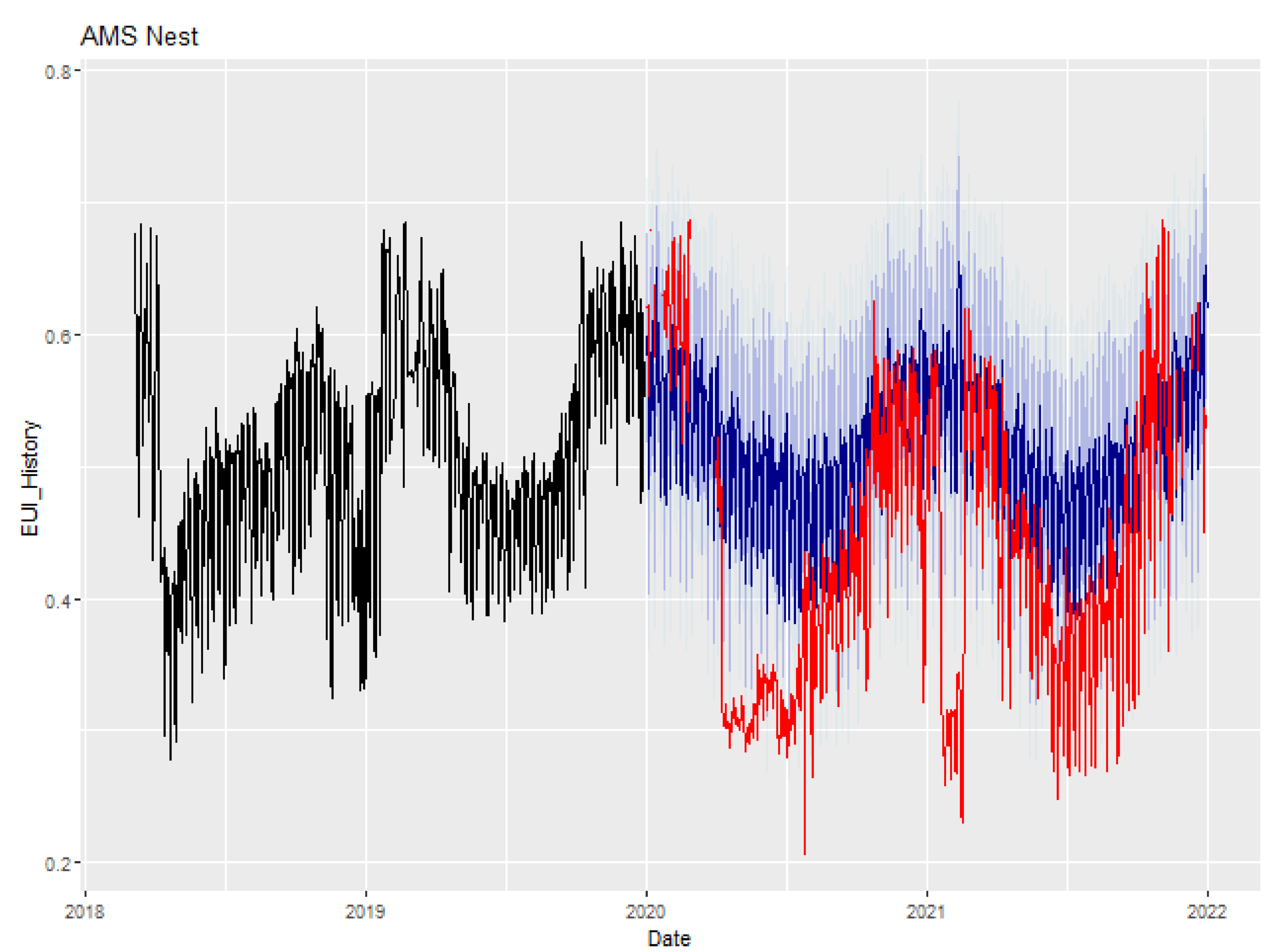


Figure 2.
AMS Nest, showing a over-prediction (in blue) over actual values (red), particularly noticeable during lock-downs. A brief under-prediction is shown in September 2021 during UBC opening.

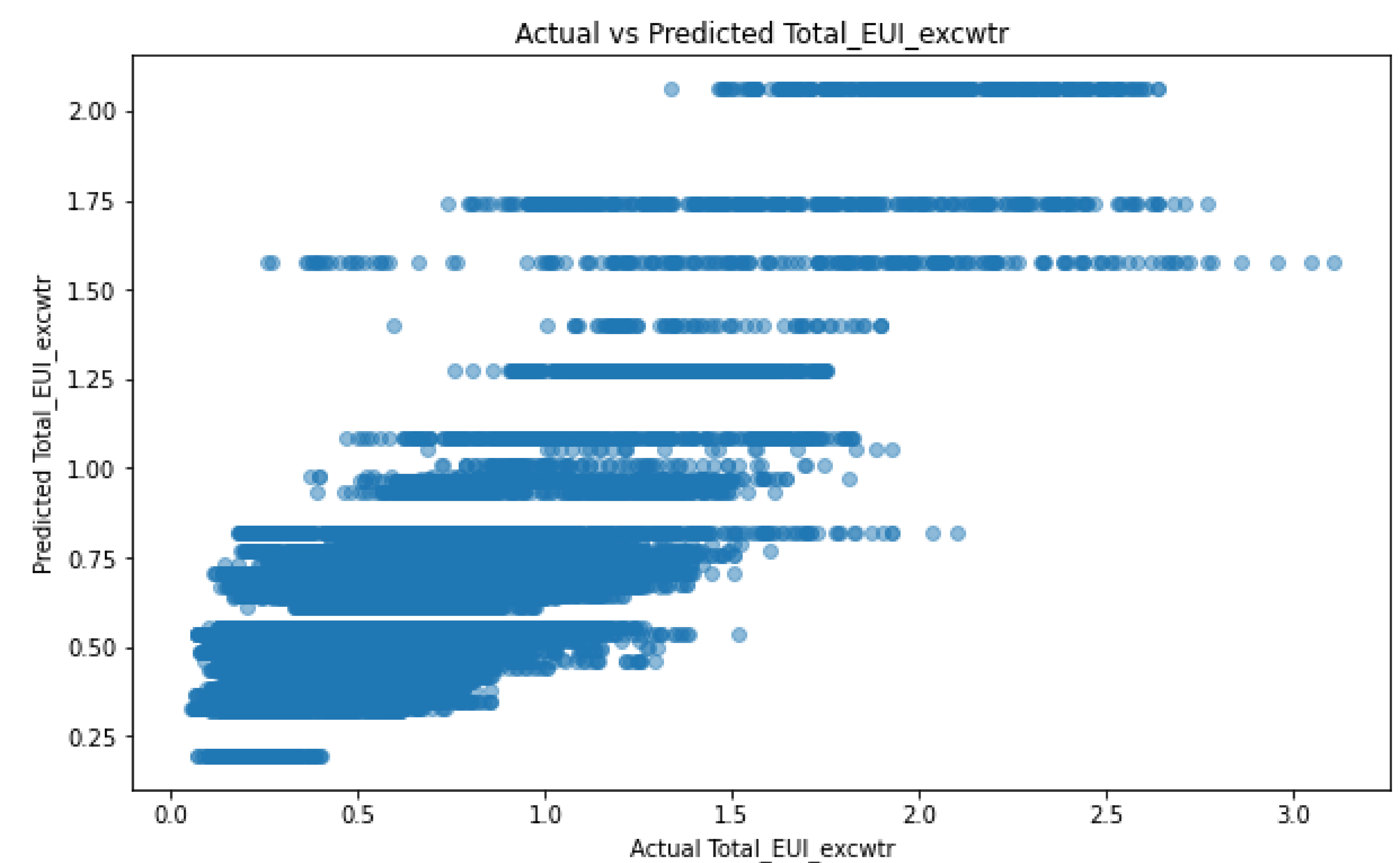
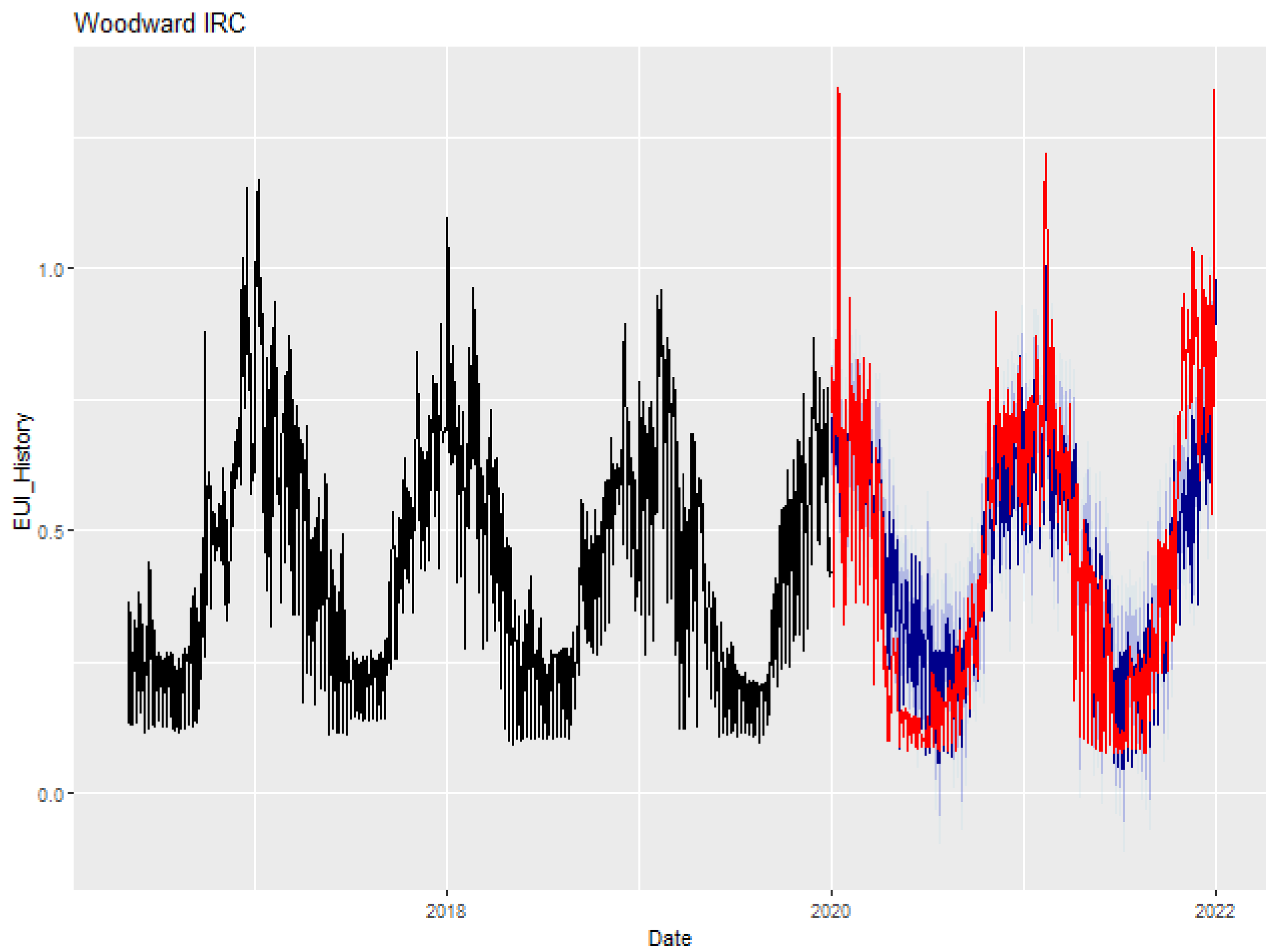


Figure 3.
Actual vs. Predicted Total EUI from ANN Model.

Artificial neural network models are applicable for building energy modelling as it allows for detection of patterns within large multivariate datasets. Input features for the ANN model was derived from the building features including height, construction type, total floors, floor area, and floor space percentage by room type. The model was configured with a rectified linear activation function (ReLU) and Adam optimizer. Construction type was manually encoded based on increasing thermal resistance of the material (wood, concrete, steel, and other mixed material in increasing order). The testing data used was 20% of the total data. The model yielded an R-squared value of 0.62 and MSE of 0.070.

DISCUSSION AND CONCLUSION



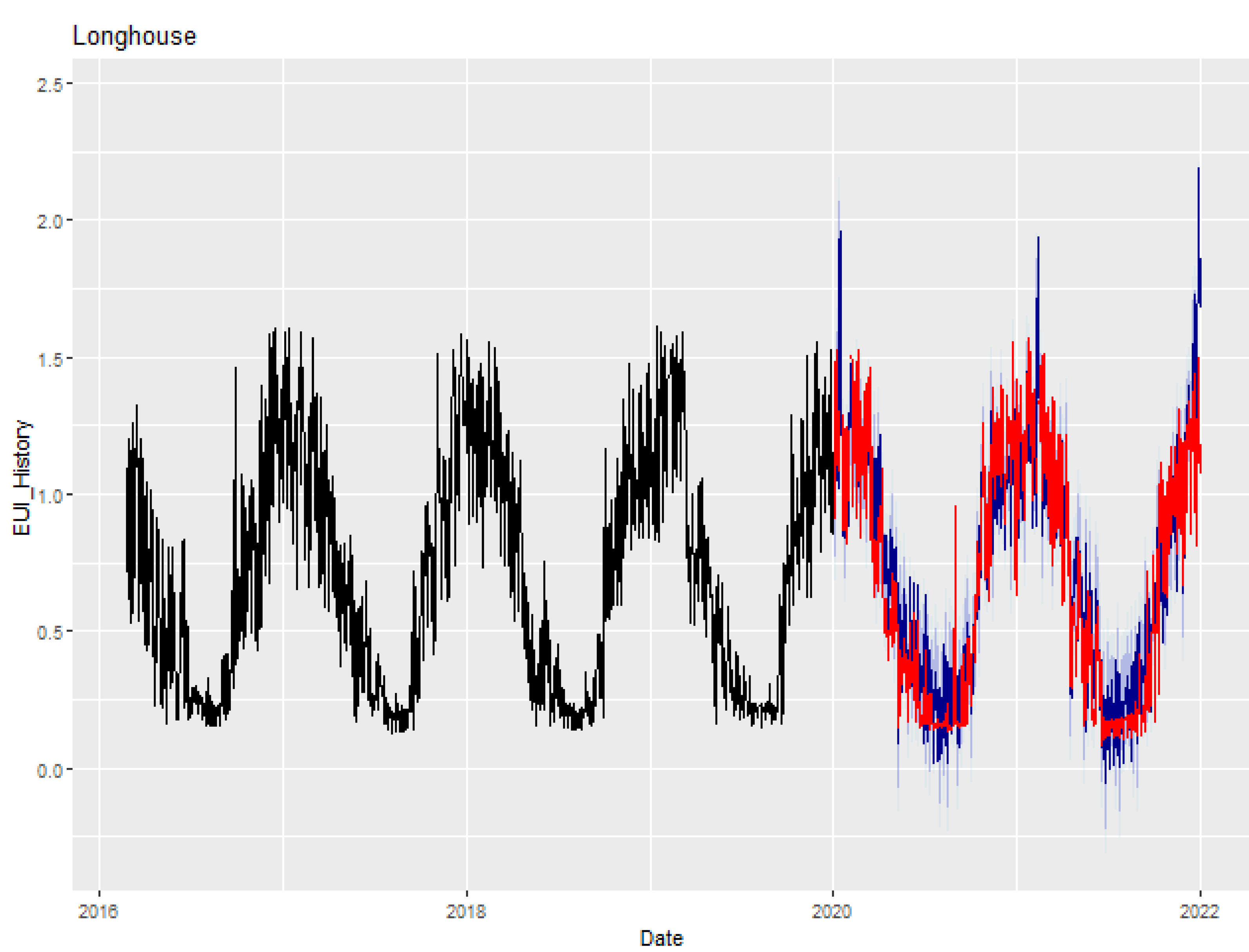


Figure 4.
Forecasted EUI for Woodward IRC (top) and Longhouse (bottom), showing regularity and high predictability.

Discussion

Residential buildings were better predicted with only weather and calendar data, and that venue buildings need additional variables which were not available, such as operational times, events and closure, in order to adequately model. Buildings with large open spaces also tend to fluctuate in energy use more based on weather conditions. Since Canada is relatively colder compared to the rest of the world, energy use is heating dominant and winter seasons consistently showed higher levels. Although the ANN model's performance is comparatively lower than other methods employed in this study, supporting research has shown its power in predictive energy modeling with sufficient inputs to accurately match the recorded EUI data.

Conclusion

Overall, the research was helpful in establishing a framework for EUI prediction and modeling. Future branches of this study may enclose in-depth analysis on the building envelope material, envelope thickness, and influence of the building geometry which may account for the errors captured in the ANN model. Weather patterns may also be explored to account for the available outside energy that can penetrate through the envelope material during different time periods. Our research could also be further developed for energy demand-supply optimization, which can be expanded upon given more time than was available to us. Weather data can also be used for renewables production such as solar and wind calculations, and an optimization model would likely be more mathematically involved, with system dynamics and agent-based modeling in effect. However, we regard this research as highly impacting as the intermittency of solar and wind power is an ongoing issue needing to be resolved.