Bridging Building Energy Performance Gap: A Hybrid Approach Integrating Deep Learning and First-Principles Models

Navid Shirzadi[1], Sara Gilani[1], Ali Seyd[1], Chris Kirney[1], Phylroy Lopez[1], Meli Stylianou[1]
[1]CanmetENERGY-Ottawa, Natural Resources Canada, Ottawa, Canada

Abstract

Given the building sector's significant influence on global energy consumption and carbon emissions, especially within existing building stock, it holds a pivotal role in shaping a more sustainable future. Accurately predicting an existing building's energy consumption becomes crucial for effective resource management, environmental impact assessment, resilience planning, and tracking progress towards sustainability goals. Bridging the gap between simulated and actual energy consumption is a key strategy in improving building energy performance. However, practically, this task proves challenging. To tackle this challenge, our research introduces a hybrid model that combines deep learning, specifically Long Short-Term Memory (LSTM), with a foundational first-principles model. This integrated approach aims to provide a comprehensive understanding of building energy consumption in daily future horizons. In our study, we delve into an extensive analysis of the impact of various building attributes on energy consumption, including cooling and heating usage patterns. Furthermore, our investigation of meteorological features such as outdoor temperature, reveals their influence on energy dynamics within the building. Our results not only demonstrate the hybrid model's improved accuracy and reduction of the building energy performance gap but also highlight the significance of each attribute in the prediction process during the feature selection. The results show that the integrated proposed approach could lower the error by about 28% compared to the single deep learning model.

Introduction

Buildings play a crucial role in global energy consumption, accounting for approximately 30% of the world's final energy consumption and contributing to approximately 26% of global emissions (*Buildings - Energy System - IEA*, n.d.). The growing trend of urbanization will lead to an increase in energy consumption within the building sector, highlighting the importance of energy management in this critical area (Zou et al., 2018).

To assess the impact of energy conservation measures (ECMs) and retrofit scenarios on a building's energy consumption, it is essential to quantify the building's energy usage. This can be accomplished through various methods, ranging from steady-state calculations to dynamic modeling techniques, which forecast the building's consumption over time (van Dronkelaar et al., 2016). Numerous previous studies have consistently demonstrated a disparity between predicted and actual energy consumption values in buildings which could be caused by several reasons (Menezes et al., 2012). Several studies have identified the design and simulation phase as the primary contributor to the performance gap (de Wilde, 2014; Imam et al., 2017). Two primary approaches are commonly employed for modeling building energy consumption: first principles, which are physics-based simulations, and data-driven approaches, which typically utilize black box techniques (de Wilde, 2014).

Building performance simulation (BPS) tools have been utilized for several decades to forecast and analyze the energy consumption of buildings, with a focus on heat transfer through building envelopes (Choi et al., 2023). EnergyPlus, TRNSYS, and e-QUEST are among the primary tools used for energy simulation. EnergyPlus, in particular, is a free and comprehensive whole-building simulation tool capable of simulating buildings using a network of nodes (Al-janabi et al., 2019). This feature provides users with flexibility to assess the impact of ECMs on the final energy performance.

Several factors influence a building's energy consumption, including the impact of neighboring buildings and microclimates, which can contribute to an energy performance gap while using energy simulation tools. Additionally, modeling the same building with different simulation tools can result in significant discrepancies in predicted energy consumption, as indicated by previous studies (Schwartz & Raslan, 2013; Zhu et al., 2013). To decrease the amount of this gap, modelers try to calibrate the simulation model. Li et al. (2018) attempted to reduce the number of parameters requiring calibration from 14 to 6 using the Morris sensitivity analysis, thereby enhancing the efficiency of the calibration process. Subsequently, they conducted stepwise calibration. Their findings indicated a substantial disparity between the uncalibrated model and the measured values. However, after calibration, improvements were observed, with significantly reduced errors. Despite

these enhancements, it appears that the calibrated model still struggles to predict most peak loads.

In (Royapoor & Roskilly, 2015), the authors attempted to utilize a collection of calibrated sensors within an office building to calibrate the simulation model and achieve an acceptable level of accuracy as per their chosen standard. While they successfully demonstrated that their method reduced the total error to an acceptable level, they did not present the accuracy of the calibrated model at higher resolutions, such as hourly intervals, in their study.

Data-driven methods have also been extensively utilized by researchers, particularly in recent years, to predict building energy performance. These methods can be categorized into two main categories, as described by (Wang et al., 2019): single models such as artificial neural networks (ANN) (Safa et al., 2017) or Decision Tree (DT) models (Yu et al., 2010) which requires parameter tunning and data preprocessing, or ensemble methods such as integrating clustering techniques with ANN or support vector machine (SVM) models to classify input variables (Ding et al., 2018).

Based on the discussion, it is evident that both the data-driven and physics-based techniques employ different approaches to predict the energy performance of buildings. However, in both cases, uncertainties remain high, and the level of the performance gap is often deemed unacceptable. To the best of our knowledge, only a few studies have focused on integrating machine learning and simulation models (Chen et al., 2022).

In this study, we propose a framework to address the initial uncertainties of the simulation model by pre-calibrating the physics-based model. Subsequently, we propose a methodology aimed at increasing the accuracy of building energy consumption prediction and reducing the performance gap by integrating a deep learning model with the physics-based simulation model. Additionally, we introduce an algorithm for ranking the importance of simulation model output variables and conducting proper feature selection. This integrated approach aims to harness the strengths of both machine learning and physics-based techniques to improve the accuracy of energy performance prediction in buildings.

The paper begins by presenting the related methodology, detailing the overall framework and associated algorithms for achieving the final integration model. Following that, it outlines the selected building case study and its measured data. Then, the results and their evaluation are discussed. Finally, the paper concludes and offers insights into future work.

Methods

In this section, the process of employing an integrated model to forecast future energy consumption in buildings is outlined, as illustrated in Figure 1.

Initially, data from the selected case study is collected and preprocessed to ensure accuracy. This involves gathering information on building energy usage, climatic conditions, and various building parameters, which are then stored in a dedicated database.

Subsequently, two main paths are pursued: one involves using the building characteristics to simulate energy usage with the EnergyPlus model, with the results of this simulation serving as input features for the LSTM model. The other path entails selecting key features from the energy and weather data, which are then utilized in training the Long Short-Term Memory (LSTM) model.

Following this, the parameters of the LSTM model are fine-tuned through a hyperparameter tuning stage to enhance its predictive capability.

Finally, the performance of the model is evaluated using different metrics, and the results are compared with actual data to validate the forecasting approach.

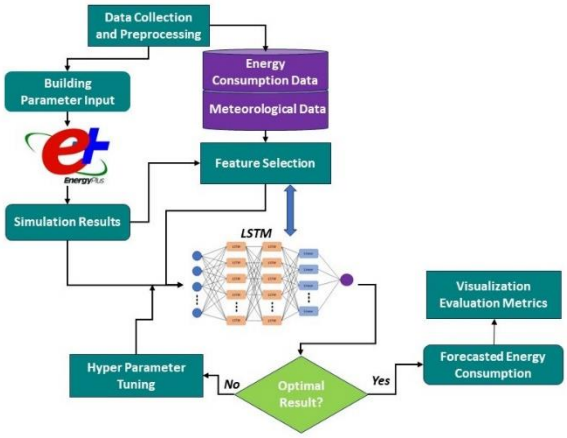

Figure 1: Proposed hybridized framework.

Preprocessing

The initial data is separated into minutely resolution energy demand and consumption data with a Unix timestep. Therefore, as the first step, the data is transformed into date-time timeseries format and then aggregated into hourly resolution.

To clean the aggregated data, the first step involves checking for missing values, which are then replaced with the average of the adjacent values.

Next, the data is visualized using boxplots to detect possible outliers. Any outliers identified are removed from the dataset by defining the first quantile (Q1) as 0.25 and the third quantile (Q3) as 0.75. The lower bound and upper bound are then calculated using the formula (1):

$$Lower\ Bound = Q1 - multiplier \times IQR$$
$$Upper\ Bound = Q3 + multiplier \times IQR \quad (1)$$

where IQR is the interquartile range, calculated as the difference between Q3 and Q1, and the multiplier is set to 1.5.

The non-time series data, encompassing descriptive information about the building such as building parameters, equipment specifications, and details regarding the heating, ventilation, and air conditioning (HVAC) system, are extracted from relevant documents. These data are essential inputs for the building simulation model, contributing to the comprehensive analysis of energy consumption and performance within the building environment.

Building Simulation Model

In this research, EnergyPlus 23.2.0 (*EnergyPlus*, n.d.) serves as the primary tool for simulating the building. The initial step in simulating a building's physics-based model involves preparing the geometry, which serves as input for the EnergyPlus model.

To accomplish this, the SketchUp 2023 (*3D Design Software | 3D Modeling on the Web | SketchUp*, n.d.) tool is utilized for creating the geometry and defining the various zones within the building. This process enables the accurate representation of the building structure and layout, facilitating comprehensive simulations of energy consumption and thermal performance.

To model the portion of the building geometry in contact with the ground, the Kiva (Kruis, 2015) tool is employed to generate the hourly temperature profiles of the wall's effective temperature. Subsequently, these results are aggregated to monthly temperatures using the average method. The entire year's ground temperature is then manually inputted into EnergyPlus to ensure accurate simulation of ground contact effects on the building's thermal performance.

The majority of the building schedules are defined based on the metadata provided by the data provider. In cases where such metadata is unavailable, schedules are created adhering to the standards outlined in the National Energy Code of Canada for Buildings 2020 (NECB) (Canadian Commission on Building and Fire Codes & Canada. Natural Resources Canada, 2020.) or the National Building Code (NBC) (Canadian Commission on Building and Fire Codes, n.d.). This approach ensures that the building schedules follow established industry standards and regulatory requirements, facilitating accurate simulation and analysis of energy consumption patterns.

The "Flow Coefficient" model is employed as the infiltration object within the EnergyPlus model. This model utilizes formula (2) (*EnergyPlus*, n.d.) to calculate the infiltration rate for each zone, facilitating accurate representation of air leakage and infiltration dynamics within the building environment.

$$Infiltration = F_{Schedule}\sqrt{(cC_s\Delta T)^2 + (cC_W(s \times WindSpeed)^{2n})^2} \quad (2)$$

where the F_Schedule represents a fraction between 0 and 1, indicating the portion of the house's volume, and remains constant throughout the year. C_S, C_W and s are stack coefficient, wind coefficient and shelter factor respectively which are extracted from ASHRAE Handbook of Fundamentals (American Society of Heating Refrigerating and Air-Conditioning Engineers Inc, 2017). The wind speed data is sourced from the EnergyPlus weather file, and it is recorded in meters per second (m/s) unit. The pressure exponent n is typically assumed to be 0.67. ΔT represents the average temperature difference between the zone's air and outdoor air. The flow coefficient c is calculated as follows:

$$c = \frac{ACH}{\Delta P^n} \quad (3)$$

where ACH represents the air changes per hour and ΔP denotes the pressure difference between indoor and outdoor. Both of these parameters are typically calculated based on the blower door test data. However, in the absence of such data, standard values are considered from the NBC.

.

Feature Importance and Selection

An essential step in the integration process is feature selection. With numerous output variables generated by the EnergyPlus model, selecting the most relevant ones is crucial for achieving optimal prediction accuracy. Utilizing a large number of features may lead to the curse of dimensionality, along with increased computational costs. Therefore, feature selection in this study is conducted in two stages.

In the first stage, a simple correlation strategy is employed to assess the initial importance of each EnergyPlus output variable. This helps in identifying the variables that exhibit significant correlations with the target outcome. Non-linear correlation is assessed using the Spearman method.

In the second stage, the importance of each feature is evaluated within the training process of the deep learning model. This step further refines the selection by considering how each feature contributes to the model's predictive performance. By employing these two stages of feature selection, we aim to identify and retain the most informative variables while mitigating the challenges associated with dimensionality and computational complexity.

Deep Learning Model

In this study, the Long Short-Term Memory (LSTM) method serves as the core prediction algorithm, utilizing the outputs of the physics-based model as predictors for the LSTM. Unlike simple neural network modules that typically include only one activation function, LSTM modules are composed of multiple layers. Each layer within the LSTM is designed to process input values differently, making it

particularly well-suited for time series prediction tasks. This multi-layered architecture allows the LSTM to effectively capture and learn complex temporal patterns present in the data, enhancing its predictive capabilities compared to simpler neural network architectures.

An LSTM module consist of input layer, forget layer, input gate layer, cell state update, output gate layer and hidden state output (Shirzadi et al., 2023). The forget layer in LSTM models plays a crucial role in addressing long-term dependencies by selectively discarding irrelevant information from the past (Gers et al., 2000). This functionality is particularly beneficial when dealing with energy consumption data, which may contain non-repetitive power outages or false data recorded by sensors. The ability of the LSTM model to selectively retain pertinent information while filtering out outdated or erroneous data enables it to dynamically adapt and sustain accurate predictions over time and mitigate the impact of the possible noises. The input gate layer within the LSTM model is also advantageous for filtering out noisy information and prioritizing relevant data. This functionality enables the model to effectively capture temporal dependencies and learn the seasonality and trends present in the data. To control the network's internal memory, the LSTM employs a cell state update layer. This layer dynamically adjusts the internal memory using the current input and the preceding cell state, enabling the model to selectively preserve pertinent information while filtering out less significant details. Ultimately, the output gate layer dictates the cell's output and decides which information should be passed to the next hidden state from the cell state. In the hidden state output, the LSTM cell's final output is calculated by passing it through the tanh function. This enables the LSTM to capture intricate temporal dynamics, ensuring that the model can effectively encode and process complex patterns in the data (Hochreiter & Schmidhuber, 1997).

The LSTM architecture employed in this study comprises an input layer, one or more LSTM layers, and a linear layer (fully connected layer) positioned before the output layer. The inclusion of the linear layer alongside the LSTM layer(s) introduces additional non-linearities, facilitating improved learning of complex relationships between the LSTM hidden state and the final output (Hochreiter & Schmidhuber, 1997). The LSTM architecture is constructed utilizing the PyTorch library (Paszke et al., n.d.), which is a popular open-source framework for deep learning.

Hyperparameter Tuning

Some of the neural network hyperparameters were tuned using the Bayesian optimization technique with an initial exploration of 5 points and a total of 10 iterations. This method utilizes a probabilistic model to estimate the objective function and select hyperparameters that strike a balance between exploration and exploitation. In this tuning scenario, the size of the hidden layer and fully connected layer, along with the learning rate, were the parameters considered for tuning.

To mitigate overfitting, the learning rate reduction schedule technique is implemented based on the previously tuned learning rate. In this approach, 20 percent of the training data is reserved for validation purposes during each iteration (epoch). A multiplier is then defined to adjust the learning rate, decreasing it if there is no improvement observed in the validation loss. This schedule is implemented using one of the PyTorch sub-packages called "ReduceLROnPlateau." Specifically, the patience parameter is set to 5, indicating the number of epochs to wait before reducing the learning rate, while the reduction multiplier is set to 0.9.

Evaluation Metrics

To assess the final results of energy demand prediction and analyze the performance gap, two primary metrics are chosen based on the characteristics of the data: root mean square error (RMSE) and mean absolute error (MAE). These metrics offer insights into prediction accuracy by measuring the disparities between predicted and actual energy demand values.

The RMSE is computed using equation 4, while the MAE is determined using equation 5.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)} \quad (4)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

where, y_i is the actual value and $\hat{y}_i$ is the predicted value for the i^{th} observation and n is the number of observations.

Case Study and Data

Building Description

The selected case study for this research is a residential building located in the greater Vancouver area of British Columbia, Canada. Constructed in 1955, the building underwent significant renovation between 2005 and 2006. It consists of one level above grade and one basement, with a rental unit situated within the basement. The total floor area of the building is 199 m^2, with approximately 99.5 m^2 per floor, while the rental unit in the basement occupies about 56 m^2. Further details regarding the building's construction materials and renovations can be found in (Makonin et al., 2016).

Given that the heat pump and gas furnace share the same loop, a control strategy is developed using EnergyPlus' energy management system (EMS), employing EnergyPlus runtime language (Erl). The gas furnace is to operate only

when the temperature falls below 2°C, while the heat pump should not operate when the outdoor temperature is less than 2°C.

Data

The electricity and natural gas data were collected from mid-2012 to mid-2014 at a minute resolution. The data is openly available and can be accessed at the following sources: (*AMPds: Almanac of Minutely Power Dataset (R2013) - Harvard Dataverse*, n.d.; Makonin et al., 2016). After preprocessing the data, the data related to the year 2013 was extracted for training and testing the model with the prediction horizon of 48 hours. The entire year, with the exception of the final two days designated as the test set, is regarded as the training set.

Results

After modeling the building in EnergyPlus, several variables were initially selected. Subsequently, correlation analysis was conducted to identify and remove non-correlated variables. The results of the Spearman correlation analyses are presented in Figure 2.

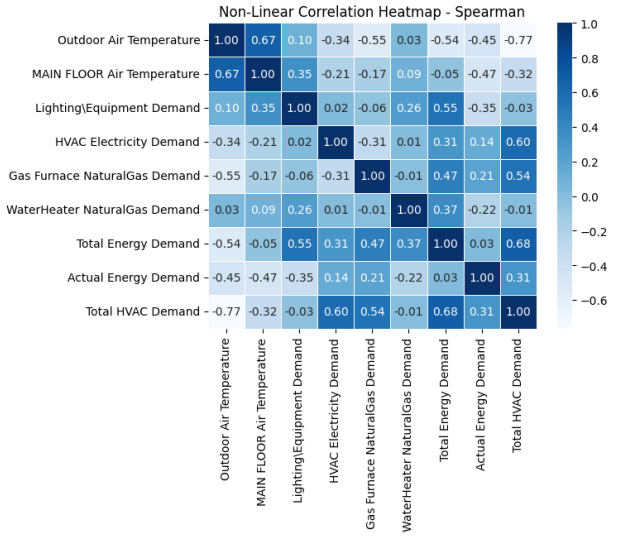

Figure 2: Correlation analysis between actual energy demand and simulation variables.

In this correlation analysis, the observed variables include outdoor air temperature and actual energy demand, while the other predictors are the outputs from the simulation model. As depicted in Figure 2, although there is no strong correlation between the actual energy demand and the simulation variables, moderate positive and negative correlations are observed with outdoor and indoor air temperature, as well as with water heater natural gas demand and gas furnace natural gas demand.

To further assess the importance of each variable, including additional variables extracted from the climatic data sourced from the EnergyPlus weather file, another importance assessment procedure is employed, as explained in the methodology section. The results of this assessment are illustrated in Figure 3.

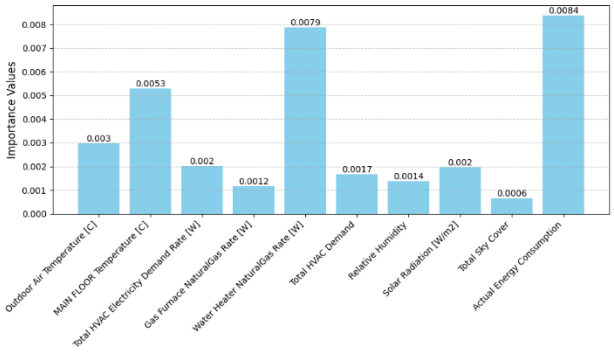

Figure 3: Importance assessment of variables, including climatic data extracted from the EnergyPlus weather file.

In Figure 3, the importance of each variable is depicted, which differs from the correlation assessment results. In this assessment, apart from actual energy consumption itself, the air temperature of the main floor of the building, along with the water heater natural gas rate, are identified as the most valuable predictors in the training process. This differs from the correlation analysis, where the natural gas rate for the water heater did not exhibit a significant correlation, However, it's important to note that the importance value of the water heater natural gas rate may be slightly exaggerated due to the presence of numerous zeros in this predictor. Even after scaling, these zeros may remain unchanged, and their sensitivity to perturbations could lead to potential distortion in the importance calculation.

After selecting the relevant features based on their importance, the chosen predictors along with the target variable (actual energy demand) are fed into the developed LSTM model.

To determine an appropriate lookback period (window size), three options were considered. Option 1 assumes a window size of one day, while for options 2 and 3, window sizes of two days and one week were considered, respectively. The impact of changing the lookback period is presented in Table 1.

Table 1: Impact of changing lookback period on training accuracy.

	Window Size (hour)	**MAE**	**Computing Time (s)**
Option1	24	3.625	452
Option2	48	1.552	886
Option3	168	1.476	12,515

Based on the results presented in Table 1, given that the prediction horizon is 48 hours, selecting a lookback window of 24 hours results in forecasting based on previous forecasts, leading to higher MAE. However, increasing the window size to match the prediction horizon (48 hours) significantly reduces the MAE, while maintaining acceptable computational time. Conversely, extending the

lookback period to one week drastically increases computational time, with only a slight improvement in error. Hence, a window size of 48 hours is chosen as the optimal selection.

The developed LSTM model was then fine-tuned using the Bayesian optimization method, as previously explained, resulting in the final tuned hyperparameters presented in Table 2. The inclusion of a second LSTM layer was not considered in the optimization process. Manual examination revealed that adding another LSTM layer could increase the complexity of the neural network without necessarily improving accuracy; in fact, it had the opposite effect. As a result, the second LSTM layer was not included in the hyperparameter tuning optimization process.

Table 2: Final tuned hyperparameters of the LSTM model

Learning Rate	LSTM Layer Size 1	LSTM Layer Size 2	Linear Layer Size
0.002787	91	NA	40

The tuned LSTM model is subsequently trained using two sets of data based on the feature importance procedure. Firstly, the single LSTM model is trained solely with the cleaned actual energy consumption values along with only outdoor air temperature as the predictor. Then, the LSTM model is trained using all the other selected predictors identified in the feature selection process. The results will then be compared with both the building simulation results and the actual energy consumption data.

The prediction results for 48 hours into the future are depicted in Figure 4. The output indicates that the integrated model outperforms the other two single models (single LSTM and EnergyPlus simulation). Notably, there are two main peaks in the actual data and several valleys, with the hybrid model consistently exhibiting superior forecasting quality. This improvement is attributed to the hybrid model's ability to learn from both actual data and simulation results. However, it's worth noting that the first peak remains significantly higher than the predicted value, suggesting that the model may require additional information to accurately forecast such extreme peaks.

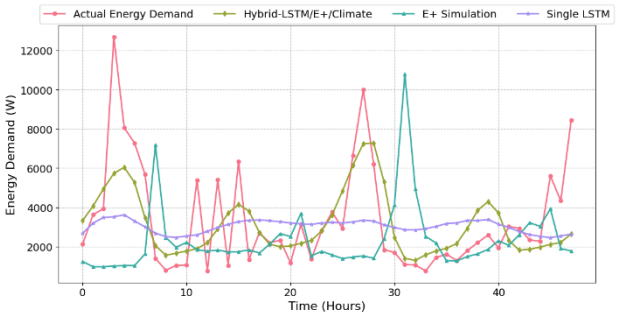

Figure 4: Prediction results for 48 hours into the future, comparing the performance of the integrated model with single LSTM and EnergyPlus simulation models.

To further evaluate the results and compare the forecasting accuracy in a more detailed manner, evaluation metrics are calculated for each model, and the results are presented in Table 3.

Table 3: Forecasting accuracy comparison between EnergyPlus, LSTM and hybrid model.

Model	MAE	RMSE	Computing Time (s)
EnergyPlus Simulation	2.517	3.692	10
Single LSTM	1.920	2.590	307
Hybrid Model	1.381	1.993	590

While utilizing a building simulation model may offer computational cost efficiency, there remains a considerable gap between the predicted energy demand and the actual values. Conversely, while employing an LSTM model can enhance prediction accuracy, it still encounters significant challenges in forecasting peaks, as evident in Figure 4. Employing a hybrid model not only aids in forecasting peaks and valleys but also reduces the MAE by 45% and 28% respectively compared to the building simulation and single LSTM model.

Conclusion

This study investigated energy consumption prediction for a residential building in the greater Vancouver area of British Columbia, Canada. Energy data, collected from mid-2012 to mid-2014 at minute resolution, was preprocessed for analysis.

The study considered the HVAC system as a critical component affecting energy consumption. The heat pump and gas furnace were incorporated into a single loop, with two separate air terminals serving the living zones. A control strategy was developed using EnergyPlus' energy management system (EMS), ensuring that the gas furnace operates only when temperatures fall below 2 °C and the heat pump is inactive when outdoor temperatures are less than 2 °C.

The research employed an integrated approach, incorporating building simulation models and deep learning techniques, specifically LSTM models, for energy demand prediction. Through feature selection, the importance of each variable was assessed, revealing that the air temperature of the main floor and water heater natural gas rate were the most influential predictors in the training process. Hyperparameter tuning further optimized the LSTM model, which was then trained using data from 2013.

Results demonstrated that while using a building simulation model offers computational cost efficiency, it fails to accurately predict energy demand. Conversely, employing an LSTM model improves energy prediction accuracy but struggles with forecasting peaks. The hybrid model, integrating actual data and simulation outputs, significantly improves forecasting accuracy, reducing MAE by 45% and

28% compared to the building simulation and single LSTM models, respectively.

Overall, this research underscores the importance of integrating actual data and simulation outputs for more accurate energy demand prediction in residential buildings, offering valuable insights for energy management and optimization in the built environment while highlighting the crucial role of HVAC systems in energy consumption prediction and optimization strategies.

Acknowledgement

Thanks to Natural Resources Canada's Office of Energy Research and Development for funding this research.

References

3D Design Software | 3D Modeling on the Web | SketchUp. (n.d.). Retrieved February 14, 2024, from https://www.sketchup.com/

Al-janabi, A., Kavgic, M., Mohammadzadeh, A., & Azzouz, A. (2019). Comparison of EnergyPlus and IES to model a complex university building using three scenarios: Free-floating, ideal air load system, and detailed. *Journal of Building Engineering*, *22*, 262–280. https://doi.org/10.1016/j.jobe.2018.12.022

American Society of Heating Refrigerating and Air-Conditioning Engineers Inc. (2017). *2017 ASHRAE® Handbook - Fundamentals / American Society of Heating Refrigerating and Air-Conditioning Engineers* (SI edition).

AMPds: Almanac of Minutely Power dataset (R2013) - Harvard Dataverse. (n.d.). Retrieved February 22, 2024, from https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/MXB7VO

Buildings - Energy System - IEA. (n.d.). Retrieved February 22, 2024, from https://www.iea.org/energy-system/buildings

Canadian Commission on Building and Fire Codes. (n.d.). *National Building Code of Canada 2020*.

Canadian Commission on Building and Fire Codes, & Canada. Natural Resources Canada. (n.d.). *National Energy Code of Canada for Buildings 2020*.

Chen, X., Guo, T., Kriegel, M., & Geyer, P. (2022). A hybrid-model forecasting framework for reducing the building energy performance gap. *Advanced Engineering Informatics*, *52*. https://doi.org/10.1016/j.aei.2022.101627

Choi, K., Park, S., Joe, J., Kim, S. I., Jo, J. H., Kim, E. J., & Cho, Y. H. (2023). Review of infiltration and airflow models in building energy simulations for providing guidelines to building energy modelers. In *Renewable and Sustainable Energy Reviews* (Vol. 181). Elsevier Ltd. https://doi.org/10.1016/j.rser.2023.113327

de Wilde, P. (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, *41*, 40–49. https://doi.org/https://doi.org/10.1016/j.autcon.2014.02.009

Ding, Y., Zhang, Q., Yuan, T., & Yang, F. (2018). Effect of input variables on cooling load prediction accuracy of an office building. *Applied Thermal Engineering*, *128*, 225–234. https://doi.org/10.1016/j.applthermaleng.2017.09.007

EnergyPlus. (n.d.). Retrieved February 13, 2024, from https://energyplus.net/

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, *12*(10), 2451–2471. https://doi.org/10.1162/089976600300015015

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Imam, S., Coley, D. A., & Walker, I. (2017). The building performance gap: Are modellers literate? *Building Services Engineering Research and Technology*, *38*(3), 351–375. https://doi.org/10.1177/0143624416684641

Kruis, N. (2015). *Development and Application of a Numerical Framework for Improving Building Foundation Heat Transfer Calculations*. University of Colorado.

Li, W., Tian, Z., Lu, Y., & Fu, F. (2018). Stepwise calibration for residential building thermal performance model using hourly heat consumption data. *Energy and Buildings*, *181*, 10–25. https://doi.org/10.1016/j.enbuild.2018.10.001

Makonin, S., Ellert, B., Bajić, I. V., & Popowich, F. (2016). Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.37

Menezes, A. C., Cripps, A., Bouchlaghem, D., & Buswell, R. (2012). Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap. *Applied Energy*, *97*, 355–364. https://doi.org/10.1016/j.apenergy.2011.11.075

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., Facebook, Z. D., Research, A. I., Lin, Z., Desmaison, A., Antiga, L., Srl, O., & Lerer, A. (n.d.). *Automatic differentiation in PyTorch*.

Royapoor, M., & Roskilly, T. (2015). Building model calibration using energy and environmental data.

Energy and Buildings, *94*, 109–120. https://doi.org/10.1016/j.enbuild.2015.02.050

Safa, M., Safa, M., Allen, J., Shahi, A., & Haas, C. T. (2017). Improving sustainable office building operation by using historical data and linear models to predict energy usage. *Sustainable Cities and Society*, *29*, 107–117. https://doi.org/10.1016/j.scs.2016.12.001

Schwartz, Y., & Raslan, R. (2013). Variations in results of building energy simulation tools, and their impact on BREEAM and LEED ratings: A case study. *Energy and Buildings*, *62*, 350–359. https://doi.org/10.1016/j.enbuild.2013.03.022

Shirzadi, N., Nasiri, F., Menon, R. P., Monsalvete, P., Kaifel, A., & Eicker, U. (2023). Smart Urban Wind Power Forecasting: Integrating Weibull Distribution, Recurrent Neural Networks, and Numerical Weather Prediction. *Energies*, *16*(17). https://doi.org/10.3390/en16176208

van Dronkelaar, C., Dowson, M., Spataru, C., & Mumovic, D. (2016). A Review of the Regulatory Energy Performance Gap and Its Underlying Causes in Non-domestic Buildings. In *Frontiers in Mechanical Engineering* (Vol. 1). Frontiers Media S.A. https://doi.org/10.3389/fmech.2015.00017

Wang, R., Lu, S., & Li, Q. (2019). Multi-criteria comprehensive study on predictive algorithm of hourly heating energy consumption for residential buildings. *Sustainable Cities and Society*, *49*. https://doi.org/10.1016/j.scs.2019.101623

Yu, Z., Haghighat, F., Fung, B. C. M., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, *42*(10), 1637–1646. https://doi.org/10.1016/j.enbuild.2010.04.006

Zhu, D., Hong, T., Yan, D., & Wang, C. (2013). A detailed loads comparison of three building energy modeling programs: EnergyPlus, DeST and DOE-2.1E. *Building Simulation*, *6*(3), 323–335. https://doi.org/10.1007/s12273-013-0126-7

Zou, P. X. W., Xu, X., Sanjayan, J., & Wang, J. (2018). Review of 10 years research on building energy performance gap: Life-cycle and stakeholder perspectives. In *Energy and Buildings* (Vol. 178, pp. 165–181). Elsevier Ltd. https://doi.org/10.1016/j.enbuild.2018.08.040