Enhancing Building Retrofit Decision-Making: A Synergistic Approach Combining Calibrated Simulations and Machine Learning

Navid Shirzadia, Meli Stylianou
CanmetENERGY-Ottawa
Natural Resources Canada,
Ottawa, Canada
e-mail: navid.shirzadi@nrcan-rncan.gc.ca

Abstract— **Achieving energy efficiency and reducing greenhouse gas (GHG) emissions are critical goals for building retrofitting. This study tackles challenges such as limited data and scenario generalizability by adapting the U.S. ComStock database for Canadian buildings using a Euclidean distance-based matching algorithm, achieving a 92% success rate for matches below a 2.43 threshold. Machine learning models, Random Forest (RF) and Extreme Gradient Boosting (XGBoost), were selected due to their effectiveness in handling high-dimensional, non-linear datasets and were applied to predict Energy Use Intensity (EUI) and GHG emissions. XGBoost, with optimized hyperparameters, outperformed RF, achieving R^2 values of 0.91 for EUI and 0.86 for GHG emissions, with lower RMSE and MAE values, showcasing its capability in handling complex, high-dimensional data. A comparative analysis highlighted significant environmental benefits of transitioning Heating, Ventilation, and Air Conditioning (HVAC) systems to cleaner fuels, such as air-source heat pumps. The proposed distribution-based method, leveraging 100 buildings across diverse climates and types, offers a robust framework for policymakers to guide energy-efficient retrofitting decisions.**

Keywords- smart building retrofitting; energy efficiency; greenhouse gas emissions; machine learning.

I. INTRODUCTION

The construction and operation of buildings contribute significantly to global energy consumption and greenhouse gas (GHG) emissions [1][2]. In Canada, existing buildings alone account for over 40% of emissions in major urban centers [3][4]. Consequently, enhancing energy efficiency and minimizing the environmental impact of existing buildings have become critical priorities within the building sector.

Despite government initiatives such as the Canada Greener Homes program providing financial support [5], retrofitting buildings faces numerous challenges and uncertainties for building owners [6][7]. Key obstacles include a lack of information and awareness, which complicates decision-making processes. Retrofitting decisions typically depend on the expertise of energy advisors and audits—often time-consuming and expensive processes designed to identify potential retrofit measures [8][9].

Another approach involves physics-based energy models that simulate building energy use and define retrofit scenarios based on these simulations. For instance, a historical building in Italy was modeled using EnergyPlus in [10], while Rahman et al. [11], simulated an office building in Australia, exploring major retrofit scenarios that achieved approximately 42% energy savings. Similar physics-based approaches have been applied in various cases [12]-[15]. However, these models often face significant uncertainties and energy performance gaps, which are rarely considered in final evaluations. Furthermore, their complexity, reliance on specialized expertise, and time-consuming processes render them less accessible for many building owners.

The rise of artificial intelligence has spurred interest in Machine Learning (ML) and data-driven approaches for building retrofits. A key challenge in applying data-driven models to retrofit scenarios is the availability of reliable retrofit data. Common issues include uncertainty and subjectivity in data quality [16] and sometimes the privacy issues about gathered measured data. As a result, researchers often rely on artificial data or surrogate models to analyze retrofit scenarios [17]. While surrogate models mitigate some expertise and computational demands associated with physics-based simulations, they do not resolve uncertainties or performance gaps inherent in these models.

Recently, a highly granular tool named ComStock [8], [18][19], developed by the National Renewable Energy Laboratory (NREL), has provided a database of over 300,000 buildings, containing detailed building characteristics and information. The simulation results in this database are highly calibrated, incorporating stochastic models to more accurately reflect occupant behavior. This level of detail addresses the data availability issue for developing data-driven models. However, ComStock is limited in its geographic scope, being tailored to locations within the United States. Extending its utility to other regions, such as Canada, requires innovative approaches to ensure compatibility and relevance.

This study addresses the challenges through an integration of data-matching techniques and machine learning models. The key contributions of this research are:

 ISBN: 978-1-68558-300-2

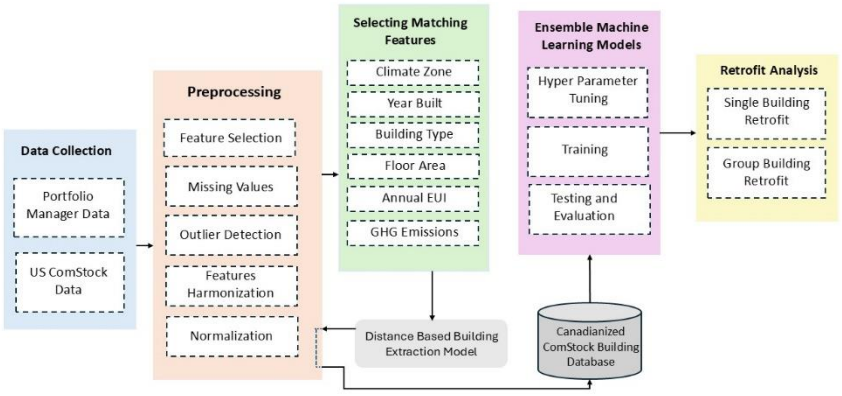

Figure 1. Workflow for developing retrofit scenarios from raw data to analysis

• Data Matching for Enhanced Model Accuracy: The study leverages a highly calibrated simulation database (ComStock) using a Euclidean distance-based algorithm to identify and extract data from buildings that closely match Canadian counterparts. This approach addresses the data scarcity challenge and ensures reliable inputs for data-driven analysis.

• Scalable Retrofit Evaluation Framework: A distribution-based method was proposed to generalize retrofit impacts across a diverse sample of buildings, considering variations in type, location, and climate zone. This method offers a robust framework for policymakers and stakeholders to make informed decisions.

• Environmental and Energy Impact Insights: The research highlights the benefits of some retrofit scenarios such as transitioning HVAC systems to cleaner fuel sources, demonstrating their potential to significantly reduce GHG emissions.

This study not only advances the application of machine learning for building retrofits but also provides a scalable framework for evaluating retrofit scenarios in diverse contexts, contributing to sustainable energy and environmental management in the building sector.

The remainder of this paper is organized as follows: Section II details the proposed methodology, outlining steps from data extraction and preprocessing to the development of data-driven models, Section III presents and discusses the results of the case study. And finally, Section IV concludes the paper and highlights directions for future research.

II. METHODOLOGY

The overall process of developing retrofit scenarios from raw data is illustrated in Figure 1. The workflow begins with data collection, utilizing real user-input data from the Energy Star Portfolio Manager (ESPM) database—a database derived from Canadian user inputs—combined with the U.S. ComStock database. The preprocessing stage involves several steps, including feature selection, handling missing values and outliers, and feature engineering. For the ESPM database, a climate zone feature was generated using the heating degree days (HDD) metric for each building. In the ComStock database, a filtration process was applied to exclude buildings located in climate zones not present in Canada.

Six features were selected for the building extraction process, implemented through a distance-based matching model. Categorical features underwent harmonization to ensure consistency between the datasets. The extracted buildings then went through the preprocessing pipeline again, starting with feature selection, followed by normalization. The preprocessed data was then used to train ensemble learning models, followed by a detailed retrofit analysis.

A. Matching Process

To adapt the ComStock database, originally created for buildings in the USA, for use in Canada, the portfolio manager database, which is based on Canadian user inputs, is employed. Six key features are selected for the matching process: climate zone, building construction year, building type, gross floor area, annual energy use intensity, and annual greenhouse gas (GHG) emissions. Categorical features in the dataset are transformed into numerical representations to ensure compatibility with the distance-based matching process. Specifically, one-hot encoding is applied to categorical variables, converting them into binary feature vectors. This transformation allows categorical attributes to be incorporated alongside numerical features without introducing ordinal biases. Once categorical variables are encoded, all features undergo normalization to eliminate discrepancies in scale and ensure that no single attribute dominates the distance calculation. The normalization process standardizes numerical features to a common range, facilitating a fair comparison between different building attributes during the matching process. Euclidean distance (Formula 1) is used to calculate the similarity between buildings in the two databases.

$$D(A,B) = \sqrt{\sum_{i=1}^{6}(a_i - b_i)^2} \quad (1)$$

 ISBN: 978-1-68558-300-2

Where a_i and b_i represent the feature values of buildings A and B, respectively, for the i-th feature.

In the matching process, each target building, buildings in the ESPM database, was iterated through, and the closest matched building in the ComStock database was identified based on pre-computed distance values. For each target building, the already matched ComStock buildings were filtered out, and the closest building was selected by finding the minimum distance. The index of the closest building was recorded as the match, and the building was marked as used to prevent it from being selected again. This process was repeated until all target buildings were matched, ensuring that each building in the ESPM database was paired with the closest available building in the ComStock.

To statistically show the accuracy of the matches, a threshold is first calculated based on the below formula using the mean and standard deviation of the minimum distances between buildings to establish a threshold. The threshold is a distance value that helps define what is considered a good match.

$$Threshold = \mu + k.\partial \quad (2)$$

Where μ is the average of the minimum distances, ∂ is the standard deviation of the minimum distances and k is multiplier which adjusts the sensitivity of the threshold.

Then to calculate the good matches by counting how many of the minimum distances between buildings are less than the threshold previously calculated.

$$\begin{aligned} Good\ Matches &= \sum_{i=1}^{n} minimum\ distance \\ &< Threshold \\ percentage\ of\ good\ matches &= \frac{Good\ Matches}{n} \times 100 \end{aligned} \quad (3)$$

Which *n* is the number of minimum distances in the minimum distance array.

B. Machine learning models

Two powerful ensemble learning techniques, Random Forest (RF) and XGBoost, were used to predict EUI and GHG emissions. RF was chosen for its ability to handle high-dimensional feature interactions and provide interpretability, while XGBoost was selected for its superior predictive performance through gradient boosting and optimized learning. RF is an ensemble method that creates multiple decision trees using random subsets of the data, then combines their predictions to enhance model robustness and accuracy [20]. This approach is particularly effective at reducing overfitting and managing high-dimensional data. In contrast, XGBoost is a boosting-based technique that trains weak learners iteratively, focusing on minimizing errors from prior iterations. Its gradient boosting framework, along with features like regularization and efficient handling of missing data, enables it to capture complex patterns within the data effectively [21].

To optimize the performance of these models, RandomizedSearchCV method, which is a package of Scikit Learn library [22], was used for hyperparameter tuning. This method efficiently explores the hyperparameter space by randomly selecting combinations and evaluating their performance based on cross-validation. Key parameters, such as the number of estimators, maximum depth of trees, and learning rate, were tuned for both models. The results of this tuning process, including the best parameters and corresponding performance metrics, will be detailed in the results section. This tuning ensured that the models were well-suited to the dataset and provided reliable predictions for retrofit scenarios.

C. Evaluation metrics

This study employs three primary evaluation metrics: R-squared (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These metrics assess the performance of the trained model on the training data, which is then tested using the testing data. The corresponding formulas are provided below:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (4)$$

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

Where n is the number of data points, y_i is the actual value, $\hat{y}_i$ is the predicted value and $\bar{y}$ is the mean of the actual values.

III. RESULTS

This section presents the key findings derived from the dataset, highlighting the steps taken to ensure its relevance and applicability to Canadian building stock characteristics.

A. Matching and data evaluation

The ComStock database was initially filtered based on climate zone, reducing the dataset from 336,149 to 193,741 buildings to better align with Canadian conditions. The

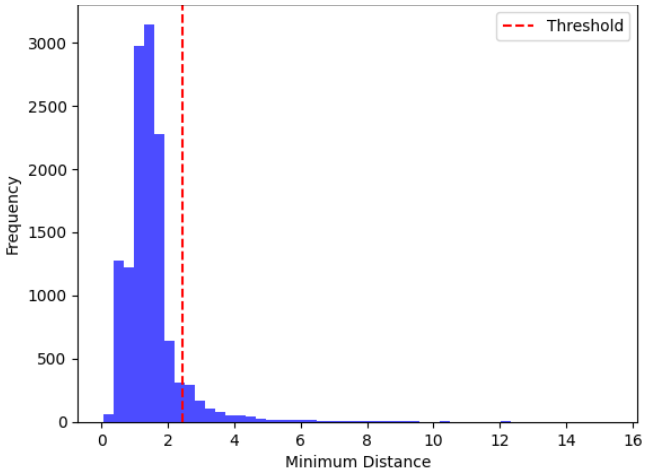

Figure 2. Distribution of minimum distances

matching process aimed to find the most similar buildings for 12,865 target buildings from the Portfolio Manager (PM) database. To achieve this, we utilized a Euclidean distance-based approach across six key features, ensuring a robust and consistent comparison. Each target building from the PM database was iteratively matched to its closest counterpart in the filtered ComStock dataset, prioritizing similarity while preventing duplicate matches. The quality of these matches was assessed using a threshold-based validation method, which determined that approximately 92% of matches were within an acceptable distance threshold of 2.43 (with the multiplier set to 1). Figure 2 presents the distribution of minimum distances, highlighting a peak in frequency just before reaching the threshold, indicating the effectiveness of the matching strategy in pairing buildings with similar characteristics.

Evaluating the extracted matched buildings, the boxplots for EUI and GHG emissions (Figure 3) reveal significant variability among building types, underscoring the influence of operational characteristics on energy consumption and greenhouse gas output. Quick Service Restaurants and Full-Service Restaurants consistently exhibit the highest median values for both EUI (above 100 kWh/ft²) and GHG emissions (20,000–40,000 kg CO_2). This is likely due to the energy-intensive nature of their operations, including frequent use of cooking equipment and extended operating hours. In contrast, building types such as small offices, medium offices, warehouses, and hospitals show narrower distributions, reflecting more uniform energy use, though the limited hospital sample size may artificially reduce observed variability.

Interestingly, while large offices have slightly higher median EUI compared to medium offices, the latter demonstrates higher median GHG emissions. This disparity may reflect differences in energy source mixes or operational efficiencies. The extended whisker lengths for Quick Service and Full-Service Restaurants in the EUI plot, along with the longer whiskers for medium offices and Full-Service Restaurants in the GHG plot, highlight significant variability

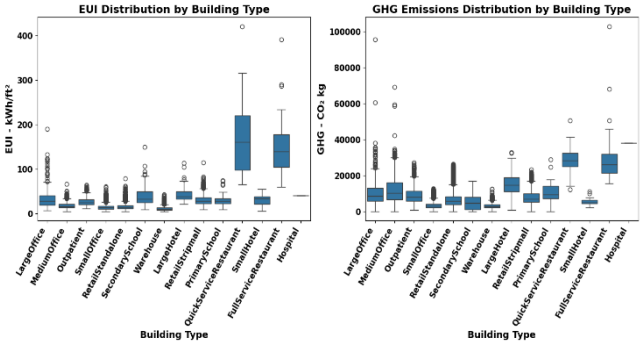

Figure 3. EUI and GHG emissions distribution across building types

within these categories, possibly due to diverse building designs or operational practices. While outliers are present in both plots, particularly for Large Offices and restaurants, they were retained as they represent simulated variations intended to capture diverse building performances. These trends emphasize the importance of tailored energy and emissions management strategies for different building types.

Based on Figure 4, Zone 7 has a medium EUI compared to other climate zones but records the lowest GHG emissions. In contrast, Zone 6 exhibits the highest average GHG emissions despite not having the highest EUI. This suggests that factors beyond energy consumption, such as the type of fuel used, play a significant role in GHG emissions. As shown in Figure 5, buildings in Zone 7 primarily use cleaner energy sources like electricity, natural gas, and district heating, which contribute to its lower GHG emissions. Conversely, a significant portion of Zone 6's GHG emissions is attributed to the use of carbon-intensive fuels like fuel oil, which lead to higher CO_2 emissions despite moderate energy usage.

The challenge with zones like Zone 7 or Zone 8 is that they do not encompass buildings with a wide variety of fuel sources, which can reduce the accuracy of the training process. To address this limitation, a key future step is to update the Portfolio Manager data used to extract the datasets and include a broader range of buildings from the source database. This would ensure more comprehensive representation and improve the reliability of the analysis.

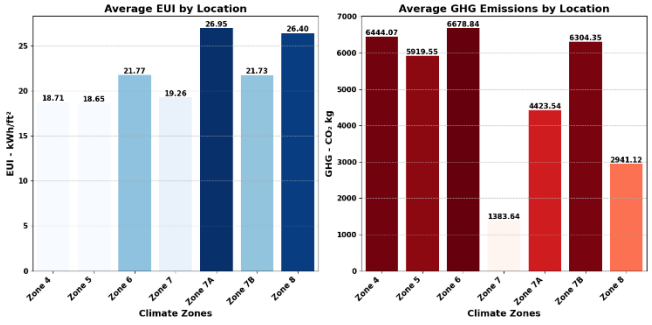

Figure 4. EUI and GHG emissions by climate zone

 ISBN: 978-1-68558-300-2

TABLE I: HYPERPARAMETER RANGES AND OPTIMIZED VALUES FOR XGBOOST AND RANDOM FOREST MODELS USING RANDOMIZEDSEARCHCV

XGB Params	Range	Optimum	RF Params	Range	Optimum
Number of estimators	[100, 200, 400]	400	Number of estimators	Randint (50, 300)	169
Estimator max depth	[3, 6, 10]	6	max depth	[None, 10, 20, 30, 40]	40
Estimator learning rate	[0.01, 0.05, 0.1, 0.2]	0.1	Min sample split	Randint (2, 10)	2
Estimator subsample	[0.7, 0.8, 0.9]	0.8	Min sample leaf	Randint (1, 10)	3
Estimator colsample bytree	[0.7, 0.8, 0.9]	0.8	max_features	['sqrt', 'log2', None]	sqrt
			Bootstrap	[True, False]	False

B.Model training

Table 1 summarizes the hyperparameter ranges and the optimized values for the two machine learning models used in this study: XGBoost (XGB) and Random Forest (RF). The purpose of this parameter tuning is to improve the predictive performance of each model by identifying the optimal combination of hyperparameters. The tuning process was conducted using RandomizedSearchCV method, which performs a randomized search over the specified parameter ranges to find the best-performing configuration.

The hyperparameters chosen for the Random Forest (RF) and XGBoost models reveal important aspects of their optimization and performance. For example, max_features in the RF model defines how many features are considered at each split, with "sqrt" being the optimal value here. This parameter contributes to the diversity of the decision trees, a crucial element in improving generalization while maintaining computational efficiency. Similarly, the number of estimators (optimal value: 169 for RF and 400 for XGBoost) governs the ensemble size, directly affecting both model accuracy and training time. The selection of a higher number of estimators in XGBoost suggests its ability to handle larger ensembles effectively, while the smaller optimal value for RF indicates a balance between computational efficiency and predictive power.

Another point of interest is the learning rate in XGBoost, which controls the step size during optimization. The optimal value of 0.1 strikes a balance between convergence speed and overfitting prevention. On the other hand, parameters like min samples split and min samples leaf in RF are essential for controlling tree growth and preventing overfitting by requiring a minimum number of data points at splits or leaves. The interplay between these parameters highlights how RandomizedSearchCV has fine-tuned the models to suit the dataset's characteristics. These optimal values reflect the need to manage trade-offs between model complexity, overfitting, and computational demands, providing an essential balance for practical implementation. The selected optimum values were then utilized during the training process to ensure the models were fine-tuned for optimal performance.

The model was trained using 90% of the data and then tested on the remaining 10%, which includes over 1,200 different buildings. The training evaluation results (Table 2) highlight that XGBoost outperforms Random Forest in predicting both EUI and GHG emissions, as shown by higher R^2 values and lower RMSE and MAE metrics. XGBoost achieves an R^2 of 0.91 for EUI prediction compared to 0.83 for Random Forest, and similarly, an R^2 of 0.86 for GHG emissions prediction compared to 0.75 for Random Forest. This superior performance underscores XGBoost's capability to handle complex patterns in the data effectively. For example, in EUI prediction, the reduction in RMSE from 5.05 (Random Forest) to 3.54 (XGBoost) reflects its ability to better capture underlying relationships, while the drop in MAE from 2.94 to 1.99 shows improved precision in its predictions.

The high-dimensional input vector, consisting of approximately 48 predictors, plays a critical role in XGBoost's superior performance. XGBoost is particularly adept at managing complex feature interactions and identifying important predictors, thanks to its gradient boosting framework and regularization techniques. This is especially beneficial when working with many predictors, as it reduces the risk of overfitting and effectively prunes less relevant splits. Random Forest, while robust, may struggle with high-dimensional data, as it treats all features more equally and lacks the inherent optimization for feature selection and interaction modeling. The results are visualized in Figure 6, using plots that show the relationship between predicted and measured values, with the red dashed line representing perfect predictions. These graphs highlight the performance of both models. The scatter plots demonstrate how closely the predicted values align with the measured values, with the red dashed line indicating the ideal prediction scenario. It is evident that XGBoost offers a more accurate fit for both EUI and GHG emissions data, as reflected by its higher R^2 values and smaller deviations from the perfect prediction line. This makes XGBoost the superior model in terms of predictive power and accuracy.

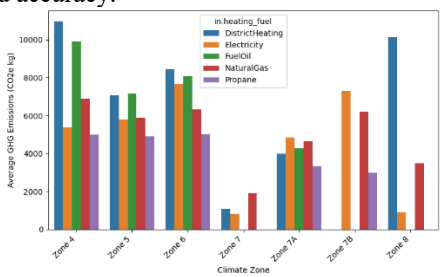

Figure 5. Average GHG emissions by heating fuel type and climate zone

 ISBN: 978-1-68558-300-2

TABLE II: TRAINING EVALUATION

	EUI Prediction			GHG Emissions Prediction		
Model	**R^2**	**RMSE (kWh/ft2)**	**MAE (kWh/ft2)**	**R^2**	**RMSE (kg-CO2)**	**MAE (kg-CO2)**
Random Forrest	0.83	5.05	2.94	0.75	2440.59	1481.82
XGBoost	0.91	3.54	1.99	0.86	1840.10	1012.21

C. Retrofit evaluation and savings

Figure 7 provides a clear comparison of how switching HVAC heating types affects both EUI and GHG emissions. For example, transitioning from a Furnace with Propane as its fuel source to an air source heat pump (ASHP) using Electricity results in a notable reduction in both metrics. The EUI decreases from approximately 29.5 kWh/ft² to around 27.5 kWh/ft², indicating improved energy efficiency. Similarly, the GHG emissions drop significantly from around 3,612 kg-CO_2 to just above 2,348 kg-CO_2, highlighting the environmental benefit of switching to a cleaner fuel source. This comparison underscores the potential of fuel changes in HVAC systems to achieve both energy and emission savings. While the evaluation presented in Figure 7 for a single building is valuable for specific retrofit scenarios, it is generally insufficient for policymakers and organizations involved in building retrofits. Decisions are rarely made based on the performance of a single building; instead, stakeholders prefer to evaluate retrofit savings across a broader portfolio of buildings [23]. To address this need, a sample of 100 diverse buildings across different climate zones was selected from the original database, and two distinct datasets—pre-retrofit and post-retrofit—were created for analysis.

The pre-retrofit dataset represents the baseline scenario, where all buildings maintain their existing features except for the retrofit-specific variables (in this case, HVAC Heating Type and Heating Fuel Type), which were standardized to Furnace and Propane. Conversely, the post-retrofit dataset includes the same buildings with identical features as the pre-retrofit dataset, but with HVAC Heating Type and Heating Fuel Type changed to ASHP and Electricity to simulate the retrofit. Additional post-retrofit scenarios were created by varying the HVAC and fuel types, consistent with the single building retrofit analysis.

The resulting distributions of EUI and GHG emissions for various retrofit scenarios are shown in Figure 8. In this context, 'positive savings' refers to reductions in EUI and GHG emissions compared to the base case. Retrofitting from Furnace/Propane to options such as Furnace/Natural Gas, ASHP/Electricity, and Electric Resistance/Electricity demonstrated more than 70% of buildings achieving positive energy savings, highlighting the reliability of these scenarios for energy efficiency improvements. On the other hand, the transition from Furnace/Propane to Furnace/Fuel Oil exhibited the lowest positive change in GHG emissions, with nearly 98% of buildings showing negative savings. This outcome aligns with expectations, as Fuel Oil typically results in higher GHG emissions. Retrofitting to ASHP and Electric Resistance systems also showed high positive changes in GHG reductions, reinforcing their effectiveness in lowering carbon emissions across a diverse building sample.

For District/District Heating, while the exact methodology for calculating heating energy and GHG emissions in the ComStock database is not explicitly detailed, the results presented in Figure 8 highlight the significant variability in district heating impacts among buildings. This variability is largely due to differences in fuel sources used within district heating networks, which can range from renewable energy sources to fossil fuels. For instance, in the case of the building depicted in Figure 7, the energy source could predominantly be renewables, resulting in very low GHG emissions alongside reasonable EUI. In contrast, other buildings, likely relying on fossil fuels for district heating, exhibit markedly different results, with higher EUI outcomes. Although, on the other hand, the positive GHG emissions observed for a group of buildings suggest that other attributes may influence district heating GHG emissions and even EUI. These attributes could not be accounted for due to a lack of available information. It is also worth noting that while 99% of the buildings using District/DistrictHeating show positive GHG emissions, the average GHG emissions of this group are 4526 kg-CO2, which is lower than buildings using Furnace/Propane (base case) with an average of 6365 kg-CO2, yet notably higher than the example building shown in Figure 7. This suggests that, despite the variability in district heating sources, the overall emissions performance of these systems may align more closely with that of fossil fuel-based systems like Furnace/Propane.

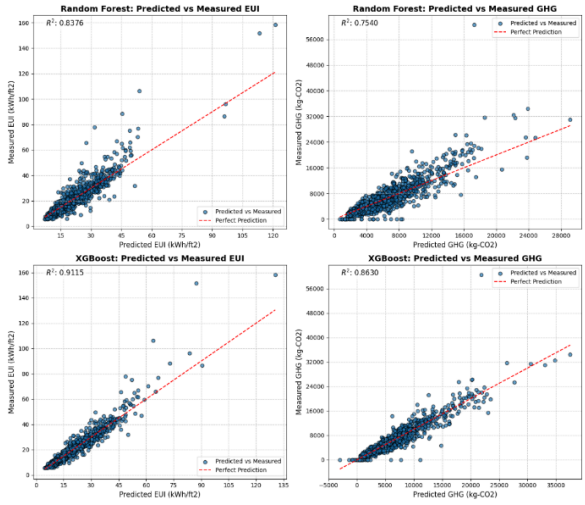

Figure 6. Comparison of predicted vs. measured EUI and GHG emissions for Random Forest and XGBoost models

 ISBN: 978-1-68558-300-2

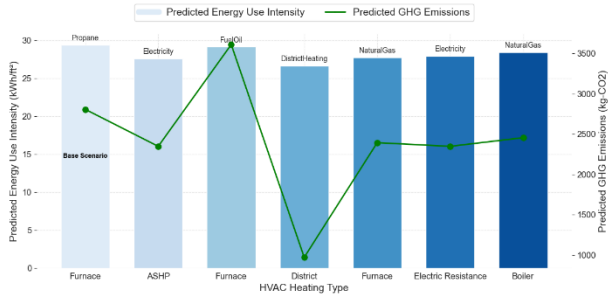

Figure 7. Comparative analysis of predicted energy use intensity (kWh/ft²) and GHG emissions (kg-CO_2) across various HVAC heating types

These findings highlight the need for further investigation into the specific factors influencing emissions variations across different building groups.

There may be some inaccuracies in the results due to the limited data used for training the regression model and its inability to accurately predict certain combinations of features. Other retrofit scenarios, such as changing window types, wall materials, or window-to-wall ratios, can also be evaluated using the same proposed methodology. While combining multiple retrofit scenarios could provide a more comprehensive evaluation, it introduces complexity, making it challenging to isolate the impact of each individual scenario.

IV. CONCLUSION AND FUTURE WORK

This study addresses the critical challenge of improving energy efficiency and reducing greenhouse gas (GHG) emissions in existing buildings, which contribute significantly to global emissions. While traditional methods such as physics-based energy simulation models offer valuable insights into retrofit scenarios, their limitations, including expertise requirements, time-consuming processes, and inherent uncertainties—can impede their practical application. Similarly, data-driven approaches using ML models face challenges due to the lack of reliable measured data and privacy concerns.

To bridge these gaps, this research integrates data-matching techniques with highly calibrated simulation databases to overcome data limitations for ML-based retrofitting analyses. By employing a Euclidean distance-based matching algorithm, this approach successfully identifies comparable buildings and extracts valuable data, achieving a high success rate. Ensemble learning models, specifically Random Forest (RF) and XGBoost, were trained on the matched data to predict EUI and GHG emissions. The optimized XGBoost model outperformed RF, demonstrating superior accuracy and robustness with R^2 values of 0.91 for EUI and 0.86 for GHG emissions.

The study further demonstrated the environmental and energy efficiency benefits of transitioning HVAC systems to cleaner fuel sources, such as air-source heat pumps. To enhance the generalizability of the findings, a distribution-based method was introduced, which analyzed retrofit impacts across a sample of 100 buildings of various types and in different climate zones. This method provides actionable insights and a scalable framework for policymakers and stakeholders to make informed, data-driven decisions on building retrofits.

By combining advanced data-matching techniques with machine learning and proposing a scalable evaluation framework, this research contributes to the growing body of knowledge on sustainable building retrofits. The scalability lies in its ability to generalize retrofit impacts across a diverse range of building types, climates, and geographic regions. Future work could focus on:

1. Adding cost as a target variable to evaluate the economic aspects of retrofits, providing a more comprehensive assessment that integrates energy, environmental, and financial impacts.
2. Updating the portfolio manager data dynamically with new inputs from users to enhance the adaptability and applicability of the proposed framework, ensuring it remains relevant and effective in real-world scenarios.

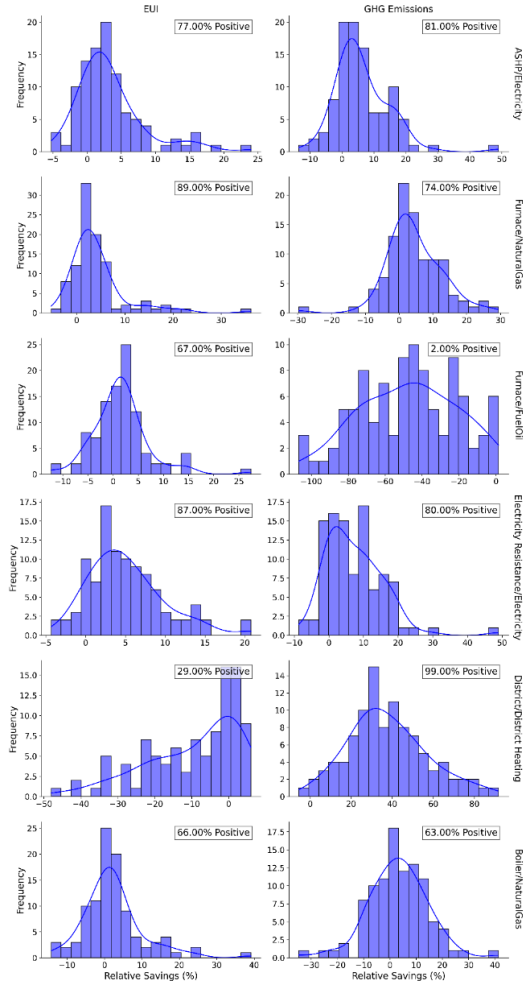

Figure 8. Distributions of EUI and GHG emissions savings for various retrofit scenarios of heating HVAC and fuel type across a 100-building sample

ACKNOWLEDGMENT

Thanks to Natural Resources Canada's Office of Energy Efficiency for funding this research.

 ISBN: 978-1-68558-300-2

REFERENCES

[1] E. Asadi, M. G. Da Silva, C. H. Antunes, and L. Dias, "Multi-objective optimization for building retrofit strategies: A model and an application," Energy Build., vol. 44, pp. 81–87, Jan. 2012, doi: 10.1016/j.enbuild.2011.10.016.

[2] Z. Ma, P. Cooper, D. Daly, and L. Ledo, "Existing building retrofits: Methodology and state-of-the-art," Energy Build., vol. 55, pp. 889–902, Dec. 2012, doi: 10.1016/j.enbuild.2012.08.018.

[3] D. Malomo, Y. Xie, and G. Doudak, "Unified life-cycle cost–benefit analysis framework and critical review for sustainable retrofit of Canada's existing buildings using mass timber," Can. J. Civ. Eng., vol. 51, no. 7, pp. 687–703, Jul. 2024, doi: 10.1139/cjce-2023-0222.

[4] N. R. C. Government of Canada, "National Energy Use Database." Accessed: July. 28, 2025. [Online]. Available: https://oee.nrcan.gc.ca/corporate/statistics/neud/dpa/data_e/databases.cfm

[5] N. R. Canada, "Eligible retrofits and grant amounts." Accessed: July. 28, 2025. [Online]. Available: https://natural-resources.canada.ca/energy-efficiency/homes/canada-greener-homes-grant/start-your-energy-efficient-retrofits/plan-document-and-complete-your-home-retrofits/eligible-grants-for-my-home-retrofit/23504

[6] A. Akhatova and L. Kranzl, "Agent-based modelling of building retrofit adoption in neighbourhoods," Energy Build., vol. 328, p. 115172, Feb. 2025, doi: 10.1016/j.enbuild.2024.115172.

[7] S. Ebrahimigharehbaghi, Q. K. Qian, F. M. Meijer, and H. J. Visscher, "Unravelling Dutch homeowners' behaviour towards energy efficiency renovations: What drives and hinders their decision-making?," Energy Policy, vol. 129, pp. 546–561, Jun. 2019, doi: 10.1016/j.enpol.2019.02.046.

[8] F. Re Cecconi, A. Khodabakhshian, and L. Rampini, "Data-driven decision support system for building stocks energy retrofit policy," J. Build. Eng., vol. 54, p. 104633, Aug. 2022, doi: 10.1016/j.jobe.2022.104633.

[9] T. Hong, L. Yang, D. Hill, and W. Feng, "Data and analytics to inform energy retrofit of high performance buildings," Appl. Energy, vol. 126, pp. 90–106, Aug. 2014, doi: 10.1016/j.apenergy.2014.03.052.

[10] F. Ascione, F. De Rossi, and G. P. Vanoli, "Energy retrofit of historical buildings: theoretical and experimental investigations for the modelling of reliable performance scenarios," Energy Build., vol. 43, no. 8, pp. 1925–1936, Aug. 2011, doi: 10.1016/j.enbuild.2011.03.040.

[11] M. M. Rahman, M. G. Rasul, and M. M. K. Khan, "Energy conservation measures in an institutional building in sub-tropical climate in Australia," Appl. Energy, vol. 87, no. 10, pp. 2994–3004, Oct. 2010, doi: 10.1016/j.apenergy.2010.04.005.

[12] B. Abu-Hijleh, A. Manneh, A. AlNaqbi, W. AlAwadhi, and A. Kazim, "Refurbishment of public housing villas in the United Arab Emirates (UAE): energy and economic impact," Energy Effic., vol. 10, no. 2, pp. 249–264, Apr. 2017, doi: 10.1007/s12053-016-9451-x.

[13] H. Abdullah and H. Alibaba, "Retrofits for Energy Efficient Office Buildings: Integration of Optimized Photovoltaics in the Form of Responsive Shading Devices," Sustainability, vol. 9, no. 11, p. 2096, Nov. 2017, doi: 10.3390/su9112096.

[14] C. Citadini De Oliveira, I. Catão Martins Vaz, and E. Ghisi, "Retrofit strategies to improve energy efficiency in buildings: An integrative review," Energy Build., vol. 321, p. 114624, Oct. 2024, doi: 10.1016/j.enbuild.2024.114624.

[15] T. D. Mora, A. Righi, F. Peron, and P. Romagnoni, "Cost-Optimal measures for renovation of existing school buildings towards nZEB," Energy Procedia, vol. 140, pp. 288–302, Dec. 2017, doi: 10.1016/j.egypro.2017.11.143.

[16] E. Thrampoulidis, G. Mavromatidis, A. Lucchi, and K. Orehounig, "A machine learning-based surrogate model to approximate optimal building retrofit solutions," Appl. Energy, vol. 281, p. 116024, Jan. 2021, doi: 10.1016/j.apenergy.2020.116024.

[17] S. A. Sharif and A. Hammad, "Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA," J. Build. Eng., vol. 25, p. 100790, Sep. 2019, doi: 10.1016/j.jobe.2019.100790.

[18] A. Parker et al., "ComStock Reference Documentation: Version 1," Renew. Energy, 2023.

[19] J. Landsman et al., "Leveraging NREL's ResStock & ComStock Dataset to Evaluate Building Stock Electrification," 2024.

[20] Z. Wang, Y. Wang, R. Zeng, R. S. Srinivasan, and S. Ahrentzen, "Random Forest based hourly building energy prediction," Energy Build., vol. 171, pp. 11–25, Jul. 2018, doi: 10.1016/j.enbuild.2018.04.008.

[21] D. Ma, X. Li, B. Lin, Y. Zhu, and S. Yue, "A dynamic intelligent building retrofit decision-making model in response to climate change," Energy Build., vol. 284, p. 112832, Apr. 2023, doi: 10.1016/j.enbuild.2023.112832.

[22] L. Buitinck et al., "API design for machine learning software: experiences from the scikit-learn project," Sep. 01, 2013, arXiv: arXiv:1309.0238. doi: 10.48550/arXiv.1309.0238.

[23] T. Walter and M. D. Sohn, "A regression-based approach to estimating retrofit savings using the Building Performance Database," Appl. Energy, vol. 179, pp. 996–1005, Oct. 2016, doi: 10.1016/j.apenergy.2016.07.087.

 ISBN: 978-1-68558-300-2