

Original Article

# Explainable Machine Learning for Policy-Driven Carbon Emission Reduction in Building Projects

Riza Suwondo<sup>1</sup>, Religiana Hendarti<sup>2</sup>, Militia Keintjem<sup>3</sup>, Chee Fui Wong<sup>4</sup>

<sup>1</sup>Department of Civil Engineering, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia.

<sup>2</sup>Department of Architectural, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia.

<sup>3</sup>Karbonara Research Institute, Jakarta, Indonesia.

<sup>4</sup>Department of Civil Engineering, Lee Kong Chien Faculty of Engineering, Universiti Tunku Abdul Rahman, Selangor, Malaysia.

<sup>1</sup>Corresponding Author : [riza.suwondo@binus.ac.id](mailto:riza.suwondo@binus.ac.id)

Received: 09 January 2026

Revised: 10 February 2026

Accepted: 11 March 2026

Published: 28 April 2026

**Abstract** - The construction industry is a major contributor to global carbon emissions; therefore, finding a way to reduce emissions to meet climate-mitigation goals is vital. The challenges that impede emissions reductions are numerous and reliant on implemented policies. This study proposes an explainable machine learning framework to predict and interpret carbon emission reduction performance in building projects by integrating project characteristics, policy intervention indicators, certification levels, lifecycle information, and emission-related variables. Three machine learning models were created-Ridge Regression, Random Forest, and Gradient Boosting-and tested on a split of the data to assess their performance. To explain emissions reductions, the models were evaluated for predictive performance using a set of statistical measures and the explainable machine learning metric, Shapley Additive Explanations (SHAP). The models demonstrated robust predictive performance, with a coefficient of determination of over 0.81 for each model, and all models had similar performance despite using different statistical techniques. Shapley values attributed most of the focus to high-level green certifications and policy measures, whereas project size had little influence on reducing emissions. The results reinforce the idea that policies have the largest impact on emissions reductions in the building sector. The suggested explainable framework provides clarity and relevant insights into policies that aid evidence-based decision-making and strategic planning for decarbonising the building sector.

**Keywords** - Carbon emission reduction, Explainable machine learning, Building sector decarbonisation, Policy intervention analysis, Shapley Additive Explanations (SHAP).

## 1. Introduction

Rapid urbanisation across the globe has led to increasing carbon emissions, which affect the environment, economy, and humanity [1-3]. If not controlled, carbon emissions lead to an increase in the greenhouse effect, which leads to climate problems and threatens human health and quality of life [4, 5]. Among these sectors, the building and construction industry plays a particularly critical role, accounting for over one-third of global energy consumption and more than 30% of energy-related carbon emissions [6-8]. Therefore, the construction industry must develop effective ways to solve the problem of carbon emissions to help achieve global goals for construction and climate change.

Owing to the complexity of building carbon emissions arising from the interaction of design parameters, operational performance, material selection, and policy context, it remains challenging to construct mechanism-based models that explicitly capture these multivariate relationships.

Consequently, data-driven approaches and machine learning techniques have been increasingly adopted to forecast building energy use and carbon emissions due to their capability to model nonlinear interactions efficiently [9,10].

The analysis of machine learning literature applied in the building industry can signal three areas. The first and most developed area predominantly concerns the use of operational energy load and demand prediction, and the use of advanced algorithms in the field of Artificial Neural Networks, Support Vector Machines, and ensemble methods. A second area can be seen as an extension of the first, where the same (or very similar) techniques are used to calculate the made (or embedded) carbon and greenhouse gas emissions of a building during all phases of its life cycle, in particular, during the design phase [11-16]. Even when both areas of study have reached remarkable levels of precision, most of the research focus remains in the domain of engineering, particularly the optimization of the algorithms. This form of research, in the



engineering domain, greatly neglects the non-technical factors, including the institutional and political circumstances, which are predominantly responsible for shaping the emission profile of a building. This signals the need for a third, more pluralistic (or more integrative) stream of research.

Regarding the policy literature, there is significant research on the effectiveness of carbon trading, carbon tax, and green certification schemes and how they contribute to emission reduction in the building sector [17-19]. Most of these studies use econometric and/or scenario-based assessments to gauge the effectiveness of these policy measures. While these contributions are valuable, they have little to no use of machine learning methodologies to address the complicated relationships that policy depth and a specific project context have. Thus, the extent to which policy and project attributes interact to influence emission reduction outcomes is still poorly understood.

Moreover, the majority of models developed in this field are ‘black boxes’, creating barriers to comprehending their predictive processes, as well as to their utility for supporting

policy initiatives [20-22]. The potential of Explainable Artificial Intelligence (XAI) is likely to address this shortcoming. Nevertheless, the systematic use of XAI approaches to investigate the carbon mitigation potential of buildings-specifically, those buildings that are policy-driven-remains underexplored in the literature.

The disparity within the literature is evident in the summary provided in Table 1. Some studies prioritize predictive accuracy, forecasting the emissions intensity of a particular firm or project. Other studies, in separate streams of research, look at policy impacts using econometrics and causal inference. While both streams have made important contributions, the degree to which they integrate is unsatisfactory. There is no framework that analytically balances the intensity of policy measures, the degree of green certification, lifecycle emissions, and project-specific technical factors. This has resulted in a fragmented and unsystematic understanding of the interplay of institutional and regulatory factors with project-specific technical factors within the context of emissions reduction.

**Table 1. Differentiation of the present study from recent research**

<b>Prior Research Focus</b>	<b>Identified Limitation</b>	<b>Contribution of the Present Study</b>
ML-based building energy or carbon emission prediction [23]	Emphasis on predictive accuracy; limited integration of policy intervention and certification variables	Integrates policy intensity, certification level, lifecycle, and project characteristics within a unified framework
ML applied to counterfactual identification in policy evaluation [22]	Focuses on improving treatment-control balancing; does not model emission reduction outcomes or interpret policy drivers	Develops an explainable ML framework to predict and interpret carbon emission reduction performance directly
Hybrid ML models for carbon emission prediction in the energy industry [24]	Focuses on improving prediction accuracy through signal decomposition and denoising; does not incorporate policy variables or interpret mitigation drivers	Develops an explainable ML framework integrating policy intervention, certification, lifecycle, and project characteristics to predict and interpret emission reduction performance
ML-based corporate carbon intensity prediction using linear and nonlinear models [25]	Focuses on emission intensity forecasting at the firm level; limited integration of lifecycle and project-specific mitigation variables	Develops a building-sector framework integrating project characteristics, policy intervention intensity, certification levels, and lifecycle context
ML-based corporate carbon intensity prediction [26]	focuses on firm-level emission forecasting; limited integration of lifecycle and project-level mitigation variables	Develops a building-sector framework integrating project characteristics, policy intervention intensity, certification levels, and lifecycle context

To address gaps in existing literature, this study develops an integrated, explainable machine learning framework capable of predicting and articulating carbon mitigation performance at the building project level. What sets this framework apart from others is its ability to incorporate, in a unified analytical framework, policy variables (e.g., degree of regulatory policy, green certification) along with technical and project-specific variables in a seamless manner. Along with enhanced predictive capability, the framework uses the SHAP method to explain policy and technical variables and quantify the impact (positive/negative) of each on emissions outcomes. As a result, this study bridges the divide between technical modelling and policy modelling to create analytical tools that can be used in the sustainable decarbonization of buildings.

## 2. Dataset and Methodology

Employing an explainable machine learning methodology, this research endeavours to determine the effectiveness of carbon mitigation across various building projects. The study attempts to capture the potential reductions of emissions, while explaining the attributable factors, by formulating a single analytical framework comprising project characteristics, policy variables, life cycle stages, and emissions data. The flow of the method consists of five systematic processes: (i) collection and exploration of data; (ii) feature engineering and pre-processing; (iii) training of a model through supervised machine learning; (iv) testing and validation of the model performance, and; (v) explaining the results through the use of XAI techniques to formulate policies that are reasonable and applicable.

### 2.1. Dataset Description

This study handled a dataset that contains 1,640 records of residential building projects retrieved from the public database Kaggle. It is critical to highlight that this dataset is a result of scenario-based synthesis rather than actual field data. The researchers created the dataset through designed and structured simulations to encompass a wide variety of combinations of technical measures, policies, and building lifecycle variables. The intention was to allow the study to target less easily identifiable contexts of the dataset to examine patterns of carbon mitigation. This dataset, due to the nature of scenario-based synthesis, must be used carefully to avoid over-generalization in this study, as it's primarily a representation of scenarios that exist in these records, rather than a true representation of the real world.

This research examines both numerical and categorical variables pertaining to building projects' data, including details on the construction projects' physical attributes, the degree of policy intervention, the status of green certification, the position in the building life cycle, and the associated emissions data. The percentage reduction in carbon emissions, which the author considers the dependent variable, is calculated using the dataset's available baseline scenario and the entire building life cycle emissions. Therefore, this

variable illustrates the instances of emissions mitigation and also acts as a variable to explain the degree to which a certain combination of the technical and institutional variables at the project level stimulates the reduction of emissions.

This study organizes predictor variables into four categories: i) attributes of the project; ii) indicators of policy and certification; iii) life cycle stages; and iv) metrics of emissions. The policy intensity variable is quantitatively defined as the total count of policy and incentive frameworks activated in a particular scenario. There are limitations to this operationalization; notably, the methodology captures the quantitative dimensions of policy instruments, but neglects the qualitative dimensions of policy instruments, such as the effectiveness of enforcement and compliance. Accordingly, in this analysis, policy variables are contextualized as perimeter variables, as opposed to causal variables. Given this limitation, researchers should refrain from inferring causal relationships between policies and mitigation outcomes.

In preparation for modelling, descriptive statistical analyses were performed to understand the distribution and variability of the numerical variables. Results are in Table 2. The analyses showed that the dataset possessed adequate variability across all the critical aspects-project size, mitigation measures, emissions, and reduction metrics-indicating high sample heterogeneity. This characteristic is important because it ensures that the developed machine learning model will not only be accurate in its predictions but also have good generalization capacity when applied to project contexts different from those in the training data.

Table 2. Summary statistics of key numerical variables

Variables	mean	std	min	max
Building Area (m <sup>2</sup> )	10312.5	5627.5	511.0	19981.0
Policy Interventions	2.0	1.4	0.0	4.0
Renewable Energy (%)	35.9	20.3	0.0	70.0
Material Reuse (%)	25.7	14.1	0.0	50.0
Waste Minimization (%)	20.3	11.6	0.1	40.0
Embodied Emissions (tCO <sub>2</sub> e)	1073.9	515.0	200.4	1998.6
Operational Emissions (tCO <sub>2</sub> e)	1632.0	757.6	300.1	2996.3
Total Lifecycle Emissions (tCO <sub>2</sub> e)	2705.9	926.8	578.4	4921.8
Emission Reduction (%)	26.2	12.6	2.0	55.0

### 2.2. Data Pre-Processing

Prior to model training, the dataset was pre-processed to ensure consistency, stability, and reproducibility. Missing numerical values were addressed using median imputation to reduce sensitivity to skewness and outliers, whereas missing categorical values were replaced with the most frequent category. Categorical variables were transformed using one-hot encoding to ensure compatibility with machine learning algorithms.

To evaluate the potential multicollinearity among the numerical variables, Pearson’s correlation analysis was conducted, as shown in Figure 1. Highly correlated aggregated and disaggregated emission indicators were not included simultaneously in the same model configuration to prevent redundancy and unstable coefficient estimation in the linear models.

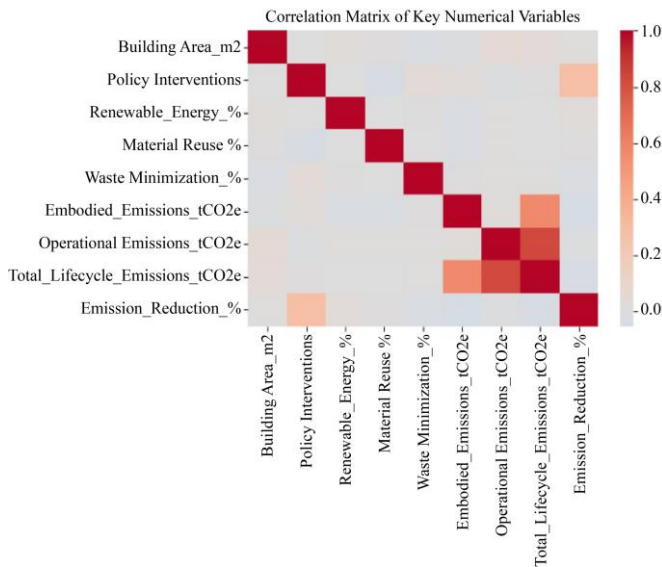


Fig. 1 Correlation matrix of key numerical variables used in the analysis

The influence of outliers was examined through distributional inspection. Because the dataset reflects structured scenario combinations rather than raw field measurements, extreme values correspond to valid high- or low-mitigation configurations and were therefore retained. For tree-based models, the variables were maintained at their original scale because these models are scale-invariant. For linear models, regularisation techniques (Ridge) were employed to mitigate scaling and multicollinearity effects. The dataset was randomly divided into training (80%) and testing (20%) subsets to ensure independent evaluation. Model performance was assessed using cross-validation within the training set to reduce the risk of overfitting.

This study adopted a predictive modelling framework rather than a causal inference approach; therefore, the SHAP-based feature attributions reflect the relative contribution to prediction accuracy and should not be interpreted as direct

causal effects. The dataset is scenario-based rather than derived from in situ field measurements, which may limit its ability to capture behavioural, compliance, or regional enforcement variability observed in real-world projects. Policy intervention intensity is represented as a structured count indicator and may simplify the differences in regulatory strength or implementation effectiveness. Additionally, green-certified configurations may inherently reflect favourable mitigation conditions, potentially introducing a selection bias within the dataset. Despite these limitations, the dataset provides controlled variability across technical and institutional conditions, enabling a systematic evaluation of emission reduction performance under diverse project-policy configurations.

### 2.3. Machine Learning

Three supervised machine learning models were employed to predict the carbon emission reduction performance: Ridge regression, Random Forest regression, and Gradient Boosting regression.

Ridge Regression is used in this study due to its usefulness in obtaining stable coefficient estimates, as well as its ability to handle potential multicollinearity issues among the predictors. This technique minimizes the following objective function:

$$\min_{\beta} (\sum_{i=1}^n (y_i - \beta^T x_i)^2 + \lambda \|\beta\|_2^2) \tag{1}$$

Where  $y_i$  refers to the observed percentage emission reduction,  $\beta$  is the coefficient vector, and  $\lambda$  denotes the regularisation parameter, which is used to balance the trade-off between the model’s bias and variance. The introduction of this penalty term not only enhances the model’s numerical stability but also diminishes the effects of multicollinearity by pushing highly correlated coefficients towards zero, without removing them entirely.

### 2.4. Model Training and Validation

At the outset, the dataset is split at random into an 80%-20% ratio, with 80% allocated to a training set and 20% to a testing set. The testing set is kept separate from training in order to assess the training set and evaluate the model objectively. Afterward, stable performance estimates and an approximate lowered risk of overfitting were achieved through five-fold cross-validation with the training set [27]. The model configuration is achieved through hyper-parameter optimization via a grid search, which is cross-validated with a grid search for coupled hyperparameter optimization and model configuration. The training set is divided into five subsets, and each subset is used for both training and validation, with the hyperparameter search conducted to cross-validate the optimal configuration. The model’s performance was assessed through various statistical parameters: coefficient of determination ( $R^2$ ), Mean Absolute

Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics are mathematically defined as follows:

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

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{4}$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{5}$$

Where  $y_i$  and  $\hat{y}_i$  are the actual and estimated values of the percentage of emission reduction, while  $\bar{y}_i$  is the average of actual values.

Using four different, but complementary, metrics ensures that model evaluation is both statistically valid and practically relevant. The coefficient of determination ( $R^2$ ) indicates the degree of variability/variation the model explains in the data. This offers an idea of the model’s substantive validity. RMSE and MAE assess absolute predictive accuracy (i.e.,  $R^2$ ); while RMSE is sensitive to large errors that may materially impact the policy estimates, MAE is informative of the averaging of the error. Meanwhile, MAPE provides inertial errors in percentage terms, which is important in policy/target unit communication (i.e, reduction of percent emission), as policymakers tend to lean on peg-based metrics during discussions of target emission reductions. In this regard, the combination of these four metrics provides an assessment that is both statistically credible and practically relevant.

**2.5. Explainable Machine Learning Analysis**

The Shapley Additive Explanations (SHAP) method was used to interpret models due to a mathematical approach based on cooperative game theory to obtain fair predictive contributions from all input variables. SHAP emphasizes fairness in measurements by assessing one variable’s contribution across various coalitional subsets of other variables, including single, main, cross, and interaction effects. The decomposition of the prediction is transparent and reveals to analysts and decision makers the key variables that affect positively or negatively, and to what extent emission reduction targets can be achieved [28-30]. The SHAP value for a particular feature is formulated as follows:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{j\}) - f(S)] \tag{6}$$

Where F represents the full set of input features, and S denotes a subset of features excluding feature j. Feature ranking and analysis of the direction and magnitude of the effects of the variables that drive carbon emission reduction performance were conducted using the Global SHAP analysis.

**3. Results and Discussion**

**3.1. Model Performance Comparison**

The cross-validation and independent test performances are summarised in Tables 3 and 4, respectively. Overall, all models exhibited notable and reliable predictive performances. In both cross-validation and independent testing, the models achieved  $R^2$  values greater than 0.81, signifying that the proposed framework successfully identified the key drivers impacting performance in the reduction of carbon emissions.

**Table 3. Cross-validation performance**

Model	CV R <sup>2</sup>	CV MAE	CV RMSE
Ridge Regression	0.821	4.206	5.346
Gradient Boosting	0.819	4.181	5.382
Random Forest	0.815	4.196	5.441

**Table 4. Test performance**

Model	Test R <sup>2</sup>	Test MAE	Test RMSE	Test MAPE(%)
Ridge Regression	0.826	4.030	5.122	27.497
Gradient Boosting	0.824	4.037	5.156	27.460
Random Forest	0.816	4.008	5.269	26.759

Of all the analysed models, Ridge Regression delivered the highest predictive accuracy, with a cross-validated mean  $R^2$  of 0.821 and a test  $R^2$  of 0.826. The MAE and RMSE values were also slightly lower than those of the ensemble models, indicating consistent and reliable generalisation performance. The close alignment of the cross-validation and test metrics suggests limited overfitting and stable generalisation.

The models that performed comparably were Gradient Boosting and Random Forest, with test  $R^2$  values of 0.824 and 0.816, respectively. Despite these models performing best as they capture more complex interactions, the incremental gain in performance was still very small compared to that of Ridge Regression. This suggests that the relationship between emission reduction outcomes and input variables is structured and largely monotonic, with no irregular relationships. As there are more or less constant relationships between the input variables and the outcome, there were no substantial accuracy improvements due to the complexity of the model being added.

The Mean Absolute Percentage Error (MAPE) metric across all models showed a range between 26.8% and 27.5%. Although this shows a low range of emission reduction values and even contains no reduction cases that would typically

increase the percent-based errors, this is an acceptable range. With all models producing similar MAPE values, this study confirms that the choice of model affects the relative error behaviour very little, thus strengthening the reliability of the framework.

In general, the comparable performances of the linear and nonlinear models illustrate that model interpretability does not trade off predictive accuracy in this case. The strong performance of Ridge Regression reinforces its appropriateness as a baseline model for policy-related analyses, whereas the ensemble models offer additional confirmation of the resilience of the findings. This result establishes a basis for applying explainable machine learning methods to identify the factors that most influence performance in reducing carbon emissions.

### 3.2. Predicted vs Actual Emission Reduction

The Ridge Regression model was used to predict carbon emission reductions and was compared with the actual reductions. As shown in Figure 2, the predictions were in a very good linear proportion to the measured values. The different data points were closely aligned with a 1:1 reference line. The model displayed the ability to capture the dominant trends governing emission reduction performance, which was evident across different scenarios.

The model performs particularly well at medium-to-high emission reduction levels (approximately above 25%), where the predicted values closely align with the observed values and exhibit limited dispersion. This indicates that the combined influence of policy, certification, and mitigation measures was consistently captured by the model in higher-reduction scenarios.

The prediction errors were more spread out at lower emission-reduction levels. At lower levels of the target variable emission volume, there is a relatively greater prediction error because there are smaller corresponding absolute values. A more pronounced prediction error results in smaller absolute values. In all models, this behaviour accounts for the moderate Mean Absolute Percentage Error (MAPE) and reflects more on the characteristics of percentage-based appreciation of error rather than on the model bias. Most importantly, their prediction range did not display systematic trends of over or under, indicating an even spread.

Additionally, these outcomes may be a result of different policies and mitigation practices rather than a continuum of policy and mitigation practices, given that clustered outcomes still exist for the predicted-actual relationship. This reinforces the view that the performance of carbon emissions reduction is a result of a more complex inter-structured set of combinations of policy interventions and strategic measures, and not merely the size of the project itself. Overall, the predicted vs. actual comparison showed that the proposed

machine-learning framework could make accurate and consistent predictions. Thus, the framework can support future explainable machine learning analyses and determine the causes of emissions reductions.

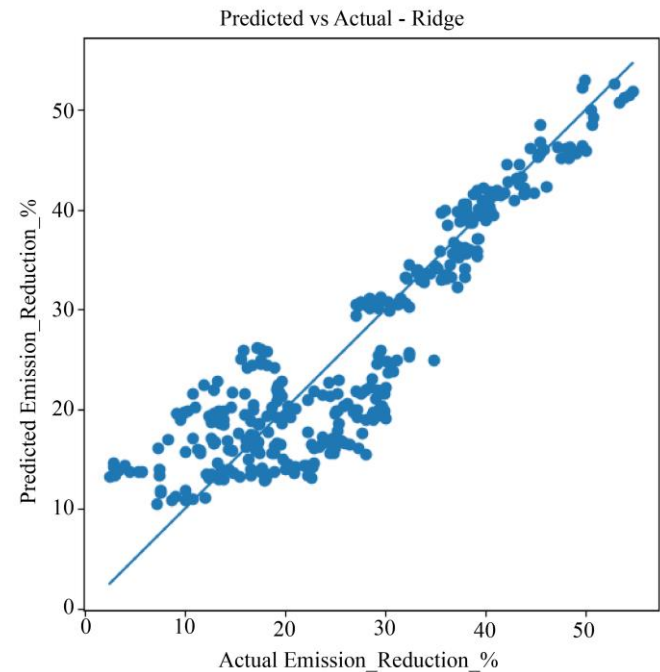


Fig. 2 Predicted and actual carbon emission reduction (%) using the ridge regression model

### 3.3. Explainable Machine Learning Results (SHAP Analysis)

Given its strong predictive performance and inherent interpretability, the Ridge Regression model was selected for the SHAP-based explanation. The global SHAP summary plots for the Ridge Regression model in Figure 3 show, among other things, the global SHAP values for the input variables and their contribution to the predicted percentage reduction in carbon emissions, as well as the increases and decreases in the predicted percentage. Features with positive SHAP values increased the predicted reduction, whereas those with negative SHAP values reduced the predicted reduction.

The findings show that most policies and certification-related variables are the most important factors driving the reduction of carbon emissions. More specifically, the green certification level (platinum and gold) had positive SHAP values that were the greatest, indicating the presence and strength of a consistent contribution of higher levels of emission reductions. Projects with higher certification levels dynamically predict larger reductions, which speaks to the rigor of sustainability standards. Similarly, there were significant positive SHAP contributions for policies, where higher levels of policy interventions were associated with greater predicted reductions, which captures the most important function of the combination of policies and direct and/or indirect incentivising of emission reductions.

In contrast, the building area, renewable energy proportion, reuse of materials, and waste reduction all had similar small SHAP values compared with the other technical and project-scale variables. These factors help reduce emissions; however, in isolation from policies and certification systems, their impact will be far less significant. The building area had a SHAP influence close to zero, which reiterates that the project scale alone will not dictate the level of reduction performance that can be achieved. Variables related to the life cycle phase and baseline emission indicators also had low SHAP values, suggesting that higher total emissions do not lead to proportionately greater reductions. This shows that the results of emission reductions are much more a result of strategies and frameworks than the magnitude of the emissions.

The SHAP analysis shows that, in general, the most important predictors of carbon emission reduction performance are developed policies and certification regimes. This association was not apparent in the simple correlation analysis. The findings of this analysis demonstrate the usefulness of explainable machine learning in identifying the multiple dimensions of a driver in mitigation efficiency and provide clear findings that are relevant to policymaking and foster evidence-based decision-making within the constraints of building decarbonisation.

### 3.4. Policy Implications and Practical Insights

The findings of this study provide important insights into the institutional and practical mechanisms driving carbon emissions reductions in the building sector. The combined predictive modelling and SHAP analysis consistently indicated that policy intervention intensity and green certification levels were the dominant contributors to emission reduction performance, whereas project scale and isolated technical measures exhibited comparatively smaller marginal effects. This suggests that emission mitigation outcomes are strongly conditioned by the regulatory and institutional environment in which projects operate, reinforcing the importance of systemic governance frameworks over purely project-level optimisation.

Policy implications regarding the importance of different levels of green certifications require attention by policymakers in several ways. First, green certification must be viewed as a holistic process rather than a mere procedural stamp, and as such, must incorporate a set of standards that address sustainability in all stages of a building’s life cycle, including the design and operational phases. Second, certification programs must go beyond checklist compliance and focus on measurable emissions reductions. Finally, to remain relevant to changing decarbonization goals and new technologies, certifiers need a systematic process for ongoing changes to certification criteria.

The important role of the intensity of policy interventions provides evidence of the importance of the strategic use of policy instruments and incentives addressing the drivers of building emissions reduction. Such mechanisms as carbon pricing, fiscal mechanisms, performance-based building codes, and mandatory disclosure have proven successful in addressing the financial and institutional barriers that have historically hindered the adoption of the ‘hard’ measures. For the first time, the positive relationship between the strength of the policy context and the projected greater building emissions reductions provides evidence that policy acts as a driver, rather than a secondary role. These results highlight the need for a more integrated approach to decarbonization that prioritizes policy as a driver for the uptake of technical measures, as opposed to just a facilitator.

The suggested framework has good computational efficiency and flexibility to manage extensive datasets, especially when paired with tree-based algorithms and ensemble approaches. This is good for handling high-dimensional data at a low computational cost. This is especially good for efforts at the regional and national levels. It will help the simulation of extensive policy scenarios across diverse portfolio projects. In addition, the SHAP attribution method enhances transparency of impact and is useful for implementation at the institutional decision-support framework. In high accountability and policy defensibility environments, the communicative framework will also be

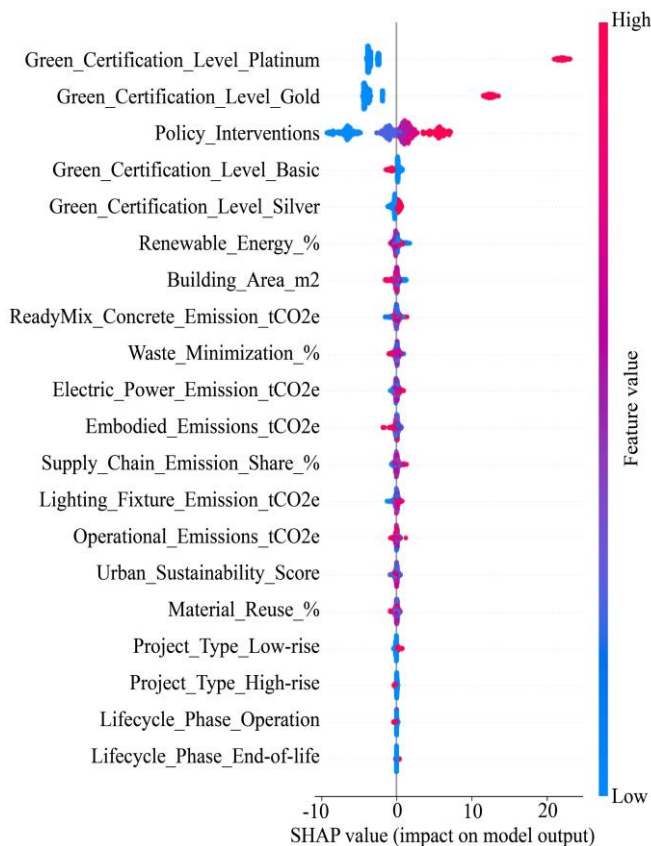


Fig. 3 SHAP summary plot of feature contributions to carbon-emission-reduction predictions

predictive and accurate. This is due to the ability to explain every factor that contributes to the predicted emission reductions. Due to these factors, the building sector has a flexible framework aiding in policy advocacy.

The developed framework has two primary functions. First, using the model as an ex-ante policy evaluation tool, policymakers can estimate the impacts of different scenarios concerning regulatory intensity and certification requirements and their corresponding impacts on emissions reduction. Second, using the framework as a decision-support tool, developers and consultants can evaluate the likely success of mitigation strategies during the pre-construction phase. To ensure the model's ongoing usefulness, it will require regular data updates and recalibrations due to changes in regulation and technology in the construction industry.

### 3.5. Limitations and Future Considerations

There are a number of drawbacks to consider for this study. First, scenario-based data, as opposed to longitudinal field data, offers a limited representation of the spectrum of true behaviours and the complex dynamics of law enforcement and jurisdictions. Second, the operationalization of policy intensity as a simple discrete count indicator neglects the qualitative complexities of the real world, including the many dimensions of enforcement and compliance. Third, the predictive (rather than causal) orientation of the model means that feature contribution analysis can only be understood in a non-direct causal way. Future studies are encouraged to include empirical field data, embrace a causal perspective, and explore the model's applicability in diverse geographical and regulatory settings.

In essence, this study creates a machine learning framework justified as a clear and adaptive decision-support tool for carbon mitigation in the building sector. With the integration of precise predictions and the interpretation of SHAP-based feature contribution analysis, the framework allows for ex-ante assessments of the effectiveness of different policy and mitigation scenario possibilities. This empirical approach facilitates the creation of evidence-based, responsive policies that comply with the desired future state of sustainably targeted goals, all while bolstering the systematic decarbonization of the building sector.

## 4. Conclusion

Based on evaluation results, the proposed framework demonstrates solid predictive capability for all tested machine learning models. The nearly identical results from the various cross-validation schemes and testing on separate data support the model's generalizability and ability to withstand overfitting. Notably, the cross-algorithm comparison discovered that the narrow performance disparity between simple linear models (e.g., Ridge Regression) and ensemble techniques (e.g., Random Forest and Gradient Boosting)

suggests that the emission reduction results are controlled by simple, systematic, and consistent processes. This suggests that the relationship between the predictor and target variables lacks extreme nonlinear complexity or erratic interactions and instead is structurally consistent to the point that even a simple parametric model can effectively capture these characteristics.

Analysis of residuals provides additional support for this finding, as it shows the lack of systematic bias and a consistent error pattern throughout the range of predictions. The remarkable consistency between the predicted and the actual values, especially at moderate to high levels of reductions, shows that the framework consistently works to predict trends of emissions reductions across varied scenarios. This stability and consistency of predictions are especially important for decision support systems, and as such, these features contribute positively to the utility of the framework.

SHAP-based XAI analysis demonstrated a distinct order of variable impacts in assessing emission reduction performance. The most notable result is the overpowering influence of policy and certification factors relative to other contextual and technical considerations. In other words, the predictive value of policy intervention intensity and green certification level is greater than any of the influencing factors alone, including project size, baseline emissions, and specific technical measures.

The substantial positive relationship that exists among higher degrees of green certification and increased policy intensity, combined with greater expected reductions in emissions, demonstrates the importance of the institutional framework in the decarbonization of the building sector. It should be noted that conventional correlation analysis and traditional statistical techniques that treat variables independently will miss the multivariate relationship described here. This further demonstrates the value of explainable machine learning in addressing the complexities due to the interconnectedness of policy, technology, and context.

From a practical perspective, this study's findings provide an understanding of the importance of policy frameworks and enhanced integrated certification schemes as opposed to guaranteed technical optimizations, supporting that without further advancements of frameworks, technical measures lack effectiveness. Meaning, technical measures without policies to incentivise, standardise, and disclose will lack an effective institutional framework.

The explainable machine learning framework created in this study provides transparency and evidence to stakeholders. It can be used by policymakers to evaluate how effective different scenario interventions are, and it can also be used by practitioners and developers to identify the effectiveness of mitigation measures during the project planning phase. Hence,

this framework balances the complexity involved in data analytics and the practical requirements of policymaking.

### Acknowledgements

The authors would like to thank the Karbonara Research Institute for all assistance and resources provided during this research. Special appreciation is given to the Civil Engineering Department of BINUS University for their consistent mentoring and for providing the means and tools required to complete this study.

### Ethical Approval

This study did not involve human participants, animal subjects, or confidential data, requiring ethical approval. All analyses were conducted with respect to academic integrity

and ethical considerations. The authors affirm that the data presented are original and have not been manipulated or fabricated.

### Funding Statement

This work was supported by Bina Nusantara University.

### Author Contribution

RS prepared the manuscript, RH reviewed the manuscript, MK performed the analysis, and CFW reviewed the manuscript.

### Data availability

The dataset and analysis scripts are available at: <https://zenodo.org/records/16757666>.

### References

- [1] Haochun Guan et al., "How Urban Street-Scape Visual Features Influence Carbon Emissions from Residents Visiting Urban Parks: A Case Study of Shenzhen, China," *Landscape and Urban Planning*, vol. 266, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Chao Shi, Ning Xu, and Zongling Niu, "Does Urban Renewal Help Reduce Carbon Emissions? Empirical Evidence from Chinese Cities," *Economic Analysis and Policy*, vol. 89, pp. 384-401, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Limao Zhang, Shiqiang Tang, and Yiheng Su, "Building Carbon Emissions Estimation in Urban Agglomerations Considering Regional Dependence," *Energy*, vol. 344, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Jaeun Koo, Sangchan Park, and Yunsung Lee, "National Circumstances Matter: How Climate Change Vulnerability and Political Instability Affect Greenhouse Gas Coverage in Nationally Determined Contributions," *Environmental Science and Policy*, vol. 175, pp. 1-10, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Linfeng Li et al., "Responses of Greenhouse Gas Fluxes to Climate Extremes in a Semiarid Grassland," *Atmospheric Environment*, vol. 142, pp. 32-42, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Ian Hamilton et al., "Global Status Report for Buildings and Construction: Towards a Zero-Emission," *Efficient and Resilient Buildings and Construction Sector*, 2020. [[Google Scholar](#)]
- [7] Nwe Ni Myint, and Muhammad Shafique, "Embodied Carbon Emissions of Buildings: Taking a Step Towards Net Zero Buildings," *Case Studies in Construction Materials*, vol. 20, pp. 1-20, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Khozema Ahmed Ali, Mardiana Idayu Ahmad, and Yusri Yusup, "Issues, Impacts, and Mitigations of Carbon Dioxide Emissions in the Building Sector," *Sustainability*, vol. 12, no. 18, pp. 1-11, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Tengfei Huo et al., "What is the Driving Mechanism for the Carbon Emissions in the Building Sector? An Integrated DEMATEL-ISM Model," *Energy*, vol. 274, pp. 1-41, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Abdulrahim Ali et al., "Machine Learning as a Surrogate to Building Performance Simulation: Predicting Energy Consumption Under Different Operational Settings," *Energy and Buildings*, vol. 286, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Xiang Li et al., "An Integrated Building Energy Performance Evaluation Method: From Parametric Modeling to GA-NN based Energy Consumption Prediction Modeling," *Journal of Building Engineering*, vol. 45, pp. 1-42, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Fuyu Wang et al., "Research on a Hybrid Model for Cooling Load Prediction based on Wavelet Threshold Denoising and Deep Learning: A Study in China," *Energy Reports*, vol. 8, pp. 10950-10962, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Hong Ye et al., "Modeling Energy-Related CO<sub>2</sub> Emissions from Office Buildings using General Regression Neural Network," *Resources, Conservation and Recycling*, vol. 129, pp. 168-174, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Yujie Cang et al., "Prediction Method of Building Embodied Carbon Emissions for Conceptual Design Stage: based on the Overall Design Parameters by Machine Learning," *Building and Environment*, vol. 289, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Xiaocun Zhang et al., "Predictive Models of Embodied Carbon Emissions in Building Design Phases: Machine Learning Approaches based on Residential Buildings in China," *Building and Environment*, vol. 258, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Lu Luo et al., "Progressive Prediction of Embodied Carbon Emissions Across Stages of Schematic Design with Machine Learning," *Journal of Cleaner Production*, vol. 517, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Kiley Feickert, and Caitlin T. Mueller, "Thin Shell Foundations: Quantification of Embodied Carbon Reduction Through Materially Efficient Geometry," *Architecture, Structures and Construction*, vol. 4, no. 1, pp. 15-36, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [18] Yi-Cheng Zhang, Jia-Bao Liu, and Shan-Yong Wang, "The Impact of Carbon Trading Policy on Embodied Carbon Emission in China's Construction Industry: Evidence from a Quasi-Natural Experiment," *Journal of Cleaner Production*, vol. 522, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Zhan-Ming Chen, Zihao Chen, and Bowei Song, "Emission Reduction Effect and Transmission Mechanism of Carbon Tax from the Perspective of Embodied Carbon Emissions," *Journal of Environmental Management*, vol. 394, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Sheng Lun (Christine) Cao, Destenie Nock, and Alex Davis, "An Analysis of Machine Learning Approaches for Enhancing Decision-Making in Complex Discrete Choice Tasks," *Decision Analytics Journal*, vol. 18, pp. 1-16, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Xingjian Zhang et al., "Machine Learning Nested MCDM Model to Enhance Decision Reliability for Transport Safety Engineering," *Results in Engineering*, vol. 29, pp. 1-20, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] F. Fantechi, and A. Cusimano, "The Counterfactual Challenge: How Machine Learning Can Enhance Policy Evaluation," *Journal of Policy Modeling*, vol. 48, no. 2, pp. 343-355, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Rui Gao, Wen Li, and Qingjun Zeng, "Synergistic Effects of New Energy and Low-Carbon Pilot Dual Policies on Pollution and Carbon Reduction: A Causal Analysis based on Double Machine Learning," *Energy Reports*, vol. 14, pp. 2435-2450, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Dongge Zhu et al., "A New Method for Predicting Carbon Emissions in Energy Industry based on CEEMD Wavelet Denoising and Hybrid Machine Learning," *Energy Reports*, vol. 14, 3132-3141, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] Yishuang Xu, Peng Wei, and Yinze Ji, "Machine Learning for Predicting Corporate Carbon Emissions: The Role of Corporate Governance," *Journal of Digital Economy*, vol. 5, pp. 1-16, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Xiaoqian Huo, and Hui Xing, "New-Type Urbanization and Corporate Carbon Performance: Evidence from Double Machine Learning Model," *Sustainable Futures*, vol. 11, pp. 1-16, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Tzu-Tsung Wong, and Nai-Yu Yang, "Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 11, pp. 2417-2427, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Dani Bromberger et al., "Interpretable Machine Learning in Industrial Water Pump Failure Prediction: The Role of Shapley Additive Explanations Values and Data Sampling Techniques," *Engineering Applications of Artificial Intelligence*, vol. 165, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Chaowei Zhang et al., "Prediction of Compressive Strength and Mix Proportion Design of Coal Gangue Mixtures using Ensemble Learning and SHapley Additive exPlanations," *Case Studies in Construction Materials*, vol. 24, pp. 1-24, 2026. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [30] M.J. Mvita, Nokubonga G. Zulu, and B.M. Thethwayo, "Artificial Neural Network Integrated SHapley Additive exPlanations Modeling for Sodium Dichromate Formation," *Engineering Applications of Artificial Intelligence*, vol. 158, pp. 1-14, 2025. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]