Machine Learning-Assisted Assessment of Low-Carbon Building Materials for Sustainable Construction Engineering

Blanaid O'Beirn
Department of Computer Science, Binghamton University, Binghamton, NY, USA.
blanaid.work@binghamton.edu

Bernarda Pintor
Department of Computer Science, University of Central Florida, Orlando, FL, USA.
pintorbernarda@ucf.edu

Safina Iswahyudi
Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA.
iswahyudi2004@ku.edu

Harshil Chakrabarti
Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.
harshil916@buffalo.edu

Abstract

Low-carbon building materials are central to reducing embodied emissions in construction, yet their assessment is complicated by coupled environmental, mechanical, durability, and cost criteria that are measured at different scales and with unequal reliability. Existing life-cycle assessment and material-selection procedures often treat carbon intensity as a scalar attribute and separate it from the statistical uncertainty of property prediction, while machine-learning studies frequently optimize predictive accuracy for a single mechanical response without embedding engineering feasibility or life-cycle boundaries. This manuscript develops an integrated, uncertainty-aware assessment framework that links material and process descriptors to low-carbon performance indicators through surrogate models, calibrates predictive uncertainty, and converts probabilistic outputs into robust design decisions. The framework combines Gaussian-process and ensemble learning for small and heterogeneous construction-material data, multi-objective optimization for carbon, strength, durability, and cost, and constraint-based decision rules that require acceptable probability of compliance rather than point predictions. It also specifies the physical meaning of key variables, including clinker substitution, recycled aggregate porosity, activator demand, thermal conductivity, and fabrication waste, and explains how these variables alter both environmental burdens and engineering reliability. The main contribution is not a claim of empirical superiority but a coherent methodological structure for assessing low-carbon building materials when data are sparse, heterogeneous, and consequential. The framework provides a defensible basis for comparing material alternatives, identifying trade-offs, and exposing the conditions under which machine-learning predictions should not be used.

Keywords

low-carbon building materials; machine learning; life-cycle assessment; uncertainty quantification; multi-objective optimization; sustainable construction

1. Introduction

Construction materials account for a large share of industrial greenhouse-gas emissions, and cementitious systems dominate that burden because clinker production combines calcination chemistry with high-temperature fuel demand [1]. Technical pathways such as clinker substitution, alternative binders, recycled aggregates, and bio-based components can lower embodied carbon, but their benefits depend on strength development, durability, resource availability, transport, and end-of-life conditions [2]. Consequently, low-carbon material assessment cannot be reduced to ranking candidates by a single emissions factor. It requires a coupled treatment of environmental performance and engineering function, with explicit attention to uncertainty in both material behavior and life-cycle accounting [3].

Life-cycle assessment provides the appropriate accounting structure for embodied energy and carbon, but construction materials introduce difficult allocation questions because the same mixture may be used in different structural roles and service environments [4]. Reviews of embodied-energy analysis show that system boundaries, data sources, and functional-unit definitions can change comparative conclusions [5]. Policy and certification frameworks have therefore moved toward multi-criteria sustainability assessment, although material-level evidence often remains incomplete or inconsistently reported [6]. Alternative material streams illustrate the same tension. Recycled ceramic waste can replace natural aggregate and reduce extraction pressure, but its porosity and absorption affect workability and durability [7]. Alkali-activated systems can reduce clinker demand, yet activator production, curing energy, and long-term performance must be included in any credible low-carbon claim [8].

Machine learning offers a way to integrate these heterogeneous relationships, but most materials-informatics work has been developed for molecular or crystalline systems where simulation and high-throughput data are more abundant than in construction materials [9]. Materials-informatics approaches can learn property-structure relationships from sparse data [10], and recent reviews document both the promise and the fragility of machine learning in solid-state materials science [11]. Concrete-specific reviews show that predictive models for mechanical properties are increasingly common, but they also reveal limited external validation, inconsistent data reporting, and inadequate treatment of uncertainty [12]. The central gap is therefore not a lack of algorithms. It is the absence of an assessment architecture that connects material descriptors, life-cycle carbon accounting, physical performance constraints, uncertainty calibration, and engineering decision rules in one coherent procedure.

This manuscript develops such an architecture for machine-learning-assisted assessment of low-carbon building materials. The contribution is a methodological framework, not a claim of new experimental measurements. It defines the target variables and functional units, specifies surrogate-model choices and data-conditioning requirements, formalizes uncertainty propagation and calibration, and embeds predictions in a multi-objective decision analysis with robustness constraints. The framework is designed to make explicit when a material is genuinely low-carbon under a defined functional unit, when apparent carbon savings are offset by durability or strength penalties, and when data limitations make machine-learning outputs unsuitable for engineering use.

2. Problem Formulation and Assessment Criteria

2.1 Functional Unit and Carbon Accounting

The assessment begins with a functional unit that fixes the service provided by the material, such as a square metre of façade, a cubic metre of structural concrete at a specified strength class, or a square metre of partition with a specified thermal resistance. Carbon intensity is computed per functional unit so that material quantity, density, transport, and replacement frequency are comparable across alternatives. A transparent baseline formulation is:

$$C_m = \sum_k f_k E_k + C_t + C_{eol}$$

Here, C_m is the embodied greenhouse-gas intensity per functional unit for material option m; f_k is the quantity of input k normalized to the functional unit; E_k is the emissions factor for that input, including extraction, processing, and manufacturing; C_t is transport-to-site emissions; and C_{eol} represents end-of-life or replacement burdens allocated to the functional unit. The expression is deliberately additive because it exposes where a low-carbon claim originates. A clinker substitute may reduce E_k for binder production while increasing transport distance or reducing early strength; a recycled aggregate may reduce extraction emissions while increasing cement demand to maintain workability. The equation also implies that emissions factors and allocation choices are not neutral metadata. They are inputs to the assessment and must carry their own uncertainty.

2.2 Material Performance and Feature Space

Low-carbon assessment requires parallel prediction of engineering performance. A material option is represented by a vector of target responses rather than by a single carbon value:

$$y_m = (S_m, \lambda_m, D_m, C_m, R_m)$$

In this vector, S_m denotes a strength or stiffness measure appropriate to the functional unit, λ_m denotes thermal conductivity where relevant, D_m denotes a durability or service-life indicator, C_m is the carbon intensity from the previous equation, and R_m is a normalized cost or resource-risk indicator. The inclusion of D_m is essential because a material with lower initial carbon but shorter service life can have higher life-cycle emissions after replacement is counted. The inclusion of λ_m reflects the role of thermal performance in building operation and envelope design. The feature vector for model input should describe mixture composition, processing, and service conditions: binder type and clinker factor, supplementary cementitious material dosage, water-to-binder ratio, aggregate type and replacement level, admixture dosage, fiber content, curing regime, intended exposure class, design service life, and transport distance. Categorical variables such as binder family should be encoded hierarchically so that the model can share information across related mixtures without treating chemically distinct systems as interchangeable.

3. Machine Learning-Assisted Assessment Framework

3.1 Surrogate Modeling and Data Conditioning

The assessment framework uses supervised surrogates to map material and process descriptors to the target vector in Equation 2. A Gaussian-process formulation is appropriate when data are sparse and uncertainty must be represented explicitly:

$$\hat{y}_j(x) = \mu_j + \sum_i \alpha_{ij} k(x, x_i)$$

Here, $\hat{y}_j(x)$ is the predicted value of target j for a material descriptor x; μ_j is a mean function; α_{ij} are weights determined by the training data; and $k(x, x_i)$ is a positive-definite kernel that encodes similarity between the query material and training observation i. The kernel can combine continuous-distance terms for dosage and curing variables with categorical terms for binder family. This formulation is useful because its predictive variance increases away from observed mixtures, which is exactly the behavior needed when a proposed low-carbon formulation lies outside the training domain. Tree-ensemble models provide a complementary option for tabular data with non-smooth interactions and mixed variable types [14], while Gaussian processes provide smoother interpolation and calibrated variance in small-data regimes [15]. A practical framework can therefore maintain multiple candidate surrogates and select among them through grouped validation rather than through in-sample fit.

Data conditioning must precede model fitting. Mixture quantities should be converted to consistent mass or volume bases, emissions factors should be aligned to a declared life-cycle inventory, and test results should be normalized to specimen geometry, age, curing, and loading rate. Missing values require explicit treatment because deletion can bias the model toward well-reported material families. Multiple imputation or probabilistic matrix completion can be used when missingness is plausibly random, but missingness related to material class must be modeled as a group effect. Data leakage is a serious risk: if specimens from the same batch, laboratory, or publication appear in both training and test partitions, validation will overstate generalization. Grouped cross-validation by mixture family, source study, or production batch is therefore necessary. A multi-target training objective can be written as:

$$L(\theta) = \sum_j w_j \ell_j(y_j, \hat{y}_j(\theta)) + \lambda \|\theta\|^2$$

In this loss, θ denotes model parameters, w_j are target-specific weights, ℓ_j is a loss function for target j, y_j is the observed response, $\hat{y}_j(\theta)$ is the prediction, λ is a regularization strength, and $\|\theta\|^2$ is a quadratic penalty. The weights should reflect engineering consequence rather than convenience. A large error in a durability indicator that controls service life may deserve more weight than a small error in a non-critical thermal property.

3.2 Uncertainty, Calibration, and Interpretability

Predictive uncertainty should be separated into aleatoric and epistemic components. Aleatoric uncertainty arises from specimen variability, measurement error, and uncontrolled curing; epistemic uncertainty arises from limited data, model-form choice, and extrapolation. Gaussian-process variance supplies an epistemic component under the kernel assumptions, while ensemble disagreement or approximate Bayesian neural-network variance can supply additional evidence [17]. Calibration is the test of whether predicted intervals have the coverage they claim. If a model reports a 90% prediction interval, approximately 90% of held-out observations should fall inside it within tolerable sampling error. Calibration should be assessed by material class and exposure condition, because a model may be well calibrated for ordinary Portland-cement concrete and poorly calibrated for alkali-activated systems.

Interpretability is not an optional reporting layer in this context. Material engineers need to know whether a low-carbon prediction is driven by credible chemical and physical mechanisms or by incidental correlations. Feature-attribution methods such as Shapley-value

explanations can identify the contribution of clinker factor, aggregate replacement, and curing variables to a prediction [16]. Attribution should be checked against known mechanisms: increasing clinker substitution often lowers embodied carbon but may slow early strength; increasing recycled aggregate absorption may reduce workability and increase required water or admixture. If an attribution contradicts established materials science, the model should be treated as suspect even if its aggregate error is low. A probability-of-compliance rule converts uncertain predictions into an engineering decision:

$$P_pass(x) = \int 1[g_i(y) \geq 0 \; \forall i] \, p(y|x) \, dy$$

Here, P_pass(x) is the probability that a material with descriptor x satisfies all engineering constraints; $g_i(y) \geq 0$ denotes the i-th constraint, such as minimum strength, maximum thermal conductivity, or minimum durability index; $1[\cdot]$ is the indicator function; and p(y|x) is the predictive distribution of performance. This equation is preferable to thresholding a point prediction because it makes risk explicit. A material with a mean strength above the required value but a wide predictive distribution may still be unacceptable for a structural application.

4. Multi-Objective Decision Analysis

4.1 Objectives and Constraints

The selection problem is inherently multi-objective. A designer seeks low carbon intensity, low cost, high strength, high durability, and appropriate thermal performance, but these objectives can conflict. A constrained Pareto formulation is:

$$min_x \, (C(x), \; -S(x), R(x)) \; s.t. \; g_i(x) \geq 0$$

In this expression, C(x) is predicted carbon intensity, S(x) is a strength or performance index, R(x) is cost or resource risk, and $g_i(x) \geq 0$ represents feasibility constraints such as minimum strength class, maximum water absorption, chloride ingress limits, or fire performance. The negative sign on S(x) converts maximization of strength into minimization. The solution is not a single material but a set of nondominated candidates whose trade-offs are transparent. A candidate that lowers carbon by sacrificing durability should not be ranked as superior merely because its carbon coordinate is lower; it belongs to a different trade-off region and may be excluded by the feasibility constraints.

4.2 Robust Decision Rules and Sensitivity

The final decision rule should combine Pareto dominance with a compliance threshold. A material is admissible only if it is nondominated in the objective space and if P_pass(x) exceeds a project-specific reliability target. This two-stage rule prevents the optimizer from exploiting uncertain predictions near constraint boundaries. Sensitivity analysis should then identify which inputs most strongly influence the carbon and performance estimates. A normalized local sensitivity can be expressed as:

$$\eta_k = (\partial \hat{C}/\partial x_k)(\sigma_{xk}/\sigma \hat{C})$$

Here, η_k is the normalized sensitivity of predicted carbon $\hat{C}$ to input k; $\partial\hat{C}/\partial x_k$ is the partial derivative of the carbon prediction with respect to that input; σ_{xk} is the standard deviation of input k; and $\sigma\hat{C}$ is the standard deviation of the carbon prediction. This quantity is useful because it combines model response with input variability. A binder with a large emissions-factor effect but tightly controlled dosage may be less influential than a recycled aggregate whose replacement level varies widely in practice. Sensitivity results should be

reported alongside the assumed life-cycle inventory and functional unit, because changing transport distance or allocation method can alter the ranking of options.

5. Technical Discussion: Error Modes, Trade-offs, and Validation Logic

5.1 Physical Extrapolation and Dataset Shift

The most serious failure mode for machine-learning-assisted material assessment is silent extrapolation. A model trained on conventional concrete may produce plausible numeric outputs for a novel alkali-activated binder, a bio-based insulation composite, or a high-volume recycled-aggregate system even though the underlying hydration, porosity, and transport mechanisms differ. Dataset shift is not only statistical; it is physical. New binder chemistry changes phase assemblage, new aggregates change interfacial transition zones, and new fabrication routes change compaction and anisotropy. The framework should therefore include a domain-of-validity check based on distance in feature space, material-class matching, and mechanistic review. Predictions outside the validated domain should be reported as hypotheses requiring experimental confirmation, not as design values.

5.2 Life-Cycle Boundaries and Measurement Error

Life-cycle assessment introduces its own error modes. Emissions factors may be regional averages, transport distances may be estimated rather than measured, and end-of-life allocation may treat demolition, recycling, or landfilling inconsistently. The ISO 14040 framework provides general principles for life-cycle assessment, but it does not resolve every allocation choice for construction materials [19]. A low-carbon comparison should therefore state the system boundary, the functional unit, the service life, and the treatment of biogenic carbon and recycling credits. Uncertainty in emissions factors should be propagated into the carbon prediction, not hidden in a single deterministic value. Alternative cements and supplementary materials can lower process emissions, but their benefits depend on activator production, grinding energy, and curing [20]. Broader industrial ecology arguments likewise show that concrete sustainability requires attention to durability, material efficiency, and replacement rates rather than clinker substitution alone [21]. Materials-selection methods that compare embodied impacts with mechanical function provide a useful cross-check on model outputs [22].

5.3 Engineering Adoption and Model Governance

For adoption in construction engineering, the model must be auditable. This requires versioned datasets, documented preprocessing, reported hyperparameters, explicit validation splits, and calibrated prediction intervals. It also requires a clear separation between screening and design. A surrogate model may be appropriate for screening candidate mixtures, identifying promising ranges for testing, and comparing trade-offs, but structural design values should come from standardized testing and code-conforming verification. The framework should therefore label each prediction with its evidence level: interpolative, near-domain, or extrapolative. Human review should focus on mechanistic consistency, life-cycle boundary assumptions, and the consequences of false acceptance. A false acceptance is more costly than a false rejection when a low-carbon material is proposed for a load-bearing or moisture-exposed application.

5.4 Validation Logic Without Empirical Overstatement

Validation should proceed from internal to external and from statistical to engineering criteria. Internal validation includes grouped cross-validation, calibration coverage, and

sensitivity stability. External validation requires held-out data from independent laboratories, material suppliers, or construction projects. Engineering validation requires comparison with standardized test results and, where relevant, physical mock-ups or field exposure. None of these steps can be replaced by high aggregate R^2. A model that predicts a suite of properties accurately on average but fails near code-defined strength or durability thresholds is unsuitable for design. The proposed framework therefore treats validation as a hierarchy of claims: predictive calibration, mechanistic plausibility, life-cycle representativeness, and engineering compliance. Each claim has different evidence requirements and different consequences for decision-making.

6. Conclusion

This manuscript has developed a machine-learning-assisted assessment framework for low-carbon building materials that integrates life-cycle carbon accounting, engineering performance prediction, uncertainty calibration, interpretability, and robust multi-objective decision analysis. The framework treats embodied carbon as a functional-unit-dependent quantity rather than an isolated material attribute and represents material performance as a vector of coupled responses, including strength, thermal conductivity, durability, carbon intensity, and cost. Gaussian-process surrogates and tree-ensemble models are positioned according to data regime and uncertainty requirements, while grouped validation, calibration checks, and feature-attribution reviews guard against the most common failure modes in construction-material machine learning. The constrained Pareto formulation and probability-of-compliance rule convert uncertain predictions into decisions that distinguish screening from design and false acceptance from false rejection. The main scientific implication is that low-carbon material assessment should be organized around defensible evidence levels, transparent life-cycle boundaries, and explicit engineering constraints. The main practical implication is that machine learning can support sustainable construction most credibly when it is used to expose trade-offs, identify data gaps, and quantify risk rather than to replace material testing or life-cycle verification.

References

1. Habert, G., Miller, S. A., John, V. M., Provis, J. L., Favier, A., Horvath, A., & Scrivener, K. L. (2020). Environmental impacts and decarbonization strategies in the cement and concrete industries. Nature Reviews Earth & Environment, 1(11), 559–573. https://doi.org/10.1038/s43017-020-0093-3

2. Scrivener, K. L., John, V. M., & Gartner, E. M. (2018). Eco-efficient cements: Potential economically viable solutions for a low-CO2 cement-based materials industry. Cement and Concrete Research, 114, 2–26. https://doi.org/10.1016/j.cemconres.2018.03.015

3. Monteiro, P. J. M., Miller, S. A., & Horvath, A. (2017). Towards sustainable concrete. Nature Materials, 16(7), 698–699. https://doi.org/10.1038/nmat4930

4. Hammond, G. P., & Jones, C. I. (2008). Embodied energy and carbon in construction materials. Proceedings of the Institution of Civil Engineers - Energy, 161(2), 87–98. https://doi.org/10.1680/ener.2008.161.2.87

5. Dixit, M. K. (2017). Life cycle embodied energy analysis of residential buildings: A review of literature to investigate embodied energy parameters. Renewable and Sustainable Energy Reviews, 79, 390–413. https://doi.org/10.1016/j.rser.2017.05.051

6. Kylili, A., & Fokaides, P. A. (2017). Policy trends for the sustainability assessment of construction materials: A review. Sustainable Cities and Society, 35, 280–288. https://doi.org/10.1016/j.scs.2017.08.028

7. Pacheco-Torgal, F., & Jalali, S. (2010). Reusing ceramic wastes in concrete. Construction and Building Materials, 24(5), 832–838. https://doi.org/10.1016/j.conbuildmat.2009.10.023

8. Provis, J. L., & van Deventer, J. S. J. (Eds.). (2014). Alkali activated materials: State-of-the-art report, RILEM TC 224-AAM. Springer. https://doi.org/10.1007/978-94-007-7672-2

9. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. Nature, 559(7715), 547–555. https://doi.org/10.1038/s41586-018-0337-2

10. Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., & Kim, C. (2017). Machine learning in materials informatics: Recent applications and prospects. npj Computational Materials, 3, 54. https://doi.org/10.1038/s41524-017-0056-5

11. Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials, 5, 83. https://doi.org/10.1038/s41524-019-0221-0

12. Chaabene, W. B., Flah, M., & Nehdi, M. L. (2020). Machine learning prediction of mechanical properties of concrete: Critical review. Construction and Building Materials, 260, 119889. https://doi.org/10.1016/j.conbuildmat.2020.119889

13. Yu, L., Luo, D., & Xu, W. (2018). Dynamic robotic slip-form casting and eco-friendly building façade design. In Robotic fabrication in architecture, Art and Design (pp. 421-433). Cham: Springer International Publishing.

14. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

15. Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. MIT Press.

16. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (pp. 4765–4774). Curran Associates.

17. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In Proceedings of the 33rd International Conference on Machine Learning (Vol. 48, pp. 1050–1059). PMLR.

18. Miller, S. A., Horvath, A., & Monteiro, P. J. M. (2016). Readily implementable techniques can cut annual CO2 emissions from the production of concrete by over 20%. Environmental Research Letters, 11(7), 074029. https://doi.org/10.1088/1748-9326/11/7/074029

19. International Organization for Standardization. (2006). Environmental management—Life cycle assessment—Principles and framework (ISO 14040:2006). ISO.

20. Gartner, E. (2004). Industrially interesting approaches to low-CO2 cements. Cement and Concrete Research, 34(9), 1489–1498. https://doi.org/10.1016/j.cemconres.2004.01.021

21. Meyer, C. (2009). The greening of the concrete industry. Cement and Concrete Composites, 31(8), 601–605. https://doi.org/10.1016/j.cemconcomp.2008.12.010

22. Ashby, M. F. (2013). Materials and the environment: Eco-informed material choice (2nd ed.). Butterworth-Heinemann.