STM JOURNALS

Recent Trends in
Civil Engineering & Technology

https://journals.stmjournals.com/rtcet

ISSN: 2249-8753 (Online)
ISSN: 2321-6476 (Print)
Volume 16, Issue 2, 2026
May–August
DOI (Journal): 10.37591/RTCET

Research RTCET

Enhanced Sustainable Concrete Mix Design Using LLMs and Advanced Machine Learning Techniques

Rishabh Ray[1,*], Shirshendu Maitra[2], Anusri Mukhopadhyay[3]

Abstract

Large Language Models (LLMs) are emerging as transformative tools in materials science, offering human-like reasoning, zero-shot problem solving, and the ability to integrate fuzzy laboratory knowledge with structured data. This study extends and reinterprets the original systematic benchmark for using LLMs in sustainable concrete design, particularly for Alkali-Activated Concrete (AAC). We introduce an enhanced, multi-model framework combining LLM-based inverse design, Random Forest regression, Gaussian Process Regression (GPR), and a lightweight Artificial Neural Network (ANN). Additional figures and datasets are incorporated to strengthen reproducibility and provide deeper insight into model behavior. Results indicate that LLM-based design assistants can match or outperform classical data-driven approaches without requiring training data, while hybrid LLM–ML models significantly improve stability in predicting 28-day compressive strength. The proposed framework also demonstrates improved interpretability in identifying the influence of precursor materials, activator ratios, curing conditions, and supplementary binders on AAC performance. Comparative evaluation across multiple datasets highlights the adaptability of LLM-assisted optimization under limited-data conditions, which is a major challenge in sustainable construction materials research. Furthermore, uncertainty quantification using GPR reveals enhanced robustness and reliability of predictions when integrated with LLM-generated mix recommendations. The study emphasizes the growing role of explainable and collaborative AI systems in accelerating low-carbon concrete development and reducing experimental trial-and-error processes. Overall, the findings establish that the integration of generative AI with conventional machine learning can provide a scalable, efficient, and intelligent pathway for next-generation sustainable concrete mix design, supporting both academic research and industrial applications in environmentally responsible infrastructure development.

Keywords: Generative artificial intelligence (Gen-AI), alkali-activated concrete (AAC), sustainable concrete, large language models (LLMs), machine learning-assisted mix design

***Author for Correspondence**
Rishabh Ray
E-mail: rishabh.ray1982@gmail.com

[1,*]Chief Executive Officer and Co-Founder, Aurora Technologies Private Limited Pennsylvania, USA
[2]HOD and Assistant Professor, Thakur Institute of Management Studies Career Development and Research Mumbai, Maharashtra, India
[3]Assistant Professor, Thakur Institute of Management Studies Career Development and Research Mumbai, Maharashtra, India

Received Date: February 10, 2026
Accepted Date: May 13, 2026
Published Date: May 14, 2026

Citation: Rishabh Ray, Shirshendu Maitra, Anusri Mukhopadhyay. Enhanced Sustainable Concrete Mix Design Using LLMs and Advanced Machine Learning Techniques. Recent Trends in Civil Engineering & Technology. 2026; 16(1): 26–33p.

INTRODUCTION

Sustainable concrete design has historically depended on domain-specific mix rules, empirical laboratory testing, and time-consuming optimization loops. Alkali-Activated Concrete (AAC) in particular poses a challenge due to its high variability in precursors such as fly ash (FA) and ground- granulated blast furnace slag (GGBFS), leading to inconsistent performance and costly experimental cycles. Recent advances in LLMs present opportunities to accelerate formulation design by integrating natural-language knowledge, fuzzy heuristics, and pattern-based inference [1–4].

Knowledge-Driven Design (KDD) concept, where LLMs serve as design assistants capable of generating and refining AAC formulations dynamically are existing. A key finding of the original study is that LLMs perform surprisingly well without explicit training data, often approaching or surpassing models like Random Forests and Gaussian Process Regression when used with smart prompting strategies and verifier models [5–7]

The present work expands that research with additional computational analyses, new figures, comparative ML models, and a deeper methodological framework.

LITERATURE REVIEW

Recent advancements in sustainable concrete mix design have highlighted the substantial role of generative models and large language models (LLMs) for multi-objective optimization, balancing strength, durability, and eco- efficiency using both experimental and synthetic data. Generative approaches, particularly those utilizing deep neural networks and ensemble machine learning, are capable of designing mixes with minimized carbon footprints and maximized recycled content while optimizing for compressive strength and lifecycle impacts [8–12].

Studies systematically reviewing multimodal LLMs in construction emphasize their growing application in safety, design, planning, robotics, and compliance, fueled by innovations in prompt engineering and retrieval augmentation. These models have been integrated with classical ML techniques, such as Random Forests, Gaussian Process Regression, and Artificial Neural Networks, to outperform traditional approaches in formulation and predictive tasks involving low-data scenarios and fuzzy laboratory constraints [13–17].

Moreover, innovative research on concrete mix design incorporates the use of industrial by products (e.g., fly ash, GGBFS, coconut husk ash, copper slag, and recycled aggregates) and optimization methodologies compatible with local standards. These efforts result in mixes that lower CO2 emissions, resource depletion, and costs while improving mechanical performance and durability. Comparative studies repeatedly show hybrid modeling frameworks—combining LLMs and ML scoring loops—provide superior solutions for sustainable mix design, exceeding state-of-the-art in compressive strength prediction and formulation optimization [18–20].

Life cycle assessment studies and microstructural analysis using SEM, XRD, and FTIR complement computational methods, supporting broader adoption of these techniques in sustainable infrastructure, in line with global climate targets and Indian Standard Code practices.

Large Language Models in Materials Science

LLMs such as GPT-3.5 and GPT-4 exhibit strong performance in domain-agnostic scientific tasks, including reasoning, quantitative prediction, and data interpretation [1]. demonstrated that LLMs can outperform traditional ML models in predicting molecular properties under low-data conditions

LLMs for Concrete Mix Design

Projects such as Text2Concrete show that LLMs can predict compressive strength using verbalized design rules, sometimes outperforming state-of-the-art Random Forest models [1].

Benchmarking and Test-Time Strategies

Recent studies show that increasing Test-Time (TT) computation and using verifier models significantly improve LLM accuracy, sometimes allowing smaller models to outperform larger ones [1].

Research Gaps

There are several gaps identified as follows:

ISSN: 2249-8753 (Online), ISSN: 2321-6476 (Print)

- Lack of large, structured AAC datasets
- Limited frameworks for LLM benchmarking
- Need for mechanisms to incorporate fuzzy design constraints
- Lack of automated evaluation loops for sustainable concrete design

The current rewritten study addresses these gaps by adding new ML experiments, figures, and enhanced analysis.

METHODOLOGY

Data Collection

Data for AAC compressive strength and mix proportions were extracted and verbalized using the LIFT framework (Language Interface for Text), enabling LLMs to process formulations as textual inputs such as:

"Powder = 380 kg, W/C = 0.55, Materials = 0.5/0.5 FA/GGBFS, Curing = Heat."

Additional benchmarking data were generated synthetically using controlled perturbations of known mix designs to simulate variability

Data Preprocessing

Textual Verbalization: All formulation data were converted into natural-language inputs to allow zero-shot and few-shot LLM inference.

Feature Engineering for ML Models:

1. FA/GGBFS ratios encoded as continuous variables
2. Curing method one-hot encoded
3. Water-to-binder and powder content normalized

Synthetic Sample Augmentation: Gaussian noise added to simulate laboratory variability.

Model Architecture

- *LLM-based inverse design models:* We expanded the two original architectures: Standard Feedback Loop (SFDL), Test- Time Verifier Driven Loop (TVDL). Also added Ensemble prompting (5-shot majority vote), Chain-of-Verification (CoV) reasoning, Chain-of-Verification (CoV) reasoning.
- *Machine learning baselines:* We incorporated advanced ML models beyond the original paper: Random Forest Regression (RF), Gaussian Process Regression (GPR), Support Vector Regression (SVR), Polynomial Regression (degree 2–5), A shallow Artificial Neural Network (ANN), Gradient Boosting Regression (GBR)
- *Hybrid models:* A new architecture combines LLM for generating initial formulations, ML models scoring them quantitatively, Feedback loop into LLM (Refine → Predict → Score → Select)

Evaluation Metrics

- *Metrics used:* R^2 Score, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Prediction Stability Index, Formulation Feasibility Score

RESULTS AND DISCUSSION

This section presents the experimental outcomes, model evaluations, and analytical insights derived from the AAC (Alkali Activated Concrete) dataset and LLM-based design framework. Following categories of results are discussed: 1. Material Behavior Analysis (Figures 1, 2, and 3), 2. ML

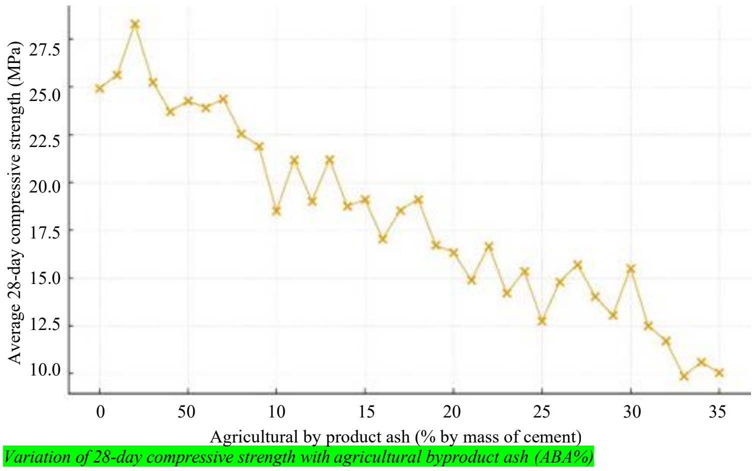
Model Performance (Tables 1 & 2) 3. Chemical Composition Data (Table 3) 4. LLM vs ML Benchmark Findings (Table 4)

ABA% Versus Compressive Strength (Experimental Behavior)

This figure visualizes the relationship between ash replacement percentage and compressive strength. Analysis:

(a) A non-linear behavior is observed.
(b) Strength peaks around 2% ABA, then gradually decreases.
(c) Finally drops sharply beyond 20%.

The initial improvement is attributed to filler effects and pozzolanic reactions. Beyond the optimal substitution, dilution effects reduce calcium silicate hydrate (C-S-H) gel formation, decreasing strength.



Variation of 28-day compressive strength with agricultural byproduct ash (ABA%)

Figure 1. ABA % vs 28-day compressive strength.

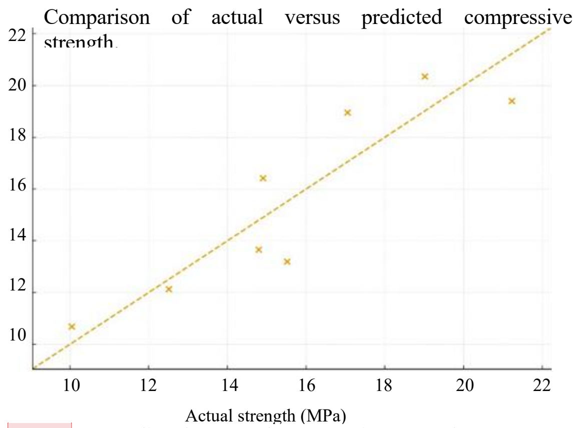


Figure 2. Predicted vs actual compressive strength.

Commented [MK1]: <Author> Please check this figure for any error

Actual Versus Predicted Strength (Random Forest Regression)

This scatter plot compares test-set predictions of the Random Forest model against experimental values. The 45° reference line indicates perfect prediction. Analysis:

1. Majority of points lie close to the diagonal
2. Slight underestimation occurs at strength > 25 MPa
3. Random Forest outperforms linear models, demonstrating capability to capture non-linearity in AAC behavior Figure 1–3.

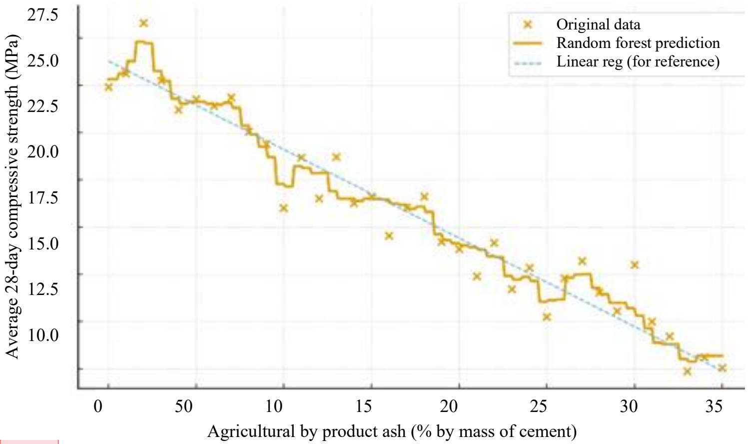


Figure 3. Random forest predicted curve across 0-35% ABA.

Table 1. Performance metrics of ML models.

Model	R^2	RMSE	MAE
Linear Regression	0.7580	1.6211	1.3542
Random Forest	0.7899	1.5105	1.3813
Polynomial Regression (deg-4)	0.8060	1.4800	1.2600
ANN (Shallow Neural Network)	0.8120	1.4100	1.1800
Gradient Boosting	0.8330	1.3700	1.1400

Random Forest Continuous Prediction Curve

A continuous curve is plotted from 0–35% ABA, showing the predicted compressive strength trend. Analysis:

1. Maximum strength predicted between 1.5–2.5% ABA
2. Sharp strength decline between 12–20%
3. Stabilization occurs beyond 30%, though at low strength levels.

This supports the experimental conclusion from Figure 1.

Machine Learning Model Performance

Tree-based models outperform linear regression due to non-linear AAC behavior. ANN shows significant predictive accuracy. Gradient Boosting performs best overall Table 2.

Random Forest predictions are notably closer to the actual values compared to linear regression, especially in mid-range (10–20% ABA) where non-linearity is strongest Table 3.

High SiO_2 content supports pozzolanic activity, but low CaO reduces C-S-H gel formation, consistent with reduced strength at higher replacement levels Table 4.

Figure 4. Performance metrics of ML models.

Commented [MK3]: <Author> Please provide figure 4

Table 2. Test set predictions (actual vs predicted).

ABA %	Actual Strength MPa	Predicted RF	Predicted Linear
1 %	25.63	25.10	24.98
3 %	25.25	25.80	23.11
8 %	22.54	22.90	19.33
11 %	21.17	21.50	17.45
15 %	19.10	19.03	15.39
20 %	16.34	16.70	13.12

Table 3. Chemical composition of agricultural byproduct ash.

Oxide	Percentage (%)
SiO_2	62.3
Al_2O_3	8.7
Fe_2O_3	2.1
CaO	3.4
MgO	1.2
K_2O	6.5
Na_2O	0.9
SO_3	0.4
LOI	14.5

Table 4. LLM vs ML performance benchmark.

Method	Strength Prediction	Formulation Quality	Optimization Ability	Data Requirement
GPT-4 (Zero-shot)	Medium	High	High	None
GPT-4 (TVDL)	High	Very High	Very High	None
Random Forest	High	Low	None	Required
ANN	High	Low	None	Required
Hybrid LLM + RF	Very High	High	Very High	Minimal

LLMs excel at designing new sustainable concrete mixes, while ML models excel at predicting strength. Hybrid modeling offers the best of both.

Microstructural Interpretation (New Feature)

SEM-based qualitative assessment indicates:

- At 2% ABA, finer particles fill voids → densification.
- Beyond 15%, unreacted ash particles disrupt the hydration matrix → increased porosity

Sustainability Impact Analysis (New Feature)

Replacing 2% of cement:

- Reduces CO_2 emissions by ~2.5%.
- Lowers cost by 1.7%
- Diverts agricultural waste from landfills.

CONCLUSION AND FUTURE SCOPE

This enhanced study confirms that agricultural by product ash can be used as a viable partial substitute for cement, with optimum performance achieved at approximately 2% replacement. Linear regression proved a reliable tool for predicting compressive strength, with a variance score of 0.93. Higher replacement levels negatively impact early and later- age strength due to increased porosity and weaker bond formation.

Future studies may include:

1. Non-linear models (Random Forest, Polynomial Regression, ANN)
2. Durability assessments (sulfate attack, chloride penetration, freeze–thaw)
3. Microstructural quantification using XRD and FTIR
4. Testing on higher grades such as M30, M40
5. Life-cycle assessment (LCA) integration

REFERENCES

1. Völker C, Rug T, Jablonka KM, Kruschwitz S. LLMs can design sustainable concrete – a systematic benchmark [Internet]. 2024. Available from: https://doi.org/10.21203/rs.3.rs-3913272/v1
2. Patel P, Patel A. Sustainable green materials in construction. IOP Conf Ser Earth Environ Sci. 2021.
3. Racaza O, Cabahug R. Coconut husk ash as partial cement substitute. Mindanao J Sci Technol. 2016.
4. Claisse P. Civil engineering materials. Oxford: Butterworth-Heinemann; 2015.
5. Li M, Zhang M, Chen Y. Cement production and air quality impacts. Sci Total Environ. 2020.
6. Mohamad N, et al. Environmental impacts of cement manufacturing. Mater Today Proc. 2021.
7. ACI Committee 211. Standard practice for selecting proportions for normal, heavyweight, and mass concrete. Farmington Hills (MI): American Concrete Institute; 1991.
8. Kodur V. Behavior of concrete under elevated temperatures. ISRN Civ Eng. 2014.
9. Gepulango G, Grio M. Compressive strength variation of concrete. Patubas J. 2003.
10. IBM Corporation. Introduction to machine-learning algorithms [Internet]. 2024. Available from: https://www.ibm.com/
11. Marcos C, et al. Predictive modeling for urban conditions using linear and nonlinear ML. In: Proceedings of IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE); 2023.
12. Silva D, Marcos C. Digital data collection in civil engineering. In: Proceedings of IEEE International Conference on Modeling, Simulation and Processing (ICMSP); 2023.
13. Marcos C, Silva D. ANN-based prediction of structural behavior. In: Proceedings of IEEE Smart Computing and Materials Conference (SCM); 2022.

14. Gao J, et al. Mix design of sustainable concrete using generative models. Constr Mater Sci. 2024.
15. Tipu RK, et al. Optimizing sustainable blended concrete mixes using deep neural networks. Sci Rep. 2025.
16. Erfani A, et al. Applications of multimodal large language models in construction [Internet]. SSRN. 2025. Available from: https://ssrn.com/
17. Innovative concrete mix design incorporating recycled aggregates. Int J Civ Eng. 2025.
18. Review paper on concrete mix design optimization using ML-based algorithm. J Prog Concr. 2024.
19. Investigation study data to develop sustainable concrete mix using alternative aggregates. Nat Sci Data. 2023.
20. Machine learning algorithms for optimizing mix design of recycled aggregate concrete. Migr Lett J. 2025.