Generative Artificial Intelligence Guided Sustainable Cementitious Material Design

Teiboklang Chyne[1] and Biranchi Panda[1*]

Indian Institute of Technology Guwahati, Assam, India

Abstract. Concrete production is increasingly criticized for its high carbon footprint, prompting the need for sustainable mix design strategies. Traditional methods rely on trial-and-error, making them time-consuming, costly, and often inefficient. This study presents a generative artificial intelligence framework integrating a Conditional Variational Autoencoder (CVAE), a Firefly-optimized CatBoost regression model, and NSGA-II for multi-objective optimization. The CVAE reduces the high-dimensional design space into a compact latent space, enabling efficient sampling of viable mixes. Strength is predicted using CatBoost, while cost and CO_2 emissions are computed directly from material quantities. Explainable AI techniques and statistical analysis ensure practical feasibility. The framework is validated on a public UHPC dataset, achieving over 50% cost reduction and over 40% reduction in embodied carbon without compromising compressive strength. This approach offers a robust, data-driven solution for sustainable and economically viable concrete mix design, with promising implications for widespread adoption in the construction industry.

Keywords: Sustainable Concrete Design, Generative Models, Multi-objective Optimization, CatBoost, CVAE, NSGA-II, Embodied Carbon

1 Introduction

Concrete is a foundational material in modern construction, but its key ingredient—Portland Cement—is a major contributor to CO_2 emissions. Addressing this environmental concern requires a shift toward low-carbon concrete technologies that balance sustainability with economic feasibility. The conventional approach used by many laboratories and companies is an empirical trial-and-error method guided by human expertise [1], making the process inefficient and often suboptimal in terms of performance and cost. Variability in raw material availability and the complex trade-offs between cost and quality further complicate the development of sustainable concrete mixtures. Designing optimal mixtures involves balancing multiple objectives, including strength, durability, workability, cost, and emissions.

*Corresponding author: pandabiranchi@iitg.ac.in

While traditional approaches rely on expert intuition and repetitive experimentation, recent advances in research aimed at accelerating the search for optimized mix designs have demonstrated that predictive models—including neural networks, gradient boosting algorithms, and neuro-fuzzy systems—are effective at estimating properties such as compressive strength and workability [2, 3, 4]. Ensemble-based methods like CatBoost and XGBoost have shown high accuracy in predicting key properties, achieving R^2 values above 0.9 in many studies [3, 4]. Additionally, generative models have been proposed to learn the underlying structure of high-dimensional mix design data [5].

Despite these advancements, the application of multi-objective optimization in concrete mix design is hindered by several critical challenges, such as the curse of dimensionality and complex ingredient interactions. Traditional optimization techniques often rely on predefined constraints and offer limited adaptability [6, 7]. To overcome these limitations, this study introduces a novel data-driven framework that combines a generative model, a regression model, and a multi-objective optimization algorithm. In contrast to conventional trial-and-error approaches, the proposed framework captures the underlying constraints and interdependencies directly from the data. This enables a more efficient exploration of the design space, allowing for the identification of mix compositions that not only satisfy structural performance criteria but also minimize both material cost and environmental impact [8, 9, 10].

2 Materials and Methods

2.1 Dataset Description and Statistical Analysis

The dataset employed in this study originates from the work of Katlav and Ergen [11], who improved the prediction of compressive strength for ultra-high-performance concrete (UHPC) using the CatBoost model optimized through various algorithms. It comprises 785 samples and 15 features, including 14 input variables and one output variable. The input features encompass both material constituents and environmental parameters influencing the performance of 3D-printed concrete.

The mix constituents include cement (kg/m^3), silica fume (SF) (kg/m^3), blast furnace slag (BFS) (kg/m^3), fly ash (FA) (kg/m^3), quarry powder (QP) (kg/m^3), limestone powder (LSP) (kg/m^3), nano-silica (NS) (kg/m^3), fiber (kg/m^3), sand (kg/m^3), gravel (kg/m^3), water (kg/m^3), and superplasticizer (SP) (kg/m^3). Environmental parameters consist of temperature (T) (°C), relative humidity (RH) (%), and curing age (days). The output feature is the compressive strength (f_c) measured in megapascals (MPa), serving as the target variable for prediction.

Figure 1 presents a correlation heatmap that visualizes the pairwise relationships among variables, especially in relation to compressive strength. The color intensity ranges from light orange (negative correlation) to brown (positive correlation), indicating the strength and direction of correlation.

The distributional analysis reveals distinct patterns across features. Cement displays a unimodal distribution centered at 800 kg/m^3 with slight negative skew.

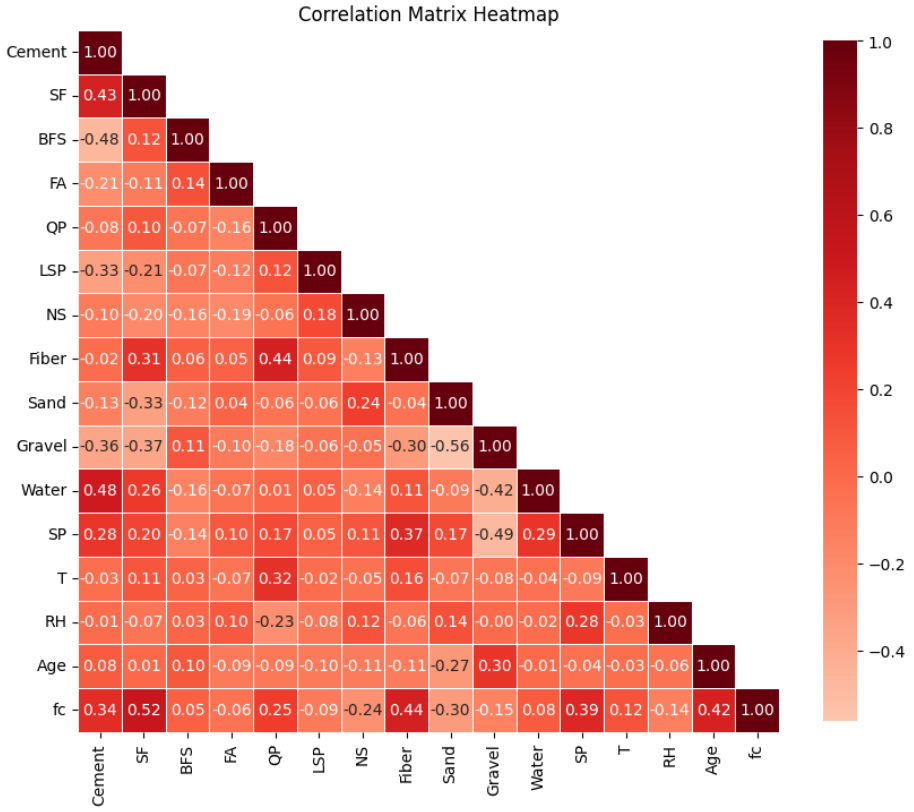

Fig. 1. Correlation Matrix Heatmap of Variables.

SF exhibits a bimodal distribution peaking at 0 and 200 kg/m^3, while BFS, FA, QP, and LSP are strongly right-skewed with peaks at 0 kg/m^3, reflecting infrequent usage. NS, fiber, and gravel are minimally used, with fiber showing a bimodal trend and gravel a right-skewed distribution. Sand has a complex multimodal pattern, and water follows a slightly right-skewed unimodal distribution centered around 175–180 kg/m^3. SP demonstrates multimodal behavior with distinct dosage peaks.

Environmental variables show similar skewness: temperature is right-skewed with a concentration between 20–30 °C, relative humidity is clustered near 100%, and age is largely below 28 days. The compressive strength exhibits a unimodal distribution peaking at approximately 125 MPa with mild right skewness.

2.2 Data Collection for Cost and Embodied Carbon Calculation

To optimize 3D concrete printing mix designs using NSGA-II, data on material costs and embodied carbon were collected. Material costs were obtained from

Table 1. Material costs and embodied carbon values

(a) Material Costs (INR/kg)

Material	Cost
Cement	6.5
Silica fume	45.7
Blast furnace slag	2.3
Fly ash	2.0
Quartz powder	3.0
Limestone powder	2.5
Nano silica	250
Fiber	135
Sand	1.6
Gravel	0.8
Superplasticizer	240

(b) Embodied Carbon (kg CO_2/kg)

Material	CO_2
Cement	0.84
Silica fume (SF)	0.03
Blast furnace slag (BFS)	0.08
Fly ash (FA)	0.004
Quartz powder (QP)	0.023
Limestone powder (LSP)	0.02
Nano silica (NS)	5.00
Fiber	2.50
Sand (fine aggregate)	0.007
Gravel (coarse aggregate)	0.016
Water	0.0002
Superplasticizer (SP)	0.94

online sources for 2024–2025, based on bulk procurement prices in the Indian market. Embodied carbon values for selected materials were sourced from the ICE Advanced Database and other references. The combined summary of both datasets is shown in Table 1 and forms the foundation of the multi-objective optimization framework, aiding in the balance between cost-efficiency and sustainability.

2.3 Implementation

This subsection explains the implementation of the framework, as depicted in Fig. 2.

Mix Design Generation Using CVAE and Virtual Test of Mix Designs To address the challenge of efficiently sampling valid and high-performing concrete mix designs, this study employs a Conditional Variational Autoencoder (CVAE) to approximate the high-dimensional design space with a low-dimensional latent space. The CVAE comprises an encoder that maps input data—conditioned on a target compressive strength—to a latent distribution, and a decoder that reconstructs high-dimensional input from this representation. Once trained, the decoder enables the generation of new mix designs by sampling latent variables.

After generating new mix compositions, their performance is virtually evaluated. CO_2 emissions and material cost are calculated directly from the ingredient quantities, while compressive strength is predicted using the FOA-CatBoost regression model. This integration of generative modeling and predictive evaluation allows rapid exploration of viable mixture candidates prior to physical testing.

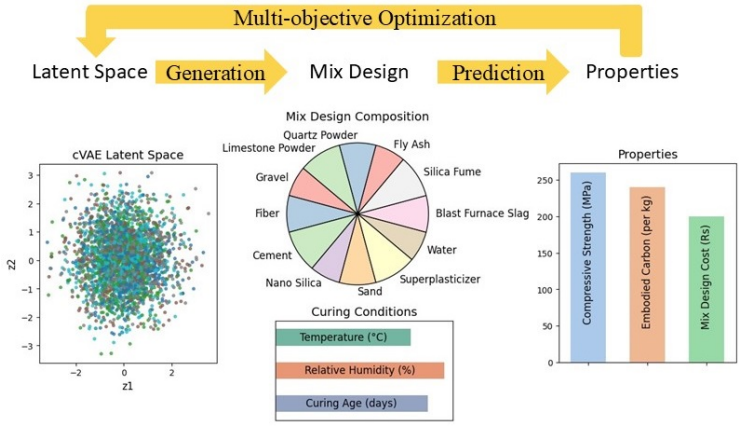

Fig. 2. The Suggested Framework: New mix designs can be generated from the latent space using a CVAE decoder. Each generated design is virtually evaluated for its performance, including CO_2 emissions, cost, and strength, through regression models. A multi-objective optimization algorithm then explores the latent space to enhance their performance.

The implementation of the CVAE includes data preprocessing, model architecture design, and training. The dataset is normalized using Min-Max scaling and partitioned into training and testing sets. The CVAE architecture adopts an encoder-decoder structure where both components are conditioned on the target compressive strength. Several layer configurations were experimented with, as shown in Table 2, to assess model performance.

The CVAE loss comprises four terms: reconstruction $\mathcal{L}_{\text{recon}}$, KL divergence $\mathcal{L}_{\text{KL}}$, material cost $\mathcal{L}_{\text{cost}}$, and embedded carbon $\mathcal{L}_{\text{carbon}}$. The objective is:

$$\mathcal{L}_{\text{CVAE}} = \mathcal{L}_{\text{recon}} + \beta\mathcal{L}_{\text{KL}} + \lambda_{\text{cost}}\mathcal{L}_{\text{cost}} + \lambda_{\text{carbon}}\mathcal{L}_{\text{carbon}}, \tag{1}$$

where $\mathcal{L}_{\text{recon}} = \frac{1}{N}\sum \|x_i - \hat{x}_i\|^2$ ensures input fidelity, and $\mathcal{L}_{\text{KL}}$ regularizes the latent space. Cost and carbon terms penalize deviations in material quantities, weighted by unit cost m_j and emission factor e_j, respectively. Hyperparameters β, λ_{cost}, and λ_{carbon} control the trade-offs weighted at 0.001, 0.1, and 0.001, respectively. This multi-objective formulation encourages the model to balance structural accuracy with sustainability considerations. Training is conducted using the Adam optimizer over 5000 epochs with a batch size of 32, ensuring robust convergence and generalization.

Multi-objective Optimization Using NSGA-II To identify concrete mixtures that balance strength, cost, and embodied carbon, a multi-objective optimization strategy is employed using the Non-dominated Sorting Genetic Algorithm II (NSGA-II). The latent space defined by the best-performing CVAE architecture (encoder-decoder layers of size (128,64) and (64,128)) serves as the

Table 2. CVAE Encoder and Decoder Design

Number of Hidden Layers	Encoder	Decoder
1	(16)	(16)
1	(32)	(32)
1	(64)	(64)
1	(128)	(128)
1	(256)	(256)
2	(32, 16)	(16, 32)
2	(128, 64)	(64, 128)
2	(256, 128)	(128, 256)
3	(32, 16, 8)	(8, 16, 32)
3	(256, 128, 64)	(64, 128, 256)
5	(32, 16, 8, 4, 3)	(3, 4, 8, 16, 32)

search domain. New mix designs are generated by sampling latent vectors and decoding them via the CVAE. These decoded mixtures are virtually evaluated using the FOA-CatBoost regression model, which estimates compressive strength, while cost and CO_2 emissions are computed from constituent quantities. Additional constraints may be applied on the lower and upper bounds of the quantity of individual components to allow for fluctuation in availability of materials. Similarly, lower and upper bounds on the curing conditions may be applied depending on the available conditions.

The optimization objective is to minimize both cost and embodied carbon, subject to the constraint $|\hat{f}_c - f_c| \leq 1$ MPa, ensuring structural performance. Bounds can also be enforced on individual material quantities and curing conditions to reflect practical limitations.

NSGA-II begins with an initial population of 1000 latent vectors and evolves solutions through 300 generations using non-dominated sorting, crowding distance, tournament selection, crossover, and mutation. The final set of Pareto-optimal designs is selected based on hypervolume contribution and retained for further analysis.

3 Results

3.1 Comparison of CVAE Architectures Based on Loss Metrics

The performance of various Conditional Variational Autoencoder (CVAE) architectures was evaluated using training and test loss metrics, as shown in Table 3. Increasing the number of hidden layers and neurons generally improved performance, with the *Encoder: (256,128) Decoder: (128,256)* configuration achieving the lowest test loss (0.044430). Single-layer architectures with fewer than 64 neurons showed limited representational power, while multi-layer models, especially those with hierarchical encoding, demonstrated better generalization. The

Table 3. CVAE Train and Test Loss for Different Architectures

Model No.	Architecture	Train Loss	Test Loss
1	Encoder: (16) Decoder: (16)	0.096568	0.142106
2	Encoder: (32) Decoder: (32)	0.051849	0.082918
3	Encoder: (64) Decoder: (64)	0.041571	0.065377
4	Encoder: (128) Decoder: (128)	0.039315	0.064780
5	Encoder: (256) Decoder: (256)	0.027361	0.046817
6	Encoder: (32,16) Decoder: (16,32)	0.034560	0.074149
7	Encoder: (128,64) Decoder: (64,128)	0.036857	0.085323
8	Encoder: (256,128) Decoder: (128,256)	0.010312	0.044430
9	Encoder: (32,16,8) Decoder: (8,16,32)	0.035585	0.062233
10	Encoder: (256,128,64) Decoder: (64,128,256)	0.011154	0.049881
11	Encoder: (32,16,8,4,3) Decoder: (3,4,8,16,32)	0.509857	0.513288

Encoder: (256,128,64) Decoder: (64,128,256) model (test loss: 0.049881) and the *Encoder: (128,64) Decoder: (64,128)* model (test loss: 0.085323) were also strong performers. In contrast, deeper architectures with excessive layers, like the *Encoder: (32,16,8,4,3) Decoder: (3,4,8,16,32)* configuration, suffered from higher loss, suggesting that excessive complexity may hinder optimization. The optimal CVAE architecture balances depth and width, with two or three hidden layers yielding the best results.

3.2 Analysis of Learning Curves Across Autoencoder Architectures

The learning curves of models 1 through 11 (Fig. 3) show a rapid decrease in both training (blue) and test (red dashed) losses during the first 1000 epochs, indicating quick pattern recognition. As training progresses, the losses converge to low values, with minimal overfitting, except in models 1 and 11, which underfit more. Models 1 to 5 show smooth trends with stable training, with model 4 (encoder/decoder dimensions of 128) outperforming the others, and model 5 (256 dimensions) offering improved generalization despite slightly worse performance. Models 6 to 11 exhibit instability, particularly with spikes in their curves, suggesting that deeper models (more layers) face training challenges. Model 7, with encoder dimensions (128, 64) and decoder dimensions (64, 128), achieves the best performance, with minimal gap between training and test losses and the least instability. Model 8, with encoder (256, 128) and decoder (128, 256), performs better than single-layer models but shows higher instability. Models 9 to 11, especially the five-layer architecture, experience the highest instability and worse performance, indicating that deeper models struggle with data pattern capture.

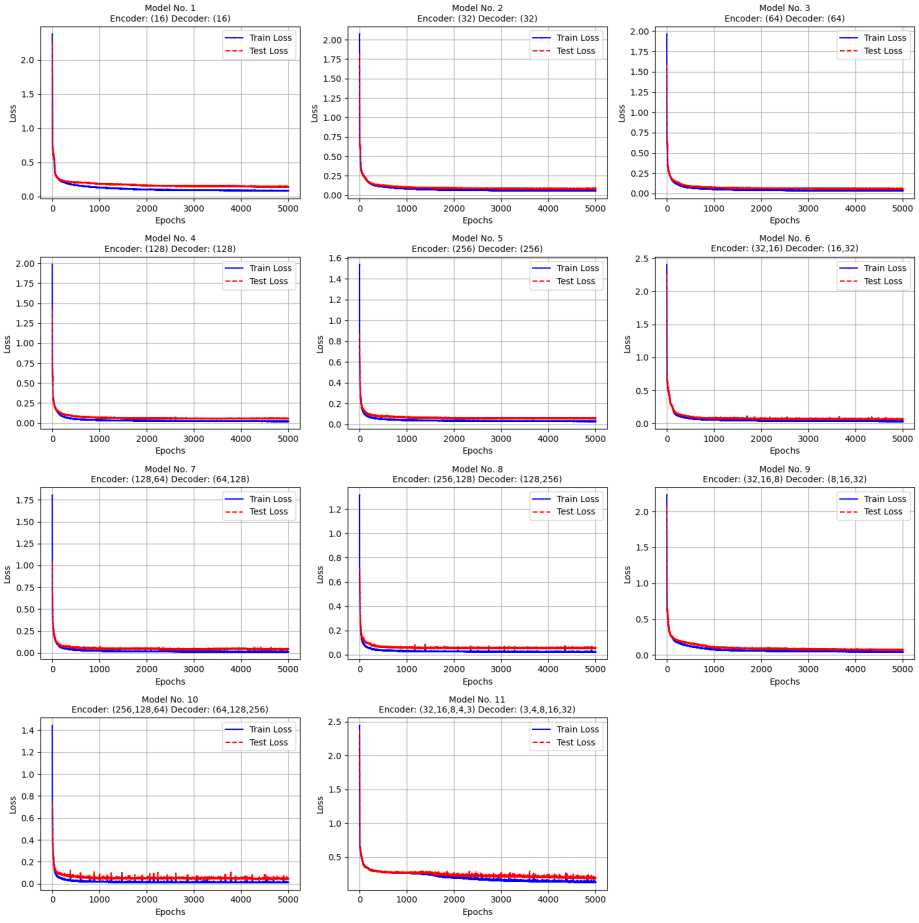

Fig. 3. Learning curve plots

3.3 Mix Design Generation and Optimization

The CVAE, NSGA-II, and FOA-CatBoost were integrated to generate and optimize concrete mix designs for six target compressive strengths. The optimization process selects designs within ±5% of the target strength based on the highest hypervolume from the Pareto frontiers. Table 4 shows the optimized mix designs, including component weights and curing conditions, for each target strength. The best-performing mix is selected from the Pareto front for each strength.

Table 5 highlights the effectiveness of the optimized mix designs in reducing both material cost and embedded carbon across all target compressive strength.

Fig 4 shows the pareto frontiers of the sampled designs and 5 compares the optimization results with the sampled designs. The Pareto frontiers for cost and embodied carbon were identified, with solutions selected based on the highest

Table 4. Optimized Mix Proportions and Environmental Conditions for Target Compressive Strengths

Component	**60 MPa**	**80 MPa**	**100 MPa**	**120 MPa**	**140 MPa**	**160 MPa**
Cement (kg/m³)	349.36	360.90	325.15	399.30	320.05	336.19
SF (kg/m³)	10.45	14.88	46.37	0.04	71.13	68.64
BFS (kg/m³)	66.83	51.73	23.32	0.00	53.00	162.85
FA (kg/m³)	0.00	0.00	0.00	0.01	0.00	0.00
QP (kg/m³)	0.00	0.00	0.00	0.00	0.00	0.00
LSP (kg/m³)	42.98	0.42	91.64	0.00	148.30	217.51
NS (kg/m³)	0.00	0.00	0.00	1.04	0.00	0.00
Fiber (kg/m³)	0.00	0.00	0.00	0.00	0.00	0.00
Sand (kg/m³)	387.04	739.88	676.40	820.85	768.93	695.10
Gravel (kg/m³)	1113.13	1021.74	1018.94	1064.89	894.18	906.14
Water (kg/m³)	156.75	161.95	137.66	104.47	139.00	145.10
SP (kg/m³)	7.16	5.47	9.51	21.44	14.27	14.67
T (°C)	20.07	20.50	20.15	20.11	20.07	20.02
RH (%)	99.53	100.00	96.50	100.00	93.68	94.44
Age (days)	5.00	53.65	36.62	73.19	72.97	130.97
Strength (MPa)	61.97	83.05	101.95	116.22	146.19	152.04
Material Cost (Rs.)	6236.92	6459.94	8694.71	10168.34	11193.99	11598.67
Carbon (kg CO_2/kg)	327.26	334.45	308.22	383.56	311.31	335.03

Table 5. Comparison of Optimized and Original Dataset Mixes for Cost and Embedded Carbon

Target Strength (MPa)	**60**	**80**	**100**	**120**	**140**	**160**
Optimized Cost (Rs.)	6236.92	6459.94	8694.71	10168.34	11193.99	11598.67
Mean Cost (Rs.)	13221.14	24942.65	25241.71	28478.27	32762.81	35642.63
Optimized Carbon	327.26	334.45	308.22	383.56	311.31	335.03
Mean Carbon	556.33	745.38	771.49	831.89	903.52	962.99

hypervolume. The optimized mixes significantly reduced both cost (by 50% to 80%) and embodied carbon (more than 50%), demonstrating the effectiveness of the optimization framework in identifying environmentally and economically efficient designs while meeting mechanical performance requirements.

4 Conclusion and Future Work

This study presents a comprehensive framework for sustainable concrete mix design that integrates predictive modeling, generative algorithms, and multi-objective optimization. By combining the CatBoost regression model, optimized through the Firefly Optimization Algorithm (FOA), with a Conditional Varia-

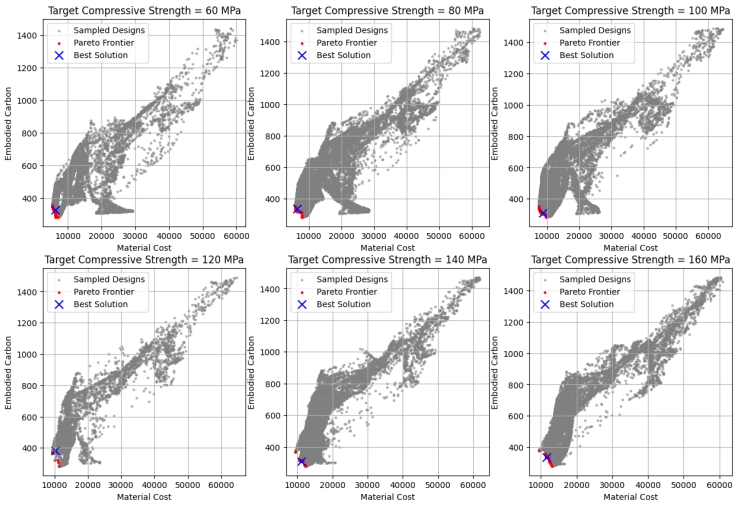

Fig. 4. Pareto Diagram

tional Autoencoder (CVAE) and the NSGA-II optimization algorithm, the proposed approach effectively addresses the limitations of traditional trial-and-error methods in mix formulation.

The use of CVAE enables the generation of realistic mix designs that adhere to underlying constraints and material interactions learned from data, thus eliminating the need for manual constraint definitions. The integration of FOA-CatBoost improves predictive accuracy for compressive strength, while NSGA-II identifies Pareto-optimal solutions balancing strength, cost, and embodied carbon.

Evaluations on the dataset demonstrate that the proposed framework significantly reduces both material costs and CO_2 emissions across various tar-

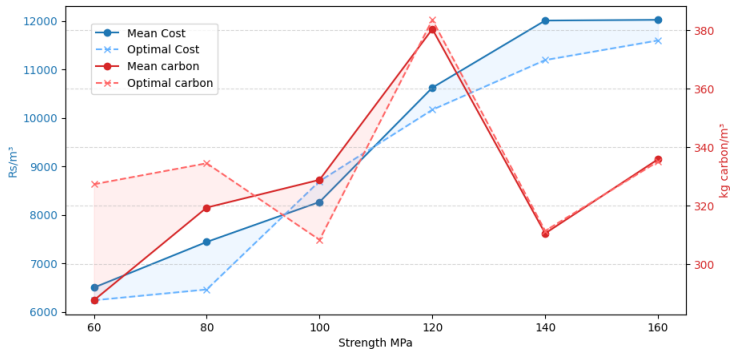

Fig. 5. Cost Comparison Diagram

get strength levels. The architecture comparison shows that two- to three-layer CVAEs offer the best generalization and stability, and the optimization results exhibit superior performance compared to baseline designs in terms of both economic and environmental metrics.

Overall, this generative AI-guided methodology establishes a robust foundation for intelligent material design in concrete engineering. Future work will focus on experimental validation through physical testing of optimized mixtures, and extending the framework to broader applications in construction materials and sustainable design practices.

References

[1] Abebe Demissew. "Comparative Analysis of Selected Concrete Mix Design Methods Based on Cost-Effectiveness". In: *Advances in Civil Engineering* 2022.1 (2022), p. 4240774.

[2] Ali Ashrafian et al. "An evolutionary neuro-fuzzy-based approach to estimate the compressive strength of eco-friendly concrete containing recycled construction wastes". In: *Buildings* 12.8 (2022), p. 1280.

[3] Mohamed K. Elshaarawy, Mostafa M. Alsaadawi, and Abdelrahman K. Hamed. "Machine Learning and Interactive GUI for Concrete Compressive Strength Prediction". In: *Scientific Reports* 14 (2024), p. 16694. DOI: 10.1038/s41598-024-66957-3.

[4] Dayou Luo et al. "Artificial Intelligence in the Design, Optimization, and Performance Prediction of Concrete Materials: A Comprehensive Review". In: *npj Materials Sustainability* 3 (2025), p. 14. DOI: 10.1038/s44296-025-00058-8.

[5] Jianhao Gao, Chaofeng Wang, and SH Chu. "Mix design of sustainable concrete using generative models". In: *Journal of Building Engineering* 96 (2024), p. 110618.

[6] Feixiang Chen et al. "Advancing Concrete Mix Proportion through Hybrid Intelligence: A Multi-Objective Optimization Approach". In: *Materials* 16.19 (2023), p. 6448. DOI: 10.3390/ma16196448.

[7] Omid Bamshad et al. "Prediction and Multi-Objective Optimization of Workability and Compressive Strength of Recycled Self-Consolidating Mortar Using Taguchi Design Method". In: *Heliyon* 9.6 (2023), e16381. DOI: 10.1016/j.heliyon.2023.e16381.

[8] Sai Yang et al. "Intelligent Multiobjective Optimization for High-Performance Concrete Mix Proportion Design: A Hybrid Machine Learning Approach". In: *Engineering Applications of Artificial Intelligence* 126 (2023), p. 106868. DOI: 10.1016/j.engappai.2023.106868.

[9] Fan Zhang et al. "Optimized Design of Low-Carbon Mix Ratio for Non-Dominated Sorting Genetic Algorithm II Concrete Based on Genetic Algorithm-Improved Back Propagation". In: *Materials* 17.16 (2024), p. 4077. DOI: 10.3390/ma17164077.

[10] Rupesh Kumar Tipu et al. "Optimizing Sustainable Blended Concrete Mixes Using Deep Learning and Multi-Objective Optimization". In: *Scientific Reports* 15 (2025), p. 16356. DOI: 10.1038/s41598-025-00943-1.

[11] Metin Katlav and Faruk Ergen. "Improved forecasting of the compressive strength of ultra-high-performance concrete (UHPC) via the CatBoost model optimized with different algorithms". In: *Structural Concrete* (2024).