https://doi.org/10.48047/AFJBS.6.15.2024.10055-10064

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

EVALUATION OF COMPRESSIVE STRENGTH OF SLAG CONCRETE WITH ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TECHNIQUES

Sunil SK[1], Dr. Aala Satyanarayana[2], Dr Pandu Kurre[3], Praveen Samarthi[4], Aryasree Madhukumar[5], R Gowrishankar[6]

[1]Assistant Professor, Department of Civil Engineering, Acharya Institute of Technology, bagalgunte, Bangalore, Karnataka, India.

[2]Assistant Professor, Department of Civil Engineering, Maturi Venkata Subba Rao (MVSR) Engineering College (A), Nadergul, Telangana – 501510, India.

[3]Lecturer, Department of Civil Engineering, Maulana Azad National Urdu University Polytechnic, Kadapa, Andhra Pradesh, India.

[4]Sr. Assistant Professor, Department of Engineering, CVR College of Engineering, Hyderabad, Telangana, India.

[5]Assistant Professor, Department of Civil Engineering, Cambridge Institute of Technology, KR Puram, Bangalore, Karnataka, India.

[6]Assistant Professor. V.S.B. Engineering College, Karudayampalayam, Karur-639111, Tamil Nadu, India.

Volume 6, Issue 15, Sep 2024

Received: 15 July 2024

Accepted: 25 Aug 2024

Published: 05 Sep 2024

doi:10.48047/AFJBS.6.15.2024.10055-10064

Abstract: This study explores the use of Steel Slag Aggregate (SSA) as a sustainable alternative in concrete, addressing environmental concerns like resource depletion and high CO2 emissions. Four modeling techniques—Gene Expression Programming (GEP), Artificial Neural Network (ANN), Random Forest Regression (RFR), and Gradient Boosting (GB)—were used to predict the compressive strength (CS) of SSA concrete using 367 datasets. Among the models, Gradient Boosting (GB) showed the best performance, with the highest R2 values and lowest error metrics, outperforming RFR, GEP, and ANN. The findings highlight GB's effectiveness in predictive modeling for sustainable construction.

Keywords: gene expression programming; artificial neural network; random forest regression; gradient boosting; soft computing; artificial intelligence; steel slag aggregate; sustainable construction; compressive strength of concrete

Introduction

The construction industry is a significant consumer of natural resources, with concrete being a key material composed of cement, water, and aggregates. Aggregates, which make up 80% of concrete's weight, provide essential structural qualities but contribute to environmental issues such as resource depletion, noise pollution, habitat loss, and increased CO2 emissions due to extensive mining, processing, and transportation. In 2018, Europe's aggregate industry was the largest non-energy mining sector, producing 3 billion tons across 39 nations. Globally, 60% of raw materials in construction come from the lithosphere, accounting for substantial energy and water use [1-2].

To reduce the industry's carbon footprint, more sustainable solutions are needed. Qatar, committed to sustainability through its Qatar National Vision 2030 (QNV 2030) and Qatar National Development Strategy (QNDS), faces local shortages of natural resources for concrete production. In response, Qatar has increased the use of green concrete, incorporating recycled materials like steel slag aggregate (SSA), wadi gravel, excavation waste, and construction and demolition wastes (CDW) [3].

Qatar's second development strategy recognized that 80% of generated solid waste, amounting to 80–100 million tons, ends up in landfills. In 2022, the country set targets to recycle 15% of all solid waste and use 20% of construction waste in building projects. The adoption of green concrete supports these goals by reducing construction waste, greenhouse gas emissions, and the reliance on imported materials, aligning with the QNV 2030 objectives.

Steel slag, a by-product of the steel industry, can be effectively used as a replacement for traditional aggregates in concrete. The iron and steel sector contributes significantly to global greenhouse emissions, with a notable portion used in construction. Using steel slag as an aggregate substitute offers financial benefits by reducing reliance on expensive imported materials like gabbro, especially in Qatar, where local natural aggregates are limited. Additionally, recycling steel slag helps mitigate environmental issues such as landfill overuse, leading to more eco-friendly concrete that lowers waste, costs, and CO2 emissions [4].

Artificial intelligence (AI), aligned with Qatar’s National Vision 2030 (QNV 2030), is revolutionizing various sectors, including construction. AI, particularly machine learning (ML), has transformed the prediction of structural properties by eliminating inefficiencies in traditional methods. Machine learning algorithms, which learn from data rather than explicit programming, can analyze and predict complex datasets, offering valuable insights into concrete performance. This can guide material selection, optimize mix designs, and enhance structural performance, promoting a more sustainable and resilient construction sector in Qatar. This research aims to advance knowledge in sustainable construction by studying the effects of steel slag aggregate (SSA) on concrete compressive properties and exploring innovative predictive modeling methods. The findings are intended to assist engineers, researchers, and policymakers in promoting sustainable construction practices. By integrating SSA, the research seeks to reduce environmental impact and address resource scarcity, aligning with Qatar's sustainability goals as outlined in the QNV 2030 and national development strategies, which emphasize recycling and minimizing construction waste [5-6].

2. Literature Review

Steel slag, a byproduct of steel production that also contains large stones and dust, is a significant industrial waste. As worldwide unrefined steel creation keeps on rising, roughly 150 kg of steel slag is produced per ton of steel, frequently winding up in open regions and presenting ecological risks. Regardless of these difficulties, steel slag has acquired consideration for its expected in substantial applications because of its extraordinary properties. A few examinations have explored the mechanical characteristics of cement containing steel slag totals (SSA) contrasted with regular totals.

For instance, Qasrawi reported that steel slag with a high Fe2O3 content improves the compressive and structural strength of concrete, surpassing the strength development of conventional concrete over time. Alizadeh et al. evaluated hardened concrete using SSA and found that it had a higher modulus of elasticity, flexural strength, and compressive strength than natural aggregate concrete. This finding is in line with that study. As to appraisals, Awwad et al. examined the replacement of SSA for sand in substantial blends in with target qualities of 25 MPa. Their outcomes showed worked on substantial strength without compromising functionality, especially eminent at a 30% substitution proportion.

Borole et al. utilized M30 grade cement to assess the impacts of to some degree subbing steel slag for regular total, finding that a 25% substitution rate ideally upgrades compressive, flexural, and rigid qualities without negative impacts. Sinha also looked into the effects of using steel slag to replace fine and coarse aggregates in conventional concrete mixes. He found that after 28 days, the concrete had improved flexural and tensile strength as well as increased compressive strength. Further upgrading substantial properties, Pushpakumara and Silva assessed the adequacy of steel slag in supplanting fine and coarse totals, confirming that substantial containing 75% steel slag shows expanded unit weight, parting elasticity, compressive strength, and consumption opposition.

Tarawneh et al. looked at environmental factors and compared SSA's physical and mechanical properties to those of conventional crushed limestone aggregate concrete. They found that steel slag had a faster rate of early strength development and was more resistant to abrasion. Nguyen et al. zeroed in on the compressive properties of steel slag concrete by supplanting it with coarse total, noticing fast strength increments inside the initial 7 days.

Aparicio et al. concentrated on the impacts of natural circumstances on concrete containing reused total or SSA, affirming prevalent compressive strength for SSA concrete at 28 days. Lately, respectful designing has experienced issues requiring instinct and gaining from previous encounters. SCT gather measurable, risky, and enhancement instruments to gain from previous encounters and utilize these discoveries to create new information, distinguish designs, or anticipate novel patterns. Different AI and delicate processing strategies, like fake brain organizations, fluffy rationale, and hereditary calculations, can take care of these issues. A few examinations have utilized ML and SCT to foresee the underlying properties of cement containing SSA.

For instance, Kumar et al.'s prediction models for fly ash concrete based on ELM, MARS, and DNN demonstrated their effectiveness in predicting compressive strength. Additionally, Kumar et al. utilized ANNs to anticipate past cement's compressive strength and penetrability with GGBS. In general, these studies demonstrated that a variety of soft computing and machine learning methods can accurately predict concrete's compressive strength. The number of experiments required to determine concrete's structural factors could be reduced by using these models and methods, saving money. The accompanying segment will give an understanding into the overall technique utilized in this examination and the standards of the various models.

3. Methodology Overview

This research employed Machine Learning (ML) and Soft Computing Techniques (SCT) to predict the compressive strength (CS) of concrete containing steel slag aggregate (SSA). The study began by compiling and pre-processing 367 datasets from the literature, standardizing parameters like cement content, SSA, water, coarse and fine aggregates, age, and superplasticizer. Data was normalized using Min-Max scaling, and an 80–20 split was applied to create training and testing sets.

3.1. Data Collection and Statistical Analysis

The research began by gathering data on steel slag concrete strengths from sources like Google Scholar and Mendeley. After preprocessing, the dataset was narrowed to 334 samples, excluding those with fly ash (FA) content and 0% SSA substitution. The final dataset includes 367 samples, covering a wide range of concrete mix compositions and experimental conditions, with SSA content varying from 0% to 100%. It also accounts for curing periods from 1 to 365 days, focusing on key intervals like 7, 28, and 90 days. Key variables such as cement content, water-to-cement ratio, and superplasticizer were included to capture complex interactions affecting compressive strength [7-8]

3.2. Data Grouping

This study evaluated the effectiveness of four machine learning and soft computing techniques: Gene Expression Programming (GEP), Artificial Neural Networks (ANN), Random Forest Regression (RFR), and Gradient Boosting (GB). Two datasets were created, with 80% (267 observations) used for model training and 20% (67 observations) for accuracy testing. The study assessed each input variable's importance using established methods, including the stepwise significance method and feature importance scores generated by ANN and tree-based models like RFR and GB. Additionally, correlation analysis was conducted to measure the strength and direction of relationships between input and output variables. These approaches ensured a thorough and transparent analysis of each variable's contribution to the predictive models [9-10]

3.3. Developing Models

This study utilized four machine learning and soft computing techniques—Gene Expression Programming (GEP), Artificial Neural Networks (ANNs), Random Forest Regression (RFR), and Gradient Boosting (GB)—to predict the compressive strength of steel slag concrete. These models were chosen for their ability to handle complex relationships, non-linear data, and large datasets while reducing overfitting. Hyperparameters for each model were carefully optimized using grid and random search techniques, ensuring robustness and generalizability. Performance was evaluated using metrics like MAE, RMSE, and R^2, with the best configurations used to train the final models.

3.3.1. Artificial Neural Network (ANN)

The purpose of ANNs is to imitate the human brain's biological nervous system's function and ability to learn, particularly in information processing. ANNs emulate the cerebrum's usefulness in two essential ways: gaining information through a growing experience and putting away or remembering data by means of the qualities of interconnected neurons, known as synaptic loads. An ANN's structure is characterized by a parallel arrangement of highly interconnected neurons that are capable of complex training. Information are handled through a progression of interconnected layers, isolated into three segments: input, stowed away layers, and result, each including a few hubs (neurons). The info layer gets and processes information prior to passing it to the following hubs. The secret layers perform complex numerical activities to remove helpful elements, while the result layer delivers the last result or expectation. An ANN's capacity to adjust to changing information and result information, perform non-straight capability planning, and catch obscure connections makes it a flexible model for resolving genuine issues [11].

During the testing stage, explicit upgrades were made to improve the presentation and exactness of the ANN model. These upgrades zeroed in on enhancing the model design, changing hyperparameters, and carrying out regularization procedures to forestall overfitting. Experiments with various numbers of hidden layers and neurons per layer were initially used to refine the model architecture. An ideal arrangement was distinguished through iterative testing and assessment, adjusting intricacy and execution. It was found that rising the quantity of secret layers and neurons improved the model's capacity to catch complex examples in the information. However, due to the risk of overfitting, an excessively complex model was avoided. To improve the model's performance, grid and random search methods were used to systematically adjust hyperparameters. A scope of values for key hyperparameters, including the learning rate, bunch size, and the quantity of ages, were investigated. The model's performance on a validation set was looked at to find the best hyperparameter combination that produced the lowest validation error and highest predictive accuracy [12].

3.3.2. Gene Expression Programming (GEP)

Hereditary calculations (GA) are one of the fundamental sorts of AI and SCT; the primary guideline of this strategy or method depends on the Darwinian standard of regular choice to take care of perplexing issues. This technique has been utilized to tackle numerous issues, zeroing in principally on enhancement issues constrained by different factors. Through the Gene Expression Program (GEP), Ferreira proposed a more advanced type of genetic programming (GP). The GEP is a learning algorithm that creates a model to explain the relationships between various variables in datasets and focuses on understanding these relationships. A type of GA called the GEP employs chromosomes and the Tree's method to solve problems. The first or initial chromosome population is constructed from these chromosomes, which contain mathematical information or functions. The fitness of each chromosome is checked, and the ones with the highest fitness are chosen for reproduction. The hereditary activities performed incorporate hybrid, change, and generation. The GA keeps on developing until a palatable arrangement is found. A relatively straightforward estimation equation is produced by this approach, and it can be utilized for practical design and hand calculation.

3.3.3. Random Forest Regression (RFR)

The irregular backwoods relapse ML strategy is known for its incredible capacity to deal with huge arrangements of information with various characteristics and give an exact or precise assessment of property significance. RFR deals with the guideline of gathering picking up, consolidating the prescient force of various choice trees to improve exactness and dependability. Each tree is made freely on an irregular subset of the preparation information, which makes difference and decreases overfitting. Through bootstrap collection (stowing), RFR can fabricate a strong model via preparing on various dataset varieties. The last not set in stone by averaging the expectations from every one of the singular trees in the woods. This system guarantees that the aggregate choice of many trees is more precise and stable than any singular tree's. This cycle is rehashed consistently until the expected level of accuracy is achieved. The RFR's singular ability is to improve its predictive power as a whole [14-15].

3.4. Statistical Indicators and Measurements

Factual measures like the mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2) are used to assess a model's accuracy. R^2 indicates how well the model's predictions match the actual data, ranging from negative infinity to 1, with 1 being ideal. MAE measures the average difference between predicted and actual values, with lower values indicating higher accuracy. RMSE, similar to MAE but more sensitive to outliers, emphasizes larger errors. Together, these metrics provide a comprehensive view of a model's performance, with a high R^2 and low MAE and RMSE indicating accurate predictions with minimal error.

Tables 1-3 summarize the evaluation metrics (R^2, MAE, and RMSE) of the four models created.

Table 1. Training set statistical measurements.

Model	**MAE (MPa)**	**RMSE (MPa)**	**Mean**	***R*2**
GEP	6.2	6.4	1.10	0.89
ANN	10.2	11.00	1.23	0.66
RFR	3.11	4.11	1.006	0.96
GB	0.77	2.33	1.01	0.99

Table 2. Testing set statistical measurements.

Model	**MAE (MPa)**	**RMSE (MPa)**	**Mean**	***R*2**
GEP	6.65	8.02	1.111	0.89
ANN	7.70	9.29	1.19	0.83
RFR	5.17	6.34	1.09	0.994
GB	3.61	4.95	1.02	0.97

Table 3. All datasets statistical measurements.

Model	**MAE (MPa)**	**RMSE (MPa)**	**Mean**	***R*2**	**STDV**	**COV (%)**
GEP	5.88	7.36	1.08	0.86	0.49	46.77
ANN	8.89	11.18	1.21	0.67	0.32	31.96

RFR	2.55	3.71	1.06	0.96	0.12	11.73
GB	1.25	2.47	1.02	0.98	0.06	5.98

4. Results and Discussion

In a study comparing four predictive models—Random Forest Regression (RFR), Gene Expression Programming (GEP), Artificial Neural Network (ANN), and Gradient Boosting (GB)—for predicting the compressive strength of steel slag aggregate (SSA) concrete, various performance metrics were evaluated, including R^2, root mean squared error (RMSE), mean absolute error (MAE), and mean values. The GB model consistently outperformed the others, with the lowest MAE (0.79), RMSE (1.90), and highest R^2 (0.99) during training, indicating near-perfect fit and minimal error. In testing, GB maintained its lead with a low MAE (1.15), RMSE (2.45), and high R^2 (0.98), demonstrating stable and accurate predictions.

In contrast, the ANN model struggled, showing the highest MAE (9.51), RMSE (12.03), and the lowest R^2 (0.61) during training, and despite some improvement in testing, it remained the least accurate with the highest variability and error. Both GEP and RFR performed reasonably well, with RFR showing strong R^2 values and low errors, though not as robust as GB. GEP also offered decent predictive accuracy but was slightly less reliable than GB. Overall, the GB model proved to be the most reliable for predicting compressive strength in SSA concrete, while ANN lagged behind in accuracy and consistency. The GB model shows the most minor changeability and most noteworthy forecast dependability, as shown by its low STDV and COV. This settles on it a dependable decision for reasonable applications where consistency is pivotal. The R2 values across datasets highlight the unwavering quality of the GB and RFR models. These models can be relied upon to give precise expectations, with the GB model being especially important for its close amazing fit. These plots show how well and accurately each model predicts CS within a 30 percent error range and visually. The GB model exhibits the nearest arrangement with the best fit line in both datasets, showing its hearty presentation and exact expectation of CS across various situations. Predictions from this model consistently match the actual experimental values closely, indicating low bias and high reliability. On the other hand, the ANN model's predictions are more variable. Some predictions deviate more significantly from the ideal line, while others are more closely aligned. This fluctuation proposes difficulties in catching all subtleties and intricacies of the information utilizing the ongoing ANN design. In order to reduce these deviations and improve its predictive accuracy, additional optimization or feature selection may be required. The predictions are tightly clustered in the RFR model, which has the smallest scatter around the mean. This shows high accuracy and consistency in anticipating CS values, reflecting powerful execution and exact demonstrating of the fundamental information designs.

The GEP model, like RFR, shows a slight spread but focuses predictions around the mean, indicating generally accurate results with some variability. In contrast, the ANN model exhibits more scattered predictions due to its struggle to grasp complex data relationships, suggesting areas for improvement. RFR and GEP models have minimal residual errors, indicating strong predictive accuracy, while ANN has larger residuals, pointing to less precise predictions. The GB model is highly robust, while ANN may need further tuning. Although GEP has a lower R^2 compared to RFR and GB, its interpretability offers a clear advantage, making it valuable for scenarios requiring transparency. The FA variable was excluded due to data quality and model constraints.

5. Sensitivity Study

A responsiveness study or examination is significant to numerous logical examinations. This boundary responsiveness investigation assists with understanding what a specific boundary could mean for the outcomes or the result of the model expectation. This study gives a comprehension of which input boundaries influence the outcomes most and which make less impacts. The GEP model was picked for the responsiveness concentrate on in light of its straightforwardness, which can be used to break down the elements impacting compressive strength (CS).

5.1. The Effects of Changing Steel Slag Aggregate Content (SSA)

This responsiveness concentrates on analyzed the connection between the level of SSA in substantial blends and the subsequent compressive strength. As referenced in the writing survey and displayed in Figure 14a, expanding the SSA content will build the compressive strength. Additionally, Tarawneh et al. uncovered that adding SSA can further develop substantial's scraped area factor, influence esteem, and compressive strength, especially during the beginning phases. When steel slag was used, Miah et al. observed a significant decrease in porosity and an increase in compressive strength. Sinha additionally affirmed the pattern by noticing an expansion in compressive, flexural, and split elastic qualities subsequent to supplanting fine total with a specific percentage of steel slag.

5.2. The Effects of Aging on the Compressive Strength of SSA Concrete

This responsiveness examination zeroed in on researching the impacts of maturing on SSA cement's compressive strength. The review's outcomes showed that the CS expanded after some time, mirroring the progressive improvement of cement's mechanical properties (as displayed in Figure 14b). This finding lines up with what Nguyen et al. [30] found. In the first place, compressive strength quickly expanded inside the 7-day relieving time of cement, trailed by a slower yet constant increment. In addition, Tarawneh et al. emphasized the beneficial effects of SSA on improving concrete properties, particularly impact strength and abrasion resistance. This suggests that the observed strength may be due to this enhancement. The review showed the ever-evolving compressive strength improvement in SSA concrete as it ages. Besides, Aparicio et al. [31] found that utilizing SSA can increment compressive strength values following 28 days of relieving, expanding with the substitution rate.

6. Conclusions

The construction industry faces challenges related to environmental sustainability and resource depletion. This research addresses these by promoting the use of eco-friendly alternatives like steel slag aggregate concrete (SSA). The study explored the predictive abilities of various machine learning and soft computing techniques, including ANN, GEP, RFR, and GB, to predict the compressive strength (CS) of SSA concrete. Using 334 datasets, the GB model showed the highest accuracy with an R^2 of 0.98 and the lowest errors, followed by RFR, GEP, and ANN.

Hyperparameter tuning was crucial for optimizing model performance, ensuring robustness and reliable predictions. These models offer practical applications, providing a reliable method for predicting the compressive strength of SSA concrete, aiding in mix design optimization, and reducing the need for extensive physical testing. This study contributes to our understanding of SSA concrete's mechanical behavior and promotes sustainable construction practices.

References

1. Martins, A.C.P.; De Carvalho, J.M.F.; Costa, L.C.B.; Andrade, H.D.; De Melo, T.V.; Ribeiro, J.C.L.; Pedroti, L.G.; Peixoto, R.A.F. Steel slags in cement-based composites: An ultimate review on characterization, applications and performance. *Constr. Build. Mater.* **2021**, *291*, 123265.

2. Oikonomou, N.D. Recycled concrete aggregates. *Cem. Concr. Compos.* **2005**, *27*, 315–318.

3. Planning and Statistics Authority. *Qatar Second National Development Strategy 2018–2022*; Planning and Statistics Authority: Doha, Qatar, 2019.

4. United Nations Environment Programme. *2022 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector*; United Nations Environment Programme: Nairobi, Kenya, 2022.

5. Piro, N.S.; Mohammed, A.S.; Hamad, S.M. Evaluate and Predict the Resist Electric Current and Compressive Strength of Concrete Modified with GGBS and Steelmaking Slag Using Mathematical Models. *J. Sustain. Metall.* **2023**, *9*, 194–215.

6. Wang, F.C.; Zhao, H.Y. Experimental investigation on blast furnace slag aggregate concrete filled double skin tubular (CFDST) stub columns under sustained loading. *Structures* **2020**, *27*, 352–360.

7. Kumar, D.S.; Priya, G.N.; Professor, A. Replacement of Coarse Aggregate using Steel Slag in Concrete. *Int. J. Eng. Res. Technol.* **2016**, *4*, 1–3.

8. Ibrahim, D. An Overview of Soft Computing. *Procedia Comput. Sci.* **2016**, *102*, 34–38.

9. Albostami, A.S.; Al-Hamd, R.K.S.; Alzabeebee, S.; Minto, A.; Keawsawasvong, S. Application of soft computing in predicting the compressive strength of self-compacted concrete containing recyclable aggregate. *Asian J. Civ. Eng.* **2023**, *25*, 183–196.

10. Penido, R.E.K.; da Paixão, R.C.F.; Costa, L.C.B.; Peixoto, R.A.F.; Cury, A.A.; Mendes, J.C. Predicting the compressive strength of steelmaking slag concrete with machine learning-Considerations on developing a mix design tool. *Constr. Build. Mater.* **2022**, *341*, 127896.

11. Kumar, M.; Biswas, R.; Kumar, D.R.; Samui, P.; Kaloop, M.R.; Eldessouki, M. Soft computing-based prediction models for compressive strength of concrete. *Case Stud. Constr. Mater.* **2023**, *19*, e02321.

12. Da Paixão, R.C.F.; Penido, R.E.-K.; Cury, A.A.; Mendes, J.C. Comparison of machine learning techniques to predict the compressive strength of concrete and considerations on model generalization. *Rev. IBRACON De Estrut. E Mater.* **2022**, *15*, e15503.

13. Wakjira, T.G.; Alam, M.S. Peak and ultimate stress-strain model of confined ultra-high-performance concrete (UHPC) using hybrid machine learning model with conditional tabular

generative adversarial network. *Appl. Soft Comput.* **2024**, *154*, 111353.

14. Farooq, F.; Amin, M.N.; Khan, K.; Sadiq, M.R.; Javed, M.F.; Aslam, F.; Alyousef, R. A comparative study of random forest and genetic engineering programming for the prediction of compressive strength of high strength concrete (HSC). *Appl. Sci.* **2020**, *10*, 7330.

15. Bentéjac, C.; Csörgo″, A.; Martínez-Muñoz, G. A Comparative Analysis of XGBoost. *Artif. Intell. Rev.* **2019**, *54*, 1937–1967.