

# An Assessment of Tree-Based Models for Predicting Indirect Tensile Stiffness Modulus in Fiber-Enhanced Cold Mix Asphalt

**Eman Abbas Abood**

Presidency of Al-Nahrain University, University of Al-Nahrain, Baghdad, Iraq  
dremancivil2019@gmail.com

**Mohammed Hazim Mohammed**

Iraqi Ministry of Higher Education and Scientific Research, Baghdad, Iraq  
mohammed.mohammed2001d@coeng.uobaghdad.edu.iq (corresponding author)

Received: 20 June 2025 | Revised: 17 July 2025 and 20 July 2025 | Accepted: 23 July 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.12840>

## ABSTRACT

This study focuses on forecasts of the Indirect Tensile Stiffness Modulus (ITSM) of fiber-reinforced Cold Mix Asphalt (CMA), using a CHAID (Chi-Squared Automatic Interaction Detector) model of tree-based machine learning methods. One of the most important advantages of the CHAID model is its simplicity when it comes to little training. This study employed 123 laboratory samples, taken on predetermined essential elements, fibre content and type, fibre length, curing time, and air voids. The CHAID model was evaluated through several performance indicators, including  $R^2$  (coefficient of determination) and RMSE (Root Mean Square Error). The proposed model had an  $R^2$  of 0.913, an Average Accuracy Percentage (AA %) of 16.16%, and an RMSE of 170.2 MPa. Feature importance showed both curing time and air voids as the main factors that affect ITSM. This study demonstrates the practicality of the CHAID model as a transparent and interpretable model to forecast stiffness, offering a wise choice in designing heavy-duty sustainable fiber-enhanced CMA mixtures for pavements.

*Keywords-gene expression programming; machine learning; fiber-reinforced cold mix asphalt; indirect tensile stiffness modulus; modeling*

## I. INTRODUCTION

CMA is used in patching applications and in remote locations due to its low energy costs and no heat requirements to resurface and maintain roadways. The ambient temperature manufacturing of CMA decreases the amount of energy consumed in the manufacturing process, reduces the amount of greenhouse gas emitted, and decreases the need for a burner facility to heat the HMA mix [1,2]. These CMA manufacturing properties benefit the environment and support policies such as the U.S. Environmental Protection Agency's Clean Air Act [3]. Economically, CMA also has many advantages, as it benefits consumers by its lower fuel requirements, leading to lower fuel costs to produce the material, allows CMA to be worked longer during the day, and uses materials that are commonly available for small works, allowing the use of local products with little or no reliance on a hot mix plant [4]. Despite these benefits, CMA has some drawbacks compared to HMA, including lower initial strength, the amount of time it takes to cure before achieving full mechanical performance, increased sensitivity to moisture, and increased susceptibility to permanent deformation (rutting) [1, 4].

A useful method to bypass these disadvantages is the addition of discrete fibers into the matrix system, including both natural (e.g., jute, coir, hemp, basalt) and synthetic (e.g., polyester, glass, polypropylene). Fibers enhance the composite matrix of CMA by improving cohesion, distributing stresses, bridging micro-cracks, and absorbing binder, thus mitigating drain-down and increasing tensile strength [2, 5]. Previous studies have shown that optimized fiber type, length, and content (typically 0.2–0.5% by aggregate weight) significantly improve key engineering properties, such as resistance to rutting and moisture damage, flexural strength, and critically, the Indirect Tensile Stiffness Modulus (ITSM) [1, 4, 6]. ITSM, a fundamental indicator of material stiffness and resistance to deformation under traffic loads, is vital for predicting pavement performance and designing durable structures [7]. A challenge in predicting the ITSM in fiber-reinforced CMA is the complex interaction of multiple variables, including fiber characteristics (type, length, and dosage, surface texture), mix composition (aggregate gradation, binder type and content, filler), curing conditions (time, temperature), and environmental factors [8]. Empirical or regression-based models are often inadequate in capturing the non-linear relationship and complex interaction embedded in these multi-factorial systems.

Machine Learning (ML) can be used to solve this multi-objective problem, which has proved to be powerful for complex prediction tasks in engineering, particularly tree-based ensemble models. These models, including Decision Trees (DT), Random Forests (RF), Gradient Boosting Trees (GBT), and eXtreme Gradient Boosting (XGBoost), are known for their proficiency in detecting non-linear patterns and interactions in high-dimensional datasets without making strict assumptions on data distribution [8-11]. A key advantage of tree-based models is their ability to provide inherent feature importance analysis, indicating the relative importance of input variables [8]. Tree-based models, such as RF, have shown better predictive performance than the traditional statistical models in the prediction of rutting depth in HMA [12], dynamic modulus [13], and fatigue life [14] within pavement materials. However, little is known about the application of such models for the prediction of ITSM of fiber-reinforced CMA [15, 16]. The existing limited predictive efforts for CMA properties have mostly focused on simpler models, unreinforced mixtures, or HMA [3, 11], highlighting the wide knowledge gap on the effectiveness of advanced ML techniques for this specific composite material.

This study aimed to provide a comprehensive performance evaluation of some of the prominent tree-based ensemble models to predict the ITSM of fiber-reinforced CMA. Using a robust experimental dataset that encompasses various types, contents, lengths, and curing times, this study aimed to: (i) assess the predictive accuracy of models; (ii) identify the most influential input features governing ITSM; and (iii) compare the performance of these ML models against traditional regression techniques. The results provide a reliable, data-driven tool for optimizing fiber-reinforced CMA mix designs, accelerating performance assessment, and ultimately promoting the broader utilization of this sustainable pavement material.

## II. PREDICTION MODELS AND TREES

In recent years, ML has become an effective tool to solve intricate civil engineering problems. Previous research attempts on predicting the ITSM of fiber-reinforced CMA exploited models such as Artificial Neural Networks (ANN), fuzzy logic, and Support Vector Machines (SVM) [17-19], which were accurate but not interpretable, thus devoid of direct applications [20]. However, the employment of evolutionary algorithms, such as Genetic Programming (GP), Genetic Algorithm (GA), and more specifically Gene Expression Programming (GEP), allowed obtaining explicit mathematical expressions to solve such problems [21, 22]. GEP, a merge between GP and GA, outputs readable formulas and eliminates redundant solutions [23], justifying its effectiveness in predicting several applicable engineering problems, including concrete behavior, soil properties, and tunnel machine performance [24-28]. This study aims to identify the superior approach in predicting the ITSM of fiber-reinforced CMA using GEP and tree-based methods on Fiber Type (FT1 and FT2), Fiber Content percentage (FC%), Fiber Length (FL), Ripping Time (RT), Curing Time (CT), and Air Void percentage (AV%) inputs.

## III. DATASET

The robustness of supervised ML models depends on the suitability of the dataset and the model's capabilities. Particularly for tree-based algorithms, their recursive partitioning can capture non-linear relationships and interactions among variables. However, this can also expose them to overfitting when the dataset lacks diversity and localness. For civil engineering, models such as DT, RF, and GBT can accurately predict the complex behaviors of materials such as soil, rock, asphalt, etc., on ITSM only if the dataset fully represents the materials and has a wide range of variation.

A dataset of 123 cold bituminous mixtures (CMA) was compiled to develop a reliable predictive model using SPSS. The mixtures were sourced from previous studies, where fibers were used as reinforcement to improve performance under varying curing and compaction conditions [1, 2, 4, 29-31]. Different fiber types and combinations were used, as each reinforcement method influences CMA behavior differently. The dataset aimed to explore the effects of different variable combinations on the ITSM of fiber-reinforced CMA.

FT1 and FT2 are categorical descriptors that specify the fibre type, with FT1 representing the main fiber used in the mix of the modified asphalt and FT2 indicating the second fiber. The values of each of the two variables ranged from 0 to 8, corresponding to fibre type indices, which are used to pinpoint a respective fibre type used in the matrices of experimental programs. As an example of the above, indices 0, 1, and 2 can represent various types of fibres such as cellulose, nylon, or glass, chosen under specific mechanical and chemical properties as part of the mixtures.

The combined effect of fiber types is calculated using a composite variable, the FT1/FT2 ratio. For each possible unique combination of FT1 and FT2, there is a unique value of the combined factors. This captures the interaction between the two fibers within the asphalt matrix. For example, if cellulose is FT1 and glass fiber is FT2, a different FT1/FT2 index is obtained compared to the case of nylon FT1 and cellulose FT2. This coding scheme takes into account all possible combinations and allows the model to evaluate the effect of these factors on ITSM performance.

The target output is the ITSM, measured in MPa. ITSM is a key parameter for CMA's deformation and fatigue resistance. Data preprocessing aimed to make the output and the variables have the same units, which is MPa. In addition, the variables in the training set need to be removed of outliers and redundancies that can lead to biased outcomes. Table I shows the input variables. All the independent variables were formed to support a structured algorithm for training, testing, and reanalysis when studying the effect of fibers on asphalt stiffness.

TABLE I. VARIABLES USED FOR FIBER-MODIFIED CMA

Variable type	Symbol	Description
Output	ITSM	Indirect Tensile Stiffness Modulus (MPa)
Input	FT1	Fiber Type 1 (0 = No fiber, 1 = Cellulose, 2 = Glass, 3 = Nylon, 4 = Polyester, 5 = Basalt, 6 = Hemp, 7 = Jute, 8 = Coir)
Input	FT2	Fiber Type 2 (0 = No fiber, 1 = Cellulose, 2 = Glass, 3 = Nylon, 4 = Polyester, 5 = Basalt, 6 = Hemp, 7 = Jute, 8 = Coir)
Input	FT1/FT2	Fiber blend ratio by mass (FT1:FT2)
Input	FC%	Fiber Content (% by weight of total aggregate)
Input	FL	Fiber Length (mm)
Input	CT	Curing Time (days)
Input	AV%	Air Void content (%)

#### IV. MODEL ARCHITECTURE

GEP is an advanced regression technique that combines genetic algorithms and expression trees to evolve mathematical models from data. Solutions are encoded as fixed-length linear chromosomes, which are later expressed as non-linear Expression Trees (ETs). Each chromosome consists of one or more genes, with every gene comprising a head, which includes functions, constants, and variables, and a tail, which includes only constants and variables [21]. This structure ensures the generation of syntactically valid expressions for accurate modeling of nonlinear systems. The initial GEP population consists of randomly generated chromosomes combining these input features using a predefined set of mathematical operators.

A chromosome consists of a single gene expressed as a linear string. This string is translated into an ET, which visually represents the hierarchical relationship among the input variables and operators. The ET is then read left to right, top to bottom, and translated into a corresponding mathematical formula. For chromosomes with multiple genes, the respective ETs are linked using operators such as addition, multiplication, or division to construct a composite predictive model.

The accuracy of each chromosome is evaluated using a fitness function based on statistical performance metrics. This study employed multiple indicators to rigorously assess model quality: Root Mean Square Error (RMSE), Normalized RMSE (NRMSE), Mean Absolute Percentage Error (MAPE), Average Accuracy Percentage (AA%), Coefficient of Determination ( $R^2$ ), and Pearson Correlation Coefficient (R), calculated as shown in (1-5). Chromosomes yielding high fitness values are selected for further evolution.

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (A_n - P_n)^2} \quad (1)$$

$$NRMSE = \frac{RMSE}{S} \quad (2)$$

$$R^2 = 1 - \frac{\sum (A_n - P_n)^2}{\sum (A_n - S_n)^2} \quad (3)$$

$$MAPE = \frac{(\sum \frac{|A-E|}{A}) * 100}{N} \quad (4)$$

$$AA \% = 100 \% - MAPE \quad (5)$$

where  $A_n$  are actual and  $P_n$  are predicted values,  $N$  is the number of points within the dataset, and  $S$  is used to normalize the RMSE against the mean of the actual values.

To enhance diversity and search efficiency across generations, several genetic operators are applied:

- Mutation introduces random alterations to gene components.
- Recombination enables the exchange of genetic material between chromosomes. Depending on the operation, this may be single-point, two-point, or entire gene recombination.
- Transportation involves moving gene segments within a chromosome. Insertion Sequence (IS) refers to movement within the gene head, Root Insertion Sequence (RIS) moves segments from the head to the tail, and Gene Transportation (GT) shifts entire genes to different locations.
- Inversion reverses a portion of the head, modifying the gene structure without syntactic errors.

As generations progress, these mechanisms guide the population toward optimal solutions. The finalized models show an advanced representation of the asphalt mixture parameters and their ITSM. The GEP models were trained on 80% of the available data, with the other 20% kept for validation. The most effective mix design parameters were chosen as the input variables to predict the ITSM of fiber-reinforced CMA.

#### V. ANALYSIS AND RESULTS

This study used tree-based predictive models for predicting the ITSM of fiber-reinforced CMA mixtures. Each model consisted of some looped expression structures designed to reflect patterns in the input data, inspired by GEP models used in previous CMA-related research. These models evolved into mathematical expressions according to their suitability in prediction tasks.

Figure 1 shows the results of a DT model. Key elements are summarized as follows. Root Node (Node 0) represents the entire dataset ( $n = 131$ ), with a mean ITSM of 496.22 MPa and high variability ( $SD = 539.23$  MPa), indicating significant heterogeneity in the samples. For the primary split, the tree bifurcates at fiber content (FC%) of 0.3% (Bonferroni-adjusted  $p < 0.001$ ;  $F = 46.251$ ), confirming FC% as the dominant predictor. In the terminal nodes, Node 1 ( $\leq 0.3\%$  FC) has a lower ITSM (176.86 MPa,  $SD = 149.36$  MPa) observed in 42.7% of cases ( $n = 56$ ), indicating inadequate fiber reinforcement, and Node 2 ( $> 0.3\%$  FC) had higher ITSM (734.67 MPa,  $SD = 599.51$  MPa) in 57.3% of cases ( $n = 75$ ), demonstrating optimal fiber efficacy.

This split explains 57.3% of the variance in ITSM values. The tree structure confirms that fiber content  $> 0.3\%$  is critical for maximizing stiffness modulus. The high standard deviations in both nodes suggest secondary influences (e.g., fiber type, curing conditions), requiring further investigation. This aligns with established principles of fiber reinforcement in asphalt composites [32].

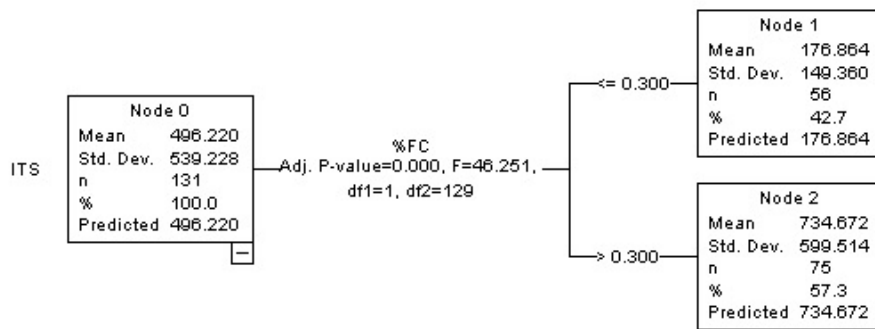


Fig. 1. Predictive model for ITSM.

Among the models evaluated, the one that incorporated eight genetic structures demonstrated superior predictive accuracy, in agreement with previous findings suggesting improved performance with more complex model architectures [33]. This model scored an  $R^2$  value of 0.913, which is an improvement from the simpler model that reached an  $R^2$  value of 0.88. This indicates that the model has a greater degree of accuracy in predicting ITSM values. The  $R^2$  is superior to previous ones, such as 0.80 and 0.90 in [34, 35], respectively. As a result, the proposed CHAID model demonstrates increased prediction capacity. Consequently, based on the structure of the tree and node analysis, the tree structure of the CHAID-based tree model was proven to be resilient and easy to interpret, and can be used as a competent framework for forecasting ITSM. The most influential predictor is FC%.

This sharp increase in the stiffness modulus for mixtures with fiber content greater than 0.3% confirms the positive effect of fiber reinforcement, especially at higher dosages. The Bonferroni-adjusted p-value (Sig.=0.000) and F-statistic (F=46.251) indicate the split is statistically significant, validating the predictive strength of FC%.

The complementary outputs of GEP and CHAID addressed different aspects of the prediction problem. GEP yielded symbolic regression equations representing non-linear interaction effects of fiber characteristics, curing conditions, and air voids. In contrast, CHAID showcased the variable hierarchy via its interpretable DT structure. Specifically, for practical purposes, GEP offered predictive accuracy with explicit mathematical forms, but CHAID provided interpretability of the prediction problem, showing that fiber content was the preeminent determinant of ITSM at a critical threshold of 0.3%. The two models complemented each other in terms of explanations. While GEP quantified the effects (magnitude), CHAID sequenced their logical order (sequence), resulting in a robust framework for accurate and interpretable prediction of ITSM.

The model's residual variance is acceptable given the simplicity (single-variable) and limited depth of the tree. Although further depth or additional variables may enhance precision, the current model remains interpretable and effective.

TABLE II. PERFORMANCE EVALUATION OF THE TREE MODEL

Indicator	Value
RMSE	170.2 MPa
NRMSE	34.9%
MAPE	83.84%
Average Accuracy Percentage (AA%)	16.16%
Coefficient of Determination ( $R^2$ )	0.913
Pearson Correlation Coefficient (R)	0.961

Quantitative evaluation metrics confirm the strong predictive capability of the tree-based model (CHAID) in estimating the ITSM of fiber-enhanced CMA. The model demonstrated an RMSE of 170.2 MPa and an NRMSE of 34.9%, indicating a relatively low level of deviation between predicted and observed values. Additionally, the MAPE was 83.84%, indicating a 16.16% or 1.91-fold AA%, which is good for practical applications. The coefficient of determination ( $R^2$ ) reached 0.913, meaning that predictor variables can provide 91.30% explanation for the differences or variabilities in the output. Finally, the Pearson Correlation coefficient (R) was 0.961, indicating a strong relationship rank between the measured (the source) and the estimated (predicted) ITSM.

These results indicate that DT algorithms are capable of modelling the complicated behaviour of BC systems that use various input features, such as fiber types (FT1, FT2), fiber content (FC%), curing time (CT), air void content (AV%), and fiber length (FL). The proposed CHAID model has great interpretability and robustness compared to other methods, which makes it preferable to use for performance prediction of FMA systems [36, 37].

The scatter plot in Figure 2 displays the predicted vs. the experimental ITSM values, showing that the data points fall around the 1:1 reference line. The data points mainly fall on the right side of the 1:1 reference line, meaning the predicted ITSM values are smaller than the experimental ones. In the mid to high stiffness range (stiffness>80 MPa), most of the data point clusters fall close to the 1:1 reference line. The spread of the data points on the bottom of the 1:1 reference line is significant in the low stiffness range (stiffness<50MPa).

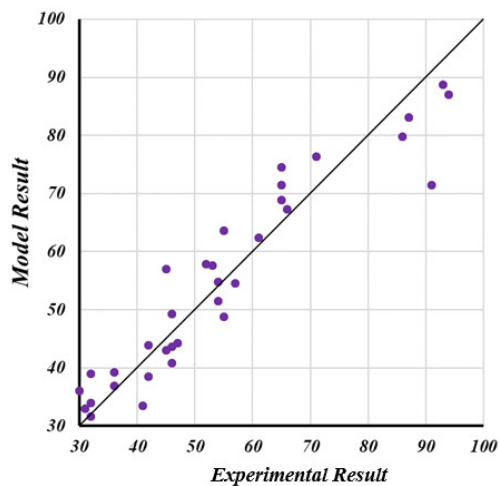


Fig. 2. Agreement between the model and experimental results.

## VI. CONCLUSIONS

After analyzing the results collected from previous studies and the proposed model, the following can be concluded:

- The CHAID model proved to be an acceptable predictor of ITSM, as illustrated by the resulting  $R^2$  of 0.913 and RMSE of 170.2 MPa for fiber-reinforced CMA.
- Among the input variables, fiber content, curing time, and air voids exhibited the highest influence on ITSM, emphasizing the critical role of mixture compaction and maturation in stiffness development.
- The model's transparency and rule-based structure offer engineers a practical, interpretable alternative to black-box algorithms in pavement material design and performance prediction.
- The results confirm the potential of tree-based approaches in advancing data-driven methods for optimizing CMA reinforced with various fiber types.
- These results support the integration of ML methods into asphalt research for the efficient evaluation of engineering performance in relation to mechanical properties and, hence, develop more sustainable approaches toward pavement engineering.

## REFERENCES

- [1] Z. Sun, Y. Ma, S. Liu, Y. Li, X. Qiu, and Z. Luo, "Evaluation of Engineering Properties of Fiber-reinforced Usual-temperature Synthetic Pitch (USP) modified cold mix patching asphalt," *Case Studies in Construction Materials*, vol. 16, Jun. 2022, Art. no. e00997, <https://doi.org/10.1016/j.cscm.2022.e00997>.
- [2] G. Ferrotti, E. Pasquini, and F. Canestrari, "Experimental characterization of high-performance fiber-reinforced cold mix asphalt mixtures," *Construction and Building Materials*, vol. 57, pp. 117–125, Apr. 2014, <https://doi.org/10.1016/j.conbuildmat.2014.01.089>.
- [3] A. Dulaimi, H. Al Nageim, F. Ruddock, and L. Seton, "New developments with cold asphalt concrete binder course mixtures containing binary blended cementitious filler (BBCF)," *Construction and Building Materials*, vol. 124, pp. 414–423, Oct. 2016, <https://doi.org/10.1016/j.conbuildmat.2016.07.114>.
- [4] H. K. Shanbara, F. Ruddock, and W. Atherton, "A laboratory study of high-performance cold mix asphalt mixtures reinforced with natural and synthetic fibres," *Construction and Building Materials*, vol. 172, pp. 166–175, May 2018, <https://doi.org/10.1016/j.conbuildmat.2018.03.252>.
- [5] H. Chen, Q. Xu, S. Chen, and Z. Zhang, "Evaluation and design of fiber-reinforced asphalt mixtures," *Materials & Design*, vol. 30, no. 7, pp. 2595–2603, Aug. 2009, <https://doi.org/10.1016/j.matdes.2008.09.030>.
- [6] H. K. Shanbara, F. Ruddock, and W. Atherton, "Predicting the rutting behaviour of natural fibre-reinforced cold mix asphalt using the finite element method," *Construction and Building Materials*, vol. 167, pp. 907–917, Apr. 2018, <https://doi.org/10.1016/j.conbuildmat.2018.02.072>.
- [7] S. M. Abtahi, M. Sheikhzadeh, and S. M. Hejazi, "Fiber-reinforced asphalt-concrete – A review," *Construction and Building Materials*, vol. 24, no. 6, pp. 871–877, Jun. 2010, <https://doi.org/10.1016/j.conbuildmat.2009.11.009>.
- [8] W. Ben Chaabene, M. Flah, and M. L. Nehdi, "Machine learning prediction of mechanical properties of concrete: Critical review," *Construction and Building Materials*, vol. 260, Nov. 2020, Art. no. 119889, <https://doi.org/10.1016/j.conbuildmat.2020.119889>.
- [9] A. M. Ahmed, M. I. Ali, and A. H. Abbas, "Recycling of Disposal Syringe Needles to Produce Fiber Reinforced Concrete," *International Review of Civil Engineering (IRECE)*, vol. 9, no. 4, pp. 148–153, Jul. 2018, <https://doi.org/10.15866/irece.v9i4.15275>.
- [10] M. I. Ali, A. A. J. Jamel, and S. I. Ali, "The hardened characteristics of self-compacting mortar including carbon fibers and estimation results by artificial neural networks," *AIP Conference Proceedings*, vol. 2213, no. 1, Mar. 2020, Art. no. 020159, <https://doi.org/10.1063/5.0000177>.
- [11] S. Rahman, A. Bhasin, and A. Smit, "Exploring the use of machine learning to predict metrics related to asphalt mixture performance," *Construction and Building Materials*, vol. 295, Aug. 2021, Art. no. 123585, <https://doi.org/10.1016/j.conbuildmat.2021.123585>.
- [12] J. Zhao and H. Wang, "Machine learning based pavement performance prediction for data-driven decision of asphalt pavement overlay," *Structure and Infrastructure Engineering*, vol. 21, no. 6, pp. 940–955, Jun. 2025, <https://doi.org/10.1080/15732479.2023.2258498>.
- [13] A. Eleyedath and A. K. Swamy, "Prediction of dynamic modulus of asphalt concrete using hybrid machine learning technique," *International Journal of Pavement Engineering*, vol. 23, no. 6, pp. 2083–2098, May 2022, <https://doi.org/10.1080/10298436.2020.1841191>.
- [14] R. Y. Liang and J. Zhou, "Prediction of fatigue life of asphalt concrete beams," *International Journal of Fatigue*, vol. 19, no. 2, pp. 117–124, Jan. 1997, [https://doi.org/10.1016/S0142-1123\(96\)00066-7](https://doi.org/10.1016/S0142-1123(96)00066-7).
- [15] A. M. Ahmed, S. I. Ali, M. I. Ali, and A. A. J. Jamel, "Analyzing Self-Compacted Mortar Improved by Carbon Fiber Using Artificial Neural Networks," *Annales de Chimie - Science des Matériaux*, vol. 47, no. 6, pp. 363–369, Dec. 2023, <https://doi.org/10.18280/acsm.470602>.
- [16] M. I. Ali and A. A. Allawi, "An Artificial Neural Network Prediction Model of GFRP Residual Tensile Strength," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18277–18282, Dec. 2024, <https://doi.org/10.48084/etasr.9107>.
- [17] A. Nazari and F. Pacheco Torgal, "Predicting compressive strength of different geopolymers by artificial neural networks," *Ceramics International*, vol. 39, no. 3, pp. 2247–2257, Apr. 2013, <https://doi.org/10.1016/j.ceramint.2012.08.070>.
- [18] M. Rathnayaka, D. Karunasinghe, C. Gunasekara, K. Wijesundara, W. Lokuge, and D. W. Law, "Machine learning approaches to predict compressive strength of fly ash-based geopolymer concrete: A comprehensive review," *Construction and Building Materials*, vol. 419, Mar. 2024, Art. no. 135519, <https://doi.org/10.1016/j.conbuildmat.2024.135519>.
- [19] M. Abdel-Mongy, M. Iqbal, M. Farag, Ahmed. M. Yosri, F. Alsharari, and S. E. A. S. Yousef, "Artificial Intelligence Prediction of One-Part Geopolymer Compressive Strength for Sustainable Concrete," *Computer Modeling in Engineering & Sciences*, vol. 141, no. 1, pp. 525–543, 2024, <https://doi.org/10.32604/cmescs.2024.052505>.
- [20] A. H. Gandomi and A. H. Alavi, "A new multi-gene genetic programming approach to nonlinear system modeling. Part I: materials and structural engineering problems," *Neural Computing and*

- Applications*, vol. 21, no. 1, pp. 171–187, Feb. 2012, <https://doi.org/10.1007/s00521-011-0734-z>.
- [21] C. Ferreira, *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*. Springer, 2006.
- [22] C. Ma, S. Wang, J. Zhao, X. Xiao, C. Xie, and X. Feng, "Prediction of shear strength of RC deep beams based on interpretable machine learning," *Construction and Building Materials*, vol. 387, Jul. 2023, Art. no. 131640, <https://doi.org/10.1016/j.conbuildmat.2023.131640>.
- [23] A. F. Ashour, L. F. Alvarez, and V. V. Toropov, "Empirical modelling of shear strength of RC deep beams by genetic programming," *Computers & Structures*, vol. 81, no. 5, pp. 331–338, Mar. 2003, [https://doi.org/10.1016/S0045-7949\(02\)00437-6](https://doi.org/10.1016/S0045-7949(02)00437-6).
- [24] M. A. Ali Khan, A. Zafar, A. Akbar, M. F. Javed, and A. Mosavi, "Application of Gene Expression Programming (GEP) for the Prediction of Compressive Strength of Geopolymer Concrete," *Materials*, vol. 14, no. 5, Feb. 2021, Art. no. 1106, <https://doi.org/10.3390/ma14051106>.
- [25] G. Abdollahzadeh, E. Jahani, and Z. Kashir, "Predicting of compressive strength of recycled aggregate concrete by genetic programming," *Computers and Concrete*, vol. 18, no. 2, pp. 155–163, Aug. 2016, <https://doi.org/10.12989/CAC.2016.18.2.155>.
- [26] V. Kaushik and M. Kumar, "Sustainable gene expression programming model for shear stress prediction in nonprismatic compound channels," *Sustainable Energy Technologies and Assessments*, vol. 57, Jun. 2023, Art. no. 103229, <https://doi.org/10.1016/j.seta.2023.103229>.
- [27] S. Jahanmiri, M. Noorian-Bidgoli, and A. Dallaki, "Optimizing tunnel boring machine penetration rate prediction using gene expression programming and whale optimization algorithms," *Modeling Earth Systems and Environment*, vol. 11, no. 2, Apr. 2025, Art. no. 133, <https://doi.org/10.1007/s40808-025-02323-0>.
- [28] A. Mollahasani, A. H. Alavi, and A. H. Gandomi, "Empirical modeling of plate load test moduli of soil via gene expression programming," *Computers and Geotechnics*, vol. 38, no. 2, pp. 281–286, Mar. 2011, <https://doi.org/10.1016/j.compgeo.2010.11.008>.
- [29] N. A. M. Shukry *et al.*, "Influence of diatomite filler on rheological properties of porous asphalt mastic," *International Journal of Pavement Engineering*, vol. 21, no. 4, pp. 428–436, Mar. 2020, <https://doi.org/10.1080/10298436.2018.1483504>.
- [30] A. Alfalah, D. Offenbacher, A. Ali, Y. Mehta, M. Elshaer, and C. Decarlo, "Evaluating the impact of fiber type and dosage rate on laboratory performance of Fiber-Reinforced asphalt mixtures," *Construction and Building Materials*, vol. 310, Dec. 2021, Art. no. 125217, <https://doi.org/10.1016/j.conbuildmat.2021.125217>.
- [31] Q. Ye, S. Wu, and N. Li, "Investigation of the dynamic and fatigue properties of fiber-modified asphalt mixtures," *International Journal of Fatigue*, vol. 31, no. 10, pp. 1598–1602, Oct. 2009, <https://doi.org/10.1016/j.ijfatigue.2009.04.008>.
- [32] J. S. Chou, T. P. T. Pham, T. K. Nguyen, A. D. Pham, and N. T. Ngo, "Shear strength prediction of reinforced concrete beams by baseline, ensemble, and hybrid machine learning models," *Soft Computing*, vol. 24, no. 5, pp. 3393–3411, Mar. 2020, <https://doi.org/10.1007/s00500-019-04103-2>.
- [33] F. Aslam *et al.*, "Applications of Gene Expression Programming for Estimating Compressive Strength of High-Strength Concrete," *Advances in Civil Engineering*, vol. 2020, no. 1, Jan. 2020, Art. no. 8850535, <https://doi.org/10.1155/2020/8850535>.
- [34] V. Bherde and U. Balunaini, "Application of Tree-Based Ensemble Techniques to Predict Subgrade Resilient Modulus of Soils," in *Proceedings of 9IYGEC 2023, Volume 1*, vol. 605, N. Satyam, A. P. Singh, and M. S. Dixit, Eds. Springer Nature Singapore, 2025, pp. 199–208.
- [35] T. Wen, H. Lang, S. Ding, J. J. Lu, and Y. Xing, "PCDNet: Seed Operation-Based Deep Learning Model for Pavement Crack Detection on 3D Asphalt Surface," *Journal of Transportation Engineering, Part B: Pavements*, vol. 148, no. 2, Jun. 2022, Art. no. 04022023, <https://doi.org/10.1061/JPEODX.0000367>.
- [36] M. I. Ali, A. A. J. Jamel, and S. I. Ali, "The hardened characteristics of self-compacting mortar including carbon fibers and estimation results by artificial neural networks," presented at the 2<sup>nd</sup> International Conference on Materials Engineering and Science (IconMEAS 2019), Baghdad, Iraq, 2020, Art. no. 020159, <https://doi.org/10.1063/5.0000177>.
- [37] C. McCue, *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis*. Butterworth-Heinemann, 2014.