ARPHA Proceedings

Proceedings CAUSummit 2024, 0254-0264
doi: 10.3897/ap.7.e0254

World Summit: Civil Engineering-Architecture-Urban Planning Congress

A Novel Prediction Technique for the Resilient Modulus of Unbound Aggregates Combined XGBoost Algorithm with Bayesian Optimization

Khaled Sandjak[1, a)], Mouloud Ouanani [2]

[1] *Faculty of Technology, University M'hamed Bougara of Boumerdès, 35000, Boumerdes, Algeria*
[2] *Faculty of Science and Technology, University Ziane Achour of Djelfa, 3117, Djelfa, Algeria*

[a)] Corresponding author: k.sandjak@univ-boumerdes.dz

Abstract. A novel prediction model for the resilient modulus of unbound aggregate bases using eXtreme Gradient Boosting (XGBoost) and Bayesian optimisation (BO) is presented in this study. First, a local dataset containing 260 samples of some common tests, including sieve analysis, Atterberg limits tests, compaction tests, and repeated load triaxial tests of unbound aggregate materials, is collected. Then, the proposed BO-XGBoost model is trained on the collected dataset. Multiple evaluation metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2), were employed to evaluate the performance of the BO-XGBoost model. The model's accuracy and generalisation capabilities were evaluated by computing these metrics for both the training and testing datasets. The findings of this study, clearly showed that the combined BO-XGBoost model outperformed other state-of-the-art machine learning or even deep learning models, in predicting the resilient modulus of unbound aggregate base. Finally, the permutation method was used to conduct a feature importance analysis and determine the impact of different input features on the resilient modulus of the unbound aggregate base. This latter was highly influenced by stress, moisture content, and liquidity limit, as demonstrated by the results of the feature importance analysis.

Keywords: Resilient modulus, unbound aggregates, XGBoost, Bayesian optimisation

INTRODUCTION

Equipped with an estimated 180,000 kilometres of road infrastructure, Algeria possesses one of the most extensive in North Africa. More than fifty percent of the road network are low to moderate-volume roads. To withstand the weight of traffic, these road structures commonly consist of asphalt pavement overlaid with substantial unbound granular base and subbase layers [1]. The resilient modulus of the unbound aggregates is a fundamental input design factor informing this pavement structure.

Three distinct tiers of reliability are defined by the Mechanistic-Empirical Pavement Design Guide (MEPDG).

To achieve the highest level of dependability, the initial tier, which corresponds to Algerian road network class 1, requires input from a resilient modulus test. Correlations and local databases may be employed for the second and third levels, which correspond to Algerian road network class 2 and secondary roads classified as medium to low volume roads [2].

In recent decades, the MEPDG, adopted by the National Road Agency in Algeria, [3] has faced considerable resistance due to its heavy reliance on laboratory analysis of pavement materials and extensive database utilisation. Due to the complex composition of the materials utilised, however, measuring the resilient modulus of unbound aggregates in road construction is a laborious undertaking that requires financial investment and specialised training [4]. Currently, the resilient modulus of locally produced unbound aggregates cannot be determined by certain regional transport agencies in Algeria due to their lack of essential testing capabilities.

Predictive models facilitate the estimation of MR values for crushed rock materials, which are frequently employed in pavement construction. Predictive models for MR typically aim to establish correlations with a diverse array of characteristics, including stress state conditions [7], California Bearing Ratio (CBR) [5], liquid limit (LL), plastic limit (PL), and plasticity index (PI) [6].

Diverse factors, including the type of aggregate, stress level, and fundamental parameters such as unit weight, moisture content, and density, can affect the resilient modulus of crushed rock materials [8]. As stated by Rada and Witczak [9], the MR value was predominantly impacted by the level of stress. The research results suggest that there is a positive correlation between the modulus of granular materials and an increase in confining stress. Furthermore, it has been observed that the resilient modulus (MR) of granular substances is intrinsically linked to deviator stress, a phenomenon caused by the grain's deflection towards a denser configuration [10,11].

The relationship between the MR of aggregate materials and other basic engineering parameters is frequently established through the use of empirical formulas by practitioners. By employing coefficients obtained from conventional regression methods [12], this is accomplished. however, empirical methods have some constraints, notwithstanding their inherent simplicity. One such limitation is the failure to consider the overall influence of all influencing factors and the degree of nonlinearity. Moreover, in the case of specific empirical relationships, traditional regression techniques might yield weak correlations [13].

Over the past few decades, the field of geotechnical engineering has utilized a vast array of machine learning models to predict the complex behaviour of a variety of geomaterials and to establish correlations between inputs and targets in a multitude of experimental datasets. To develop predictive models for unbound granular materials, including aggregate materials [7,17], numerous sophisticated techniques and other machine learning methods, such as artificial neural networks [14,15,16], have been utilized to estimate resilient modulus (MR), according to a review of the specialized literature.

The enhancement of MR prediction model precision has been the uniformly principal aim of a considerable number of researchers. Prior research has neglected the use of hybrid models in favour of focusing primarily on the application of artificial neural networks (ANN) and ensemble models to outperform conventional regression models.

The prediction accuracy of the resilient modulus model is enhanced through the implementation of a Bayesian Optimization Algorithm in conjunction with eXtreme Gradient Boosting (BO-XGBoost) in this study. Selecting the right hyperparameters for eXtreme Gradient Boosting (XGBoost) is crucial, as incorrect values can lead to overfitting or underfitting, compromising the model's performance. To address this, we employ a Bayesian Optimization (BO) integrated with cross-validation. This approach allows for the systematic optimization of hyperparameters, ensuring that XGBoost operates with the most effective settings for enhanced predictive performance.

The objective of this investigation is to ascertain the resilient modulus of aggregate materials that are frequently employed as base and subbase materials in the Northern region of Algeria. We evaluate the performance of the hybrid Bayesian Optimization Algorithm coupled with eXtreme Gradient Boosting (BO-XGBoost) model using widely recognized evaluation metrics such as the coefficient of determination (R^2), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). This model's development is facilitated by a local database that contains 260 experimental samples.

The results of the simulation reveal that the hybrid BO-XGBoost model demonstrates a high level of accuracy in predicting the resilient modulus of aggregate materials. Furthermore, comparative analysis with other models, including standard Multilinear Regression (MLR), Random Forest (RF), and Deep Neural Network (DNN) models, shows that the hybrid BO-XGBoost model significantly outperforms these alternatives in terms of predictive accuracy and reliability.

MATERIALS AND METHODS

The proposed methodology is implemented in the present investigation, which comprises three primary phases: (1) data preparation, (2) model development, and (3) model validation. During the initial phase, the data obtained from laboratory tests is employed to produce two datasets: the training and testing datasets. 80 percent of the total data is utilised to construct the initial dataset, while 20 percent of the residual data is utilised to construct the subsequent dataset. The training dataset is examined in order to develop the hybrid BO-XGBoost model. During this phase, an examination is conducted to determine the extent to which hyperparameter values influence the hybrid model's accuracy. In the third phase, the performance of the proposed model is evaluated against well-established models, such as MLR, Random Forest, and deep neural network models, by employing the testing dataset. This evaluation is conducted using statistical indicators including R^2, RMSE, and MAE.

Data Preparation

A multitude of variables can affect the resilient modulus of unbound aggregates (MR). In contrast, this investigation will focus on the primary variables that have a substantial impact on the resilient modulus to simplify the model. By utilising a local experimental database, the hybrid (BO-XGBoost) model is implemented to predict the aggregate materials resilient modulus. Aggregate geologic origin, basic engineering characteristics, and loading conditions are all considered by this model. The laboratory of the Central Transportation Agency in Algiers, Algeria conducted Repeated Load Triaxial (RLT) tests, which resulted in the acquisition of 260 experimental measures for the database [18]. The RLT tests were conducted on a variety of unbound aggregates that were generated through the quarry crushing of three distinct varieties of massive rocks: diabase, limestone, and granite. These rocks were present in a diverse array of deposits in the central region of northern Algeria.

From the experimental database, A multitude of input parameters are chosen, including the following: Moisture content of the sample (SMC), Rock Type (RT), Coefficient of Uniformity (Cu), Coefficient of Curvature (Cc), Fine content (Fc), Liquidity Limit (LL), Plasticity Index (PI), and Maximum Dry Density (MDD). Furthermore, two loading components are integrated: the deviator of stress, represented as σ_d (SIGMAD), and the confining pressure, represented as σ_3 (SIGMA3). The result of modelling with these parameters is the MR value. The references [18,19] provide detailed explanations and instructions regarding the procedure for deriving input variables from laboratory tests. It is noteworthy to mention that except for the RT variable, which takes on three values (granite, limestone, and diabase) and is categorical, all of these variables are numerical. Table 1 provides a comprehensive description of the continuous input and output variables, including their symbol, unit, and statistical analysis.

The Pearson correlation coefficient among all variables is illustrated in the heatmap of the correlation matrix in Figure 1.

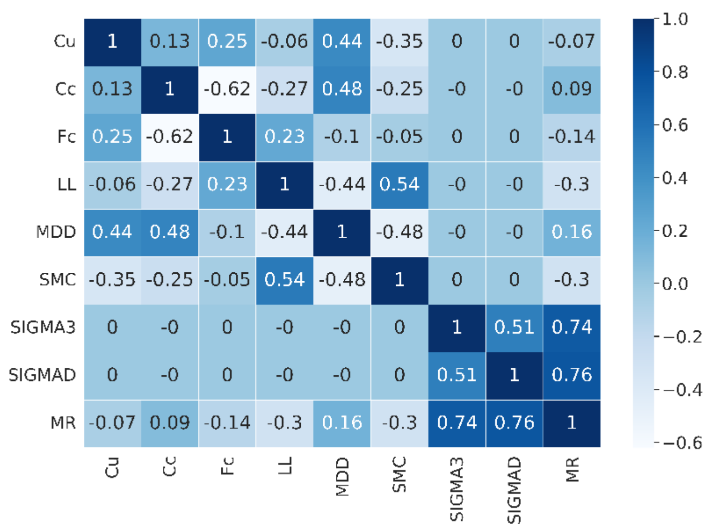

FIGURE 1. Matrix of correlation heatmap

TABLE 1. Statistics that delineate the input parameters employed in this investigation

Variable	Mean	StDev	Min	Max
Cu	58.28	29.22	4.17	112.00
Cc	3.87	1.86	0.88	6.92
Fc (%)	9.46	1.99	8.00	14.08
LL (%)	28.45	4.93	19.00	37.00
PI (%)	7.08	2.61	2.00	12.00
MDD (g/cm^3)	2.23	0.07	2.11	2.34
SMC (%)	6.54	1.13	3.60	8.50
SIGMA3 (kPa)	102.85	71.34	10.00	250.00
SIGMAD (kPa)	335	160.40	150.00	600.00
M_R (MPa)	267.23	128.20	46.60	590.31

Extreme gradient boosting (XGBoost)

As an ensemble learning technique, the boosting algorithm constructs a robust learner from a series of progressively weaker learners. Gradient boosting is a highly effective alternative method for constructing forecasting models. It operates as a stochastic gradient descent technique, incorporating weak learners in a way that minimises data loss. The additive strategy, loss function, and weak learner are its three primary components. In contrast to the random guess, a weak learner makes a prediction with reduced effectiveness. To enhance the accuracy of predictions, the strategy of addition is gradually integrated with the weak learner through a sequential and incremental process.

The training dataset may be rapidly overfitted by gradient boosting due to the algorithm's greedy nature. As a consequence, the XGBoost algorithm [20] is employed in conjunction with the regularisation term to improve the accuracy of predictions by reducing overfitting. Primarily, it comprises the following enhancements: XGBoost operates on the principle of gradient boosting, which entails the fusion of weak predictive models to construct a more robust model. By adding models iteratively to minimise the loss function via gradient descent, gradient boosting is implemented. 2) Regularisation: XGBoost incorporates a variety of regularisation methods to enhance generalisation and mitigate the risk of overfitting. The regularisation technique is employed to regulate the model's complexity and diminish the impact of individual features. 3) Gradient-Based Optimization: To optimise the objective function, XGBoost implements gradient-based optimization methodologies. The algorithm minimises the loss by adjusting the model parameters following the model predictions and determining the gradients of the loss function. 4) Feature Importance: XGBoost offers a metric for determining the relative significance of individual features within the model, known as "feature importance." This can prove advantageous in the processes of feature selection and discerning the primary contributors to the predictions. 5) Ensuring Data Integrity: XGBoost incorporates robust functionalities to manage instances of missing values. Throughout the training procedure, it acquires the ability to handle missing values automatically. XGBoost incrementally generates the trees while facing the dataset (xi, yi). In Equation (1), the sum of the predictions from each tree is possible. By aggregating weak learners in series, it successfully strikes a balance between bias and variance.

$$\hat{y}_i = \sum_{k=1}^{K} f_k\,(x_i) \tag{1}$$

Where K is the number of DT models, f_k denotes the individual DT model.

By utilising the additive strategy, the loss value resulting from the newly added and learned trees is diminished. It also prevents the simultaneous learning of all DT models and increases the efficiency of the model. The objective function (Obj) is simplified and defined as follows with the aid of the second-order Taylor expansion for the aforementioned model:

$$Obj = \sum_{i=1}^{n} \left[L\,(y_i - \hat{y}_i^{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2\,(x_i) \right] + \Omega(f_t) \tag{2}$$

$$g_i = \partial_{\hat{y}_i^{t-1}} L\,(y_i - \hat{y}_i^{t-1}) \tag{3}$$

$$h_i = \partial^2_{\hat{y}_i^{t-1}} L\left(y_i - \hat{y}_i^{t-1}\right) \tag{4}$$

where, L(,) refers to the squared loss function between predictions and ground truth, $\Omega(f_k)$ is the term of regularization, g_i represents the gradients of the loss function at the first order, while *hi* represents the gradients of the loss function at the second order.

Bayesian optimization for hyperparameters search

Optimized hyperparameters ensure good prediction for XGBoost model. Frequently, the objective function and the complex non-convex surface used to locate the optimal hyperparameters of ML models are undefined. Exploring the maximum value at the sampling point for an unknown function can be reduced to the process of identifying the optimal hyperparameters.

$$s^* = argmax_{s\in\Omega} \sum_{i=1}^{n} (s) \tag{5}$$

where (s) represents the sampling data, Ω designates its search space and s^* indicates the point at which the objective function attains its maximum value.

The Bayesian optimization technique is a reliable algorithm for determining the optimal set in scenarios where the objective function is undefined and the calculation surface is quite intricate [21]. By utilising sampling data and prior knowledge, it is possible to approximate the unknown objective function to derive the posterior distribution using the Bayesian theorem. As the amount of sampling data increases, posterior knowledge gradually becomes more comprehensive, enabling the determination of global optimal hyperparameters.

The Gaussian process (GP) and the acquisition function are the two primary components of Bayesian optimization (Z). GP is employed to generate a multidimensional Gaussian distribution that can be used to model the objective function flexibly. Conversely, the subsequent set of sampling data is determined by Z, and its posterior distribution is assessed through proxy optimization. Z is dependent on the model's provision of the mean μ(s) and standard deviation σ(s), and it utilises information from prior sampling data to identify low-confidence areas and determine global optimal hyperparameters. To balance alternatives, the weight parameter (ω) is implemented.

$$Z(s) = \mu(s) - \omega\, \sigma(s) \tag{6}$$

Performance evaluation indices

The proposed models were assessed using three statistical performance evaluation indices in the current study. MR values are predicted and measured, and the mean absolute error (MAE) quantifies the discrepancy. To ascertain the average magnitude of the errors, the RMSE was applied with greater significance to larger errors. Using the coefficient of determination, the precision of the predicted MR values was evaluated (R^2). Statistical indices are comprised of the subsequent equations:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left|y_i - t_i\right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(y_i - t_i\right)^2} \tag{12}$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}(t_i-\bar{t})\,(y_i-\bar{y})}{\sqrt{\left[\sum_{i=1}^{N}(t_i-\bar{t})^2\right]\left[\sum_{i=1}^{N}(y_i-\bar{y})^2\right]}}\right)^2 \tag{13}$$

Where y_i and t_i are predicted and measured values, $\overline{y}$ and $\overline{t}$ are the mean of the predicted and measured values of MR respectively, and N is the number of data records.

RESULTS AND DISCUSSION

Hybrid BO-XGBoost Model Construction

The BO-XGBoost model's accuracy and computational time are significantly influenced by the hyperparameters that are adjusted during its development. There are a multitude of well-known methods for tuning hyperparameters in the literature, such as the grid search algorithm, the random search algorithm, the Bayesian optimization algorithm, and the particle swarm optimization system. Additionally, it is important to note that the random search and grid search algorithms require numerous iterations and can be time-consuming. The conventional method of particle swarm optimization is well-known and time-consuming. In contrast, the BO is a state-of-the-art optimization framework that employs an acquisition function to compute the subsequent point to evaluate, thereby determining the optimal parameters significantly faster than the alternatives. The BO is utilised in the present study as a result of its systematic adjustment process, which eliminates the need for derivatives [22]. In addition, the BO proved to be more effective than alternative methodologies.

To mitigate overfitting, the 10-fold cross-validation strategy was chosen for its minimal root mean square error (RMSE). The hyperparameters of XGBoost were optimised through the implementation of the BO technique, and the models' predictive accuracies were evaluated. Figure 3 and Figure 4 show the progress of the XGBoost hyperparameter optimization and the importance of each hyperparameter in the prediction process respectively. The minimum objective score of 15.12 MPa was observed in 333 trials (iterations). Table 2 provides the tuned parameters that were employed to determine the optimal model.

TABLE 2. The optimal set of hyperparameters for the XGBoost model

Parameters	Explanation	Range	Optimized parameters
n_estimators	Number of iterations (or trees)	100-2000	1698
learning_rate	Shrinkage parameter	0-1	0.1064
max_depth	Maximum depth of a tree	1-10	3
subsample	Subsample ratio of the training instances	0.5-0.9	0.6682
reg_alpha	L1 regularization parameter on weights	0-1	0.07
reg_lambda	L2 regularization parameter on weights	0-1	0.0234

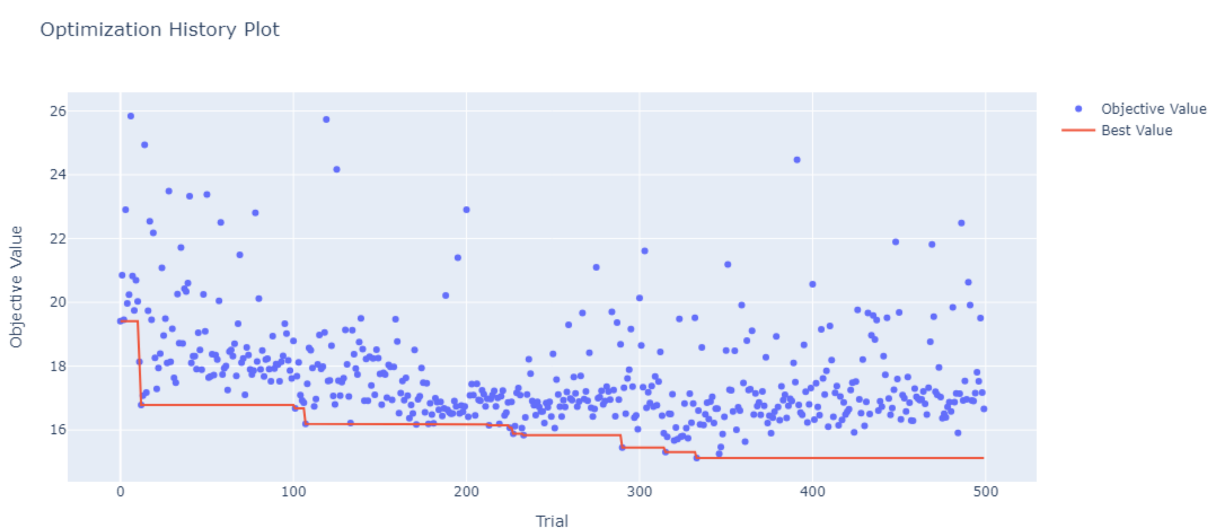

FIGURE 3. The progression of Bayesian hyperparameter optimization curve.

The applied BO–XGBoost model was developed using real-world laboratory trials, as intelligence-based models are data-dependent. Both the XGBoost model and the experiments were able to achieve the descriptive statistics and distribution of results. The statistical analyses of the resilient modulus measured values and the values estimated using the BO–XGBoost model are presented in Table 3.

Conversely, a boxplot is a frequently employed method for illustrating the distribution of data through a five-number summary, which includes the median, maximum, and minimum scores, as well as the first and third quartiles. The boxplots of the experimentally measured and predicted resilient modulus of unbound aggregates are depicted in Figure 4. The results indicate that the output distributions that were measured and predicted were similar.

The datasets did not exhibit any discrepancies, as all of the median lines were contained within the boxes. This suggests that the findings have not been widely disseminated, as the interquartile ranges for both the measured and predicted values were not significantly fluctuating.

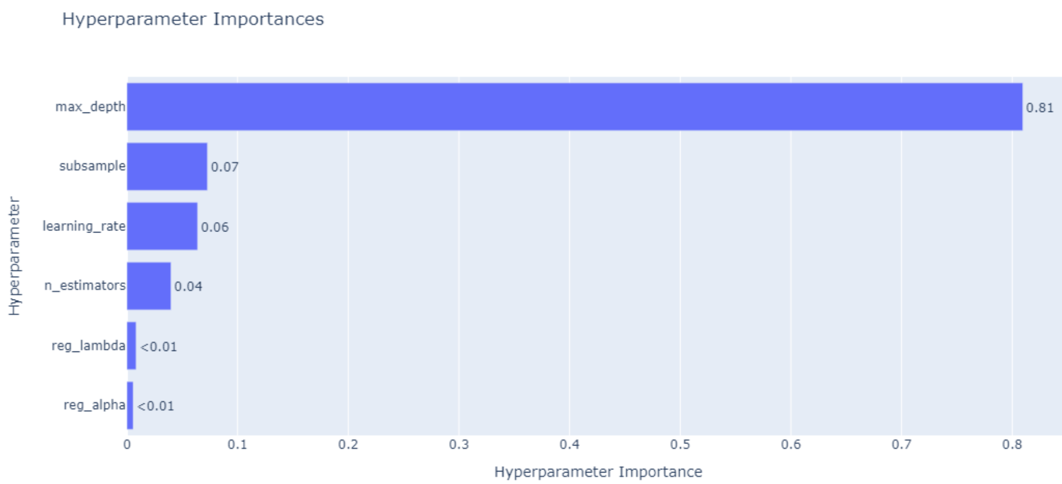

FIGURE 4. The importance of model hyperparameters in the prediction progress

TABLE 3. Statistical analysis of the measured values and the predictions of the BO–XGBoost model

Parameters	Measured	BO-XGBoost
Mean	303.12	302.90
Std. Deviation	126.08	125.59
Min	102.46	102.72
Max	590.31	586.77

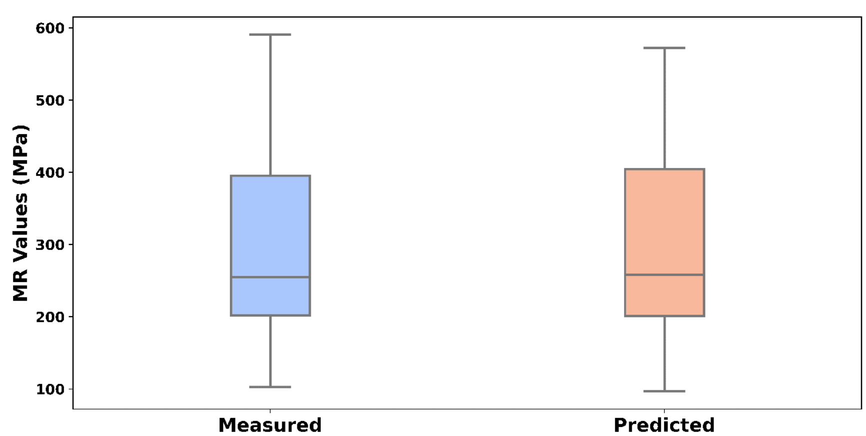

FIGURE 4. Boxplot for the resilient modulus that was predicted and measured.

Evaluation of the Hybrid BO-XGBoost Model

The resilient modulus of unbound aggregates was estimated using the hybrid BO-XGBoost model that was developed for this investigation. The relationship between the experimental and model-predicted resilient moduli is illustrated in Figure 5 by the fitted line, which is denoted as the 45∘ line. The 45° line is evidently in close proximity to the fitted line. A coefficient of determination (R^2) of 99.18 percent indicates a proper fit for the fitted curve on the test dataset. Consequently, it can be inferred that the experimental laboratory data closely resembles the predicted outcomes. The model employs the aforementioned techniques in addition to k-fold cross-validation to avoid overfitting and automatic hyperparameter tuning to optimise results close to the experimental data. The outcome was the attainment of more accurate forecasts.

Furthermore, the residual analysis results are illustrated in Figure 6 to evaluate the model's compatibility. It is important to note that the residual plot is a scatter plot. If the residual data points are dispersed around the zero line, the model is considered acceptable. The reliability and validity of the model employed are substantiated by the fact that the major residual values are within an acceptable range, as illustrated in Figure 6.

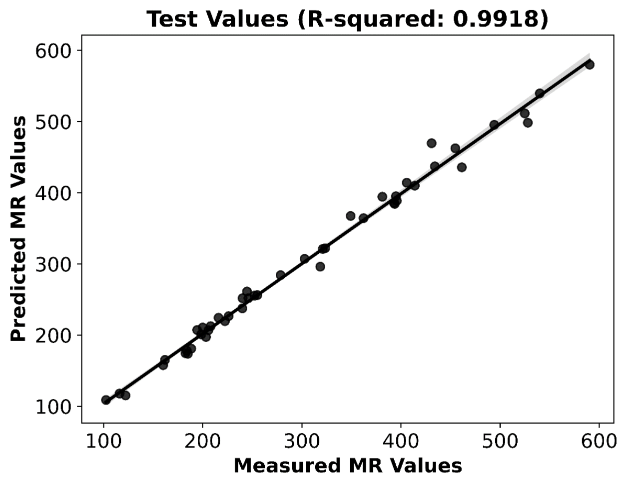

FIGURE 5. scatterplot of measured and predicted Resilient Modulus

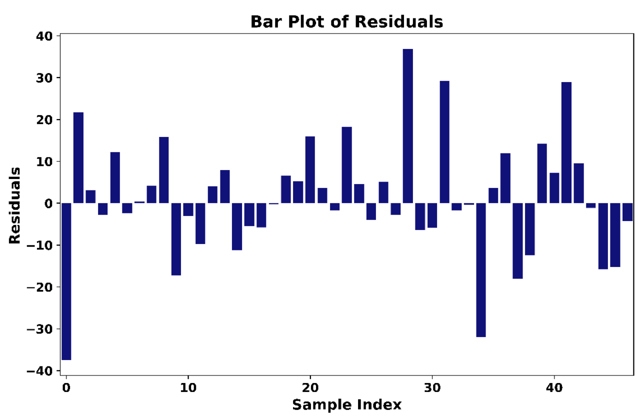

FIGURE 6. Bar plot of residuals

The statistical error parameters (MAE, RMSE) were all found to be low. The resilient modulus of unbound aggregates is strongly predicted by the model employed in this research, as evidenced by the results presented in Table 4. It is important to note that the hybrid BO-XGBoots model predictions are dependable when the coefficient of determination and error values are near 1 and 0, respectively. It is evident that the proposed model meets this criterion.

TABLE 4. Performance indices for the training and testing datasets of BO-XGBoots model

	Train data	Test data
MAE	0.46	8.03
RMSE	0.60	11.29
R^2	0.9999	0.9918

Performance Comparison of Hybrid Model vs other AI models

All performance indices derived from the hybrid model are evaluated in comparison to those associated with the conventional Multiple Linear Regression (MLR), Random Forest, and Deep Neural Network (DNN) models to enable a direct comparison between the BO-XGBoots model and the aforementioned Artificial Intelligence models. In Table 5, the MAE, RMSE, and R^2 values are summarised. As anticipated, the hybrid BO-XGBoots model demonstrated significantly superior performance in comparison to MLR, RF, and DNN models, as demonstrated by significantly lower MAE and RMSE and significantly higher R^2. The results are presented in Table 5.

TABLE 5. Comparative analysis of the models' performance on the test dataset

	BO-XGBoots	MLR	RF	DNN
MAE	8.03	27.33	22.78	10.52
RMSE	11.29	36.16	26.40	14.62
R^2	0.9918	0.9024	0.9300	0.9862

Feature Importance Analysis

It is advantageous to develop a model that accurately forecasts outcomes; however, in the majority of instances, it is also necessary to interpret the model's output. Feature importance is the most valuable instrument for interpretation. Model parameters are routinely examined by data scientists to identify significant features.

Permutation feature importance is a model inspection technique that quantifies the statistical performance of a fitted model on a given tabular dataset by evaluating the contribution of each feature. A technique that is highly beneficial when applied to non-linear or opaque estimators is the random shuffling of the values of a single feature and the subsequent decrease in the model's score [23,24]. The model's degree of dependence on a particular feature is determined by severing the connection between the feature and the target. Figure 7 illustrates the significance of features in the prediction of resilient models using the BO-XGBoost model.

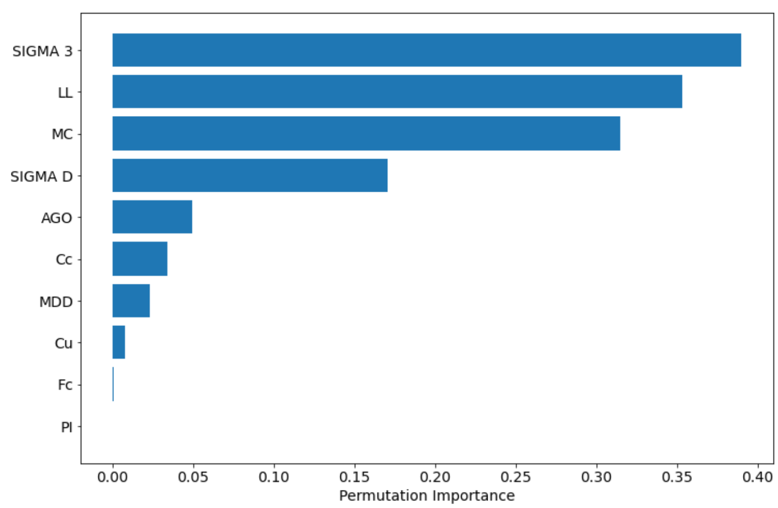

FIGURE 7. Feature importance analysis results

The most influential factors in predicting the resilient modulus are stress conditions (SIGMA 3, SIGMA D), the liquid limit (LL), and moisture content (MC), as illustrated in Figure 7.

CONCLUSION

In this research article, a model was created to estimate the resilient modulus of unbound aggregates by utilising XGBoost and a Bayesian optimization algorithm. A 10-fold cross-validation strategy was implemented to avoid overfitting. The laboratory of the Central Transportation Agency conducted a comprehensive collection of 260 experimental measures involving crushed rock materials, which facilitated the development of the BO-XGBoost model. The model's performance was evaluated using a variety of performance-measuring indicators. Additionally, the anticipated outcomes were evaluated in comparison to several AI models. From the simulation results, the following conclusions can be drawn:

1- The experimental results are closely matched by the resilient modulus of unbound aggregates prediction using the hybrid BO–XGBoost model. Additionally, the mean, median, standard deviation, and interquartile ranges of the predicted and measured results are exceedingly similar.
2- The laboratory results and estimated results are in agreement as a result of the R^2 value being extremely close to 1 between the predicted and measured values.
3- The proposed model's residual data points are observed to be located near the zero line, which serves as further confirmation of the model's dependability.
4- Comparatively, the mean, standard deviation, and coefficient of variation of the proposed BO–XGBoost model are the lowest of all the AI models discussed in this article.
5- The resilient modulus of unbound aggregates was significantly influenced by stress, liquid limit, and moisture content, as evidenced by the results of the permutation importance method of feature importance analysis.
6- The BO–XGBoost model can be employed to construct a dependable soft computing tool for pavement design, particularly for roads with low to moderate traffic volume.
7- The generalizability of the proposed model could be improved by incorporating supplementary data from other regions, as it is constructed using a regional database.

REFERENCES

1. F. Mama, "Réseau routier et autoroutier Algérien: consistances et perspectives", (In French), *Conference on Road Safety,*18 September 2017*,* Algiers, Algeria.
2. AASHTO, *Mechanistic–Empirical Pavement Design Guide: A Manual of Practice*, (American Association of State and Highway Transportation Officials, USA, 2008), ISBN: 1-56051–423-7.
3. CTTP, *Catalogue de Dimensionnement des Chaussées*, (In French), (Direction des Routes, Ministère des Travaux Publics, Algiers, Algeria, 2001).
4. T. Nguyen, H.B. Ly and B. T. Pham, "Backpropagation neural network-based machine learning model for prediction of soil friction angle", Mathematical Problems in Engineering, ID 8845768 (2020), https://doi.org/10.1155/2020/8845768.
5. C. Pati and M. Tsakoumaki, "A critical comparison of correlations for rapid estimation of subgrade stiffness in pavement design and construction", Constr. Mater. 2023, 3(1), 127-142, https://doi.org/10.3390/constrmater3010009.
6. A.M. Rahim, "Subgrade soil index properties to estimate resilient modulus for pavement design", Int. J. Pavement Eng. 6 (3), 163–169 (2005).
7. H.H. Titi and M.G. Matar, "Estimating resilient modulus of base aggregates for mechanistic-empirical pavement design and performance evaluation", Transportation Geotechnics (17), 141–153 (2018).
8. M.Y. Abu-Farsakh, A. Mehrotra, L. Mohammad and K. Gaspard, "Incorporating the Effect of Moisture Variation on Resilient Modulus for Unsaturated Fine- Grained Subgrade Soils", Transportation Research Record, 2510, (1), 44-53 (2015).
9. G. Rada and M. Witczak, "Comprehensive evaluation of laboratory resilient moduli results for granular material", Transp. Res. Rec. J. Transp. Res. Board., 810 (1), 23–33 (1981).
10. F. Lekarp, U. Isacsson and A. Dawson, "State of the art. I: Resilient response of unbound aggregates", J. Transp. Eng., 126 (1), 66–75 (2000) .

11. S. Saha, F. Gu, X. Luo and R.L. Lytton, “Use of an Artificial Neural Network Approach for the Prediction of Resilient Modulus for Unbound Granular Material”, Transportation Research Record, 2672 (52), 23–33 (2018).
12. L.N. Mohammad, B. Huang, A.J. Puppala and A. Allen, “Regression Model for Resilient Modulus of Subgrade Soils”, Transportation Research Record, 1687 (1), 47–54 (1999).
13. A. Alnedawi, M. Al-Ameri and K.P. Nepal, “Neural network-based model for permanent deformation of unbound granular materials”, Journal of Rock Mechanics and Geotechnical Engineering, vol. 11,1231-1242 (2019).
14. A.R. Ghanizadeh, R. Morteza, “Application of artifitial neural network to predict the resilient modulus of stabilized base subjected to wet-dry cycles”, Comput. Mater. Civ. Eng. 1 (1), 37–47 (2016).
15. M.R. Kaloop, A.R. Gabr, S.M. El-Badawy, A. Arisha, S. Shwally and J.W. Hu, “Predicting resilient modulus of recycled concrete and clay masonry blends for pavement applications using soft computing techniques”, Front. Struct. Civ. Eng. 13 (6), 1379–1392 (2019).
16. K. Sandjak, M. Ouanani and T. Messafer, “Bayesian Regularized Backpropagation Neural Network Model to Estimate Resilient Modulus of Unbound Granular Materials for Pavement Design” in *Advanced Computational Techniques for Renewable Energy Systems*, edited by M. Hatti (Springer, 2023), pp. 457-468.
17. F. Gu, H. Sahin, X. Luo, R. Luo and R.L. Lytton, “Estimation of Resilient Modulus of Unbound Aggregates Using Performance-Related Base Course Properties”, J. Mater. Civ. Eng. 27, (6), 04014188 (2015).
18. K. Sandjak, and M. Ouanani, “Experimental characterisation and numerical modelling of the resilient behaviour of unbound granular materials for roads”, J.Building Materials and Structures, 7 (2), 159-177 (2020).
19. Y. H. Huang, *Pavement Analysis and Design* (Pearson India, Delhi, 2003), pp. 270.
20. Chen, T.Q., C. Guestrin, and M. Assoc Comp, XGBoost: A Scalable Tree Boosting System. Kdd’16: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016: p. 785–794.
21. Zhang Y, Dong Y, Feng R. Bayes-informed mixture distribution for the EVD estimation and dynamic reliability analysis. Mech Syst Signal Process 2023;197: 110352.
22. D.T. Chang, “Bayesian hyperparameter optimization with BoTorch, GPyTorch and Ax”, Computer Science-Machine learning, (2019), https://doi.org/10.48550/arXiv.1912.05686.
23. L. Breiman, “Random Forest”, Machine Learning, (2001), 45, 5-32, https://dx.doi.org/10.1023/A:1010933404324
24. https://scikit-learn.org/stable/modules/permutation_importance.html