SPECTRUM OF ENGINEERING SCIENCES
ISSN (E) 3007-3138 (P) 3007-312X VOLUME 4, ISSUE 3, 2026

COMPUTING THE DYNAMIC MODULUS OF ASPHALT CONCRETE CONTAINING RAP BY USING AI MODEL: DEEP NEURAL NETWORKS WITH BAYESIAN OPTIMIZATION

Rawid Khan[*1], Fawad Mehmood[2], Sahibzada Hamza Jan[3]

[*1,2,3]*Department of Civil Engineering University of Engineering and Technology Peshawar, Pakistan*

[*1]dr.rawid@uetpeshawar.edu.pk, [2]17pwciv4843@uetpeshawar.edu.pk, [3]shamzajan@uetpeshawar.edu.pk

DOI: https://doi.org/10.5281/zenodo.22706683

Keywords
Dynamic Modulus, Hot Mix Asphalt, Reclaimed Asphalt Pavement, Deep Neural Network, Bayesian Optimization.

Article History
Received: 01 January 2026
Accepted: 16 March 2026
Published: 31 March 2026

Corresponding Author: *
Rawid Khan

Abstract

The dynamic modulus (|E|) of Hot Mix Asphalt (HMA) is a fundamental mechanistic property that characterizes the viscoelastic behavior of asphalt mixtures under varying loading rates and temperatures and serves as a critical input in mechanistic empirical pavement design. Laboratory determination of |E*| is time-consuming, costly, and requires specialized equipment, limiting its practical applicability. Predictive models are widely used, but conventional regression-based approaches often lack accuracy and generalizability. To address these limitations, this study develops a machine learning (ML) based approach using Artificial Neural Network (ANN) to estimate the |E*| of HMA mixtures containing reclaimed asphalt pavement (RAP). The present study employed a dataset compiled from recently published research, comprising a total of 160 data points, which encompass cumulative percent retained on 3/8 inches sieve (P38), cumulative percent retained on No. 4 sieve (P4), percent passing No. 200 sieve (P200), air voids (Va), voids in mineral aggregate (VMA), voids filled with asphalt (VFA), effective asphalt content (V_{beff}), asphalt content (AC), dust to binder ratio (DP), testing temperature (T), loading frequency (F), binder replacement type RAP (BR_Type_RAP), and binder replacement percent (BR%) as predictor variables to develop a Deep Neural Network (DNN) through BO for estimating the |E*| of HMA mixtures. The developed DNN model exhibited high prediction accuracy across all data subsets, with R^2 exceeding 0.99 across all data subsets, while MAE remained under 486 MPa and MAPE was below 9%. Sensitivity assessment identified that T and the P4 had the greatest influence on the predicted |E*| values.*

INTRODUCTION

Good road quality is very important for a country's economic growth. Well-designed roads support the development of any region. To build a strong and durable road, it is important to understand different types of road distress. It is also necessary to choose suitable materials that can resist these stresses and last for a long time [1]. The dynamic modulus (|E*|) of asphalt concrete is an important property that shows the stiffness of Hot Mix Asphalt (HMA) mixtures under different loading rates and temperatures. It is measured using a standard procedure in which stress and strain are recorded under repeated sinusoidal

loading at various temperatures and loading frequencies. This parameter is widely used in the Mechanistic-Empirical Pavement Design Guide (MEPDG) to estimate pavement response and predict its performance [2].

It is also used to develop performance based models for HMA mixtures, including models that predict fatigue cracking and rutting behavior [3]. The |E*| is found by dividing the maximum applied stress by the corresponding recoverable strain [4]. Although laboratory testing is the most reliable method for determining the |E*|, it is costly and time-consuming. In some cases, it is also difficult or impractical to conduct tests under extreme temperature and loading conditions [5]. The use of reclaimed asphalt pavement (RAP) in HMA mixtures is common in road construction and rehabilitation, especially in North America and Europe. RAP offers several benefits, including reduced waste, conservation of natural resources, preservation of existing pavement geometry, lower life-cycle costs, and energy savings. Studies have shown that recycled HMA mixtures can perform as well as or even better than conventional HMA mixtures in both laboratory and field conditions. In addition, the use of RAP in HMA mixtures can improve resistance to permanent deformation and fatigue cracking [6].

Previous studies have shown that several factors affect the |E*| of HMA mixtures. Nguyen et al. reported that additive type, aging, temperature, and loading frequency have a strong influence on the |E*|. Harvey et al. found that air voids and the level of compaction can significantly affect the stiffness of HMA mixtures. Other studies have shown that the |E*| also depends on factors such as aggregate gradation, binder type, aging level, and volumetric properties. In addition, Islam et al. observed that the |E*| increases with higher effective asphalt content, air voids, and void-filled asphalt, but decreases with an increase in voids in mineral aggregate and asphalt content [7]. Studies have shown that adding RAP significantly affects the |E*| of asphalt concrete mixtures. Mixtures with higher RAP content generally have higher stiffness, especially at lower loading frequencies or higher temperatures [8].

In recent years, improvements in computer programming have increased the use of Machine Learning (ML) techniques to predict the performance of HMA. These methods are now widely used in many civil engineering applications. One of the earliest ML techniques is the Artificial Neural Network (ANN). Over the past two decades, ANN have been widely used to develop prediction models because they can learn patterns in data and capture hidden relationships. They have also been commonly used to predict the |E*| of HMA mixtures [9]. Many studies have used ANN in different areas of pavement engineering. These include predicting pavement distress, modeling pavement structures, analyzing and designing pavement performance, planning rehabilitation and maintenance, pavement materials and properties, and pavement [10].

A Deep Neural Network (DNN) is an advanced type of ANN. Like other neural networks, it has three main parts: an input layer, one or more hidden layers, and an output layer. The main difference between DNN and a basic ANN is that DNN has many hidden layers. Because of this, DNN is more powerful and can learn complex patterns in data. It is widely used in areas such as computer vision, speech recognition, pattern recognition, natural language processing, and financial analysis [11]. Previous studies have shown that ML models perform better when their hyperparameters are optimized using systematic methods, rather than relying on manual trial-and-error [1]. The performance and prediction accuracy of an ANN model depend on its structure and training settings, known as hyperparameters. These need to be carefully selected to improve the model performance [12].

Bayesian optimization (B0) is a powerful method for tuning hyperparameters, especially when the problem has many variables. Traditional methods like grid search are slow and inefficient because they test all possible combinations of hyperparameters. In contrast, BO uses a probabilistic model to identify the most promising areas to explore. This allows the method to focus on better solutions instead of searching randomly. As a result, it finds optimal hyperparameters more quickly. It also automatically balances exploration and exploitation [13].

In this study, a DNN optimized through BO was developed for estimating the dynamic modulus of HMA mixtures containing RAP. The input variables included mixture volumetric characteristics, aggregate gradation parameters, and testing conditions. Before developing the model, the dataset was carefully prepared by eliminating outliers, scaling the input and output variables, converting binder replacement categories into one-hot encoded variables, and dividing the data through a balanced stratified splitting process. BO was employed to determine the most suitable DNN hyperparameter combination based on validation performance. The final model was then assessed in terms of prediction accuracy, error distribution, and variable influence using statistical evaluation metrics, residual plots, and sensitivity assessment.

Literature Review

This section reviews previous studies related to the prediction of |E*| of HMA mixtures using ANN models. The reviewed studies mainly focus on the use of ANN for improving prediction accuracy compared with conventional regression based models. In addition, this section also discusses recent applications of BO for hyperparameter tuning of ANN models in civil engineering and pavement related problems.

Santos et al. [14] developed an ANN model to predict the |E*| of Brazilian HMA mixtures using a database of 1,806 laboratory measured values. Aggregate gradation parameters, mixture volumetric properties, and binder properties were utilized as inputs. The study showed that ANN provided high prediction accuracy across different network topologies, with the best model achieving an R^2 of 0.98, confirming the suitability of ANN for estimating the |E*| of Brazilian asphalt mixtures. The study also showed that ANN produced more accurate predictions than the original Bari and Witczak model.

Gong et al. [2] developed an ANN model to improve prediction of the |E*| of HMA mixtures using 7,400 laboratory records. Their results showed that the proposed ANN models significantly outperformed conventional Witczak predictive equations, with the best model achieving an R^2 of 0.99 and demonstrating superior accuracy and generalization. The findings also showed that binder properties and mixture volumetrics had greater influence on |E*| prediction than aggregate gradation parameters.

Heidaripanah and Hassani [15] developed an ANN model to predict the |E*| of HMA mixtures. Using 1,320 laboratory test records and 12 input variables related to binder properties, mixture volumetrics, aggregate gradation, loading frequency, temperature, and confining stress, the proposed ANN model achieved excellent predictive accuracy, with an R^2 value of 0.99. Their sensitivity analysis further showed that temperature and confining stress were the most influential factors affecting |E*|, with temperature producing the strongest decreasing effect and confining stress the strongest increasing effect.

Barugahare et al. [16] developed an ANN model to predict the |E*| of HMA mixtures containing RAP. The model used aggregate gradation parameters, mixture volumetric properties, and binder properties as input variables. The results showed that the ANN model significantly outperformed conventional regression models and achieved excellent prediction accuracy, with R^2 values greater than 0.99 and MAPE lower than 9% on both training and testing datasets. These findings confirmed that ANN provides a reliable and accurate approach for prediction.

Li et al. [17] developed a BO optimized 1D-CNN model for structural anomaly detection using vibration-based monitoring data. Their results showed that BO efficiently tuned model hyperparameters and significantly improved detection performance, with the proposed model achieving a high identification accuracy of 99.85%. Their findings demonstrated that 1D-CNN can provide accurate and robust structural condition prediction while reducing the limitations of manual hyperparameter tuning.

Gao et al. [18] developed a BO optimized ANN model to predict homogenized Mohr–Coulomb parameters of heterogeneous soils. Their results showed that BO improved ANN hyperparameter tuning and enhanced prediction efficiency by reducing manual trial-and-error calibration. The proposed ANN model achieved low prediction errors and provided an accurate and computationally

efficient approach for geotechnical parameter prediction.

Elshaboury et al. [19] applied an ANN model with BO to improve local road pavement condition prediction. Their findings showed that BO effectively optimized ANN hyperparameters and improved prediction reliability compared with conventional ANN tuning. Their results demonstrated that the ANN–BO approach provided robust pavement condition prediction, although its performance was lower than the best-performing ensemble models.

METHODOLOGY

This section presents the methodology adopted in this study, including dataset collection, data preprocessing, model development, performance evaluation, residual and sensitivity analysis. To maintain consistency and allow reproducible results, fixed random seed were applied at certain locations throughout the workflow.

A. Dataset Collection and Description

The dataset used in this study includes 160 data points of |E*| obtained from laboratory tested HMA mixtures consisting of both conventional HMA mixtures and mixtures containing RAP. The dataset was collected from previously published studies [20]. The mixtures were prepared with different RAP contents of 0%, 15%, 25%, and 40%. The input variables included cumulative percent retained on 3/8 inches sieve (P38), cumulative percent retained on No. 4 sieve (P4), percent passing No. 200 sieve (P200), air voids (Va), voids in mineral aggregate (VMA), voids filled with asphalt (VFA), effective asphalt content (V_{beff}), asphalt content (AC), dust to binder ratio (DP), testing temperature (T), loading frequency (F), binder replacement type (BR_Type), and binder replacement percent (BR%). These variables were used to predict the |E*| of HMA mixtures. The |E*| was selected as the target or output variable in the dataset.

B. Data Processing and Preparation

The dataset was imported into MATLAB and randomly shuffled to remove any ordering bias. Then outliers were removed using the Interquartile Range (IQR) method, applied only to the target variable, |E*|. The datapoints that did not fall within the limits specified in (1) were treated as outliers and removed from the dataset to improve its quality and minimize unwanted variation, where Q1 and Q3 are the first and third quartiles, and IQR is the interquartile range as shown in (2).

$$Q_1 - 1.5 \times IQR \leq x \leq Q_3 + 1.5 \times IQR \quad (1)$$

$$IQR = Q_3 - Q_1 \quad (2)$$

After cleaning, basic statistical properties such as minimum, maximum, standard deviation, skewness, and kurtosis were checked, and histograms and boxplots were used to visually assess the data quality. In addition, correlation analysis was performed on the continuous input variables to examine the relationships between the input variables. The categorical variable (BR_Type) was converted into numerical form using one hot encoding by creating binary variables for each binder replacement type. The dataset was then divided into training (60%), validation (20%), and testing (20%) sets using combined stratification approach based on BR_Type and the target variable |E*| to maintain consistent data distribution. Finally, all input variables and the target variable were normalized to a range of [0, 1] using minimum-maximum normalization by using (3), where x_{min} and x_{max} are the minimum and maximum values of the variable, respectively.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3)$$

C. Bayesian Optimized Deep Neural Network Model Development

Following the data preparation stage, a DNN model was developed to predict the |E*| of HMA mixtures. The DNN model consists of an input layer, multiple hidden layers, activation functions, and a single output layer. To achieve optimal performance, BO was used to tune the hyperparameters of the DNN model. These hyperparameters and their corresponding ranges used in BO process are provided in Table I. The optimization was implemented using MATLAB's bayesopt function.

Table I. DNN Model hyperparameter ranges used for the Bayesian Optimization

Hyperparameter	Range
Activation function	ReLU, tanh, sigmoid, leaky ReLU
Optimizer	Adam, SGDM, RMSProp
Number of hidden layers	1 – 5
Number of neurons per layer	10 – 30
Initial learning rate	10^{-4} – 10^{-2} (log scale)
Maximum training epochs	50 – 3000
Acquisition function	Expected Improvement (EI)
Random Iterations or Trials	20
Guided Iterations or Trials	80

For each iteration of the optimization process, the BO built a DNN using a specific set of hyperparameters from the available ranges, trained it on the training data, and evaluated its performance on the validation data. The objective function of the BO was configured to minimize the validation MAE, and its goal was to achieve the lowest possible Mean Absolute Error (MAE) on the validation set within a fixed number of trials.

A total of 100 trials were conducted, starting with 20 random runs to explore different combinations, followed by 80 guided runs using an Expected Improvement (EI) acquisition function. The goal was to find the best model by minimizing the validation Mean Absolute Error (MAE). In each iteration, the model was trained using the training dataset and evaluated on the validation and testing datasets to ensure good prediction performance. The BO mechanism utilized within this study is illustrated in Fig. 1.

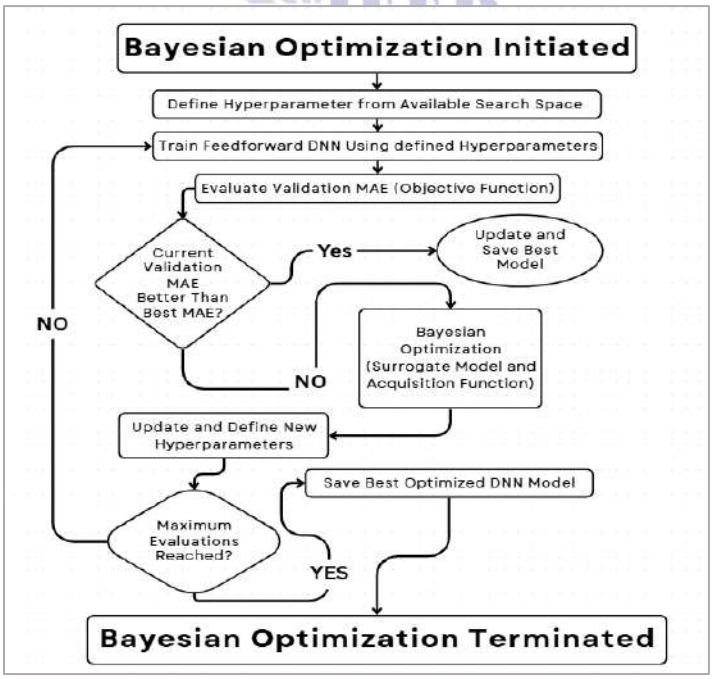

Fig. 1. Bayesian optimization framework for tuning DNN hyperparameters

The performance of each trial was evaluated using the Coefficient of Determination (R^2), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Mean Absolute Error

(MAE), which are defined in (4) – (7), where y_i and $\hat{y}_i$ represent the actual and predicted values, respectively, and n represents the number of samples.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (4)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (7)$$

D. Evaluation of DNN Structure, Predictive Performance, and Residuals Analysis

The final DNN model was selected based on its performance on the validation MAE to ensure good generalization. The architecture of the final DNN model was visually presented through a schematic representation showing the arrangement of layers, neurons, and their connections. The agreement between the measured and predicted |E*| values was examined using scatter plots for the training, validation, and testing datasets. These plots helped assess how well the DNN model captured the variation in |E*| across its full observed range. The training and validation loss plot are prepared by using the training history of the DNN model to evaluate the model learning behavior. The residual analysis was performed using the test dataset to evaluate prediction errors and assess the consistency of model performance across different binder replacement types. This analysis included examining residual behavior, absolute error distribution, and mean absolute error (MAE) variation among binder replacement types.

E. Sensitivity Assessment

A parametric sensitivity assessment was carried out to assess the influence of each continuous input variable on the predicted |E*|. Each input variable was changed within its range using only the training data, while keeping others variables constant at their mean values, and the change in output was observed. The sensitivity was evaluated using the mean based baseline values to ensure consistency and robustness of the results. Finally, the input variables were ranked based on their level of influence on the predicted |E*|. The sensitivity percentage was determined for each variable by using (8), where $\hat{y}_{max}$ and $\hat{y}_{min}$ correspond to the predicted responses at the upper and lower values of the input variable, respectively, and $\hat{y}_{\text{base}}$ is the model prediction at the mean values of all variables.

$$Sensitivity(\%) = \frac{|\hat{y}_{max} - \hat{y}_{min}|}{\hat{y}_{\text{base}}} \times 100 \quad (8)$$

RESULTS

F. Data Distribution, Preparation, and Processing Results

The statistical summary of the raw and processed datasets used in this study is presented in Table II. The Interquartile Range (IQR) method detected no outliers in the target variable, |E*|. Therefore, no observations were removed from dataset during preprocessing. The processed dataset remained identical to the raw dataset, and all 160 data points were retained for DNN model development.

Most input variables exhibit low to moderate skewness and kurtosis values, indicating relatively balanced distributions with limited asymmetry. The aggregate gradation variables (P38, P4, and P200), mixture volumetric properties (Va, VFA, and V_{beff}), T, AC, BR%, and DP all show skewness values close to or near zero, and kurtosis values between 1.44 to 1.93, suggesting near-symmetric distributions. The variables such as F and VMA show higher skewness and kurtosis values as compared to other input variables, indicating moderate deviation from normality. Overall, these trends remain within acceptable statistical limits for laboratory generated engineering datasets.

Table II. Descriptive statistics of the original and refined datasets

Input Variable	Original Dataset				
	Min	***Max***	***Std***	***Skewness***	***Kurtosis***
T (°C)	-10.00	30.00	14.19	0.00	1.70
F (Hz)	0.10	20.00	6.56	1.45	3.76
P38 (%)	22.50	27.10	1.73	-0.07	1.62
P4 (%)	48.70	51.20	0.93	0.27	1.76
P200 (%)	3.00	3.80	0.30	0.19	1.59
Va (%)	3.90	5.00	0.43	0.61	1.80
VMA (%)	16.30	17.40	0.44	-1.07	2.27
VFA (%)	69.80	77.00	2.75	-0.68	1.93
V_{beff} (%)	12.30	13.40	0.43	0.61	1.80
AC (%)	5.20	6.00	0.32	0.27	1.44
BR% (%)	0.00	40.00	15.20	0.19	1.59
DP	0.55	0.76	0.08	0.00	1.48
|E*| (MPa)	207.00	22934.50	6649.22	0.36	1.83
Input Variable	**Refined Dataset**				
	Min	***Max***	***Std***	***Skewness***	***Kurtosis***
T (°C)	-10.00	30.00	14.19	0.00	1.70
F (Hz)	0.10	20.00	6.56	1.45	3.76
P38 (%)	22.50	27.10	1.73	-0.07	1.62
P4 (%)	48.70	51.20	0.93	0.27	1.76
P200 (%)	3.00	3.80	0.30	0.19	1.59
Va (%)	3.90	5.00	0.43	0.61	1.80
VMA (%)	16.30	17.40	0.44	-1.07	2.27
VFA (%)	69.80	77.00	2.75	-0.68	1.93
V_{beff} (%)	12.30	13.40	0.43	0.61	1.80
AC (%)	5.20	6.00	0.32	0.27	1.44
BR% (%)	0.00	40.00	15.20	0.19	1.59
DP	0.55	0.76	0.08	0.00	1.48
|E*| (MPa)	207.00	22934.50	6649.22	0.36	1.83

The statistical properties of the target variable |E*|, remained unchanged after IQR screening, indicating that no outliers were detected or removed. The identical minimum, maximum, standard deviation, skewness, and kurtosis values before and after preprocessing confirm that the dataset was internally consistent and exhibit no abnormal extreme target values, as shown in Table II and Fig. 2.

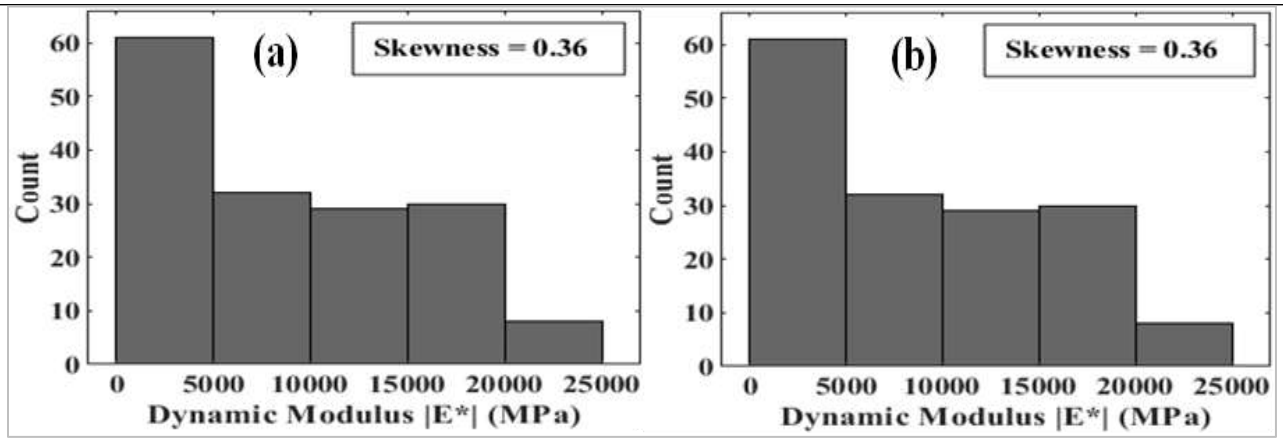

Fig. 2. Dynamic modulus |E*| distribution (a) before and (b) following outlier treatment

This observation is further supported by the boxplots shown in Fig. 3, where the distribution of |E*| remains unchanged before and after preprocessing. The boxplots confirm the absence of abnormal outliers and show that the spread, median, and interquartile range of the |E*| remained stable.

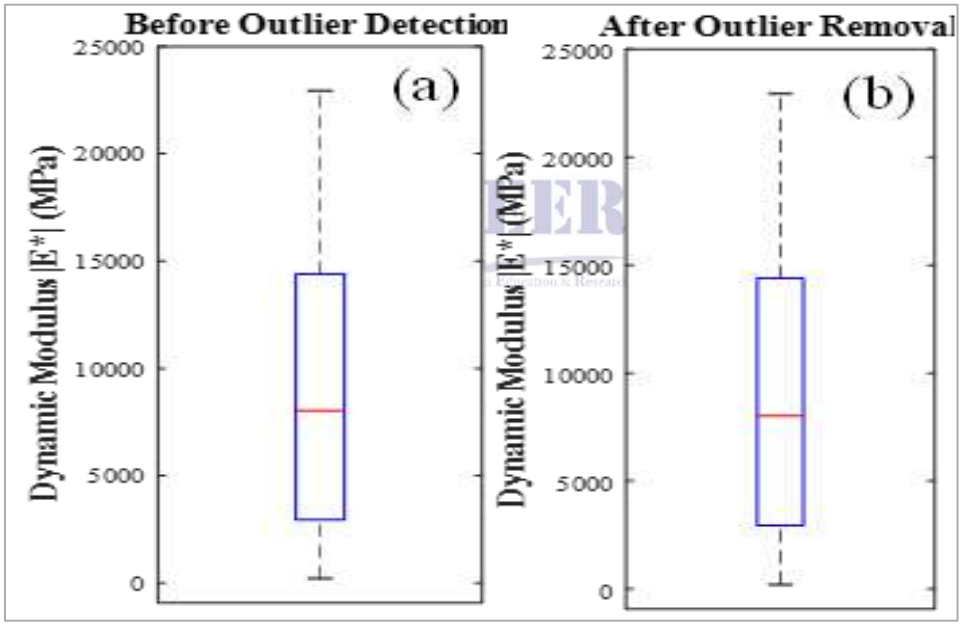

Fig. 3. Evaluation of outliers in dynamic modulus |E*| (a) before and (b) following IQR based refinement

The correlation analysis in Fig. 4 indicates strong relationships among several input variables. Strong positive correlations were observed between P38, P4, V_{beff}, and AC, indicating that these variables vary consistently with changes in aggregate structure and asphalt binder content. Strong negative correlations were observed between P200 and P38, and P4 variables, as well as between VFA and both Va and P200 variables. In contrast, T and F variables show negligible correlation with other input variables. Although several input variables were strongly correlated, all were retained for modeling to preserve predictive information, as the primary objective was to maximize prediction accuracy rather than reduce the number of input variables. This is because neural network can effectively handle correlated inputs.

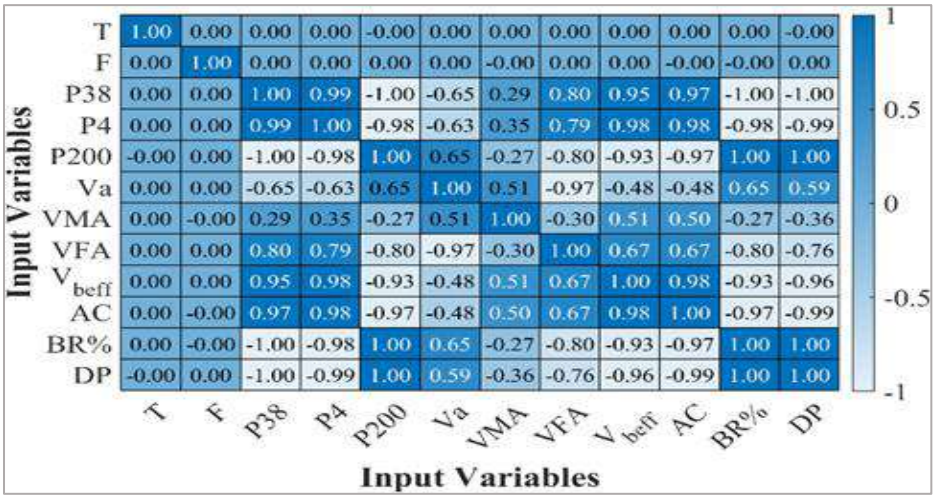

Fig. 4. Correlation analysis of input features in the refined dataset

The distribution of binder replacement type across the training, validation, and testing subsets is summarized in Table III. The proportion of each binder replacement type remained nearly consistent across all three subsets, confirming that stratified sampling preserved a balanced distribution during data splitting. Table IV shows that the target variable |E*| maintained comparable statistical properties, including minimum-maximum values, mean, and standard deviation, across each subset. This consistency confirms that the dataset was split appropriately and that each subset is representative of the overall dataset.

Table III. Binder replacement type distribution across the data subsets

Binder Replacement Type	Training Subset Size	Validation Subset Size	Testing Subset Size	Training Subset Percentage	Validation Subset Percentage	Testing Subset Percentage
Virgin	25	8	7	25.77	25.81	21.88
RAP	72	23	25	74.23	74.19	78.13

Table IV. Statistical characteristics of dynamic modulus across all subsets

Dataset	Dynamic Modulus (|E*|) Statistics Across Subsets			
	Min (MPa)	*Max (MPa)*	*Mean (MPa)*	*Std (MPa)*
Training	207	22934.5	9126.60	6663.93
Validation	378.33	21350.25	9408.68	6440.14
Testing	415.33	22587.5	8622.52	6983.95

G. Convergence Behavior and Bayesian Optimization Results

The convergence pattern of the BO procedure is presented in Fig. 5. It shows how the objective function (validation MAE) changes as the number of function evaluations increases. At the beginning, the validation MAE stays almost constant, indicating little improvement. It then slightly decreases between 20 and 28 evaluations. A major improvement occurs around 58 evaluations. After this point, the validation MAE remains stable, showing that the model has reached its optimal solution.

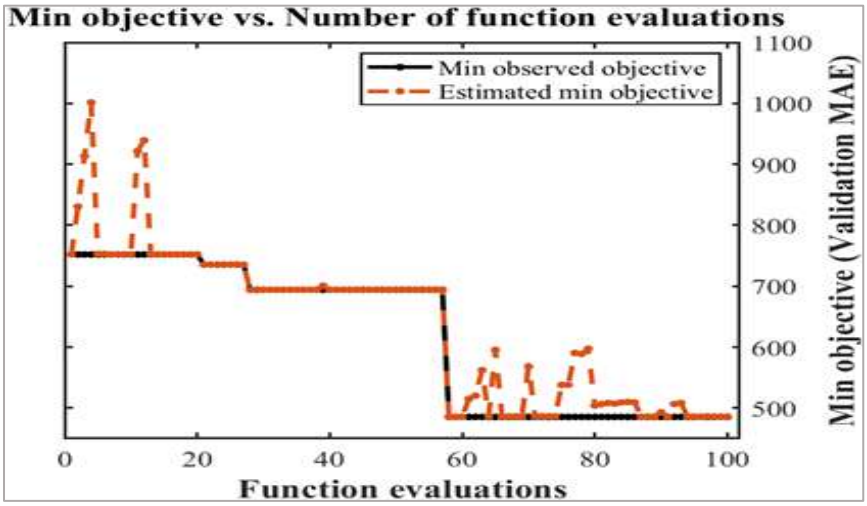

Fig. 5. Bayesian Optimization convergence behavior demonstrating validation MAE decrease across evaluations

The variation in model performance across all BO trials using the evaluation metrics MAE, R^2, RMSE, and MAPE, is illustrated in Fig. 6. The most effective model configuration was identified at the 59th iteration, where the minimum validation MAE was achieved. As shown in Fig. 7, the DNN model hyperparameter combination selected at 59th iteration also produced the lowest validation RMSE and MAPE, together with the highest validation R^2, indicating superior overall predictive performance compared with all other hyperparameter settings explored during the BO process.

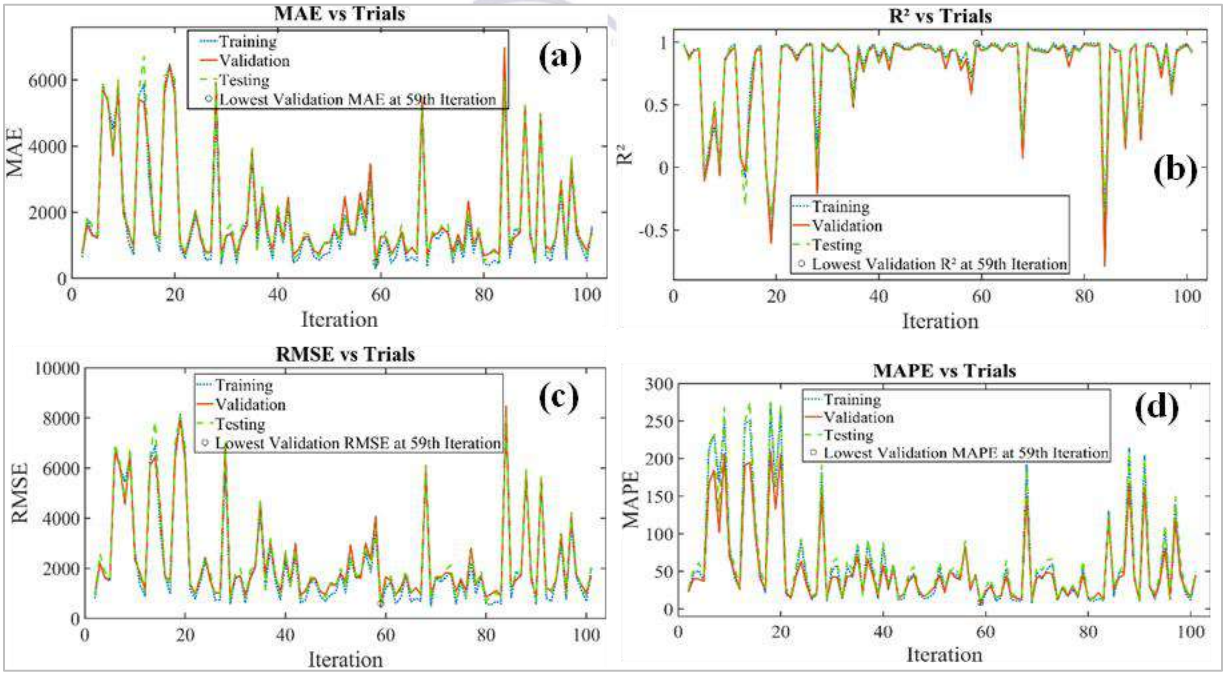

Fig. 6. Bayesian Optimization evaluations history showing training, validation, and testing performance using (a) MAE (b) R^2 (c) RMSE and (d) MAPE

The effectiveness of BO in iteratively refining the DNN hyperparameter search space through continuous surrogate model updates is shown in Fig. 6. As the optimization progresses, the acquisition guided search converges toward more optimal hyperparameter configurations, resulting in lower validation RMSE, MAE, and MAPE and greater validation R^2.

The finalized hyperparameter combination associated with the optimal DNN model identified through BO is listed in Table V. These settings correspond to the trial that achieved the minimum validation error.

Table V. Optimal DNN hyperparameters obtained through Bayesian Optimization

Hyperparameter	Value
Activation function	ReLU
Output activation	Linear
Optimizer	Adam
Hidden layers	3
Input layer neurons	14
Neurons per hidden layer	23
Output layer neurons	1
Learning rate	0.00848754
Maximum epochs	2811

H. Performance and Architecture of the Optimal DNN Model

Fig. 7 compares the actual and predicted values for the training, validation, and testing datasets. The results show a good agreement between the predicted and actual values, with high R^2 values of 0.995, 0.991, and 0.993 for the training, validation, and test sets, respectively. The error values are also low, with all MAE and RMSE below 486 MPa and 596 MPa, and MAPE below 9%, across all subsets, which indicates good prediction accuracy. The performance slightly deceased from training to testing. Overall, the DNN model performs well and gives reliable results for all datasets.

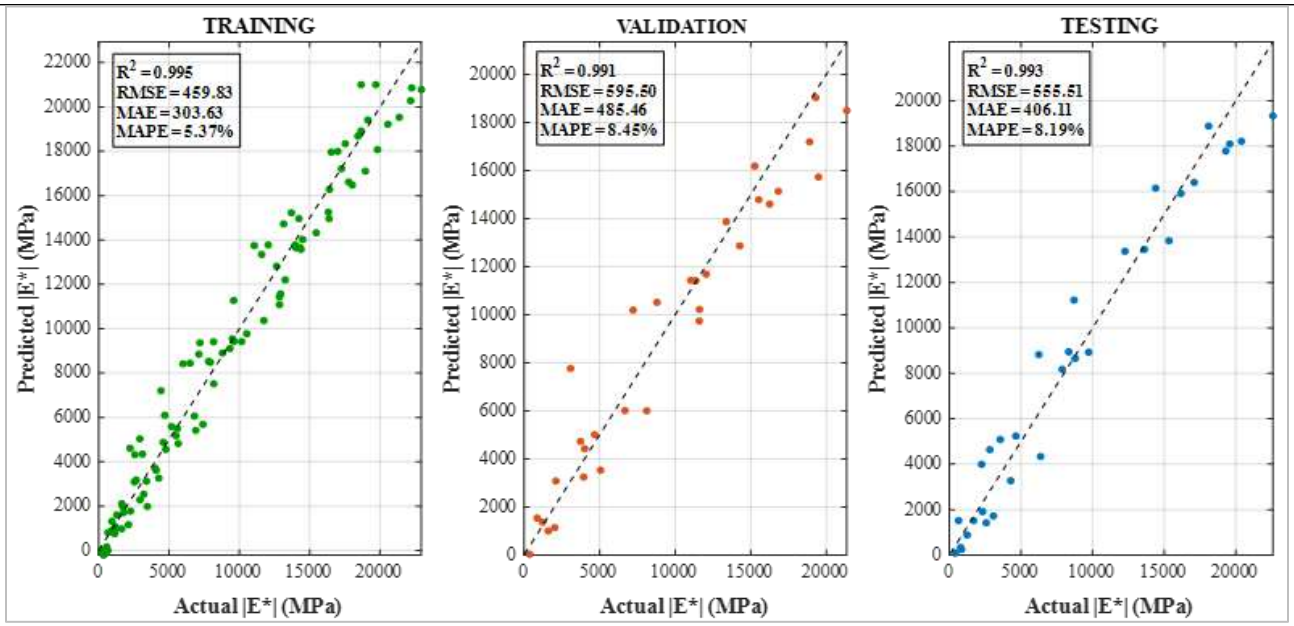

Fig. 7. Scatter plots of measured and predicted dynamic modulus values for (a) training, (b) validation, and (c) testing subsets

Table VI summarizes the predictive performance of the developed DNN model across the training, validation, and testing datasets at both the global level and for binder replacement categories. The Virgin and RAP binder showed strong predictive performance across all data subsets. However, RAP binder exhibited comparatively lower performance in the validation set with an R^2 value of 0.93, along with slightly higher MAE and RMSE values, indicating greater prediction uncertainty in that subset. In contrast, the highest MAPE was observed for Virgin binder in the testing set. Overall, the model demonstrated reliable predictive capability across all binder replacement categories.

Table VI. Evaluation of model performance across data partitions at overall and binder replacement types

Dataset	Level	Category	R^2	MAE	RMSE	MAPE
Training	Global	ALL	1.00	303.63	459.83	5.37
	Binder Type	Virgin	0.97	794.76	1029.31	20.34
		RAP	0.97	999.20	1229.17	21.06
Validation	Global	ALL	0.99	485.46	595.50	8.45
	Binder Type	Virgin	0.96	1242.13	1624.70	32.13
		RAP	0.93	1274.49	1672.28	22.29
Testing	Global	ALL	0.99	406.11	555.51	8.19
	Binder Type	Virgin	0.98	807.58	987.10	52.31
		RAP	0.95	1213.65	1468.53	21.95

Fig. 8 shows how the training and validation loss curves change during the training process. Both losses decrease quickly in the beginning, which means the model learns effectively in the early stage. The loss decreases very slightly afterward. The training and validation curves remain very close to each other, which indicates that there is no overfitting and the model performs well.

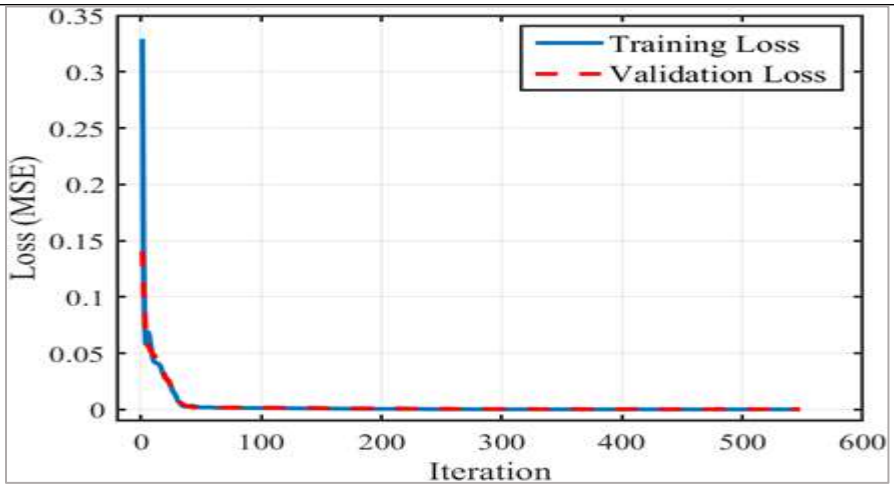

Fig. 8. Training and validation loss behavior of the developed deep neural network model

The structural configuration of the optimal DNN model selected through BO is shown in Fig. 9, and the corresponding hyperparameter details of the model are summarized in Table V.

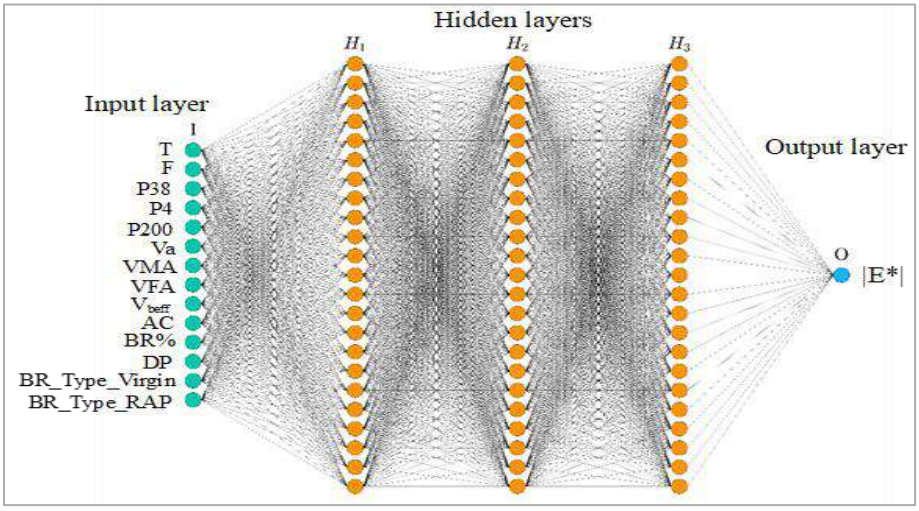

Fig. 9. Schematic representation of the developed deep neural network model architecture

I. Analysis of Residual Behavior and Prediction Errors

Fig. 10 shows the residuals versus predicted values for the test dataset. The points are spread on both sides of the zero line, which means the model does not have any strong bias. Most of the points are close to the zero line, indicating small errors and good predictions. It is observed that RAP data points generally exhibit higher residuals, while Virgin data points showing lower errors and better prediction accuracy. This pattern is consistent with the trend previously observed in Fig. 7. Overall, the model performs better for Virgin samples, while RAP samples show slightly more variation. However, since most of the errors are small, the model can be considered reliable.

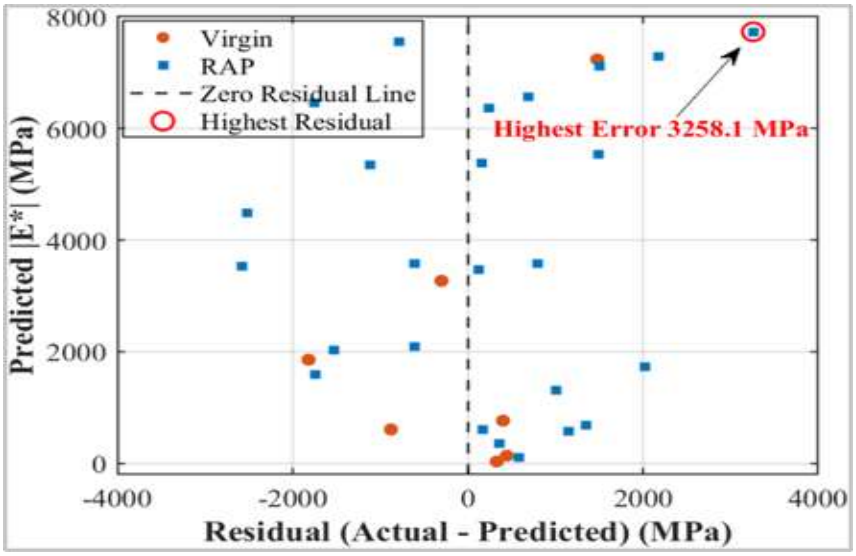

Fig. 10. Distribution of prediction residuals for the optimized deep neural network model on the testing subset

Fig. 14 presents the top 10 individual absolute errors for the test dataset. The results show that the RAP binder having the highest absolute error of approximately 3258.10 MPa, with most of the top errors exceeding 1500 MPa. It can be clearly seen that most of these high-error cases belong to RAP binder, while only one is from a Virgin binder. Overall, even though a few RAP binder have large errors, they are limited in number.

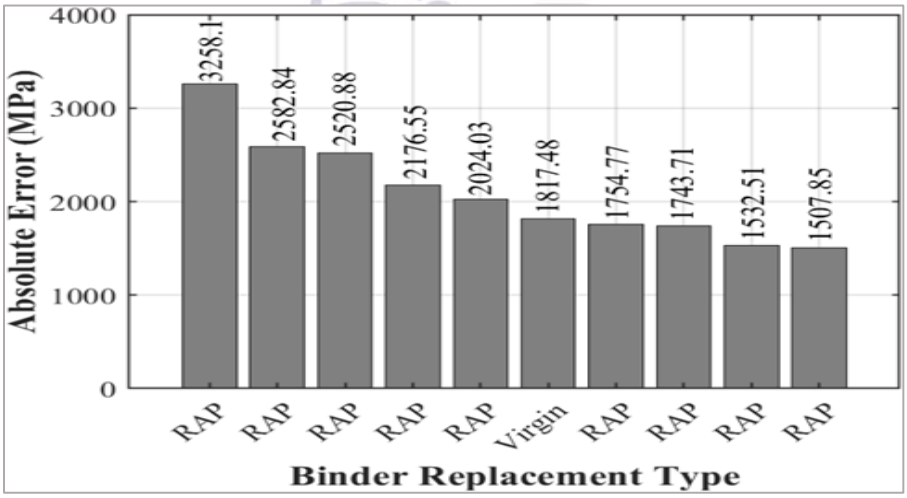

Fig. 11. Highest absolute estimation errors identified in the testing subset

Fig. 12 shows the Mean Absolute Error (MAE) for different binder replacement types. It can be seen that RAP binder have a higher MAE as compared to Virgin binder, which is consistent with the slightly lower predictive performance previously observed for RAP binder in Table 6. Overall, the residual analysis suggests that the model predictions are reasonably consistent, with the majority of errors concentrated close to zero and only a small number of cases showing comparatively larger deviations.

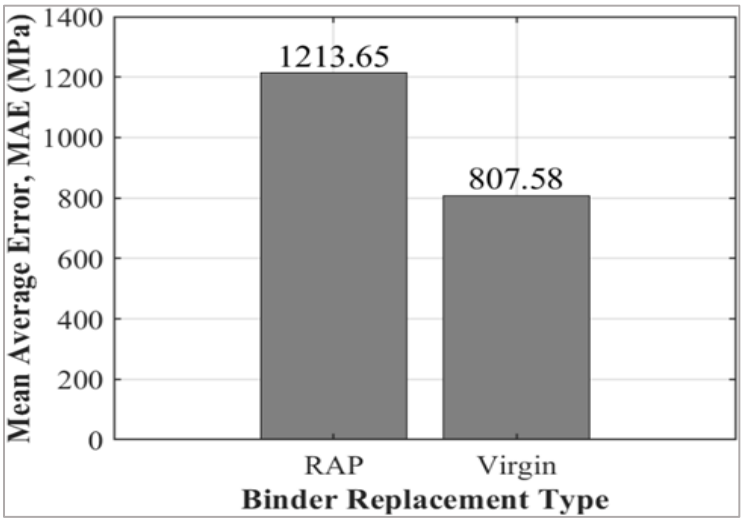

Fig. 12. Comparison of mean absolute error values across different binder replacement types

J. Sensitivity Assessment of Input Variables

Fig. 13 presents the parametric sensitivity analysis results using mean baseline approach. The T is the most important parameter, with a sensitivity of about 23.7%. The P4, DP, VMA, and Va also strongly affect the model output. A moderate effect was observed for V_{beff} and VFA, whereas BR%, AC, and F show lower influence. P38 and P200 showed comparatively least effect. Overall, the results indicate that testing temperature and mixture volumetric properties are the most important factors for predicting |E*|. The finer gradation parameters have small effect on model output.

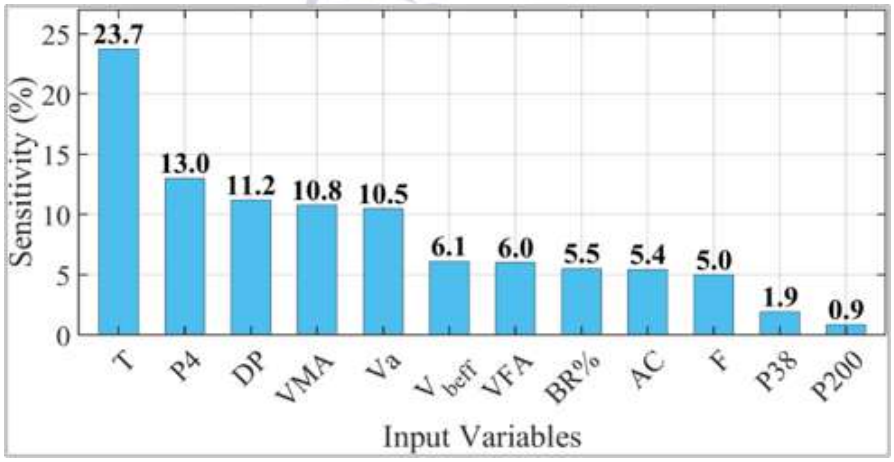

Fig. 13. Sensitivity ranking of input features contributing to dynamic modulus estimation

K. Practicality and Drawbacks of the DNN Model

The developed DNN model is recommended for prediction only when the input data fall within the limits covered by the training subset. Predictions for values outside the training limits may exhibit reduced reliability, as the model has not been trained for such conditions. Therefore, to ensure accurate results, all input features should remain within the minimum and maximum limits of the training subset.

Before applying the developed DNN model to new data, the input features must be scaled according to the lowest and highest values established from the training subset adopted during DNN model training. Similarly, the estimated outputs should be converted

back to their original scale through denormalization using the associated lowest and highest output values from the training subset.

The performance of the developed model largely depends on the quality, completeness, and distribution of the training subset. Estimation accuracy may decrease when the input data contain excessive noise, absent values, or exhibit a distribution that differs substantially from the data used during model training. Therefore, appropriate data preprocessing, quality checking, and evaluation of new input data are suggested before applying the model in practical applications.

CONCLUSION

The BO approach effectively selected the most suitable DNN configuration by automatically searching the hyperparameter space. This eliminated the need for repeated manual tuning. The final optimized network included 3 hidden layers, 23 neurons in each hidden layer, using ReLU as the activation function and Adam as the optimizer. The selected model was trained for nearly 2811 epochs.

The developed BO-DNN model produced highly accurate predictions of |E*|. Strong agreement was observed between measured and predicted values for the training, validation, and testing datasets, with R^2 values close to 0.99. The testing dataset also showed low prediction errors, with a MAE of 406.11 MPa and RMSE of 555.51 MPa, confirming that the model performed well on unseen data. The MAPE values were below 9% for all datasets, further supporting the accuracy and reliability of the model. The binder replacement type specific results showed that the model predicted both Virgin and RAP binder with acceptable accuracy. However, RAP binder showed slightly greater prediction uncertainty in the validation dataset, as reflected by a lower R^2 value of 0.93 and higher error values. For Virgin binder, the highest MAPE was observed in the testing dataset. The residual analysis also supported these findings, showing that most errors were small and centered near zero, while only a few larger deviations were observed, mainly at medium to high |E*| levels.

The sensitivity assessment revealed that testing temperature (T) had the strongest influence on predicted |E*|, followed by P4, DP, VMA, and Va. Variables such as V_{beff} and VFA showed moderate influence, while BR%, AC, and F had relatively lower effects. The least influential variables were P38 and P200.Overall, the findings indicate that the BO-DNN model is an effective data-driven tool for estimating the |E*| of HMA mixtures containing RAP. The framework can reduce dependence on costly and time-consuming laboratory testing, provided that future predictions are made within the limits of data used during model training.

ACKNOWLEDGMENT

The authors sincerely thank the University of Engineering and Technology (UET), Peshawar, for making available the resources and academic environment required to conduct this research. The lead author further extends his appreciation to his research supervisor, senior researchers, and fellow colleagues for their insightful advice, constructive support, and motivation throughout the course of this work.

REFERENCES

[1] I. Y. Amir, A. M. Yusuf, And I. Uwanuakwa, "A Metaheuristic Approach Of Predicting The Dynamic Modulus In Asphalt Concrete," *Eng. Technol. Appl. Sci. Res.*, 2024, Doi: 10.48084/Etasr.6808.

[2] H. Gong *Et Al.*, "Improved Estimation Of Dynamic Modulus For Hot Mix Asphalt Using Deep Learning," *Constr. Build. Mater.*, Vol. 263, P. 119912, Dec. 2020, Doi: 10.1016/J.Conbuildmat.2020.119912.

[3] D. Zhang, B. Birgisson, And X. Luo, "A New Dynamic Modulus Predictive Model For Asphalt Mixtures Based On The Law Of Mixtures," *Constr. Build. Mater.*, Vol. 255, P. 119348, Sep. 2020, Doi: 10.1016/J.Conbuildmat.2020.119348.

[4] G. G. Al-Khateeb, A. Shenoy, N. Gibson, And T. Harman, "A New Simplistic Model For Dynamic Modulus Predictions Of Asphalt Paving Mixtures," *J. Assoc. Asph. Paving Technol.*, Accessed: May 05, 2026. [Online]. Available: Https://Www.Researchgate.Net/Publication/301676481_A_New_Simplistic_Model_For_Dynamic_Modulus_Predictions_Of_Asphalt_Paving_Mixtures

[5] X. Shu And B. Huang, "Predicting Dynamic Modulus Of Asphalt Mixtures With Differential Method," *Road Mater. Pavement Des.*, Doi: 10.1080/14680629.2009.9690198.

[6] Z. Shahadan, M. O. Hamzah, A. S. Yahya, And A. Jamshidi, "Evaluation Of The Dynamic Modulus Of Asphalt Mixture Incorporating Reclaimed Asphalt Pavement," *Indian J. Eng. Mater. Sci.*, Vol. 20, No. 5, Pp. 376–384, 2013.

[7] H. Zhao *Et Al.*, "Study On The Influence Factors Of The Dynamic Property Of The Polyurethane Mixture With Dense Gradation," *Coatings*, Vol. 13, No. 8, P. 1465, Aug. 2023, Doi: 10.3390/Coatings13081465.

[8] B. Asadi, R. Hajj, And I. L. Al-Qadi, "Asphalt Concrete Dynamic Modulus Prediction: Bayesian Neural Network Approach," *Int. J. Pavement Eng.*, Doi: 10.1080/10298436.2023.2270569.

[9] G. S. Moussa And M. Owais, "Modeling Hot-Mix Asphalt Dynamic Modulus Using Deep Residual Neural Networks: Parametric And Sensitivity Analysis Study," *Constr. Build. Mater.*, Vol. 294, P. 123589, Aug. 2021, Doi: 10.1016/J.Conbuildmat.2021.123589.

[10] M. Y. Taha, "Using Radial Basis Function Artificial Neural Network For Predicting Asphalt Content Of Asphalt Paving Mixture," *Iraqi J. Stat. Sci.*, 2013, Doi: 10.33899/Iqjoss.2013.081298.

[11] V. Asghari, Y. F. Leung, And S.-C. Hsu, "Deep Neural Network Based Framework For Complex Correlations In Engineering Metrics," *Adv. Eng. Inform.*, Vol. 44, P. 101058, Apr. 2020, Doi: 10.1016/J.Aei.2020.101058.

[12] A. T. Asad, B. Kim, S. Cho, And S.-H. Sim, "Prediction Model For Long - Term Bridge Bearing Displacement Using Artificial Neural Network And Bayesian Optimization," *Struct. Control Health Monit.*, Doi: 10.1155/2023/6664981.

[13] S. Soulaimani, A. Soulaimani, K. Abdelrahman, A. Miftah, M. Fnais, And B. K. Mondal, "Advanced Machine Learning Artificial Neural Network Classifier For Lithology Identification Using Bayesian Optimization," *Front. Earth Sci.*, 2024, Doi: 10.3389/Feart.2024.1473325.

[14] I. G. Santos, R. P. Espindola, And F. T. S. Aragão, "Prediction Of The Dynamic Modulus Of Brazilian Asphalt Mixtures Using Artificial Neural Networks," In *Proceedings Of The Xli Ibero-Latin-American Congress On Computational Methods In Engineering*, Accessed: May 05, 2026. [Online]. Available: Https://Www.Researchgate.Net/Publication/355982506_Prediction_Of_The_Dynamic_Modulus_Of_Brazilian_Asphalt_Mixtures_Using_Artificial_Neural_Networks

[15] A. Heidaripanah And A. Hassani, "A Fundamental Multiple-Stage Ann-Based Sensitivity Analysis To Predict The Dynamic Modulus Of Hot Mix Asphalt Considering The Effect Of Confining Stress," *Mater. Struct.*, Vol. 54, No. 1, P. 15, Feb. 2021, Doi: 10.1617/S11527-020-01581-X.

[16] J. Barugahare, A. N. Amirkhanian, F. Xiao, And S. N. Amirkhanian, "Evaluation Of Ann-Based Dynamic Modulus Models Of Asphalt Mixtures," *J. Mater. Civ. Eng.*, Vol. 33, No. 6, P. 04021099, Jun. 2021, Doi: 10.1061/(Asce)Mt.1943-5533.0003721.

[17] X. Li, H. Guo, L. Xu, And Z. Xing, "Bayesian-Based Hyperparameter Optimization Of 1d-Cnn For Structural Anomaly Detection," *Sensors*, Vol. 23, No. 11, P. 5058, Jan. 2023, Doi: 10.3390/S23115058.

[18] Y. Gao, G. Huang, Y. Li, J. Zhang, Z. Yang, And M. Wang, "The Data-Driven Homogenization Of Mohr–Coulomb Parameters Based On A Bayesian Optimized Back Propagation Artificial Neural Network (Bp-Ann)," *Appl. Sci.*, Vol. 13, No. 21, P. 11966, Jan. 2023, Doi: 10.3390/App132111966.

[19] N. Elshaboury, M. S. Yamany, S. Labi, And O. Smadi, "Enhancing Local Road Pavement Condition Prediction Using Bayesian-Optimized Ensemble Machine Learning And Adaptive Synthetic Sampling Technique," *Int. J. Pavement Eng.*, Doi: 10.1080/10298436.2024.2365957.

[20] S. J. Hall, "Determining The Effective Pg Grade Of Binder In Rap Mixtures," Master's Thesis, University Of New Hampshire, Durham, New Hampshire, Usa, 2009. Accessed: May 15, 2026. [Online]. Available: Https://Scholars.Unh.Edu/Thesis/531