Proceedings of 15th IOE Graduate Conference
Peer Reviewed
Year: 2024 Month: May Volume: 15
ISSN: 2350-8914 (Online), 2350-8906 (Print)

Evaluating Multiple Linear Regression Prediction Model for Optimum Bitumen Content in Marshall Mix Design

Moti Ram Giri [a], Gautam Bir Singh Tamrakar [b]

[a, b] *Department of Civil Engineering, Pulchowk Campus, IOE, Tribhuvan University, Nepal*
[a] 078mstre007.moti@pcampus.edu.np, [b] gbstamrakar@gmail.com

Abstract

Over time, considerable efforts have been committed to enhancing the prediction of Optimum Bitumen Content in Marshall Mix Design to reduce the time and effort required in the conventional Marshall mix design procedure, employing a wide range of factors. Prediction of Optimum Bitumen content continues to be a difficult task due to interactions among numerous variables that are challenging to gather comprehensively. The present study explores the utilization of Multi Linear Regression (MLR) for predicting Optimum Bitumen Content (OBC) in Marshall Mix Design, crucial for optimizing asphalt pavement construction. After the collection of 148 Marshall mix design forms from various projects, the uniqueness of the data is maintained initially; then an outlier test is conducted for the output set, and 141 sets of data were taken for the descriptive statistics analysis to identify the mean, median, mode, and standard deviation of the dataset. After that, a multiple linear regression (MLR) model was developed using Microsoft Excel. For the developed MLR model, the training R-value and R-squared value are 0.849 and 0.7209, respectively, indicating moderate predictive capacity and a strong correlation between independent and dependent variables. Validation on an independent dataset confirms the model's reliability, with an R-value and R-squared value of 0.6734 and 0.4535, respectively and,Chi-Square test suggested that there is no difference between the Actual OBC and Predicted OBC values. ANOVA analysis underscores significant relationships between independent variables and OBC, supported by an F-value surpassing critical thresholds. Sensitivity analysis emphasizes the influential role of aggregate gradations in OBC prediction. By integrating MLR, this research introduces an innovative approach to streamline asphalt mix design processes, offering cost-effective and durable solutions for pavement construction. The findings advocate for widespread adoption of MLR in industry practices to enhance efficiency and resource optimization.

Keywords
Optimum Bitumen Content, Multi Linear Regression, Marshall Mix Design, Sensitivity Analysis

1. Introduction

1.1 Background

Asphalt concrete pavements find extensive application in advanced highways, runways, and parking lots, where the cost of bitumen plays a significant role in project economics. Knowledge of the optimal binder content is crucial for achieving higher Marshall stability values, ensuring superior performance of asphalt concrete (AC) paved surfaces [1]. Asphalt mix design aims to accurately estimate the ideal aggregate and bitumen characteristics within the mix to meet this objective [2]. Typically, this process involves conducting laboratory tests known as Marshall Mix design to determine the optimum binder content [3]. Introduced in 1939 by Bruce Marshall of the Mississippi State Highway Department [4], the Marshall mix design method remains widely used, especially in South Asia [5]. In Nepal, it was introduced under the Standard Specification of Road and Bridge Work (SSRBW) in 2058 BS.

The Marshall test procedure entails preparing at least fifteen samples for five different asphalt contents, followed by drawing design curves [6] to estimate the Optimum Bitumen Content (OBC) that meets specific criteria governed by SSRBW 2073[3]. However, this method requires considerable time for sample preparation and testing[7], prompting research efforts to explore alternative approaches for time-saving Marshall tests. Therefore, this study aims to evaluate the efficacy of employing Multi Linear Regression (MLR) to accurately and efficiently predict OBC, thereby enhancing laboratory mix design processes.

2. Literature Review

Several studies have investigated diverse methodologies and considered various parameters to forecast the optimum bitumen content in Marshall Mix Design and Marshall properties in asphalt mixtures. Typically, aggregates constitute approximately 95% of the asphalt mix and are deemed the most crucial element of asphalt concrete[8], thus influencing the mix's characteristics primarily reliant on the aggregate used and its gradation[9].Aggregate properties significantly impact mix properties and Hot Mix Asphalt (HMA) Aggregate properties significantly impact mix properties and Hot Mix Asphalt (HMA) [7].

Saltan et al. and Androjic & Dimter employed the Marshall mix design method to ascertain the optimum bitumen content by evaluating mixtures with varying bitumen contents [10, 11]. Setiawan et al. utilized Multiple Polynomial Regression (MPR) models to establish relationships between aggregate gradation, bitumen content, and Marshall properties [12]. Khuntia et al. employed artificial neural networks (ANNs) to predict optimum bitumen content and Marshall parameters efficiently, minimizing the need for extensive experimental tests [13]. Moreover, Awan et al., Androjić & Marović, and Othman & Abdelwabah explored the use of ANNs for predicting Marshall properties in asphalt mixtures [14, 9, 15]. Baldo et al. successfully forecasted stiffness property of asphalt concretes using machine learning models, with bitumen content as a crucial variable [16]. Additionally, Baldo et al. demonstrated the application of ANNs to model Marshall parameters of hot mix asphalts [17]. These methods have paved the way for effective and precise techniques to anticipate bitumen content in Marshall Mix Design, employing a range of prediction models like MLR, ANN, MPR, and others. This has significantly contributed the progress of asphalt engineering practices.

3. Objectives of the Research

The aim of this study is to construct a multiple linear regression (MLR) model to forecast the Optimum Bitumen Content by analyzing the interrelationships among Specific Gravities of materials, Aggregate Gradations, and Practical Consideration for Aggregate across four bins (20 Down, 16 Down, 10 Down, and 4.75 Down) of Marshall Mix Design. Additionally, the research seeks to assess the precision and reliability of the MLR model.

4. Methodology

4.1 Data Collection and Extraction

This study relies primarily on secondary data obtained from various sources. A total of 148 data samples were gathered meticulously from multiple laboratories, including the Quality Research and Design Center (QRDC) under the Department of Roads (DOR) Nepal, Visow Lab Kathmandu, Everest Lab Kathmandu, and Meh Geo Lab Lalitpur. Additionally, extensive review of theses related to Marshall Mix design at the university level in Nepal was conducted. Secondary data were also sourced from documents obtained from project offices and various construction sites.

The collected data encompassed details such as batching proportions, gradations details, specific gravity of materials, mechanical properties, and volumetric properties at optimum bitumen content. Furthermore, information on mechanical and volumetric properties at five different bitumen contents was compiled in a structured format. Each Marshall mix design was validated for its uniqueness using project names and contract IDs associated with the respective projects.

4.2 Model Development

Regression analysis is a widely employed technique in research for examining the relationships between variables [18].Its popularity stems from the ease of interpretation and construction of regression models. Various types of regression models exist, including linear, nonlinear, simple, multiple, parametric, non parametric, and logistic regression [19]. In this study, the multiple linear regression (MLR) model is utilized, represented by the equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \tag{1}$$

Where:

$$\begin{aligned} y &: \text{Dependent variable} \\ x_1, x_2, \ldots, x_n &: \text{Independent variables} \\ \beta_0, \beta_1, \beta_2, \ldots, \beta_n &: \text{Regression coefficients} \\ \varepsilon &: \text{Residual error} \end{aligned}$$

The effectiveness of the models was assessed by calculating several statistical error metrics. These statistical errors included for the assessment of the models were Correlation Coefficient R, Coefficient of Determination R-Square, Adjusted R-Square, mean absolute error (MAE) and Mean Absolute percentage Error (MAPE).

4.3 Data Sampling for sensitivity Analysis

An extensive search was made within all the inputs to choose the most effective variables with the highest impact on the outputs. In a study involving multiple inputs, sensitivity analysis can be considered as one of the rational tools to determine the most important and least important parameters. Sensitivity analysis helps to ascertain the sensitivity of the output based on the change in the corresponding input values or input ranges of values. There are various methods and tools in order to carry out the sensitivity analysis for multiple variables as inputs, one of which is the sensitivity analysis proposed by Chang and Liao [20]. Based on Chang and Liao, the sensitivity index can be ascertained with the help of

$$SI = \frac{(O_2 - O_1)}{(I_2 - I_1)} \cdot \frac{IAVG}{OAVG} \tag{2}$$

Where:

$$\begin{aligned} I_1 &= \text{Smallest Input Value} \\ O_1 &= \text{Output Corresponding to Smallest Input Value} \\ I_2 &= \text{Largest Input Value} \\ O_2 &= \text{Output Corresponding to Largest Input Value} \\ IAVG &= \text{Average of all non-zero inputs} \\ OAVG &= \text{Average of outputs corresponding to} \\ &\quad\ \text{non-zero inputs} \end{aligned}$$

5. Data Processing and Descriptive Statistic of the Dataset

5.1 Data Processing

Initially, a total of 148 Marshall Mix design data points were collected, as shown in the provided picture. Subsequently, outlier detection was performed using a 95% confidence level, resulting in the removal of outliers and leaving 141 Marshall Mix designs for further analysis.

5.2 Descriptive Statistics of the dataset

Descriptive statistics offer a concise summary of the characteristics and distribution of values within one or more datasets [21], as illustrated in the following Table 1. Classical descriptive statistics enable analysts to quickly assess the central tendency and variability of values in datasets. Common measures of central tendency include the mean, median, and mode, each representing different typical values within the data. Measures of dispersion, also known as variability, encompass the minimum and maximum values, range, quantiles, and standard deviation,variance, distribution skewness, and kurtosis [22].

6. Result and Discussion

6.1 Correlation Analysis

In this study, correlation analysis utilizing Pearson's correlation coefficient reveals the relationships among various variables as shown in Table 2, including Specific Gravity of Aggregate and Bitumen, Percentage Down through Bins, and Percentage Passing. Strong positive correlations are observed among the eleven variables of percentage passing, while moderate positive to no correlations is found between percentage passing and Specific Gravity of Materials, as well as Percentage Down in Bins. Furthermore, Percentage Down through Bins shows a negative correlation with specific gravity of materials. Consequently, these independent variables collectively offer valuable insights for predicting the Optimum Bitumen Content in Marshall Mix Design.

6.2 Multiple Linear Regression and Statistical Significance

Multiple linear regression model is developed with 75% of total data set i.e. 106 number of data set (Training Dataset) to predict the Optimum Bitumen Content with four Percentage Down through Bins, Five Specific Gravity and Eleven Aggregate Gradation Percentage Passing independent variables. The multiple regression equation developed to predict OBC is shown in Equation 2. Also, Regression Statistics of MLR to Predict OBC is given in the Table 3.

$$\begin{aligned} OBC = & -3.4541 + 0.0010 \times PD_{19} + 0.0021 \times PD_{16} \\ & - 0.0002 \times PD_{13} - 0.0005 \times PD_{4.75} \\ & - 0.9050 \times SG_{20} + 1.8146 \times SG_{16} \\ & - 1.1295 \times SG_{10} + 1.1655 \times SG_{4.75} \\ & + 0.3884 \times SG_{B} + 0.0374 \times PP_{26.50} \\ & + 0.0245 \times PP_{19} - 0.0168 \times PP_{13.20} \\ & + 0.0055 \times PP_{9.50} + 0.0114 \times PP_{4.75} \\ & - 0.0051 \times PP_{2.36} - 0.0004 \times PP_{1.18} \\ & + 0.0237 \times PP_{0.60} - 0.0185 \times PP_{0.30} \\ & + 0.0323 \times PP_{0.15} - 0.0004 \times PP_{0.075} \end{aligned}$$

Regression statistics shows that the R-squared values for OBC Prediction model has moderate strength. Also, the Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) suggested more significant errors with 2.4169% as lower MAPE percentage suggests that is close to zero, on average, the predictions are very close to the actual values, indicating a high level of precision and reliability in the model's predictions [23].

Furthermore, Table 4 shows that the F observed value of 10.976 is greater than the F critical values of 1.695 and 1.50 at a 5% level of significance and at a 10% level of significance respectively. Hence, the dependent variable Optimum Bitumen Content is significantly related to the independent variables at a 5% level of significance.

The equation suggests that larger-sized aggregates lead to increased voids in the mix, thus requiring more optimum bitumen content (OBC). Conversely, smaller-sized aggregates offer greater surface area resulting higher OBC. The relationship between percentage passing and OBC involves both positive and negative correlations, which stem from their strong interdependence. This is largely because a well-graded aggregate mix is necessary. Additionally, the voids left by larger aggregates are filled by smaller aggregates during interlocking in the mix. Specifically, the percentage passing (PP) from the 0.075 sieve size has a negative relation to OBC. An increase in dust content limits the penetration of bitumen in the mix, resulting in a stiffer mix.

6.3 Model Validation

For MLR model testing purposes, 25% of independent dataset i.e.35 number of mixes was used. Figure 1 shows the comparison between the tested and predicted values of the OBC on an independent dataset that was used for evaluation purposes of the MLR model.

From Figure 2, it can be concluded that the correlation coefficient between the tested and predicted values of the OBC on the additional set of samples (35 independent samples) amounts to a low of 0.4535 compared to the tested set of samples. Removing 3 mixes from the basic dataset (rough error) leads to an increase of the correlation coefficient of 0.57 [15].

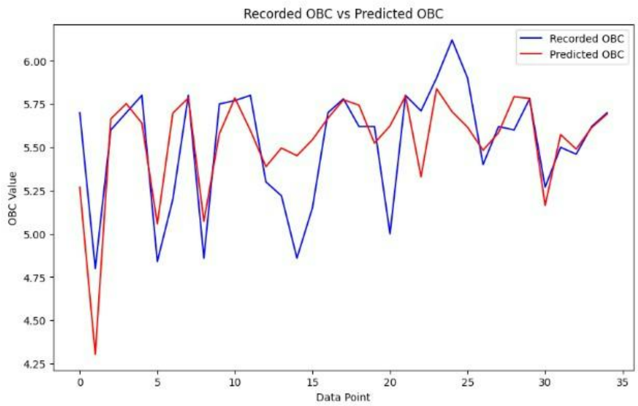

Figure 1: Comparison between the tested and predicted values of the OBC

Furthermore, The Chi-square test is then carried out in order to test the difference between the Actual OBC and the Predicted OBC values. The test is presented along with hypothesis testing and data frequencies are recorded for 0.3% OBC interval.Then, Hypothesis testing is carried out to perform this test in which the null hypothesis and the alternative hypothesis are defined.

Null hypothesis (H0): There is no significant difference between the actual OBC as expected frequencies and predicted OBC as observed frequencies.

Table 1: Descriptive Statistics of Input and Output Datasets

	Mean	Median	Mode	Standard Error	Standard Deviation	Minimum	Maximum
PD 19	17.08	20	0	1.32	15.68	0.00	51.00
PD 16	12.77	15	0	1.03	12.24	0.00	35.00
PD13	26.73	30	30	0.83	9.80	0.00	65.00
PD 4.75	42.37	43	45	0.58	6.83	15.00	60.00
SG20	2.65	2.655	2.627	0.00	0.04	2.56	2.78
SG16	2.64	2.643	2.627	0.00	0.06	2.26	2.76
SG10	2.64	2.64	2.68	0.01	0.06	2.26	2.76
SG4.75	2.66	2.664	2.686	0.01	0.06	2.32	2.84
SG B	1.03	1.028	1.028	0.00	0.01	1.01	1.09
PP 26.50	99.82	100	100	0.07	0.84	91.50	100.00
PP 19	97.00	100	100	0.43	5.08	75.30	100.00
PP 13.20	80.84	78.69	91.1	0.82	9.72	60.80	99.81
PP 9.50	69.60	70.9	81.97	1.11	13.24	39.10	88.83
PP 4.75	48.57	46.8	59.68	0.73	8.63	35.10	69.70
PP 2.36	37.35	34.79	45.11	0.54	6.46	24.10	54.10
PP 1.18	25.50	23.62	23.62	0.69	8.23	8.60	41.79
PP 0.60	21.52	19.94	19.1	0.51	6.03	8.60	34.90
PP 0.30	14.33	13.54	13.16	0.46	5.42	5.40	25.07
PP 0.15	9.64	8.8	7.68	0.25	2.92	4.60	15.94
PP 0.075	6.67	6.7	6.4	0.11	1.31	2.02	9.97
OBC	5.50	5.6	5.7	0.03	0.32	4.79	6.12

Table 2: Pearson's correlation coefficient among considered variables

Variables	**PD 19**	**PD 16**	**PD13**	**PD 4.75**	**SG20**	**SG16**	**SG10**	**SG4.75**	**SG B**	**PP 26.50**	**PP 19**	**PP 13.20**	**PP 9.50**	**PP 4.75**	**PP 2.36**	**PP 1.18**	**PP 0.60**	**PP 0.30**	**PP 0.15**	**PP 0.075**
PD 19	1.0																			
PD 16	-0.8	1.0																		
PD13	-0.5	0.1	1.0																	
PD 4.75	-0.2	0.0	-0.3	1.0																
SG20	0.1	-0.2	0.0	0.1	1.0															
SG16	-0.1	0.0	0.1	0.1	0.7	1.0														
SG10	-0.1	0.0	0.1	0.1	0.5	0.8	1.0													
SG4.75	0.1	-0.2	0.1	0.0	0.2	0.1	0.2	1.0												
SG B	0.3	-0.2	-0.1	0.0	0.2	0.1	0.0	0.1	1.0											
PP 26.50	-0.2	0.2	0.2	0.1	0.0	0.3	0.3	0.0	-0.2	1.0										
PP 19	-0.6	0.3	0.4	0.2	-0.1	0.2	0.3	0.0	-0.1	0.6	1.0									
PP 13.20	-0.5	0.2	0.3	0.4	-0.1	0.1	0.2	0.0	-0.2	0.3	0.6	1.0								
PP 9.50	-0.7	0.3	0.5	0.2	-0.2	0.1	0.2	0.0	-0.3	0.4	0.7	0.8	1.0							
PP 4.75	-0.5	0.1	0.3	0.5	0.0	0.1	0.2	0.0	-0.1	0.2	0.5	0.7	0.7	1.0						
PP 2.36	-0.5	0.2	0.2	0.5	-0.1	0.0	0.1	-0.1	-0.1	0.1	0.4	0.7	0.6	0.8	1.0					
PP 1.18	-0.4	0.1	0.1	0.5	-0.1	0.0	0.1	-0.1	-0.1	0.3	0.4	0.7	0.6	0.6	0.8	1.0				
PP 0.60	-0.5	0.1	0.2	0.5	-0.1	0.0	0.2	0.0	-0.1	0.3	0.5	0.7	0.6	0.7	0.8	0.9	1.0			
PP 0.30	-0.1	-0.1	-0.2	0.4	0.0	0.0	0.1	-0.1	0.0	0.1	0.1	0.5	0.2	0.5	0.6	0.8	0.8	1.0		
PP 0.15	-0.2	0.1	0.0	0.4	0.0	0.1	0.2	-0.1	-0.1	0.2	0.3	0.7	0.4	0.5	0.7	0.8	0.8	0.8	1.0	
PP 0.075	-0.2	0.1	0.1	0.2	0.1	0.1	0.1	0.0	-0.3	0.1	0.2	0.2	0.2	0.3	0.2	0.1	0.2	0.1	0.4	1.0

Table 3: Regression Statistics of Prediction Model

R	0.849
R-Squared	0.7209
Adjusted R-Squared	0.6552
MSE	0.0333
MAPE	2.4169

Table 4: Analysis of Variance (ANOVA) of MLR to Predict OBC

	d.f.	SS	MS	F	p-value	Fc at 5% LOS	Remarks
Regression	20	7.3207	0.366	10.9761	0	1.695	F>Fc
Residual	85	2.8346	0.0333				
Total	105	10.1554					

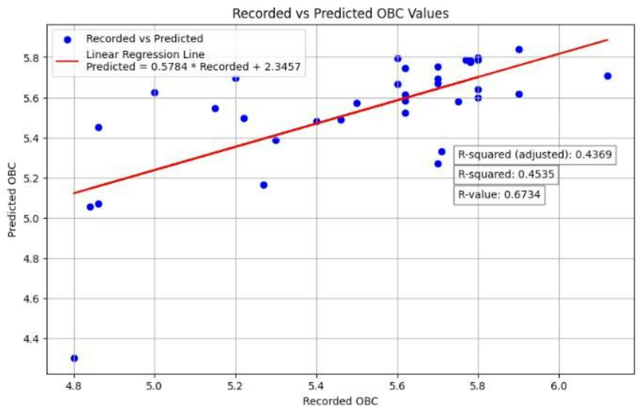

Figure 2: Correlation coefficient between the tested and predicted values of the OBC

Table 5: Chi square test Calculation Table

OBC interval	Observed Frequency (O)	Expected Frequency (E)	(O-E)2/E
4.8-5.1	2	3	0.33
5.1-5.4	4	5	0.20
5.4-5.7	16	9	5.44
5.7-6.0	10	15	1.67
Sum	32	32	7.64

Alternative Hypothesis (H1): There is significant difference between the expected frequencies and the observed frequencies.

The level of significance is set as 0.05.

Chi Square calculated = 7.64 Degree of freedom = 4-1 =3 Chi-Square critical value for Degree of freedom 3 and significance level 5% =7.815

The Chi-Square calculated is smaller then the critical value of Chi-Square suggests there is no significant difference between the observed values and the expected values and therefore, the null hypothesis is accepted. Thus, we can say that, the OBC obtained from MLR model can be used as prediction of OBC in Marshall Mix design method.

6.4 Sensitivity Analysis

The sensitivity analysis is carried out for Optimum Bitumen Content (OBC) datasets in order to determine the sensitivity index. From the results, it can be seen that the most sensitive parameter are aggregate gradations which brings about the most change in the OBC. The sensitivity Index calculation of various parameters of OBC are shown in Table 5 and Figure 3 Graphical Representation of Sensitivity Analysis respectively.

7. A new approach for the Marshall test

The normal design procedures that require the preparation and testing of 15 mix samples, which is time consuming, the MLR can be employed for estimating the OBC if accuracy of the model can be improved in future time, then only three specimens are prepared and tested to estimate the design parameters and make sure they match the design criteria. This approach saves time, resources, and the required effort to estimate the OAC.

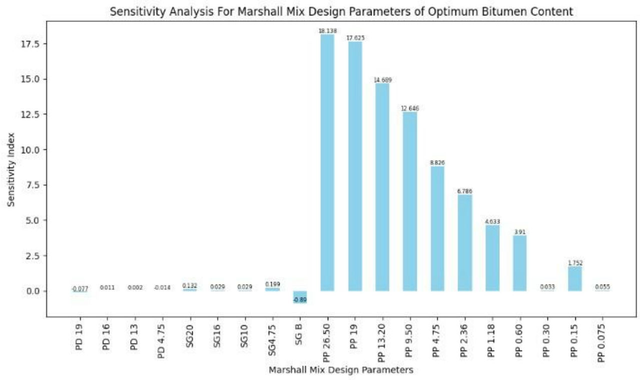

Figure 3: Graphical Representation of Sensitivity Analysis

8. Conclusion

In conclusion, this study assessed the effectiveness of Multi Linear Regression (MLR) for predicting Optimum Bitumen Content (OBC) in Marshall Mix Design, aiming to streamline asphalt mix design processes. The MLR model showed moderate strength, with low Mean Absolute Percentage Error (MAPE), indicating close alignment between predicted and actual values. Sensitivity analysis revealed the significant influence of aggregate gradations on OBC. The study suggests a novel approach using MLR to potentially reduce time and resources required for OBC estimation compared to traditional methods. However, further refinement of the MLR model is warranted to enhance its accuracy. Overall, this research contributes to advancing asphalt engineering practices by offering insights into efficient and accurate OBC prediction, potentially leading to more durable and cost-effective asphalt pavements.

9. Recommendation

The reliability and accuracy of the model can be improved through the following recommendations:

I. Expanding the modeling approach to incorporate Artificial Neural Networks (ANN) can enhance reliability, as machine learning techniques excel in capturing non-linear relationships between variables.

II. Future research endeavors should broaden the scope by including additional variables such as aggregate water absorption, aggregate flakiness, angularity properties, and bitumen viscosity, alongside factors like gradation and specific gravity of materials. These enhancements can provide a more comprehensive understanding of the factors influencing Marshall Mix design, thus leading to more accurate predictive models.

10. Abbreviations

MLR Multiple Linear Regression,
OBC Optimum Bitumen Content,
PD Percentage Down,
PP Percentage Passing,
SG Specific Gravity

Table 6: Sensitivity Analysis of OBC parameters

	IAVG	I1	I2	OAVG	O1	O2	SI
PD 19	27.4	4.7	51.0	5.4	5.5	4.8	-0.1
PD 16	22.5	5.0	35.0	5.6	5.5	5.6	0.0
PD13	27.3	5.0	65.0	5.5	5.2	5.2	0.0
PD 4.75	42.4	15.0	60.0	5.5	5.7	5.6	0.0
SG20	2.7	2.6	2.8	5.5	4.8	4.9	0.1
SG16	2.6	2.3	2.8	5.5	4.8	4.9	0.0
SG10	2.6	2.3	2.8	5.5	4.8	4.9	0.0
SG4.75	2.7	2.3	2.8	5.5	5.5	5.7	0.2
SG B	1.0	1.0	1.1	5.5	5.5	5.1	-0.9
PP 26.50	99.8	91.5	100.0	5.5	91.5	100.0	18.1
PP 19	97.0	75.3	100.0	5.5	75.3	100.0	17.6
PP 13.20	80.8	60.8	99.8	5.5	60.8	99.8	14.7
PP 9.50	69.6	39.1	88.8	5.5	39.1	88.8	12.6
PP 4.75	48.6	35.1	69.7	5.5	35.1	69.7	8.8
PP 2.36	37.4	24.1	54.1	5.5	24.1	54.1	6.8
PP 1.18	25.5	8.6	41.8	5.5	8.6	41.8	4.6
PP 0.60	21.5	8.6	34.9	5.5	8.6	34.9	3.9
PP 0.30	14.3	5.4	25.1	5.5	5.4	5.7	0.0
PP 0.15	9.6	4.6	15.9	5.5	4.6	15.9	1.8
PP 0.075	6.7	2.0	10.0	5.5	5.3	5.6	0.1

Acknowledgments

The authors are thankful to Er. Kshitiz Dhakal and Er. Neeva Dahal for assisting in collection of data. The authors are equally thankful to all individuals who supported directly and indirectly throughout the completion of this research.

References

[1] Kareem Othman. Prediction of the hot asphalt mix properties using deep neural networks. *Beni-Suef University Journal of Basic and Applied Sciences*, 11(1):40, 2022.

[2] JL Feiteira Dias, LG Picado-Santos, and SD Capitão. Mechanical performance of dry process fine crumb rubber asphalt mixtures placed on the portuguese road network. *Construction and Building Materials*, 73:247–254, 2014.

[3] Serkan Tapkın, Abdulkadir Çevik, and Ün Uşar. Prediction of marshall test results for polypropylene modified dense bituminous mixtures using neural networks. *Expert Systems with Applications*, 37(6):4660–4670, 2010.

[4] Asphalt Institute. *Mix design methods for asphalt concrete and other hot-mix types*. Number 2. Asphalt Institute, 1998.

[5] Hande Isik Ozturk, Ahmet Saglik, Birol Demir, and AG Gungor. An artificial neural network base prediction model and sensitivity analysis for marshall mix design. In *Proceedings of the 6th Eurasphalt & Eurobitume Congress, Prague, Czech Republic*, 2016.

[6] Ahmad H Aljassar, Mohammed Ahmed Ali, and Aref Alzaabi. Modeling marshall test results for optimum asphalt-concrete mix design. *Kuwait J Sci Engrg*, 29(1):181–195, 2002.

[7] Nicola Baldo, Evangelos Manthos, Marco Pasetto, et al. Analysis of the mechanical behaviour of asphalt concretes using artificial neural networks. *Advances in Civil Engineering*, 2018, 2018.

[8] Ebrahim Sangsefidi, Hasan Ziari, and Maryam Sangsefidi. The effect of aggregate gradation limits consideration on performance properties and mixture design parameters of hot mix asphalt. *KSCE Journal of Civil Engineering*, 20:385–392, 2016.

[9] Kareem Mohamed Mousa Othman and Hassan Abdelwahab. Prediction of the optimum asphalt content using artificial neural networks. *Metallurgical and Materials Engineering*, 27(2):227–242, 2021.

[10] Mehmet Saltan, Betül Öksüz, and Volkan Emre Uz. Use of glass waste as mineral filler in hot mix asphalt. *Science and Engineering of Composite Materials*, 22(3):271–277, 2015.

[11] Ivica Androjić and Sanja Dimter. Properties of hot mix asphalt with substituted waste glass. *Materials and structures*, 49:249–259, 2016.

[12] Arief Setiawan, Latif Budi Suparma, and Agus Taufik Mulyono. Modelling effect of aggregate gradation and bitumen content on marshall properties of asphalt concrete. *International Journal on Advanced Science, Engineering and Information Technology*, 7(2):359–365, 2017.

[13] Sunil Khuntia, Aditya Kumar Das, Monika Mohanty, and Mahabir Panda. Prediction of marshall parameters of modified bituminous mixtures using artificial intelligence techniques. *International Journal of Transportation Science and Technology*, 3(3):211–227, 2014.

[14] Hamad Hassan Awan, Arshad Hussain, Muhammad Faisal Javed, Yanjun Qiu, Raid Alrowais, Abdeliazim Mustafa Mohamed, Dina Fathi, and Abdullah Mossa Alzahrani. Predicting marshall flow and marshall stability of asphalt pavements using multi expression programming. *Buildings*, 12(3):314, 2022.

[15] Ivica Androjić and Ivan Marović. Development of artificial neural network and multiple linear regression models in the prediction process of the hot mix asphalt properties. *Canadian journal of civil engineering*, 44(12):994–1004, 2017.

[16] Nicola Baldo, Matteo Miani, Fabio Rondinella, Jan Valentin, Pavla Vackcová, and Evangelos Manthos. Stiffness data of high-modulus asphalt concretes for road

pavements: predictive modeling by machine-learning. *Coatings*, 12(1):54, 2022.

[17] Nicola Baldo, Evangelos Manthos, and Matteo Miani. Stiffness modulus and marshall parameters of hot mix asphalts: Laboratory data modeling by artificial neural networks characterized by cross-validation. *Applied Sciences*, 9(17):3502, 2019.

[18] Omer F Cansiz, Mustafa Calisici, and M Melik Miroglu. Use of artificial neural network to estimate number of persons fatally injured in motor vehicle accidents. In *Proceedings of the 3rd International Conference on Applied Mathematics, Simulation, Modelling, Circuits, Systems and Signals*, pages 136–142, 2009.

[19] Omer F Cansiz and Said M Easa. Using artificial neural network to predict collisions on horizontal tangents of 3d two-lane highways. *International Journal of Civil and Environmental Engineering*, 5(10):491–500, 2011.

[20] Chia-Ling Chang and Chung-Sheng Liao. Parameter sensitivity analysis of artificial neural network for predicting water turbidity. *International Journal of Geological and Environmental Engineering*, 6(10):657–660, 2012.

[21] Julie MacLeavy. School of geographical sciences, university of bristol, university road, bristol, bs8 1ss, uk forthcoming in the international encyclopedia of human geography, elsevier, edited by audrey kobayashi et al. e-mail: Julie. macleavy@ bristol. ac. uk telephone:+ 44 (0) 117 928 8901. *International Encyclopedia of Human Geography, Elsevier*, 2020.

[22] Venkat N Gudivada. Data analytics: fundamentals. In *Data analytics for intelligent transportation systems*, pages 31–67. Elsevier, 2017.

[23] Davide Chicco, Matthijs J Warrens, and Giuseppe Jurman. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *Peerj computer science*, 7:e623, 2021.