PREDICTING MARSHAL STABILITY AND MARSHAL FLOW USNIG ARTIFICIAL INTELLIGENCE: A COMPARISON OF GENE EXPRESSION PROGRAMING AND ARTIFICIAL NEURAL NETWORK

NASIR KHAN

(00000318805)

A thesis submitted in partial fulfilment of

the requirements for the degree

Master of Science
in
Transportation Engineering

MILITARY COLLEGE OF ENGINEERING (MCE)
RISALPUR
DEPARTMENT OF CIVIL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY
(NUST)
SECTOR H-12, ISLAMABAD,
PAKISTAN. (2022)

THESIS ACCEPTANCE CERTIFICATE

It is certified that final copy of MS thesis written by **ENGR. NASIR KHAN Reg. No. 00000318805** of **Military College of Engineering**, The National University of Sciences & Technology, Risalpur Campus, has been witnessed by the undersigned and is accepted as partial fulfilment for the award of an **MS degree in Transportation Engineering.** It has been found to be complete in all respects and in accordance with NUST Statutes and Regulations. It is also validated that the scholar's thesis has been updated with the essential changes that were identified by the GEC members.

Signature: ______________________________

Name of Supervisor: **Brig Dr Wasim Irshad ul Haq Kayani**

Date:__________________________________

Signature **(HOD)**: _________________________

Date: __________________________________

Signature **(Dean/Principal)**: __________________

Date: _____________________________________

PREDICTING MARSHAL STABILITY AND MARSHAL FLOW USNIG ARTIFICIAL INTELLIGENCE: A COMPARISON OF GENE EXPRESSION PROGRAMING AND ARTIFICIAL NEURAL NETWORK

by

NASIR KHAN
(00000318805)

A Thesis
of
Master of Science

Submitted to

Department of Civil Engineering
Military College of Engineering (MCE) Risalpur
National University of Sciences and Technology (NUST)
Islamabad

In partial fulfillment of the requirements for the degree of
Master of Sciences Transportation Engineering
2022

DEDICATION

This work is the outcome of many challenging sacrifices. This work is gratefully and proudly dedicated to those who have inspired it. From parents and guardians to classmates and acquaintances who offered assistance when encountered difficulties while undertaking this work.

To the academic and administrative employees of the Military College of Engineering, Rislapur. Above all, we give thanks to God Almighty, who has abundantly blessed us in every aspect of our life, especially for giving us the strength, bravery, patience, wisdom, time, and direction to complete this endeavor.

ACKNOWLEDGEMENT

I am grateful to Almighty Allah, who empowered me to complete this research and extend much appreciation and gratitude to my advisor, Brig Dr. Wasim Irshad ul Haq Kayani, whose countless inspirations and guidance made it possible to complete my research work. He always gave me advice and feedback throughout the thesis process. I would like to pay gratitude to the academic members of the Military College of Engineering who provided a lot of knowledge during the academic session in the postgraduate program.

Finally, I want to sincerely thank my parents and, in particular, my loving brothers for their support, prayers, and well wishes for the successful completion of both my academic career and this research project. I want to thank my family from all my heart for always being with me and for motivating me at every point of my life. Without them, nothing would have been possible.

(*Engr. Nasir Khan)*

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

AI	Artificial Intelligence
GEP	Gene Expression Programming
ANN	Artificial Neural Network
ETs	Expression Trees
MS	Marshal Stability
MF	Marshal Flow
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
Ps%	Aggregate Percentage
Pb%	Asphalt Content Percentage
Pba	Absorbed Asphalt Content
Pbe	Effective Asphalt Content
Gmb	Bulk Specific Gravity of Aggregates.
Gmm	Maximum Specific gravity of Mix.
Gse	Aggregate Effective Specific Gravity.
Va	Air Voids
VMA	Voids in Mineral Aggregates
VFA	Voids Filled by Bitumen

ABSTRACT:

The type and proportion of materials being used in the construction of a highway facility, as well as many other criteria, influence its comfort, ride quality, and service life. While constructing the bituminous layers of a highway facility, the type and composition of mixes must be carefully considered. These layers must be built with certain care since they are directly impacted by the applied load and environmental conditions. Properties of these layers are affected by number of factors which are described in the form of Marshal Quotient of hot mix asphalt. Calculating the marshal quotient leads the project to an uneconomical as calculating this parameter is based on trails and error and requires skilled labor and extensive time for calculation. A computer-based model has been developed using a data set composed of 110 lab experiments collected from a construction firm working on Jehagira to Risalpur road (Khyber Pakhtunkhwa, Pakistan) that can predict the values of marshal stability and marshal flow. The collected data was first screened and all the inappropriate data points were removed. Prior to modeling, insignificant variables were removed to generate a better model. Models were developed using GEP and ANN for both Marshal Stability and Marshal Flow using seven input variables. The performance of the models developed has been validated using coefficient of determination, RMSE, MAE and Adjusted R^2. Results shows that the GEP model performs better than ANN and has more better predicting power than ANN. Performance of the developed model was also validated using unseen data collected from N95 Swat (Khyber Pakhtunkhwa, Pakistan) where the models performed significant and were able to predict the output quiet accurately. A sensitivity analysis has also been performed to access the relative contribution of every variable in predicting the outputs. It was also concluded that the marshal stability increases with the increase in air voids and reduction in bitumen content while marshal flow increases upon increase in bitumen content and decrease in air voids.

CHAPTER 1

INTRODUCTION

The effectiveness of the mix design process and the quality of the asphalt concrete have a significant impact on the quality of a highway facility. Experts across the world use a variety of methodologies, including Superpave mix design, Marshall mix design, Hveem mix design, Hubbard field Asphalt Institute Triaxial and others. The Superpave, Marshall, and Hveem mix design approaches are the most popular ones. [1]. All these methods require intense care and enough time to perform certain set of laboratory test to get to the desired results. This process is repeated several times to get an optimum value for the parameters of asphalt mix design.

In Pakistan, Marshall Mix design and modified Marshall Mix design have been widely used. Bruce Marshall, an engineer in the highway department of Mississippi developed the Marshall Mix design concept in 1939. USA Corps of engineers modified this method in 1948 while the American Society for Testing and Materials (ASTM) standardized it in 1958 under the designation ASTM D1559 [2, 3]. This procedure was divided into ASTM D6926 and ASTM D6927. The former was framed to mold the asphalt concrete while the later was framed to assess the marshal stability and flow of asphalt [4]. Stability is defined as the amount of load sustained by a specimen in marshal apparatus before failure while flow is the total deformation recorded in the sample at the time of failure or at the time of recording stability under specified conditions of temperature. These two properties governs many properties of asphalt mix, therefore [5] claims that mix with high flow value have a greater tendency to deform as compare to a mix with lower flow value and a mix with high stability and low flow will behave as a brittle material and may crack under load.

1.1. Research Problem Statement:

Marshall Mix Design is time consuming and requires skilled operators to handle equipment, and there is not any mathematical relationship for parameters of Marshall Mix Design to predict the values of Marshall stability and flow, so therefore researchers use different Artificial Intelligence (AI) techniques to predict the Marshall stability and flow which will be the output parameters of this study.

It is therefore aimed to develop a computer-based model to predict the parameters (Stability and Flow) for Marshal Mix Design. This equation will be used as replacement of the Marshal Test. Different statistical tools like calculating Correlation Coefficient, Mean Absolute Error (MAE) Root Mean Squared Error (RMSE), Coefficient of determination (R^2) and Adjusted Coefficient of determination $\bar{R}^2$ will assess the significance of the developed model.

1.2. Literature Review:

The performance and strength of highways is represented in terms of stability of asphalt concrete. A facility with asphalt concrete of low stability may experience various distresses like rutting, creep etc. [6, 7]. Stability of asphalt concrete is effected by, traffic conditions, climate conditions, viscosity of bitumen and softening point of bitumen etc. [8]. Therefore, selection of type and optimum bitumen content is very important to construct a highway that is more resistive to pavement distress. The experimental procedure for determining the optimum asphalt content is a time-consuming procedure and requires intense care. Therefore, skilled workers are required to determine the Marshal Stability and Flow values by following the procedure of marshal test. These values are then plugged into various equations to calculate the values for voids in mineral aggregate, theoretical specific gravity, voids filled with asphalt and air voids and specific gravity of mixture. Therefore, if we are able to develop a model for prediction of flow and stability of asphalt, rest of the values can be calculated by following certain mathematical steps [9].

In this regard, Gene expression programming (GEP) can be a very convenient way to model the outcomes of Marshall Test procedure and help engineers to find stability and flow without carrying out destructive tests. The major rationale for adopting the GEP model is because it generates prediction equations without assuming the prior nature of the underlying connection. As a result, when compared to other Machine Learning (ML) approaches such as neural networks, the GEP model might give more insight into important linkages while still being capable of modelling complicated non-linear relationships. In addition, practitioners can easily use the generated functional relationship. GEP is an advance version of genetic programming (GP) which has been widely adopted by the civil engineering researchers [10].

Gene expression programming was suggested as a mechanism for developing a viable solution for forecasting [11]. GEP is a specialized kind of genetic programming (GP) that may also be characterized to as a sort of genetic algorithm since it is fundamentally comprised of a collection of mathematical solutions that eventually progresses the selection of good solution through an optimization procedure. Individuals in a genetic algorithm are computer programmes in a GP, which was initially proposed by [12], These computer programmes are evolved by GP using expression trees and a fitness criteria.

The first step of GEP is to select a function set (+, -, *, /) and a terminal set (X_1, X_2, 1, 2). It then proceeds towards loading data set to the model to investigate the fitness function created and generate an arbitrary population of chromosomes. For every chromosome, an expression tree is created to validate the fitness criteria. After evaluation, one program is selected and is replaced with the whole set of population. The same procedure is iterated until a best program is generated. [13].

Various researchers have adopted the GEP technique to predict different parameter in civil engineering. [13] Adopted GEP to predict the pavement roughness using a hybrid GEP-ANN model. [14] Evaluated the behavior of aggregate angularity on permanent deformation of asphalt mixture. [15] Assessed the rutting resistance of in service middle asphalt layer using GEP. [16] generated a prediction model to assess the rutting depth of asphalt mixture using GEP and ANN, where GEP outperforms the ANN model. [17] also did a comparative study of both GEP and ANN for the prediction of atmospheric temperature in Tabuk, Saudi Arabia. GEP has several advantages over other artificial inelegance technique. The first advantage is that GEP is not a black box because its outcomes are straightforward mathematical equations. Another feature of GEP is its capacity to obtain exact connections without taking previous patterns of existing associations into account. Unlike ANN, GEP does not have a problem of over training and it generates a model in each case which can be adopted by the practitioners for future use [16].

1.3. Research Aim and Objectives:

A lot of research work is available to calculate the parameters of Mix Design experimentally but there is a need of generating a mathematical model to predict the values of marshal mix design without performing laboratory testing. This research is aiming to predict the Marshal Stability and Flow values for a data set collected from a highway project in Risalpur, Pakistan. Keeping these considerations in mind, the following goals have been developed:

- Using the GEP techniques, creating a prediction model for forecasting the values of Marshal Stability and Flow. It is aimed to predict these values based on input parameters like Aggregate percentage (Ps), Binder content (Pb), Air voids (Va), Maximum Specific gravity of Mix (Gmm), Bulk specific gravity of mix (Gmb), Voids in mineral aggregates (VMA) and Voids filled with asphalt (VFA) etc.
- Developing a prediction model using Artificial Neural Network with the same set of conditions as adopted for developing a GEP model. Separate models will be developed for both Marshal Stability and Flow.
- To evaluate the validity of both GEP and ANN model by means of the validation set data by utilizing various statistical tools like R^2, RMSE, MAE, and Adjusted R^2.
- Comparing the prediction power of both GEP and ANN and to make recommendation about both the tools.

1.4. Research Methodology

1.4.1. Data Collection

Experimental data from Khattak Allied Construction Company has been collected for a highway project from Rislapur to JehanGira (Nowshehra) Pakistan. The data set is consisting of 110 data points, which will be screened before going into modeling phase.

1.4.2. Data Screening

- Extracting the data from the collected data sheets and formatting into an easy-to-read format.
- Eliminating those data sets where a data point is missing.
- Removing repeating data points if any.

1.4.3. Input Variable Selection

- To select only the significant variables for model development, a sensitivity analysis will be performed to check whether the variables are significant enough to be placed into a model. Only those variables which have more than 95% of significance will be considered as significant.

1.4.4. Model Development

- GEP Model: GEP model will be developed based on the significant variables selected in the previous step to predict the values for Marshal Stability and Flow. The data set will be divided into two groups. Trailing data set, composed of 80%

of the total data will be used to train the model and 20% will be used to check the validity of the generated model

- ANN Model: Techniques of artificial neural network will be adopted to develop NN model to predict the same output using the same set of conditions as adopted for GEP modeling. The performance of this model will be compared with GEP model developed in the previous step.

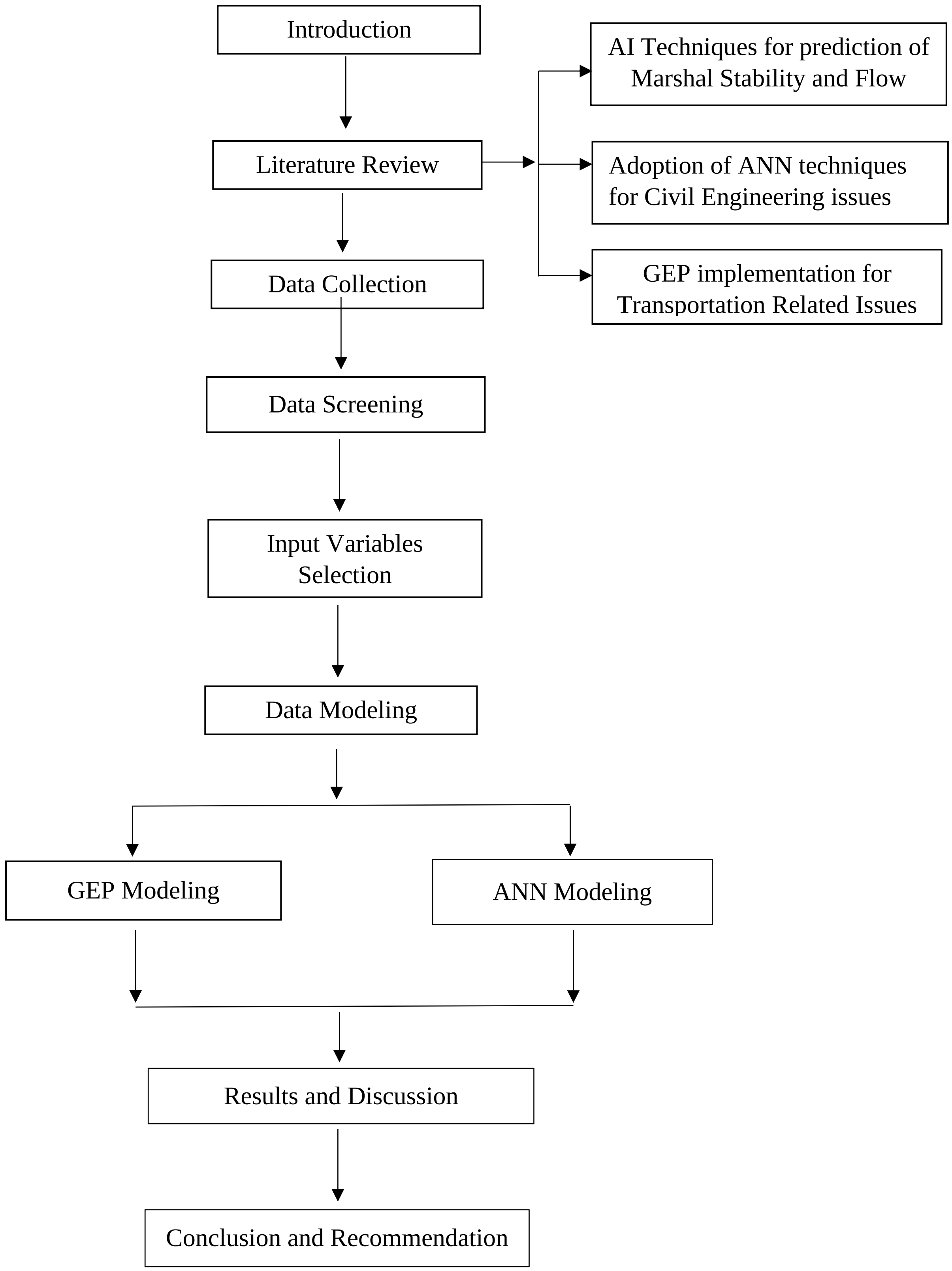
1.5. Results and Discussion

This section will present the results obtained at the end of this study. It is planned to develop a model using different artificial intelligence techniques for predicting marshal stability and marshal flow.

1.6. Conclusion and Recommendation.

Several recommendations will be made based in the conclusions of this study and will be lift open for future researchers to work on.

1.7. Flow Chart



CHAPTER 2

LITERATURE REVIEW

2.1. Application of Artificial Intelligence Techniques in Civil Engineering

Several transportation and pavement-related challenges have been solved using soft computing approaches throughout the last decade.[18] coupled Simulated Annealing and Genetic Programing to predict the flow number of asphalt mixture. The researcher generated a prediction model based on a data set composed of 118 test results of uniaxial dynamic creep tests with input variables as percentage of bitumen, percentages of filler, voids in mineral aggregate and marshal stability. A sensitivity analysis was also done to investigate the significance of every input variable. It was concluded that percentage of filler material is most significant variable among all. The model generated in this study was highly recommended to the practitioners due its accurate prediction capabilities. The accuracy of the generated model was assessed by measuring the RMSE and MSE. R^2 recorded for the testing data set was 0.979, which is why the generated model was highly recommended to the construction practitioners. [19] predicted the flow number of asphalt mixture by adopting Gene Expression Programing. The R^2 recorded for the model generated with 118 data points of uniaxial dynamic creep tests was 0.955. Genetic Programing was adopted by [20], to model the effect of filler material on the performance of the HMA. [14] Evaluated the behavior of aggregate angularity on permanent deformation of asphalt mixture. 98 samples used in this study were prepared in laboratory by the researchers with varying percentage of angular, sub angular, rounded and sub rounded aggregates. Aggregate Imaging System (AIMS) assessed the angularity of course aggregate. The generated GEP model showed R^2= 0.92 which was obtained after 377974 trails. The coefficient of determination for GEP model was close enough to that of Regression model (0.891) but the regression model overestimated permanent deformation for those samples which have higher actual permanent deformation.

[15] Assessed the rutting resistance of in-service middle asphalt layer using GEP by assessing the compound creep (CCR) of middle layer based on softening point of bitumen, proportion of rainy days, bitumen content, aggregate gradation, overloading rate, ESALs and mean temperature. 59 field core samples extracted from seven highways in Jiangsu China were preconditioned for four hours to achieve a constant temperature of 58°C which was termed as the worst high temperature condition for the middle layer in Jiangsu. The

developed model performed better and showed R^2= 0.912 for the testing data. A multi linear regression model was also developed to compare it with the GEP model and it showed R^2=0.862. [21-23] Adopted ANN and Fuzzy logic algorithm to predict the pavement roughness while [13] did the same job by using a hybrid GEP-ANN model based on data extracted from long-term pavement performance (LTPP) database of US. The generated GEP-ANN hybrid model showed R^2=0.9941 for validation data set. [24] Assessed the same parameters using ANN keeping in view the effect of temperature and exposure time. A total of 65 asphalt core samples collected from the department of state highway (D100-11) turkey were divided into 5 groups and were placed under different temperature with different exposure time where 88.66% decrease in marshal stability was observed after keeping the samples under 50 C° for six hrs. The researcher generated a prediction model using ANN based on five input variables. The significance of the generated model was assessed by calculating the R^2 value for the model. The generated model performed well and showed R^2= 0.933 for validation data set.

2.2. GEP vs. other AI Techniques

ANNs are mathematical models based on simulations of biological nerve systems. These methods might be used to solve complicated nonlinear models and supervised learning issues, but techniques like ANN, ANFIS and fuzzy logic algorithms have a drawback of not generating a mathematical equation which can be adopted by the practitioners [16]. Beside this [25] called ANN as a black box with a drawback of overfitting while on the other hand, GEP offers a more clear representation of the final model in the form of Expression Trees and Mathematical model. [16] generated a prediction model to assess the rutting depth of asphalt mixture using GEP and ANN based on 96 Hamburg Wheel Tracking Test Results, where GEP outperforms the ANN model. The researcher developed both GEP and ANN model based on 13 input variables. Both these models were validated by assessing the coefficient of determination. The value of R^2 in case of GEP model was recorded, as 0.93 while in case of ANN model it was recorded as equal to 0.84. The sensitivity analysis to measure the significance of input variables showed that percentage of asphalt binder was having the higher impact on the output variables as compared to other input variables. [17] Also compared both GEP and ANN by predicting the atmospheric temperature in Tabuk, Saudi Arabia based on the data of 30 years collected from Saudi Authority of Meteorology and Environmental Protection where GEP showed better results as compare to ANN. A total of 360 data sets were collected to develop these models in which 288 (80%) were used to train the model while the remaining 20% were used to test the model. The accuracy of the generated models was checked by calculating the co-efficient of determination and the root mean squared error for both the models by comparing the observed temperature and the predicted temperature. The results obtained in both the cases showed that the GEP model outperforms the ANN model, as the R^2 in case of GEP model equals 0.91 while R^2 in case of ANN model was 0.67. RMSE recorded for ANN model was 20.179 and for GEP model RMSE recorded was 0.44.

The multi-layer perceptron (MLP) is one of the most often used neural networks which is made up of three layers: an input layer (consisting of independent variables), a hidden layer (consisting of a number of hidden neurons, there can be one or more hidden layers), and an output layer (which contains the target values). These variables are linked together using weighted connections. The network's optimal solution is discovered by forward feeding the initial solutions, back-propagating the errors across the network, and modifying the weights of the connections [26].

The adaptive neuro-fuzzy inference system (ANFIS) is a Sugeno-type fuzzy inference system that also integrates neural network concepts [27]. The method of fuzzy inference entails modelling a collection of outputs from a set of inputs using particular membership functions, if-then rules and logical operations [28]. Any proposition in fuzzy logic is not totally true or untrue, so there is always a proportion of truthfulness or falseness. The complication of the membership functions and if-then rules that compose the final model is a disadvantage of ANFIS models.

2.3. Artificial Neural Network

ANNs is a form of artificial intelligence which is inspired by the human nervous system that tries to adopt the pattern of a particular activity after studying its behavior [29]. Such algorithm can be used to solve complex non-linear models. ANN is the collection of simple processing elements which are arranged into input, output and hidden layers (one or may be more) [17]. Researcher have adopted ANN for several application of civil engineering [30]. [31] adopted ANN for image processing and crack detection in pavement structures. [32] utilized ANN for soil classification using Liquid limit, clay content, plasticity index and water capacity as input parameters. ANN has been utilized by [33] as a pavement structure analysis tool to predict the defection profile and critical response of full depth pavement against heavy loads.

2.4. Gene Expression Programming

Gene expression programming was suggested as a mechanism for developing a viable solution for forecasting [11]. GEP is a specialized kind of genetic programming (GP) that may also be characterized to as a sort of genetic algorithm since it is fundamentally comprised of a collection of mathematical solutions that eventually progresses the selection of good solution through an optimization procedure. Individuals in a genetic algorithm are computer programmes in a GP, which was initially proposed by [12], These computer programmes are evolved by GP using a fitness criteria and expression trees.

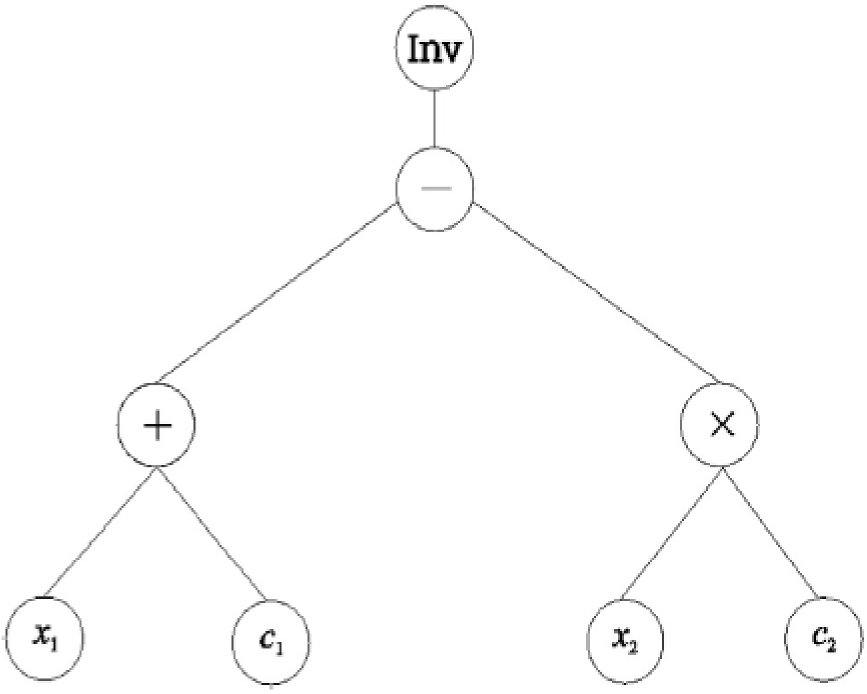


Figure 1: Representation of Expression Trees In GEP. [13]

Fig 1. Shows an example of expression that define the mathematical equation shown below. The head and intermediate nodes represent the mathematical functions while the tail nodes represent constants and variables. The mathematical equation represented in Fig. 1 as an expression tree is as follows

$$\frac{1}{(X_1 + C_1) - (X_2 x C_1)}$$

The first step of GEP is to select a function set (+, -, *, /) and a terminal set (X_1, X_2, 1, 2). It then proceeds towards loading data set to the model to investigate the fitness function created and generate an arbitrary population of chromosomes (i.e., computer programs). For every chromosome, an expression tree is created to validate the fitness criteria. After evaluation, one program is selected and is replaced with the whole set of population. The same procedure is iterated until a best program is generated. [13].

2.5. Marshal Stability and Flow

Stability is defined as the amount of load sustained by a specimen in marshal apparatus before failure while flow is the total deformation recorded in the sample at the time of failure or at the time of recording stability under the specified temperature conditions. These two properties governs many properties of asphalt mix, therefore [5] claims that mix with high flow value have a greater tendency to deform as compare to a mix with lower flow value and similarly a mix with high stability and low flow will behave as a brittle material and may crack under load.

The performance of the highway pavement is determined by the stability of the asphalt concrete. Asphalt concrete with low stability can cause number of problems with asphalt pavements [6, 7]. Rutting, Cracking, particularly fatigue cracking caused by repetitive loads, has been identified as a significant distress in asphalt concrete pavements. [34]. The stability of asphalt concrete pavements depends on viscosity of bitumen, softening point of bitumen, bitumen content, stiffness of the mix, construction practice, climate and traffic conditions and gradation of aggregate etc. [34].

[35] found that the stability of asphalt concrete reduces with environmental temperature while investigating the effect of bitumen rheology on low temperature behavior of bitumen in laboratory. Ductility and Viscosity of asphalt concrete increases with the increase of temperature. Experts [36-39] believes that the mixing and compaction temperature of the asphalt has a great influence on the performance of Hot Mix Asphalt. [40] Found that the environmental temperature and exposure time of asphalt concrete to a specific temperature increases the ductility of asphalt cement while investigating the effect of increase in exposure time and temperature on asphalt concrete.

CHAPTER 3

METHODOLOGY

3.1. Data Collection

The research, "Prediction of Marshal Stability and Flow using Gene Expression Programing" is based on a data set of 110 data points collected from "Khattak Allied Construction Company" working on "Rehabilitation & Improvement of Risalpur to Jehangira road (37 km)". The collected data was first extracted and compiled into excel sheets with all the input and output variables. Total input variables and statistics of the collected data are shown in table 1.

Table 1: Descriptive Statistics of Data

	Mean	Median	Maximum	Minimum	Range	Standard Deviation
P_s	96.379	96.403	96.891	95.731	1.16	0.282
P_b	3.619	3.598	4.269	3.109	1.16	0.284
P_{ba}	0.19	0.228	0.651	-0.112	0.763	0.124
P_{be}	3.444	3.4	4.904	2.85	2.054	0.342
G_{mb}	2.422	2.431	2.467	0.425	2.042	0.192
G_{mm}	2.56	2.552	2.608	2.521	0.087	0.025
G_{se}	2.714	2.719	2.794	2.578	0.216	0.03
V_a	4.706	4.713	6.084	3.765	2.319	0.506
VMA	13.3	13.118	23.973	11.931	12.042	1.163
VFA	64.52	65.312	70.662	55.07	15.592	3.604
Stability	3072.81	3176.15	3655.06	1430.1	2224.96	680.253
Flow	14.877	15.167	16.6	11	5.6	1.437

P_s = Aggregate Percentage

P_b = Asphalt Content Percentage

P_{ba} = Absorbed Asphalt Content

P_{be} = Effective Asphalt Content

G_{mb} = Bulk Specific Gravity of Aggregate.

G_{mm} = Maximum Specific Gravity of Mix

G_{se} = Aggregate Effective Specific Gravity

V_a = Air Voids

VMA = Voids in Mineral Aggregate

VFA = Voids Filled by Bitumen

Properties of the material used in this mix design are as follows:

Aggregate:

Source of Aggregate = Babuzai (Swat).

Table 2: Aggregates Properties

Name of Test	Actual Test Result	Project Specification
Los Angeles Abrasion, (%)	19.8	30 % Max
Soundness Loss by Sodium Sulphate, (%)	2.68	12 % Max
Flat and / or Elongated Particles (%)	7.70	15 % Max
Coating and stripping of bitumen - Aggregate Mixtures	Above 95 %	Above 95 %
Sand Equivalent, (%)	61.0	45 % Min
Plasticity Index	3.1	4 % Max

Bitumen:

Source of Bitumen = Pak-Arab Refinery Company Limited (PARCO)

Table 3: Bitumen Properties

Name of Test	Actual Test Result	Project Specification
Penetration Grade	66	60 ~ 70
Flash Point, °C	312	232 Min
Fire Point, °C	296	~
Specific Gravity	1.02	-

3.2. Data Screening

The data set is composed of 110 data points in which few data points were marked as inappropriate for the proposed research and were removed from the data set. Details of the data screening are given as follows.

Table 4: Data Screening

Total Data Points Collected	110
Inappropriate Data Points	8
Total remaining Data Points	102

The data set composed of 102 data points has been divided into training and testing samples with 80 and 20% of the total data respectively. [41] suggests that the data points should be at least 3 times more than the number of input variables. In this study, seven number of input variables and 102 data points will be used for model development which are sufficient as based on results produced by [41] at least 21 data points are required for this study.

3.3. Amputation of insignificant variables

After splitting the screened data set into training and testing sample with proportion of 80 and 20 percent respectively, a sensitivity analysis will be performed using Gene Expression Programming (GEP). A GEP model will be developed for both Marshal Stability and Marshall Flow which will be stretched further to run a sensitivity analysis to investigate the significance of input variables used to predict the output variables. All the variables with significance lesser than 5 percent will be eliminated from the list and the rest will be used for developing GEP and ANN model.

3.4. Modeling using GEP

After finalizing the number of input variables, a GEP model will be developed using GeneExpro5.0 Tool. The total data will be divided into 80 percent of training data and 20 percent testing data. The developed model will be deployed into an excel sheet to make it easy to use for the construction practitioners. To validate the performance of the developed model, R^2, MAE, RMSE and Adjusted R^2 will be used as validation criterion. This model will also be presented graphically in the form of expression trees to make it easier to understand.

3.5. Modeling using ANN

ANN model will be developed parallel with the GEP model using nntool (Matlab). The same set of data division and validation will be followed for ANN modeling as well. The best ANN model will be compared with GEP models generated by adopting Levenberg-Marquardt algorithm with five number of neurons in hidden layer.

3.6. Results Comparison

After successfully completing the modeling using both GEP and ANN, the results obtained in both the cases will be compared for Marshall Stability and Marshall Flow respectively. The validation criteria in both cases will define how precisely the generated model is able to predict the output variables.

3.7. Model Validation using New Data

As it is intended to develop the model using data from one project, then the main concern about its performance is whether this model will perform better in new environment or not. For answering this concern, the models developed will be validated using new set of data collected from another project with different conditions.

3.8. Sensitivity Analysis

After developing a prediction model for both marshal stability and flow, it is intended to analyze the contribution of every input variable on the output. Similar mechanism as mentioned in section 3.3 will be adopted to investigate this relation.

3.8. Effect Of Variation in Input Variables on the Output

A parametric study will be performed to access the effect of change in bitumen content, aggregate percentage and air voids on marshal stability and flow while keeping the other input variables constant. Different combinations (as shown in table) of bitumen content and air voids will be investigated in this section.

Table 5: Parametric Study Variable's Combination

Air Voids	Aggregate Percentage %	Bitumen Content %
2.5	97.0	3.0
	96.5	3.5
	96.0	4.0
	95.5	4.5
	95.0	5.0
	94.5	5.5

	94.0	6.0
Air Voids	Aggregate Percentage %	Bitumen Content %
3.0	97.0	3.0
	96.5	3.5
	96.0	4.0
	95.5	4.5
	95.0	5.0
	94.5	5.5
	94.0	6.0
3.5	97.0	3.0
	96.5	3.5
	96.0	4.0
	95.5	4.5
	95.0	5.0
	94.5	5.5
	94.0	6.0
4.0	97.0	3.0
	96.5	3.5
	96.0	4.0
	95.5	4.5
	95.0	5.0
	94.5	5.5
	94.0	6.0
4.5	97.0	3.0
	96.5	3.5
	96.0	4.0
	95.5	4.5
	95.0	5.0
	94.5	5.5
	94.0	6.0

CHAPTER 4

RESULTS

4.1. Variable Reduction

Every input variable used for model development has an effect on the output variable. This effect is sometime more significant while sometime it is less significant. The variable with high significance enhances the prediction ability of the model while on the other hand, insignificant variables effect the performance of the model in a negative way. To develop a model with more accurate prediction power, it is essential to remove the insignificant variables from the model.

In this regard, this research develops a GEP model to do a sensitivity analysis to assess the significance of every input variable in case of Marshal Stability and Marshal Flow respectively. For this purpose, after dividing the data into training and testing dataset, a GEP model was constructed with 10 input variables in both cases. Input variables used in this modeling are as follows:

P_s = Aggregate Percentage

P_b = Asphalt Content Percentage

P_{ba} = Absorbed Asphalt Content

P_{be} = Effective Asphalt Content

G_{mb} = Bulk Specific Gravity of Aggregate.

G_{mm} = Maximum Specific Gravity of Mix

G_{se} = Aggregate Effective Specific Gravity

V_a = Air Voids

VMA = Voids in Mineral Aggregate

VFA = Voids Filled by Bitumen

4.1.1 Marshal Stability

In this case, a GEP model was constructed with 10 above mentioned input variables using 82 (80%) data points for training the model and 20 (20%) data points to test the model for predicting the Marshal Stability. The result obtained are as follows.

Table 6: Marshal Stability Sensitivity Analysis

Marshal Stability Sensitivity Analysis		
	Training	Testing
Input Variables	10	10
Data Points	82 (80%)	20 (20%)
R^2	0.838	0.926
MAE	193.94	215.66
RMSE	256.87	255.8
Adjusted R^2	0.812	0.844

Sensitivity analysis for this model reveals that three input variables shows lesser significance than 5 percent and that is why they were marked as insignificant for the GEP and ANN modeling. The results of the sensitivity analysis are shown in table 6 and figure 2.

Table 7: Marshal Stability Input Variable's Effectiveness

Input Variable	Effectiveness (%)
Aggregate Percentage	16.67
Asphalt Content	16.67
Absorbed Asphalt Content	7.14
Effective Asphalt Content	14.29
Bulk Specific Gravity of Aggregate	4.76
Max. Specific Gravity of Mix	11.9
Aggregate Effective Specific Gravity	7.14
Air Voids	14.29
Voids In Mineral Aggregates	4.76
Voids filled by Bitumen	2.38

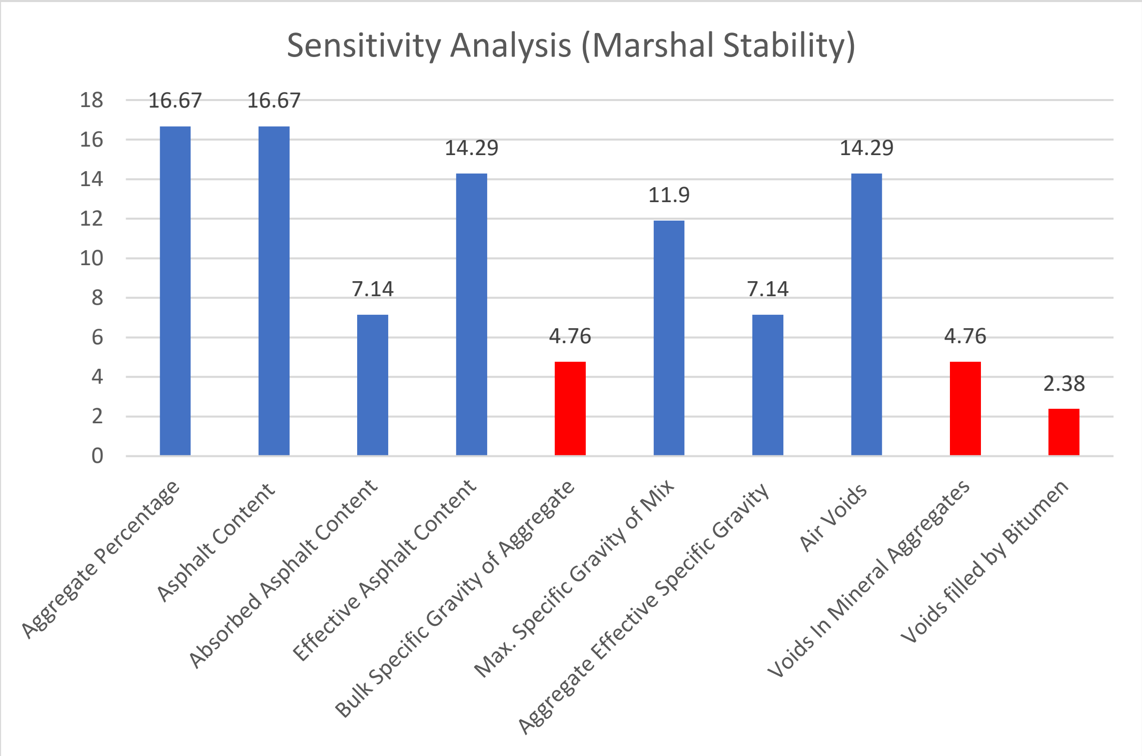


Figure 2: Marshal Stability Input Variable's Effectiveness

4.1.2. Marshal Flow

The same procedure as followed in case of Marshal Stability was adopted for Marshal Flow to reduce the number of input variables and remove the insignificant variables prior to the GEP and ANN modeling.

4.1.2.1 First Sensitivity Analysis (Flow)

Unlike the case of Marshal Stability, the sensitivity analysis results for Marshal Flow showed that only one input variable out of ten was not significant for modeling as only Effective Asphalt content showed significance of 2.38% which was lesser than the threshold mark of 5%. Therefore, it was decided to remove only the one insignificant variable and repeat the sensitivity analysis with nine input variables instead of ten. The results of sensitivity analysis for Marshal Flow are shown as follows.

Table 8: Marshal Flow Sensitivity Analysis (1st)

Marshal Flow First Sensitivity Using GEP		
	Training	Testing
Input Variables	10	10
Data Points	82 (80%)	20 (20%)
R^2	0.85	0.902
MAE	0.4185	0.448
RMSE	0.5524	0.6030
Adjusted R^2	0.828	0.793

Table 9: Marshal Flow Input Variable's Effectiveness (1st)

Input Variable	Effectiveness (%)
Aggregate Percentage	5.41
Asphalt Content	18.92
Absorbed Asphalt Content	5.41
Effective Asphalt Content	2.7
Bulk Specific Gravity of Aggregate	13.51
Max. Specific Gravity of Mix	16.22
Aggregate Effective Specific Gravity	10.81
Air Voids	5.41
Voids In Mineral Aggregates	13.5
Voids filled by Bitumen	8.11

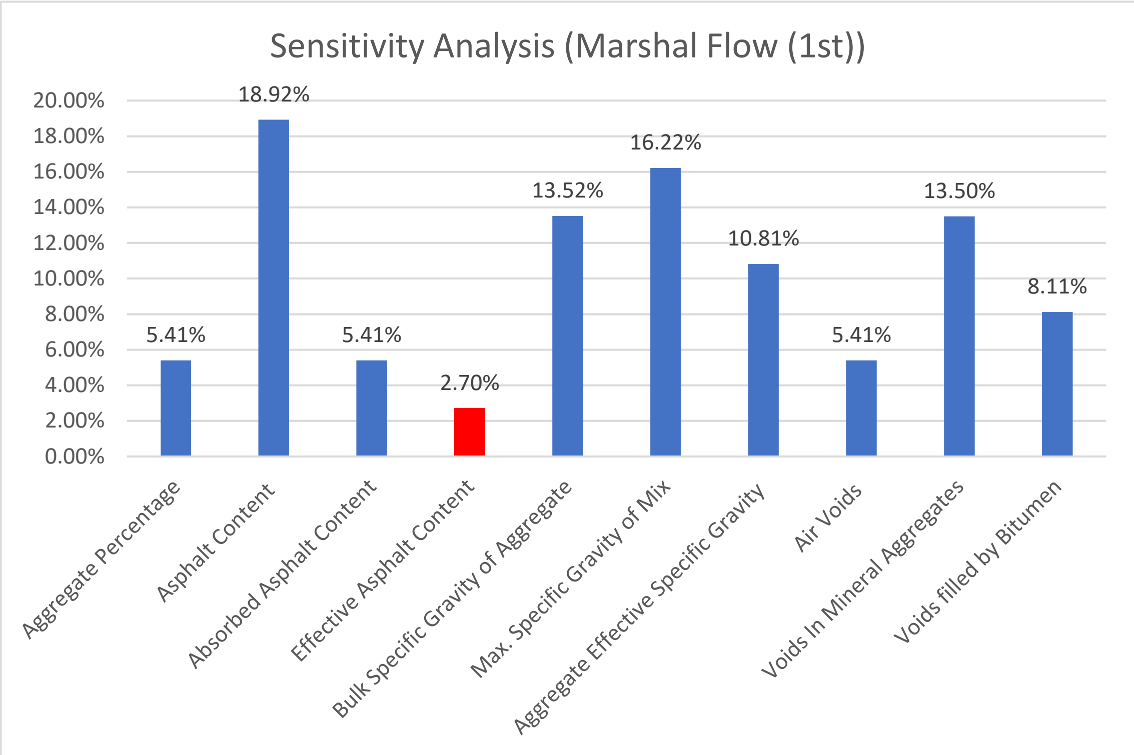


Figure 3: Marshal Flow Input Variable's Effectiveness (1st)

4.1.2.2. Second Sensitivity Analysis (Flow)

The second sensitivity analysis is a follow up of the first sensitivity analysis in case of Marshal Flow as in the first case, only one input variable was found to be insignificant. In this case, the insignificant variable was removed and the sensitivity analysis was performed again with nine variables. The results showed extremely different behavior as the variables like "Aggregate Percentage, Absorbed Asphalt Content and Air Voids" which were hardly crossing the threshold line in first case were found to be pretty significant this time. Max. Specific Gravity of Mix and Aggregate Effective Specific Gravity were found as insignificant variables this time and were removed from the data set prior to GEP and ANN modeling. The possible reason behind this behavior could be the multicollinearity of input variables. The results of second sensitivity analysis are shown below.

Table 10: Marshal Flow Sensitivity Analysis (2nd)

Marshal Flow Sensitivity Analysis (2nd)		
	Training	Testing
Input Variables	09	09
Data Points	82 (80%)	20 (20%)
R^2	0.883	0.876
MAE	0.390	0.525
RMSE	0.470	0.655
Adjusted R^2	0.868	0.752

Table 11: Marshal Flow Input Variable's Effectiveness (2nd)

Input Variable	Effectiveness
Aggregate Percentage	12.50%
Asphalt Content	12.50%
Absorbed Asphalt Content	8.33%
Bulk Specific Gravity of Aggregate	16.67%
Max. Specific Gravity of Mix	4.17%
Aggregate Effective Specific Gravity	4.17%
Air Voids	20.83%
Voids In Mineral Aggregates	8.33%
Voids filled by Bitumen	12.50%

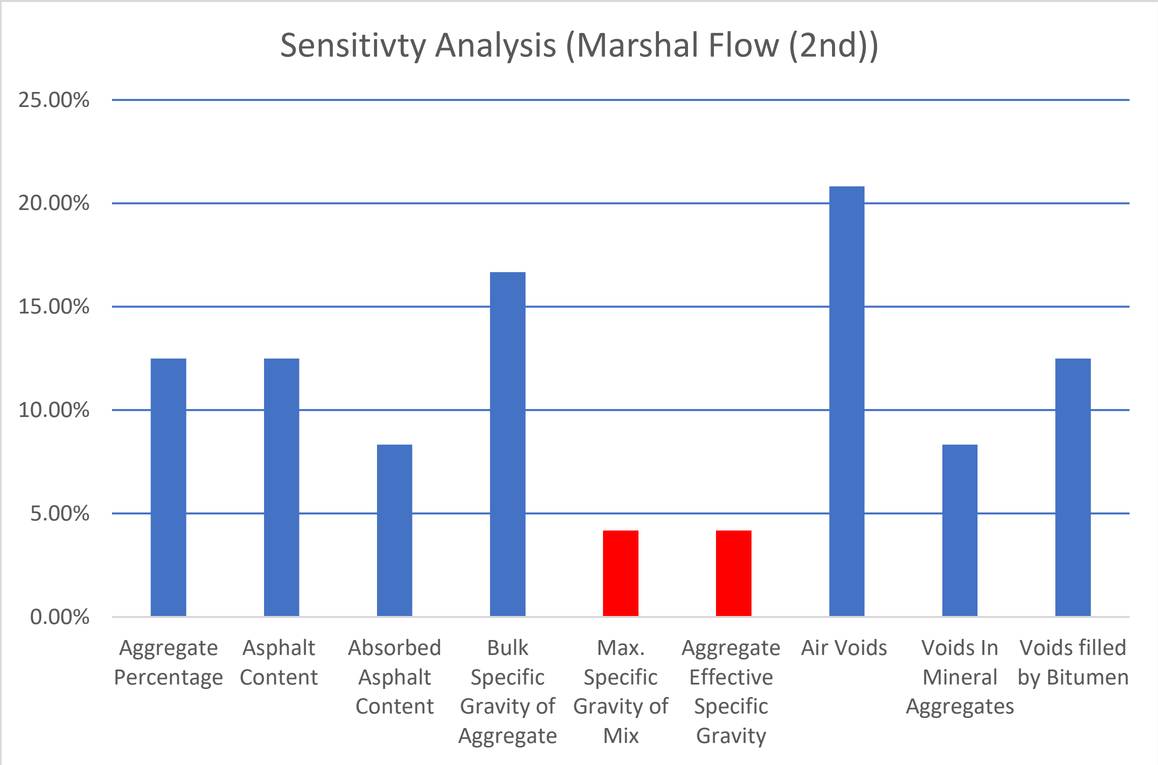


Figure 4: Marshal Flow Input Variable's Effectiveness (2nd)

4.2. GEP Modeling:

After removing the insignificant variables, the data set is now ready to be used for generating models using Gene Expression Programming (GEP). Marshal Stability and Marshall Flow models will be generated using their corresponding input variables.

4.2.1. Marshall Stability

Total number of input variables were reduced from ten to seven and were used to generate the model for Marshal Stability. List of input variables used in this case is given as follows:

1. P_s = Aggregate Percentage
2. P_b = Asphalt Content Percentage
3. P_{ba} = Absorbed Asphalt Content
4. P_{be} = Effective Asphalt Content
5. G_{mm} = Maximum Specific Gravity of Mix
6. G_{se} = Aggregate Effective Specific Gravity
7. V_a = Air Voids

Table 11. shows the parameters used for generating the Marshall Stability model in which the number of chromosomes are the number of programs that will be evolved by GEP during training, Number of genes governs the total number of sub expression trees,

complicity of expression trees is governed by the head size, and linkage function sets the relation of every sub expression tree with each other.

Table 12: Parameter setting for GEP Algorithm for Marshall Stability.

Parameter	**Setting**
Number of Chromosome	20
Number of Genes	5
Head Size	20
Linkage Function	Addition
Fitness Function Error Type	MSE
Function Set	+, -, /, *, $\sqrt{x}$, $\sqrt[3]{x}$, ^2, ^3

Total of 82 data points were used to train the model while 20 data points were used to test the functionality of the developed model. Equation 1. Shows the model developed while table 12 shows the model performance in terms of validation criteria and figure 6-10 shows the graphical representation of the model in the form of Expression Trees.

Marshal Stability= $\left((G_1C_1d_0 - 2d_0)x\left(G_1C_0d_0 + {d_0}^2\right) - G_1C_0d_0 + {d_0}^3x\, {G_1C_0}^{1/3} - d_0{d_3}^2\right)^{1/3} + Sin\left(\sqrt{d_2} + {d_3}^2 - e^{{d_3}^{1/3}}\right) - d_0 - \ln(d_2G_2C_1) - d_4 - d_0 - G_2{C_0}^2 - {d_0}^2 - G_2C_0xG_2C_1 + \left(\sqrt{G_3C_1} - \left(\frac{d_6{d_0}^3{d_3}^2}{G_3C_1 - d_6}\right)\right)^{1/3} x\left(d_3 - G_3{C_1}^2 - tan^{-1}d_4\right)^{1/3} + d_0 + \left(d_3{d_0}^2x\,(d_0 - d_6)x(d_3 - G_4C_0)x(lnd_4 + sind_1)^{1/3} + sind_1\right)^{1/2} - d_6 + d_5 - d_6 + G_5C_1d_6 + \left({d_1}^2{d_3}^2 + d_6\right)xd_2 + \frac{G_5C_1x{d_2}^{1/3}}{d_3d_2}$ - - - - - - - - - - - (1

Where:

G1C0 = 9.21

G1C1 = -6.43

G2C1 = 6.05

G2C0 = -8.54

G3C1 = 2.83

G4C0 = 0.049

G5C1 = -9.59

d0 = Aggregate Percentage

d1 = Asphalt Content Percentage

d2 = Absorbed Asphalt Content

d3 = Effective Asphalt Content

d4 = Maximum Specific Gravity of Mix

d5 = Aggregate Effective Specific Gravity

d6 = Air Voids

Table 13: Validation Criteria for Marshall Stability Model

Marshal Stability Model Using GEP		
	Training	**Testing**
Input Variables	7	7
Data Points	82(80%)	20 (20%)
R^2	0.890	0.942
MAE	148.2	131.2
RMSE	216.5	198.2
Adjusted R^2	0.880	0.908

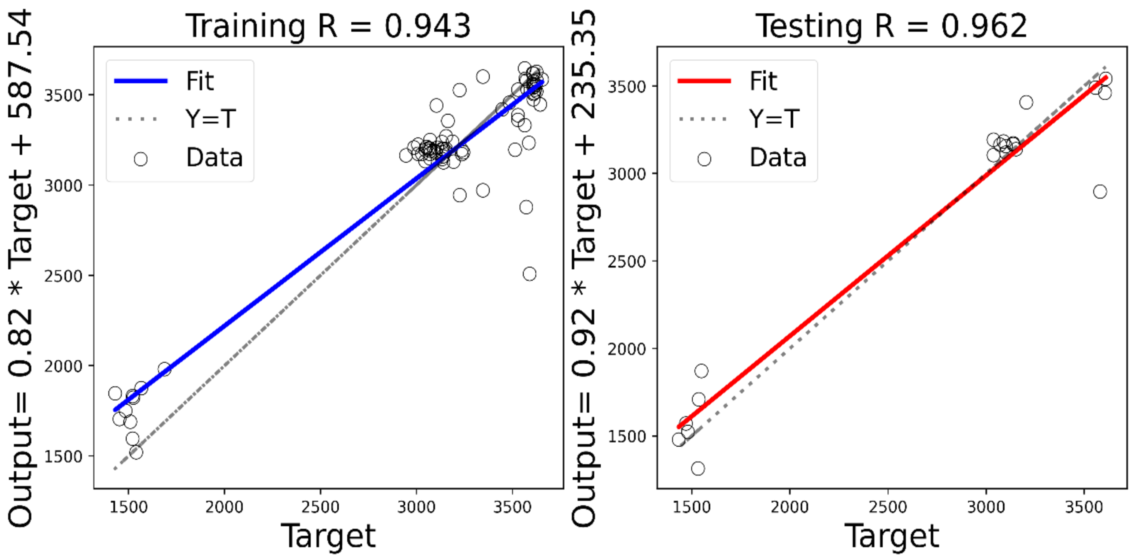


Figure 5: Regression plot for Marshal Stability Actual and Predicted values

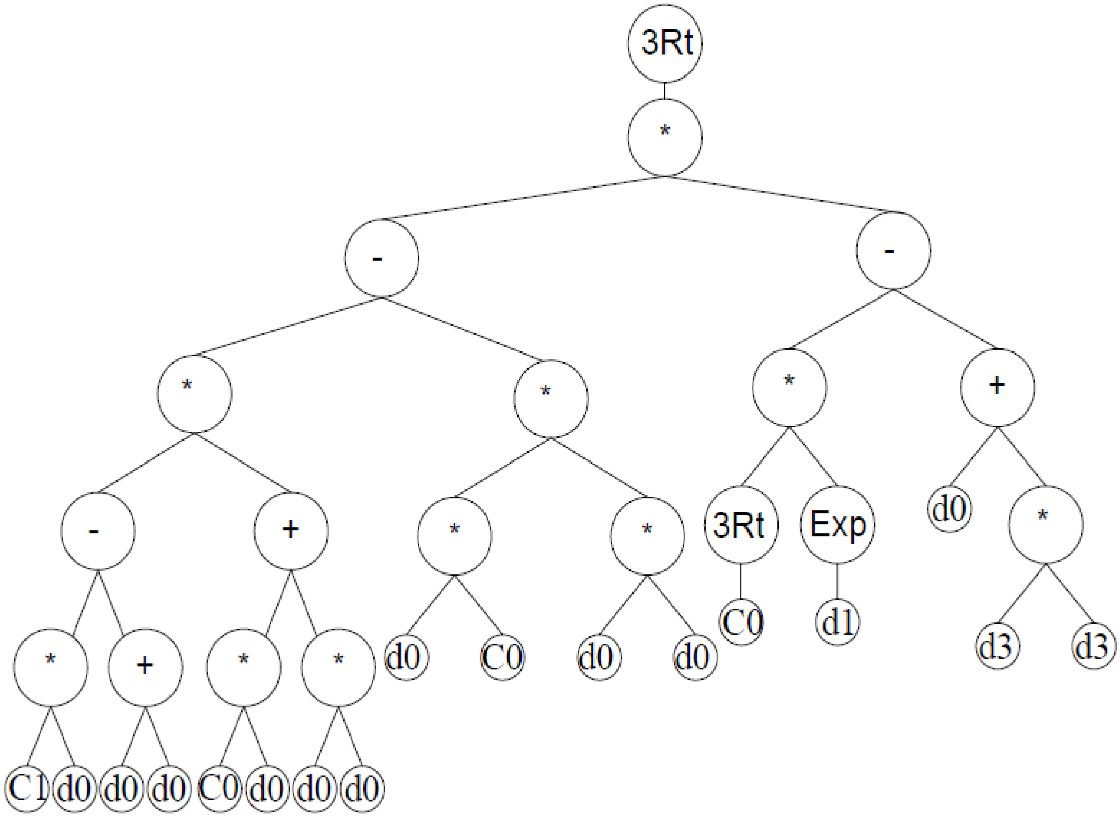


Figure 6: Marshal Stability Sub Expression Tree 1

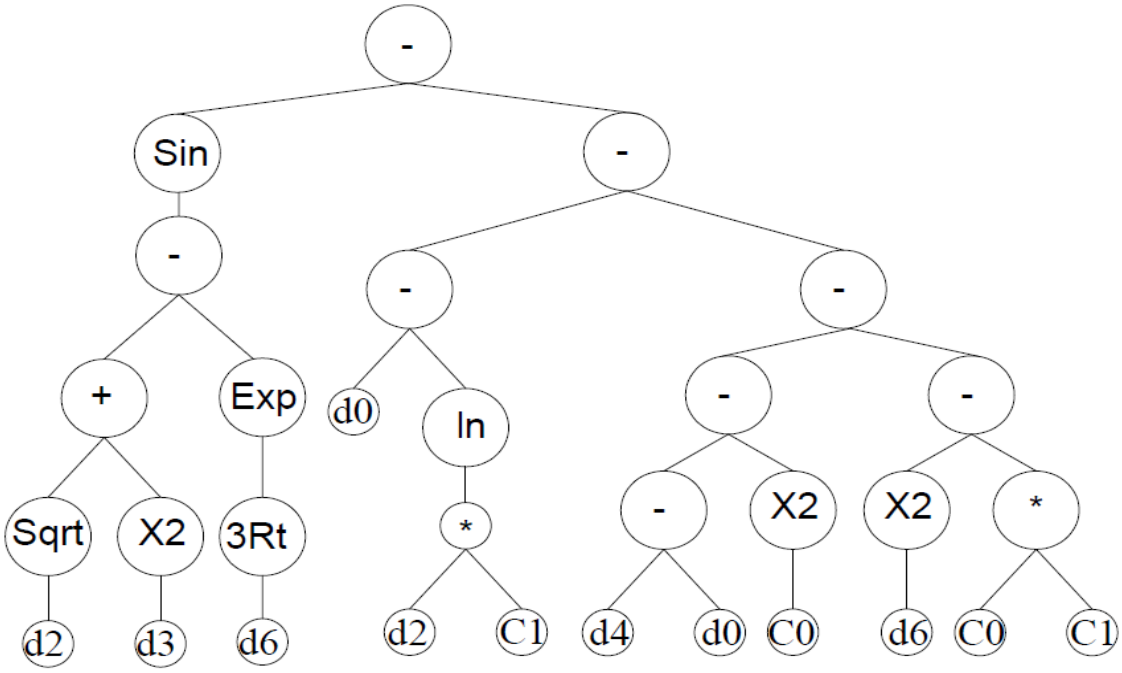


Figure 7: Marshal Stability Sub Expression Tree 2

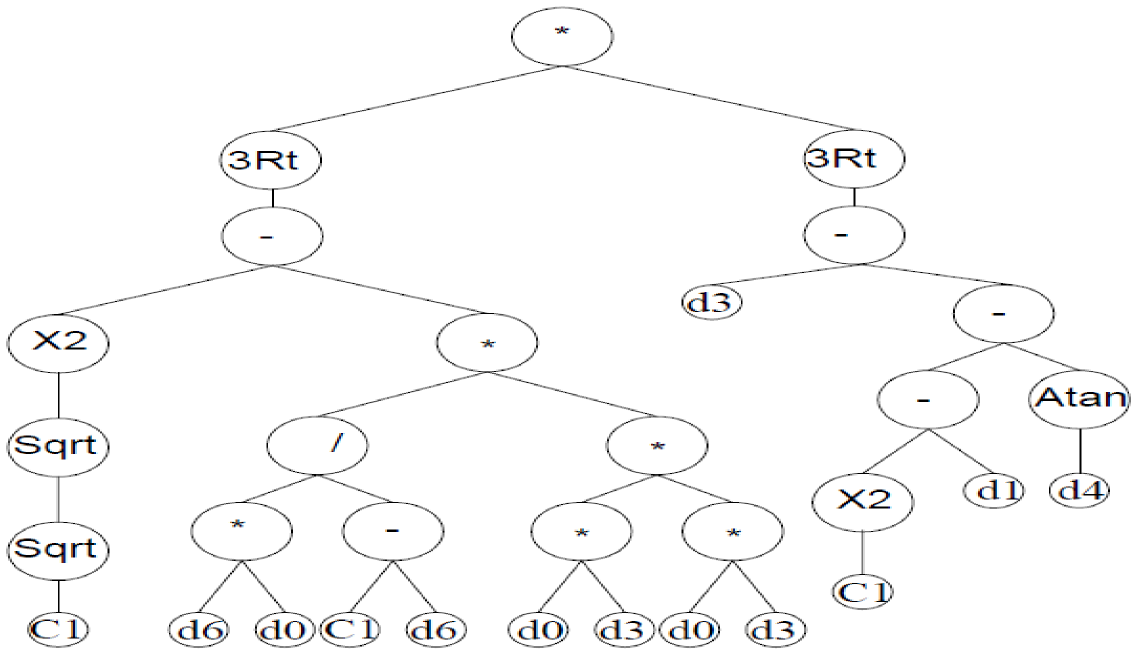


Figure 8: Marshal Stability Sub Expression Tree 3

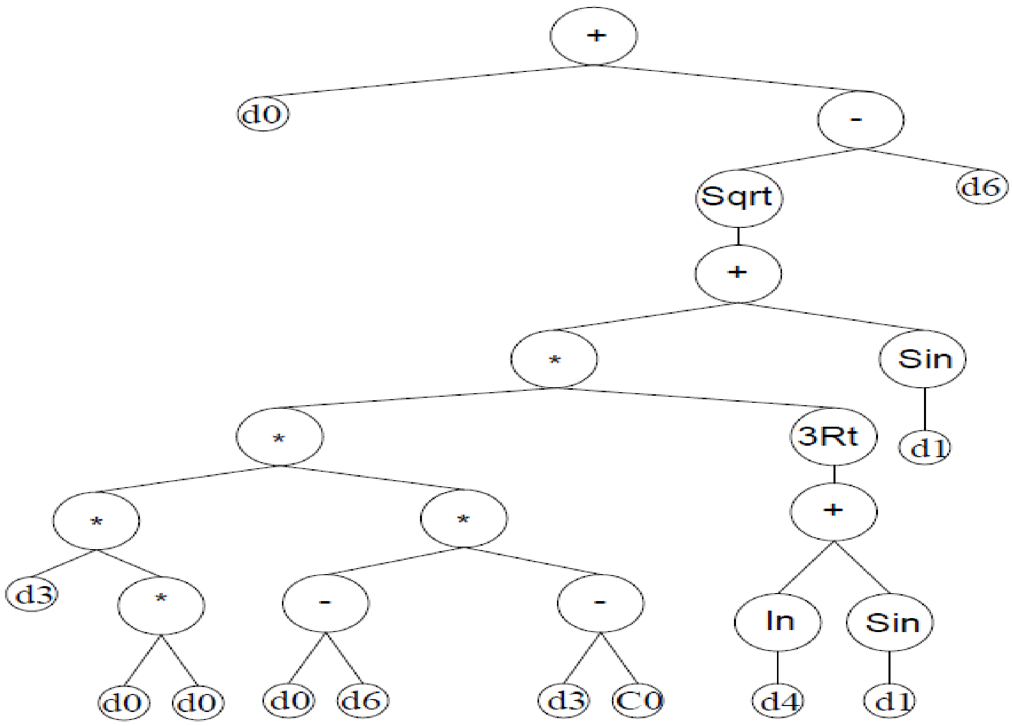


Figure 9: Marshal Stability Sub Expression Tree 4

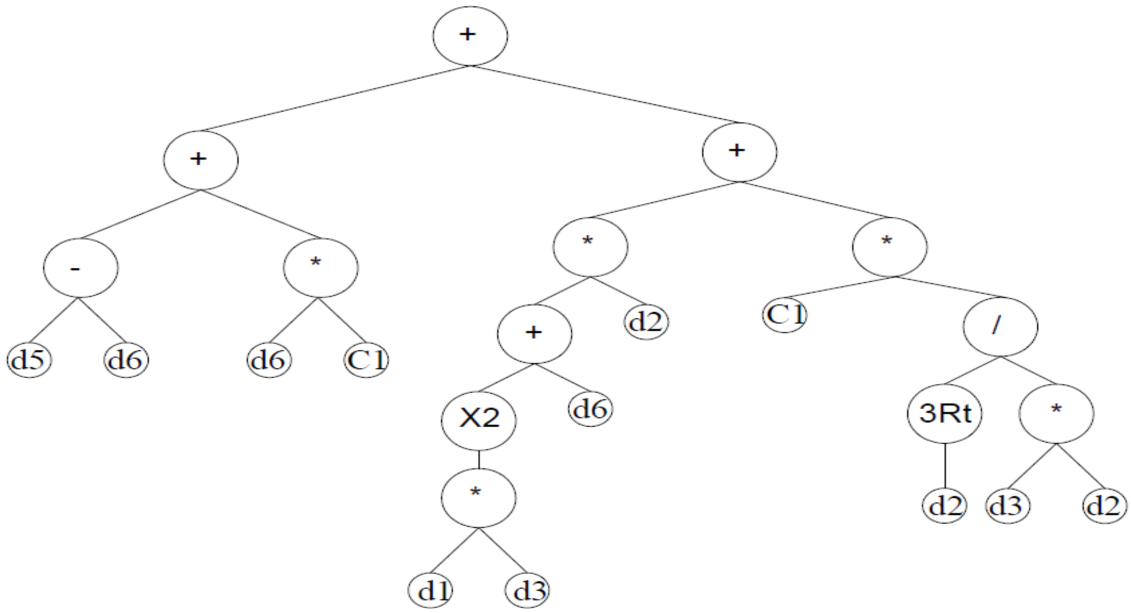
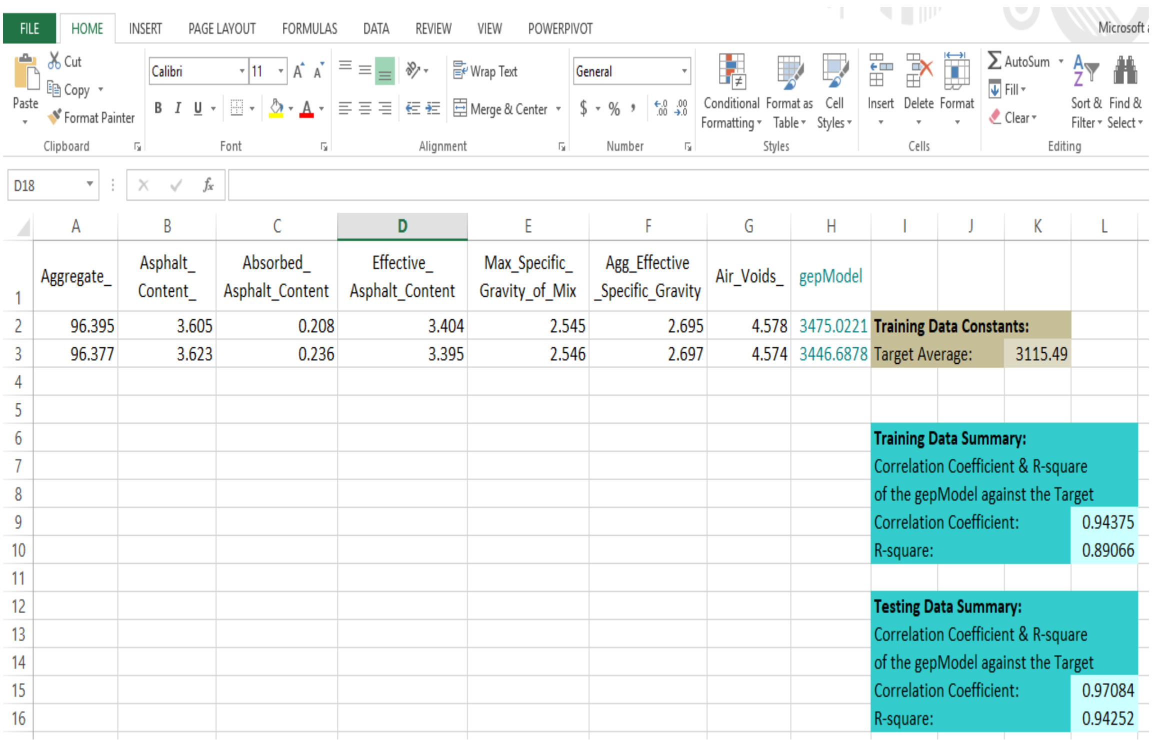


Figure 10: Marshal Stability Sub Expression Tree 5

Fig 6-10 presents the Marshal Stability model graphically in the form of expression trees which gives a much easier approach to the practitioner to understand the model. Expression trees should be read from left to right and from top to bottom. Beside this, GEP enables the practitioners to use the model deployed into an excel sheet where end user will simply plug in the input variables and the excel sheet will automatically predict the output values. Fig 11 shows a sample of the deployed Marshal Stability model.



	A	B	C	D	E	F	G	H	I	J	K	L
1	Aggregate_	Asphalt_ Content_	Absorbed_ Asphalt_Content	Effective_ Asphalt_Content	Max_Specific_ Gravity_of_Mix	Agg_Effective _Specific_Gravity	Air_Voids_	gepModel				
2	96.395	3.605	0.208	3.404	2.545	2.695	4.578	3475.0221	Training Data Constants:			
3	96.377	3.623	0.236	3.395	2.546	2.697	4.574	3446.6878	Target Average:		3115.49	
6									Training Data Summary:			
7									Correlation Coefficient & R-square			
8									of the gepModel against the Target			
9									Correlation Coefficient:			0.94375
10									R-square:			0.89066
12									Testing Data Summary:			
13									Correlation Coefficient & R-square			
14									of the gepModel against the Target			
15									Correlation Coefficient:			0.97084
16									R-square:			0.94252

Figure 11: Deployed Marshall Stability Model (GEP)

4.2.2. Marshall Flow

Total number of input variables were reduced from ten to seven and were used to generate the model for Marshal Flow. List of input variables used in this case is given as follows:

1. P_s = Aggregate Percentage
2. P_b = Asphalt Content Percentage
3. P_{ba} = Absorbed Asphalt Content
4. P_{be} = Bulk Specific Gravity of Aggregates
5. G_{mm} = Air Voids
6. VMA = Voids in Mineral Aggregate
7. VFA = Voids Filled by Bitumen

Table 14. shows the parameters used for generating the Marshall Stability model in which the number of chromosomes are the number of programs that will be evolved by GEP during training, Number of genes governs the total number of sub expression trees, complicity of expression trees is governed by the head size, and linkage function sets the relation of every sub expression tree with each other.

Table 14: Parameters setting for GEP Algorithm for Marshal Flow

Parameter	**Setting**
Number of Chromosome	20
Number of Genes	4
Head Size	20
Linkage Function	Addition
Fitness Function Error Type	MSE
Function Set	+, -, /, *, $\sqrt{x}$, $\sqrt[3]{x}$, ^2, ^3

Total of 82 data points were used to train the model while 20 data points were used to test the functionality of the developed model. Equation 2. Shows the model developed while table 14 shows the model performance in terms of validation criteria and figure 13-16 shows the graphical representation of the model in the form of Expression Tree.

$$\text{Marshal Flow} = \text{Cos (d5)} + \left(\left(\text{Cos}\sqrt[3]{e^{G_2C_1} x e^{d_5} - \left(d_1 x d_2 x {d_6}^2\right) + 2d_0{d_4}^2}\right) x d_5 + G_2C_0 + d_0\right)^{1/3} + Cos(d_1G_3C_0 - Cos(G_3C_0)x\left(d_2x(d_3 - d_4)\right)^2 + \frac{G_3C_0}{d_5} - Cosd_4 + d_0xG_3C_1 + \left(Sin\left(G_4C_0 - e^{\left(Cosd_6 + G_4C_1 + d_2\frac{d_2}{d_0}xG_4C_0\right)} x\ e^{\sqrt{G_4C_0}/d_1d_1}\right)^2 - d_0\right) - G_4C_1 \text{ - - - - - - - - - - - - (2}$$

Where:

G2C0 = -15.88

G2C1 = -1.85

G3C0 = -3.87

G3C1 = 0.80

G4C1 = -9.58

G4C0 = 10.61

d0 = Aggregate Percentage

d1 = Asphalt Content

d2 = Absorbed Asphalt Content

d3 = Bulk Specific Gravity

d4 = Air Voids

d5 = Voids in Mineral Aggregates

d6 = Voids Filled by Bitumen

Table 15: Validation Criteria for Marshall Flow Model

Marshal Flow Model Using GEP		
	Training	**Testing**
Input Variables	7	7
Data Points	82	20
R^2	0.891	0.895
MAE	0.389	0.448
RMSE	0.463	0.607
Adjusted R^2	0.881	0.834

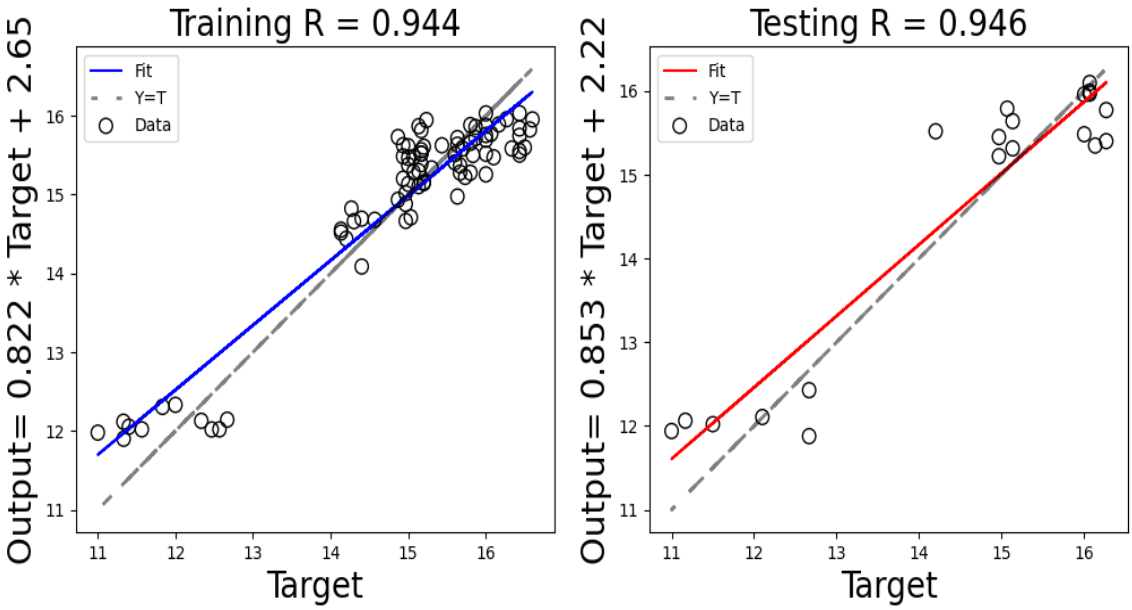


Figure 12: Regression plot for Marshal Flow Actual and Predicted values

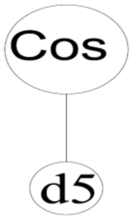


Figure 13: Marshal Flow Sub Expression Tree 1

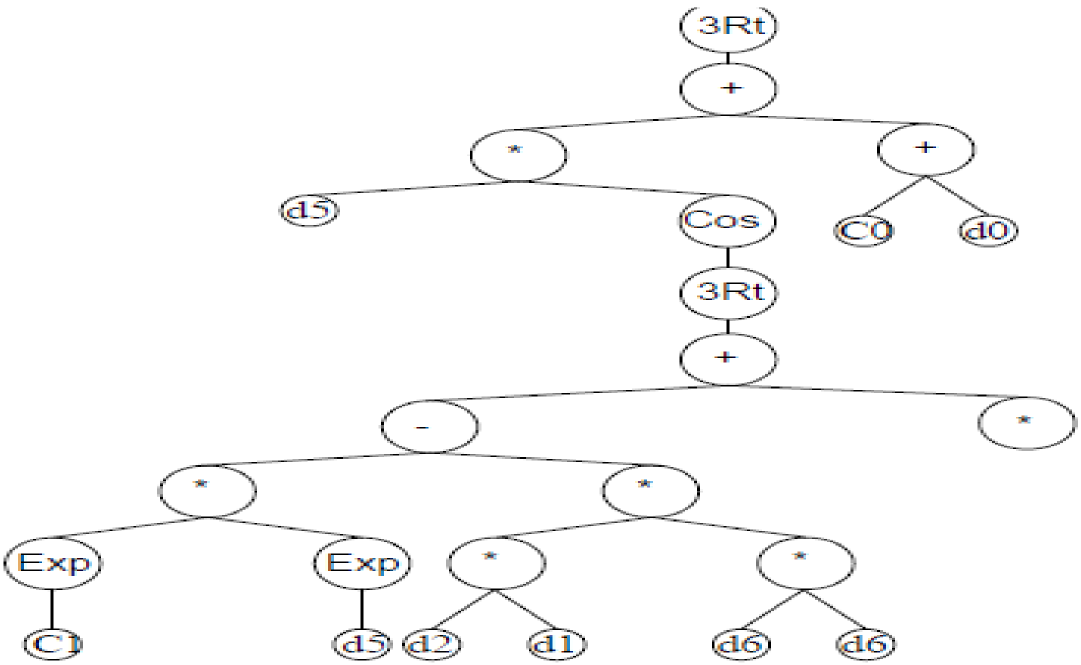


Figure 14: Marshal Flow Sub Expression Tree 2

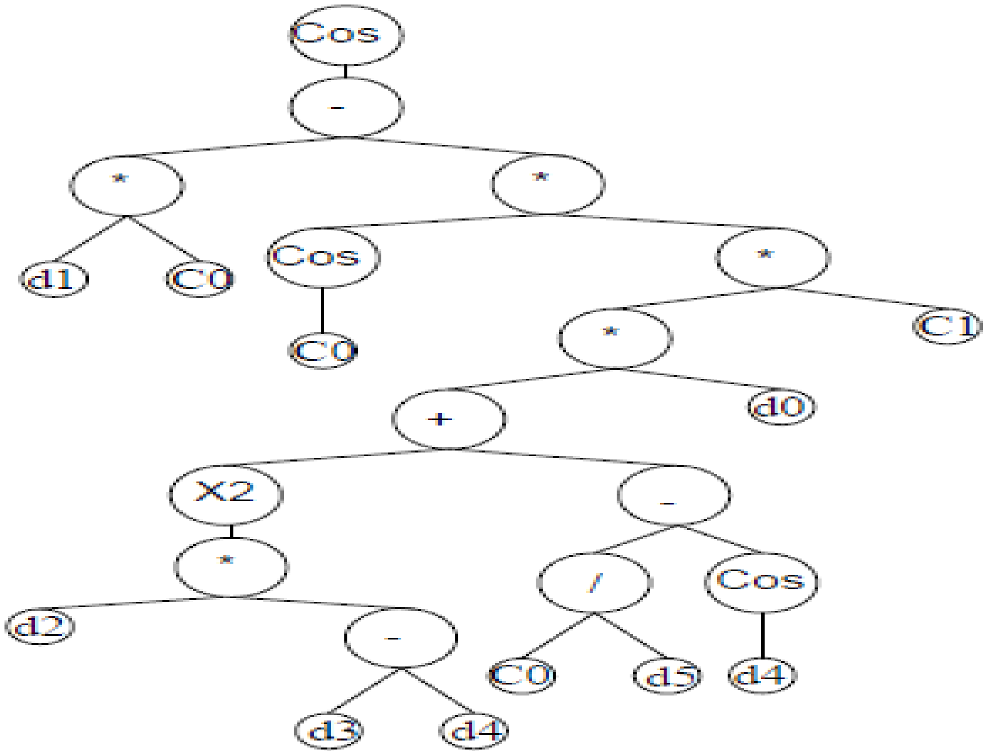


Figure 15: Marshal Flow Sub Expression Tree 3

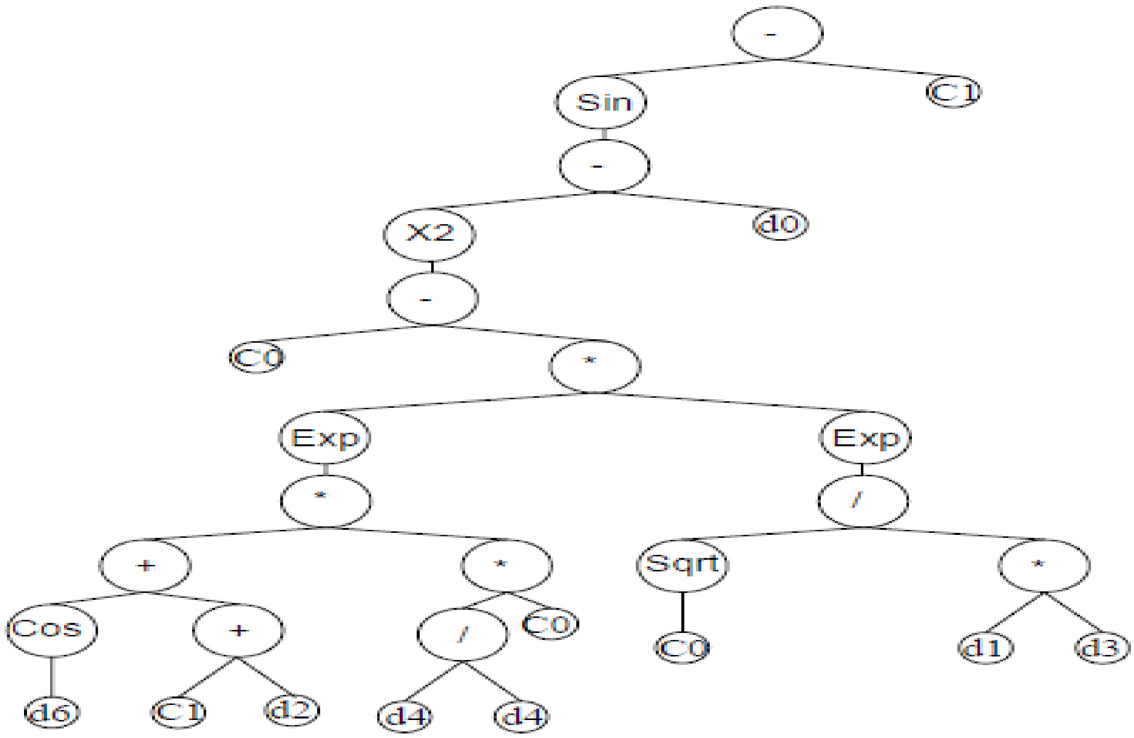
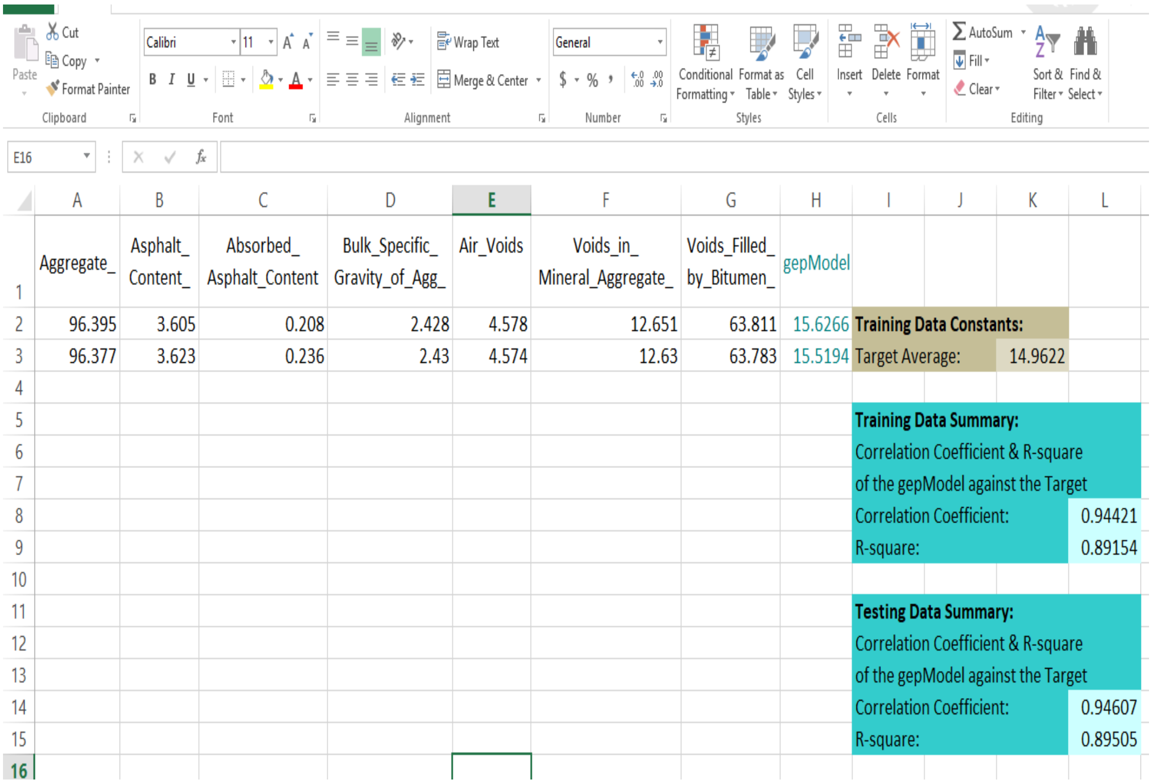


Figure 16: Marshal Flow Sub Expression Tree 4

Fig 13-16 presents the Marshal Flow model graphically in the form of expression trees which gives a much easier approach to the practitioner to understand the model. Expression trees should be read from left to right and from top to bottom. Beside this, GEP enables the practitioners to use the model deployed into an excel sheet where end user will simply put the input variables and the excel sheet will automatically predict the output values Fig 17 shows a sample of the deployed Marshal Flow model.



	A	B	C	D	E	F	G	H	I	J	K	L
1	Aggregate_	Asphalt_ Content_	Absorbed_ Asphalt_Content	Bulk_Specific_ Gravity_of_Agg_	Air_Voids	Voids_in_ Mineral_Aggregate_	Voids_Filled_ by_Bitumen_	gepModel				
2	96.395	3.605	0.208	2.428	4.578	12.651	63.811	15.6266	Training Data Constants:			
3	96.377	3.623	0.236	2.43	4.574	12.63	63.783	15.5194	Target Average:		14.9622	
4												
5									Training Data Summary:			
6									Correlation Coefficient & R-square			
7									of the gepModel against the Target			
8									Correlation Coefficient:			0.94421
9									R-square:			0.89154
10												
11									Testing Data Summary:			
12									Correlation Coefficient & R-square			
13									of the gepModel against the Target			
14									Correlation Coefficient:			0.94607
15									R-square:			0.89505
16												

Figure 17: Deployed Marshall Flow Model (GEP)

4.3. Comparison of results before and after variable reduction

Existence of insignificant variables in the model might affect the performance of the model in negative way. Therefore, insignificant variables were removed from the data prior to the modeling phase for both Marshal Flow and Marshal Stability. This section compares the results for both models before and after variables reduction. Results are presented in table 15, 16. Table 15 shows the comparison for Marshal Stability where improvement in all the validation criterion can be observed in modeling after variables reduction. Table 16 shows the results for Marshal Flow showing a slight improvement in every validation criterion except for R^2 in case of testing and the reason for this behavior is that the formula for calculating the coefficient of determination does not counts the number of input variables.

Table 16: Marshal Stability Model Before and After Variable Reduction

Marshal Stability Model Before and After Variable Reduction				
	Training		Testing	
	Before	After	Before	After
Input Variables	10	7	10	7
Data Points	82	82	20	20
R^2	0.838	0.890	0.926	0.942
MAE	193.94	148.2	215.66	131.2
RMSE	256.87	216.5	255.8	198.2
Adjusted R^2	0.815	0.880	0.844	0.908

Table 17: Marshal Flow Model Before and After Variable Reduction

Marshal Flow Model Before and After Variable Reduction				
	Training		Testing	
	Before	After	Before	After
Input Variables	10	7	10	7
Data Points	82	82	20	20
R^2	0.85	0.891	0.902	0.895
MAE	0.4185	0.389	0.448	0.488
RMSE	0.5524	0.463	0.603	0.607
Adjusted R^2	0.829	0.881	0.793	0.834

4.4. Artificial Neural Network (ANN) Modeling

Similar to GEP modeling, prediction models for Marshall Stability and Marshall Flow were created using the same set of data with same input variables by adopting Artificial Neural Network technique. Levenberg-Marquardt algorithm was adopted with five number of neurons in hidden layer. Selecting the number of neurons in the hidden layer is one of the most technical tasks to be performed, as a small mistake might lead to either over fitting or under fitting of the model. As no specific formula has been developed till today that can be used by the researchers to select the appropriate number of neurons in the hidden layer, that is why the number of neurons in the hidden layer were decided based on the following rule of thumb [42].

1. Number of hidden layer's neurons in the should be lesser than twice the number of neurons in the input layer.
2. Neurons in the hidden layer should be from 70 to 90% of the number of neurons in the input layer.
3. Neurons in the hidden layer should be more than number of neurons in the output layer and lesser than the number of neurons in the input layer.

After analyzing the data in this research, it was concluded that the number of neurons in the hidden layer should be 5 as it fulfills all the above thumb rules.

4.3.1. Marshall Stability Model

The Marshal Stability model generated using ANN was based on the same input variables that were used for developing Marshal Stability model using GEP technique. 82 data points were used to train the model and 20 were used to test the validity of the model by following Levenberg-Marquardt algorithm with 5 number of neurons in hidden layer. The model showed R^2 of 0.869 and adjusted R^2 of 0.857 in case of testing and 0.774 and 0.643 respectively in case of testing. The overall performance is shown in table 17 and figure 18.

Table 18: Performance of Marshal Stability Model Using ANN

Marshal Stability Model Using ANN		
	Training	Testing
R^2	0.869	0.774
RMSE	234.38	398.55
MAE	149.68	259.55
Adjusted R^2	0.857	0.643

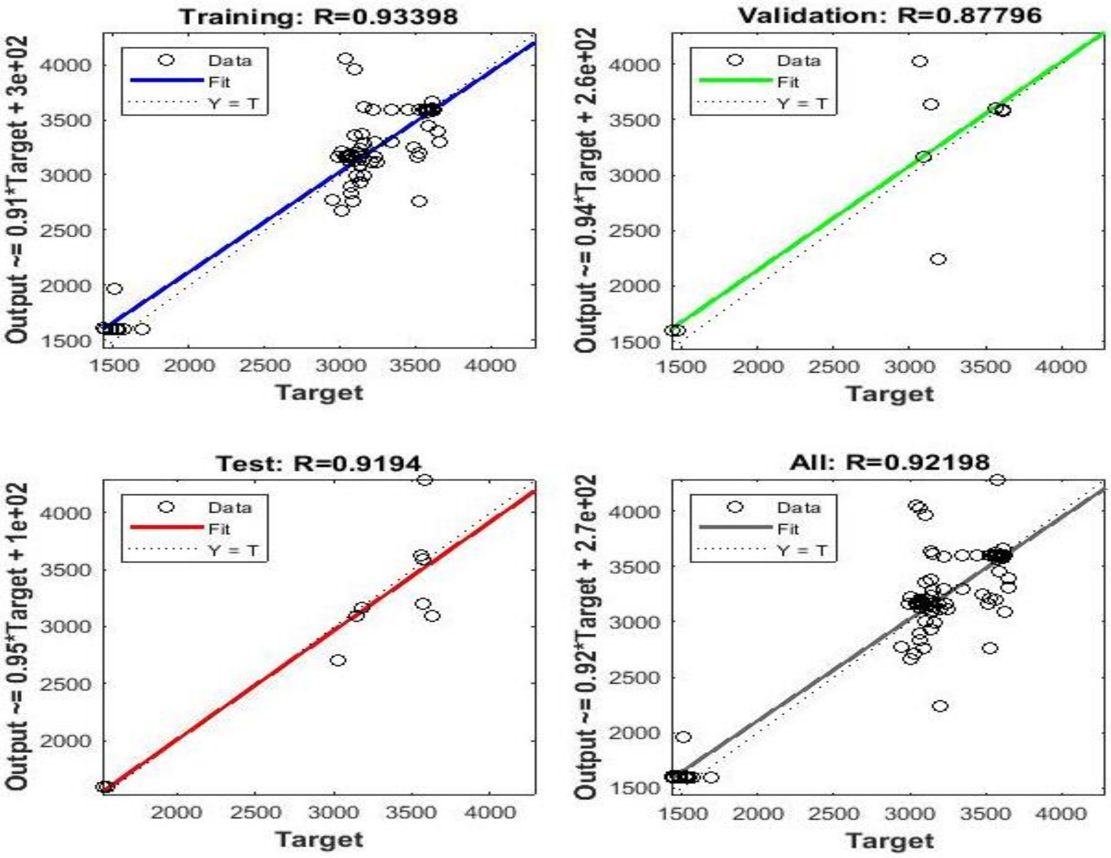


Figure 18: Regression for Marshall Stability Model Developed by using ANN

4.3.2. Marshall Flow Model

The same setup was followed for developing Marshall Flow model using ANN. Levenberg-Marquardt algorithm with five neurons in hidden layer was followed for developing this model. Coefficient of determination recorded in case of training was 0.7496 and 0.813 in case of testing. Similarly, Adjusted R^2 recorded in case of training was 0.725 and 0.704 in case of testing the model. The detailed performance of the model is shown in table 18 and figure 19.

Table 19: Performance of Marshal Flow Model Using ANN

Marshal Flow Model Using ANN		
	Training	**Testing**
R^2	0.7496	0.8137
RMSE	0.7413	0.6550
MAE	0.6177	0.5434
Adjusted R^2	0.7259	0.7041

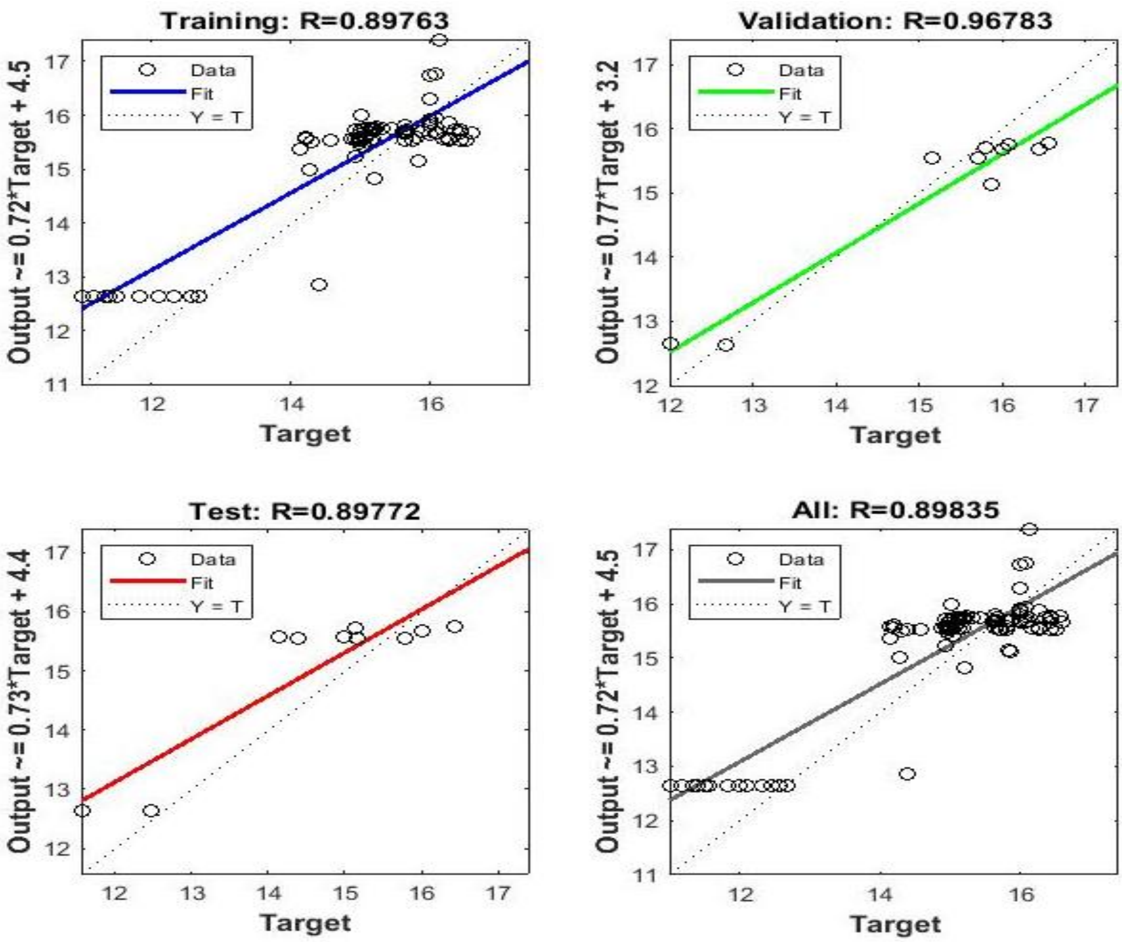


Figure 19: Regression for Marshall Flow Model Developed by using ANN

4.4. Comparison of Results

This section of the report shows the comparison of results obtained for Marshall Stability and Marshall Flow using GEP and ANN. This comparison sheds light on the performance of both GEP and ANN utilized for predicting output using the same set of data and input variables in both cases. Results shows that GEP out performs the ANN model in terms of performance validated based on the following validation criterion.

Table 20: Comparison of Marshal Stability Model using GEP and ANN

Marshal Stability Model						
	Training		**Testing**		**Difference (%)**	
	GEP	**ANN**	**GEP**	**ANN**	**Training**	**Testing**
R^2	0.890	0.869	0.942	0.774	2.410	21.70
MAE	148.2	149.68	131.2	259.55	0.990	97.80
RMSE	216.5	234.38	198.2	398.55	8.250	101.1
Adjusted R^2	0.880	0.857	0.908	0.643	2.680	41.20

Table 21: Comparison of Marshal Flow Model using GEP and ANN

Marshal Flow Model						
	Training		**Testing**		**Difference (%)**	
	GEP	**ANN**	**GEP**	**ANN**	**Training**	**Testing**
R^2	0.891	0.7496	0.895	0.8137	18.80	9.990
MAE	0.389	0.6177	0.448	0.5434	58.79	21.30
RMSE	0.463	0.7413	0.607	0.6550	60.00	7.900
Adjusted R^2	0.881	0.7259	0.834	0.7041	21.30	18.45

From table 19, it can be observed that the MS model developed using GEP out performs the ANN model in all the validation conditions. Coefficient of determination observed in case of MS model developed using GEP is 0.89 while in case of ANN the same value observed is 0.869. Comparing the RMSE for GEP and ANN shows that the error recorded for ANN model is 8.25% greater than that of GEP model. By comparing the Adjusted R-Square for the same model in case of training data set shows that the GEP model performs better than ANN model. Same behavior is observed for marshal stability model in case of testing data set. Figure 20 shows the overall comparison of GEP and ANN for Marshal Stability Model.

Table 20 also reveals the same behavior for marshal flow model where the coefficient of determination shows that GEP model performs 18.8% better than ANN model in case of training and about 10% in case of testing data set. Adjusted R-Square recorded in case of GEP is 21.3% higher than that of ANN model. Same performance is observed for marshal flow model in case of testing data set. Figure 21 shows the overall comparison of GEP vs ANN for Marshal Flow.

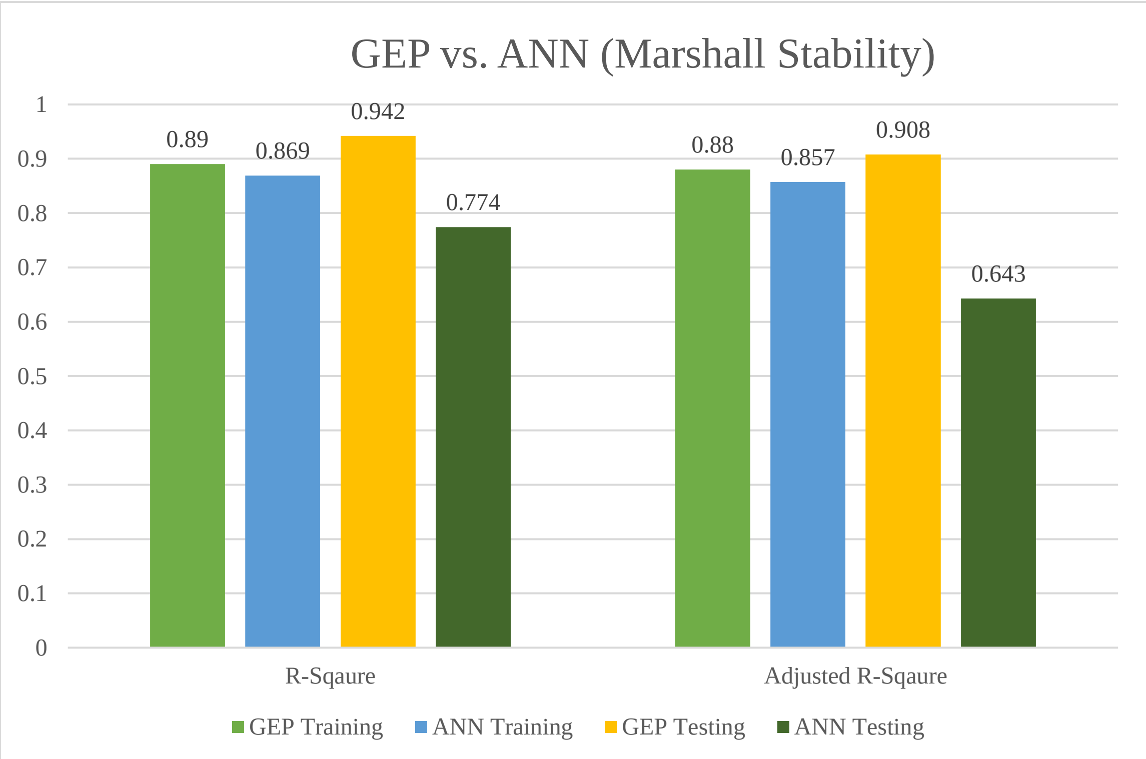


Figure 20: GEP vs. ANN (Marshall Stability)

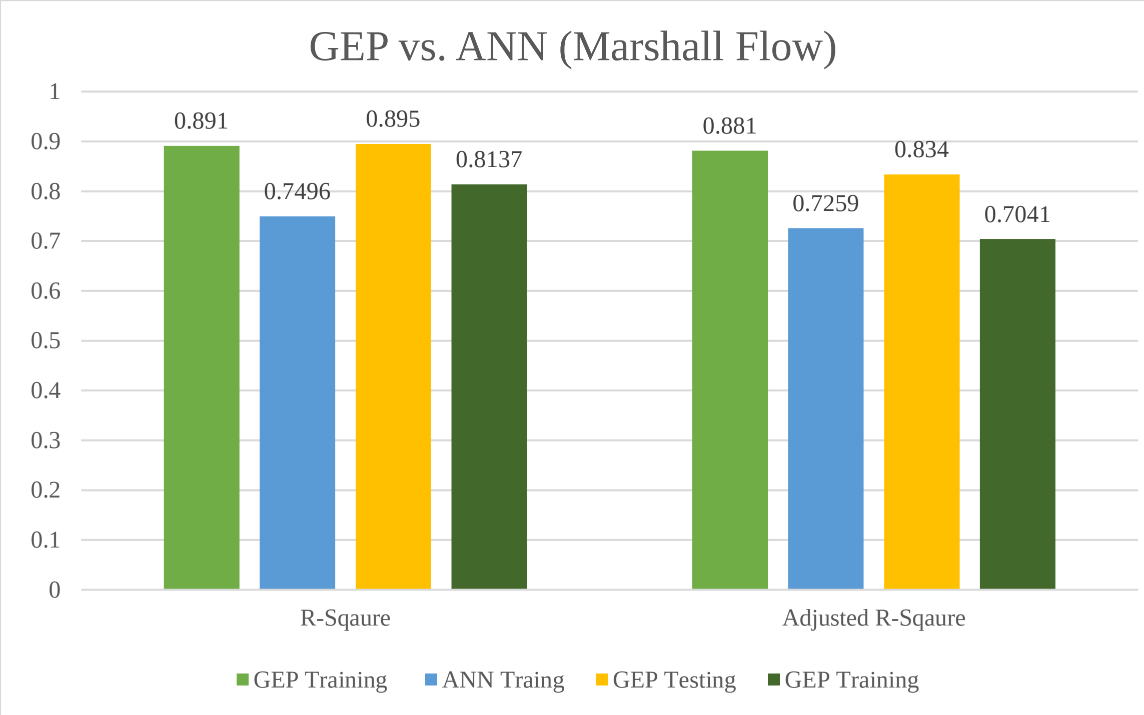


Figure 21: GEP vs. ANN (Marshall Flow)

Figure 20 and 21 presents the comparison of R^2 and adjusted R^2 observed against GEP and ANN model for Marshal Stability and Marshal Flow respectively. The prior figure compares the validation criterion for Marshall Stability and the later shows the same results recorded for Marshal Flow. The overall picture shows that the model generated for both these outputs using GEP performs better than those developed with ANN. Both these developed models are capable of predicting the Marshal Stability and Marshal Flow without going into traditional destructive testing approach adopted for calculating the values of Marshal Stability and Flow. Models developed in this study can be introduced with new data (only inputs), outputs will be predicted based these inputs with an accuracy of 94 %.

4.5. Sensitivity Analysis:

To investigate the effect of every input variable on the output, a sensitivity analysis was performed for both MS and MF. In case of MS, percentage of aggregate was found to be the most effective input variable followed by effective asphalt content with effectiveness of 36.54 and 19.23% respectively. Similarly in case of MF, the effectiveness of aggregate percentage with 24% effectiveness was found to be the most effective input variable followed by voids in mineral aggregates with effectiveness of 16% in total. The detailed result of sensitivity analysis can be observed in fig 22 and 23.

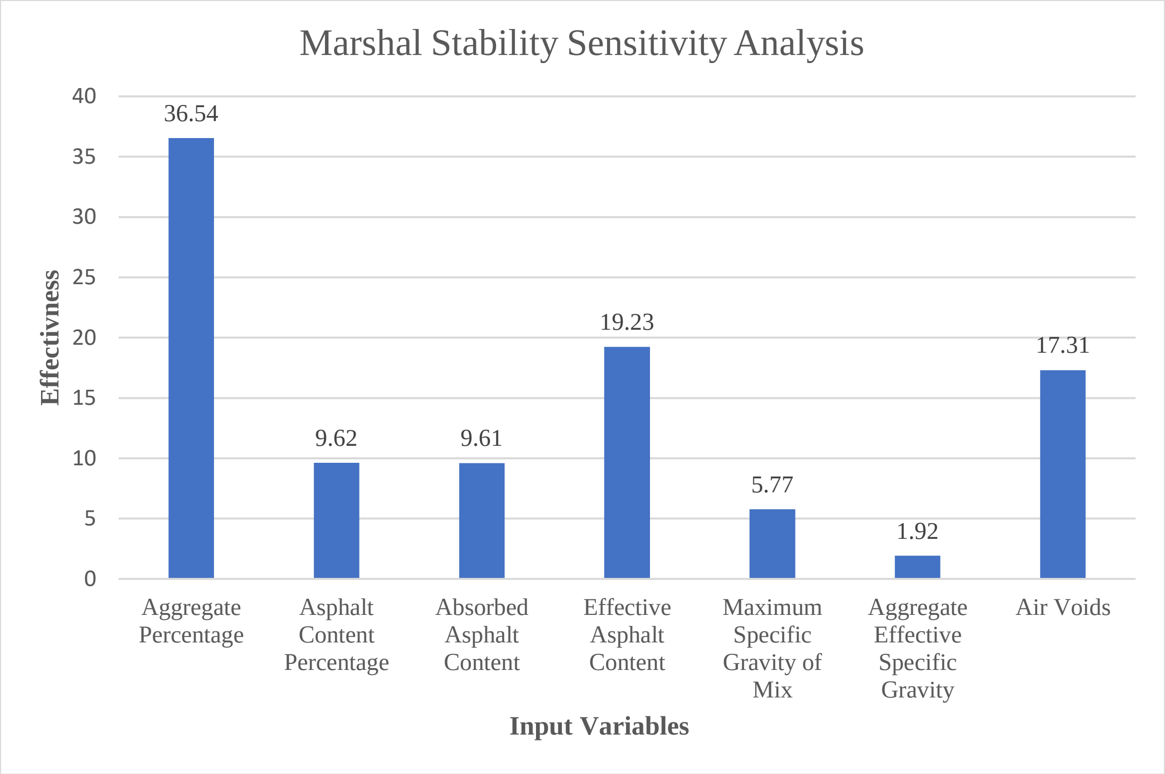


Figure 22: Marshal Stability Sensitivity Analysis

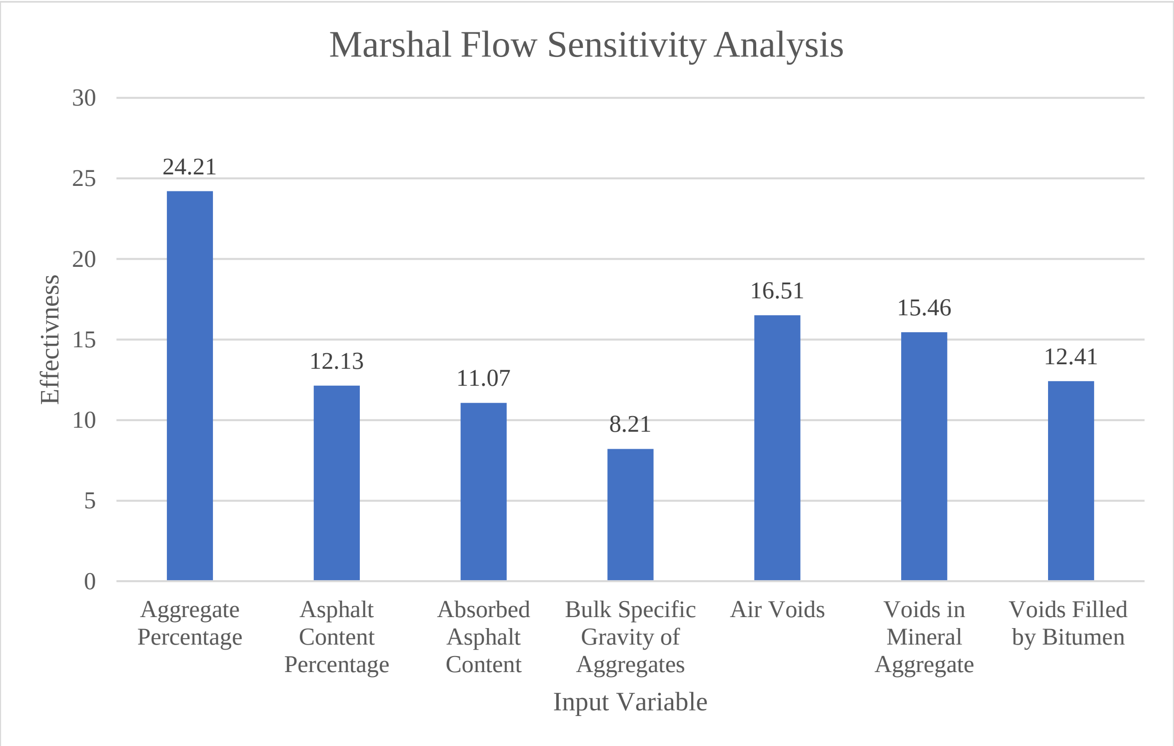


Figure 23:Marshal Flow Sensitivity Analysis

4.6. Performance Validation Using Unseen Data

Models developed using GEP and ANN for MS and MF has been introduced with a new data set collected from a construction firm working on N-95 (Swat 82km). This data set was composed of 25 screened data points using which the performance of the developed model was validated by using them as a testing data. The performance of the models developed for a highway project in Risalpur for the new dataset was almost the same as the data used for training. The detailed performance of the models can be observed from table 21 to table 25.

Table 22: MS Model Validation using new data with GEP

Marshal Stability Model using GEP		
Data	**Old Testing**	**Unseen**
R^2	0.942	0.915
MAE	131.2	194.71
RMSE	198.2	244.238
Adjusted R^2	0.908	0.910

Table 23: MS model Validation using new data with ANN

Marshal Stability Model ANN		
Data	**Old Testing**	**Unseen**
R^2	0.774	0.884
MAE	259.55	103.72
RMSE	398.55	216.7
Adjusted R^2	0.643	0.877

Table 24: MF Model Validation using new data with GEP

Marshal Flow using GEP		
Data	**Old Testing**	**Unseen**
R^2	0.895	0.874
MAE	0.448	0.496
RMSE	0.607	0.593
Adjusted R^2	0.834	0.865

Table 25: MF Model Validation using new data with ANN

Marshal Flow using ANN		
Data	**Old Testing**	**Unseen**
R2	0.8137	0.876
MAE	0.5434	0.371
RMSE	0.6550	0.471
Adjusted R2	0.7041	0.868

From table 21 to 25, it can be observed that the model is capable of explaining the new conditions with up to 91.5 % accuracy. Comparing the performance of ANN using unseen data with the old training data reveals that model is explaining the new conditions better than the older data. In case of GEP, the model shows somewhat equal performance for both old and new data. After observing the performance, it is recommended to use these developed models for new data but the input variables should fall within the maximum and minimum limits of the data which was used for developing the model. For environment where the data is not falling within the limits of the input variables o training data, then the model should be re-tuned by adjusting its parameter by using a 15 % of the new data for tuning.

4.7. Parametric Analysis.

Results of sensitivity analysis reveals that bitumen content and aggregate percentage is having the maximum impact on the output prediction. Therefore, it was intended to investigate the values of MS and MF by changing the values of bitumen content and aggregate percentage from 3 to 6 percent with a constant interval of 0.5 percent for the prior one and similarly from 94 to 97 percent for the later one. These values were changed in coordination with percentage of air voids from 2.5% to 4.5%. Results obtained against different combination of air voids and bitumen content can be observed in table 25.

Table 26: Effect of Bitumen Content and Air voids on Marshal Stability

Marshal Stability					
Air Voids / Bitumen	2.5 %	3.0 %	3.5 %	4.0 %	4.5 %
3.0 %	2303.8	2449.3	2614.1	2792.6	2980.6
3.5 %	2719.1	2805.4	2905.3	2993.5	3066.2
4.0 %	1401.3	1432.2	1455.4	1475.5	1494.9
4.5 %	1258.1	1261.4	1263.7	1265.3	1266.7
5.0 %	1253.4	1254.3	1257.6	1258.3	1259.4
5.5 %	1250.9	1252.4	1254.1	1254.7	1255.1
6.0 %	1250.3	1251.3	1251.4	1251.4	1251.3

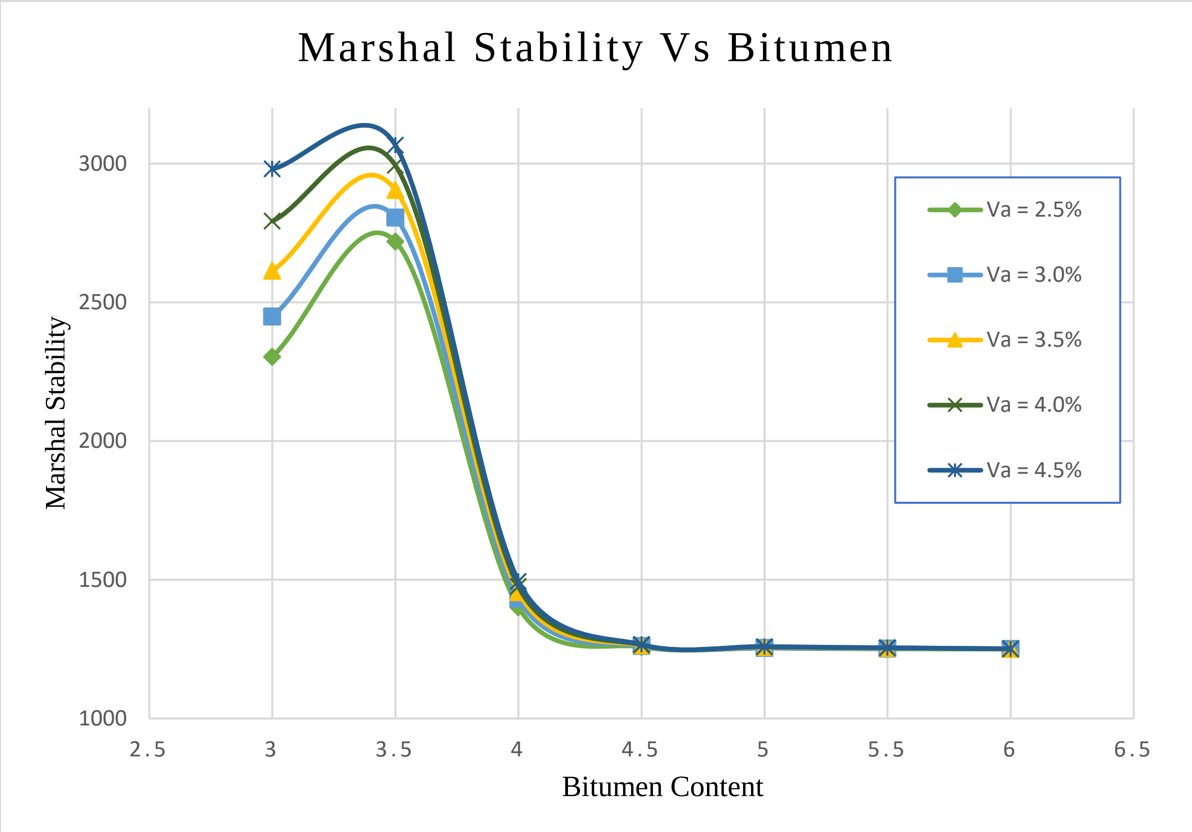


Figure 24: Marshal Stability against different Bitumen content and Air Voids.

From fig. 24, it can be observed that as the air voids increases, marshal stability also increase. Effect of air voids on marshal stability up to 4.5% has been observed in this study which reveals an increase in marshal stability upon the increase in air voids but after having a microscopic look, results show that by increasing the bitumen content while keeping the air voids constant increases the marshal stability up to a certain degree but marshal stability starts decreasing after a certain range of increase in bitumen content. Maximum values of marshal stability can be observed between 3 to 3.5% of bitumen against every percentage of air voids. Effect of variation in air voids has been observed to be more effective when the bitumen content has been kept between 3 and 3.5% while these variations are not very much effective as the bitumen content exceeds 4% by weight. Increase in bitumen content from 3.5 to 4 % reduces the stability by almost 50% as recorded against the previous combination of bitumen, aggregate and air voids. Rate of change of marshal stability against a periodic variation in both bitumen content and air voids approaches zero as bitumen content exceeds 4.5%. A similar mechanism was adopted for investigating the effect of these variations on marshal flow as well. Results obtained in case of marshal flow can be observed from table 27 and fig. 25

Table 27: Marshal Flow against Bitumen content and Air Voids

Marshal Flow					
Bitumen \ Air Voids	2.5 %	3.0 %	3.5 %	4.0 %	4.5 %
3.0 %	17.10	16.70	16.80	16.95	16.75
3.5 %	19.00	18.15	17.75	17.00	15.60
4.0 %	22.20	20.75	18.90	16.85	14.55
4.5 %	24.20	23.20	21.80	20.09	18.45
5.0 %	31.15	30.45	29.30	27.85	25.90
5.5 %	31.70	31.00	29.70	28.55	26.85
6.0 %	32.30	32.25	30.15	28.75	26.95

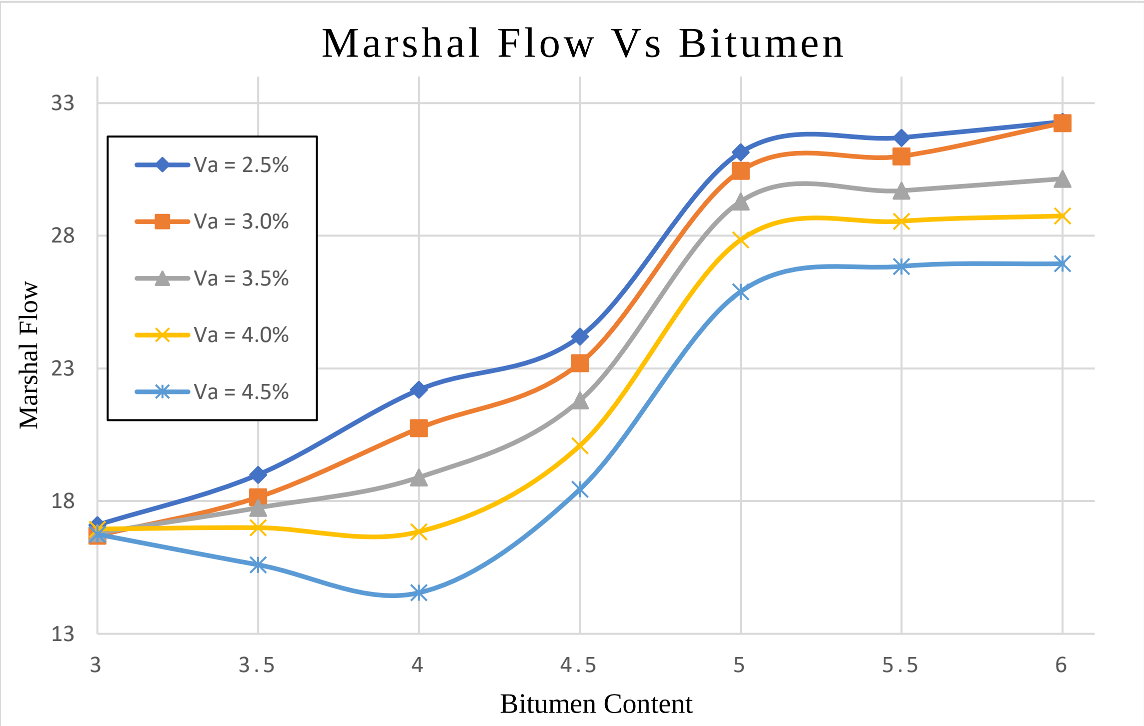


Figure 25: Effect of Bitumen content and Air Voids on Marshal Flow

From fig. 25, it can be observed that the effect of variation in air voids increases as the bitumen content increases as it can be observed that marshal flow obtained against 3% bitumen content is almost same for all combinations of air voids. 4% of bitumen content and 4.5% air voids results in minimum flow while maximum flow was observed at 6%

bitumen content and 2.5 % air voids whereas maximum rate of change in marshal flow can be observed from 4 to 4.5 % bitumen content irrespective of air voids. Rate of change in marshal flow approaches zero as the bitumen content goes beyond 5.0%. Marshal flow increases with the increase in bitumen content. This relation has been observed to be more sensitive for bitumen content from 3.5 to 5% but this relation between bitumen content and marshal flow turns to be indirect as the percentage of air voids is increased up to 4 % or higher and keeping the bitumen content lesser than 4 %. Overall, it can be observed that higher the air voids, lesser will be the flow values and vice versa.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

Marshal Stability and Flow of hot mix asphalt are derived based on laboratory testing that requires due care and time to derive these values. The procedure is consisting of sample preparation and testing during which, different temperatures are maintained for a number of times during the whole procedure. To reduce the human error, loss of human efforts, time and money, this study presents two models developed using different algorithm to bypass the need of these tests for both marshal stability and flow. These models have been deployed into excel sheets (fig. 11 and fig 17) to make it easy for future use. The performance of developed models shows that it can predict these values with up to 94% accuracy. From this research, the following points can be concluded.

1. Marshal stability increases as the air voids increases while keeping the bitumen content constant and marshal stability starts decreasing after a certain increase in bitumen content while keeping the air voids constant.
2. Effect of variation in air voids has been observed to be more effective when the bitumen content has been kept between 3 and 3.5%.
3. Marshal Stability is not affected significantly by the variation in air voids as the bitumen content exceeds 4% by weight.
4. Marshal Stability reduces by almost 50% as the aggregate percentage reduces from 96.5 to 96%.
5. Rate of change of marshal stability against a periodic variation in both bitumen content and air voids approaches zero as bitumen content exceeds 4.5%.
6. As per the specification of National highway authorities, percentage of air voids should be 3.5 % or more but in this study, the marshal stability observed against 2.5 and 3.0% of air voids using 3.0 % bitumen content is almost twice the minimum marshal stability requirement (1200 kg).
7. Marshal flow decreases as the air voids increases while keeping the bitumen content constant and it starts increasing as the bitumen content increases while keeping the air voids the same.
8. Marshal flow has an inverse relation with the percentage of air voids as the flow increases as a result of decrease in air voids.

9. Marshal flow increases with the increase in bitumen content. This relation has been observed to be more sensitive for bitumen content from 3.5 to 5%.
10. The direct relation between bitumen content and marshal flow turns to be indirect as the percentage of air voids is increased up to 4 % or higher and keeping the bitumen content lesser than 4 %.
11. Rate of change in marshal flow is higher for 3.5 to 5% bitumen content, this rate has been observed to be lesser for bitumen content falling on either side of these limits.
12. Performance of the models developed in this study shows that models are able to predict the output with the same or better accuracy in a new environment as shown in section 4.6.
13. These models are not universal and cannot be implemented everywhere until and unless the data used for making predictions is falling within the range of data used for training the model as shown in table 1. If the data used for making predictions is not falling within these limits, then the developed models should be re-tuned by adjusting the parameters shown in table 13.
14. Before going into the modeling, researcher should select the input variables based on the significance of every variable and should remove those variables that may worsen the performance of the model. Improvement in the performance of models developed in this study has been observed as a result of variable reduction.
15. Adjusted R^2 explains the performance of the model more accurately as compare to R or R^2. Researchers should use adjusted R^2 along with other criterions for accessing the performance of models.
16. This research concluded that the overall performance of GEP model is better than ANN. GEP revealed better results for all the performance validation criterion as compare to ANN in both training and testing datasets for both Marshal Stability and Marshal Flow. Beside this, following are some of the key points observed during this study.
    1. Calculating the number of hidden layers, and number of neurons in hidden layer is a very critical part of the modeling using ANN, because a slight error may lead the model to overfitting or underfitting while GEP does not possess such problem.

2. GEP offers deploying the model in the form of an equation which can be adopted by the practitioners in future while ANN does not offer a model every time.
3. GEP can predict only one output variable at a time, and the user needs to develop a separate model for each output variable in case it is intended to predict more than one variable while ANN can predict more than one variable in a single go.
4. GEP offers exporting the model graphically in the form of Expression Trees which makes it easy to understand the model while ANN does not offer such feature.
5. GEP let the users to deploy the model into an excel sheet which can be used in future to predict the output parameters by simply putting the input variables without regenerating the model (a sample has been shown in figure 11 and 17). On the other hand, ANN does not have such option.

5.2. Recommendations

From this study it is recommended to the researchers to check the significance of all the input variables before putting them into a model. To enhance the performance of the model all the insignificant variables should be removed from the data. To check the validity of model's performance, adjusted R^2 should be adopted instead of relying on coefficient of determination as adjusted R^2 measures the performance of the model better than R^2.

For future studies, it is suggested to work on developing a mechanism for calculating the number of hidden layers and number of neurons in the hidden layer for developing a model using ANN and a similar mechanism needs to be developed for calculating the head size and number of chromosomes in one gene in case of modeling with GEP. After this study, it is recommended to work on the integration of GEP and ANN with each other to develop an GEP-ANN hybrid model for more accurate results.

REFERENCES

[1] H. U. Bahia, "Bibliographies for physical properties of asphalt cement," 1993.

[2] I. P. B. Eng, "The design of bituminous concrete mixes," *Philosophy,* 1991.

[3] E. Brown, P. Kandhal, and J. Zhang, "Performance testing for hot mix asphalt. NCAT Report 01-05," *National Center for Asphalt Technology,* 2001.

[4] R. N. Hunter, A. Self, J. Read, and E. Hobson, *The shell bitumen handbook*. ICE Publishing London, UK:, 2015.

[5] T. Officials, *Hot-mix Asphalt Paving Handbook*. Transportation Research Board, 1991.

[6] S. F. Kalyoncuoglu and M. Tigdemir, "An alternative approach for modelling and simulation of traffic data: artificial neural networks," *Simulation Modelling Practice and Theory,* vol. 12, no. 5, pp. 351-362, 2004.

[7] M. Tigdemir, M. Karasahin, and Z. Sen, "Investigation of fatigue behaviour of asphalt concrete pavements with fuzzy-logic approach," *International Journal of Fatigue,* vol. 24, no. 8, pp. 903-910, 2002.

[8] K. Copper and P. S. Pell, "The effect of mix variables on the fatigue strength of bituminous materials," 1974.

[9] S. Tapkın, A. Çevik, and Ü. Uşar, "Prediction of Marshall test results for polypropylene modified dense bituminous mixtures using neural networks," *Expert Systems with Applications,* vol. 37, no. 6, pp. 4660-4670, 2010.

[10] A. H. Gandomi, S. Sajedi, B. Kiani, and Q. Huang, "Genetic programming for experimental big data mining: A case study on concrete creep formulation," *Automation in Construction,* vol. 70, pp. 89-97, 2016.

[11] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems," *arXiv preprint cs/0102027,* 2001.

[12] J. R. Koza, *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*. Stanford University, Department of Computer Science Stanford, CA, 1990.

[13] M. Mazari and D. D. Rodriguez, "Prediction of pavement roughness using a hybrid gene expression programming-neural network technique," *Journal of Traffic and Transportation Engineering (English Edition),* vol. 3, no. 5, pp. 448-455, 2016.

[14] L. P. Leon and D. Gay, "Gene expression programming for evaluation of aggregate angularity effects on permanent deformation of asphalt mixtures," *Construction and Building Materials,* vol. 211, pp. 470-478, 2019.

[15] L. Yao, Z. Leng, J. Jiang, F. Ni, and Z. Zhao, "Nondestructive prediction of rutting resistance of in-service middle asphalt layer based on gene expression programing," *Construction and Building Materials,* vol. 293, p. 123481, 2021.

[16] H. Majidifard, B. Jahangiri, P. Rath, L. U. Contreras, W. G. Buttlar, and A. H. Alavi, "Developing a prediction model for rutting depth of asphalt mixtures using gene expression programming," *Construction and Building Materials,* vol. 267, p. 120543, 2021.

[17] H. M. Azamathulla, U. Rathnayake, and A. Shatnawi, "Gene expression programming and artificial neural network to estimate atmospheric temperature in Tabuk, Saudi Arabia," *Applied Water Science,* vol. 8, no. 6, pp. 1-7, 2018.

[18] A. H. Alavi, M. Ameri, A. H. Gandomi, and M. R. Mirzahosseini, "Formulation of flow number of asphalt mixes using a hybrid computational method," *Construction and Building Materials,* vol. 25, no. 3, pp. 1338-1355, 2011.

[19] A. H. Gandomi, A. H. Alavi, M. R. Mirzahosseini, and F. M. Nejad, "Nonlinear genetic-based models for prediction of flow number of asphalt mixtures," *Journal of Materials in Civil Engineering,* vol. 23, no. 3, pp. 248-263, 2011.

[20] M. Mazari and Y. Niazi, "Modeling the effect of filler materials on performance of hot mix asphalt using genetic programming," in *Airfield and highway pavements 2015*, 2015, pp. 107-119.

[21] J. H. Choi, T. M. Adams, and H. U. Bahia, "Pavement Roughness Modeling Using Back-Propagation Neural Networks," *Computer-Aided Civil and Infrastructure Engineering,* vol. 19, no. 4, pp. 295-303, 2004.

[22] S. Terzi, "Modeling for pavement roughness using the ANFIS approach," *Advances in engineering software,* vol. 57, pp. 59-64, 2013.

[23] K. Ozbay and R. Laub, "Models for pavement deterioration using LTPP," 2001.

[24] E. Ozgan, "Artificial neural network based modelling of the Marshall Stability of asphalt concrete," *Expert Systems with Applications,* vol. 38, no. 5, pp. 6025-6030, 2011.

[25] S. K. Babanajad, A. H. Gandomi, and A. H. Alavi, "New prediction models for concrete ultimate strength under true-triaxial stress states: An evolutionary approach," *Advances in Engineering Software,* vol. 110, pp. 55-68, 2017.

[26] J. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation*. CRC Press, 2018.

[27] M. Sugeno, *Industrial applications of fuzzy control*. Elsevier Science Inc., 1985.

[28] L. A. Zadeh, "Is there a need for fuzzy logic?," *Information sciences,* vol. 178, no. 13, pp. 2751-2779, 2008.

[29] S. A. Bunyamin, T. S. Ijimdiya, A. O. Eberemu, and K. J. Osinubi, "Artificial neural networks prediction of compaction characteristics of black cotton soil stabilized with cement kiln dust," *Journal of Soft Computing in Civil Engineering,* vol. 2, no. 3, pp. 50-71, 2018.

[30] S. Tapkın, "A recommended neural trip distributon model," 2004.

[31] M. S. Kaseko and S. G. Ritchie, "A neural network-based methodology for pavement crack detection and classification," *Transportation Research Part C: Emerging Technologies,* vol. 1, no. 4, pp. 275-291, 1993.

[32] Y. Cal, "Soil classification by neural network," *Advances in Engineering software,* vol. 22, no. 2, pp. 95-97, 1995.

[33] H. Ceylan, A. Guclu, E. Tutumluer, and M. R. Thompson, "Backcalculation of full-depth asphalt pavement layer moduli considering nonlinear stress-dependent subgrade behavior," *International Journal of Pavement Engineering,* vol. 6, no. 3, pp. 171-182, 2005.

[34] R. Y. Liang and J. Zhou, "Prediction of fatigue life of asphalt concrete beams," *International journal of fatigue,* vol. 19, no. 2, pp. 117-124, 1997.

[35] U. Isacsson and H. Zeng, "Cracking of asphalt at low temperature as related to bitumen rheology," *Journal of materials science,* vol. 33, no. 8, pp. 2165-2170, 1998.

[36] H. U. Bahia, D. Hanson, M. Zeng, H. Zhai, M. Khatri, and R. Anderson, *Characterization of modified asphalt binders in superpave mix design* (no. Project 9-10 FY'96). 2001.

[37] J. S. Coplantz, M. T. Yapp, and F. N. Finn, "Review of relationships between modified asphalt properties and pavement performance," 1993.

[38] M. Stroup-Gardiner and E. R. Brown, *Segregation in hot-mix asphalt pavements* (no. 441). Transportation Research Board, 2000.

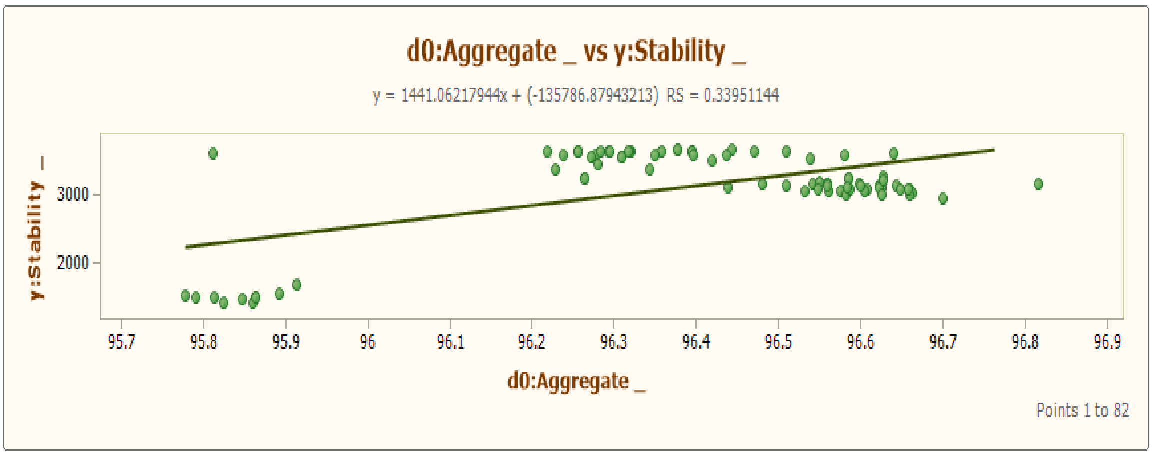
[39] M. Huner and E. Brown, "Effects of re-heating and compaction temperature on hot mix asphalt volumetrics," NCAT report2001.

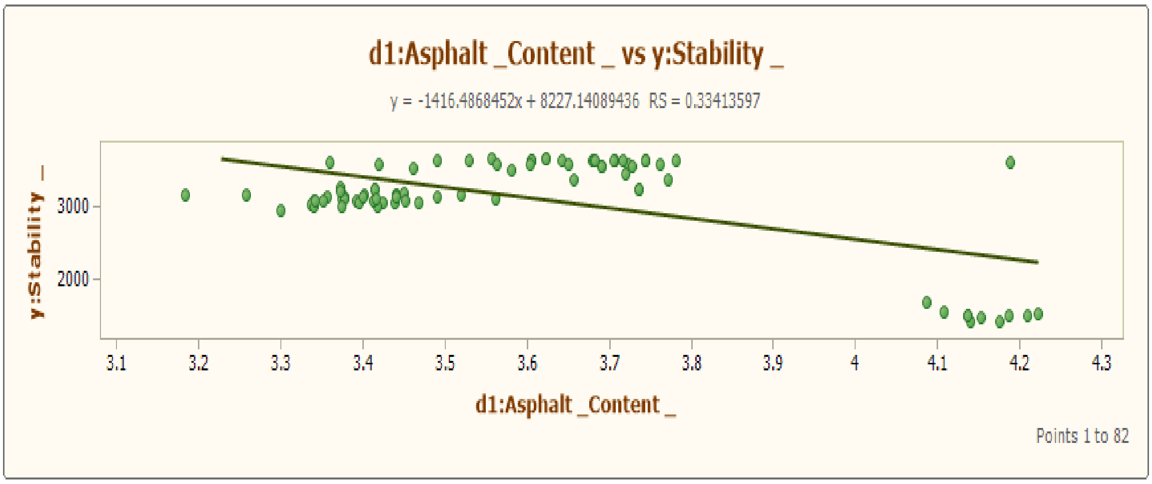
[40] E. Özgan and T. Kap, "The effects of exposure time and environment temperature to the ductility on asphalt cements," *Journal of the institute of science and technology of Sakarya University,* vol. 9, no. 1, pp. 43-47, 2005.

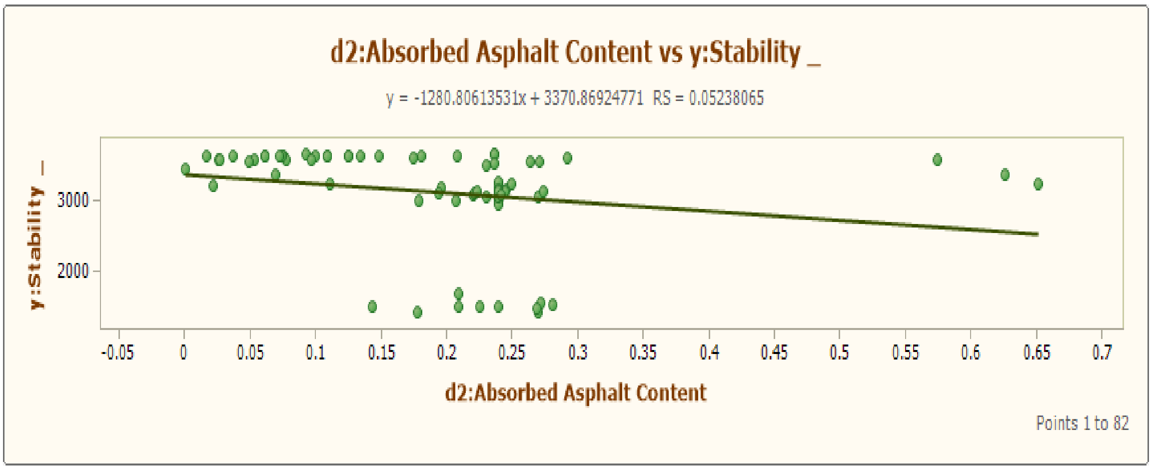
[41] P. T. Barrett, P. J. P. S. Kline, and G. Behaviour, "The observation to variable ratio in factor analysis," 1981.

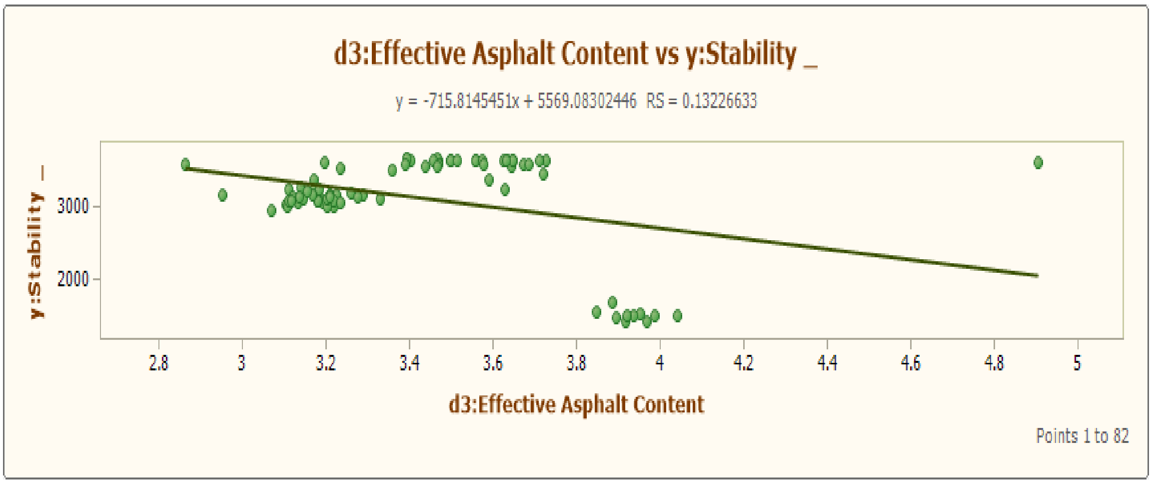
[42] S. Karsoliya, "Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture," *International Journal of Engineering Trends and Technology,* vol. 3, no. 6, pp. 714-717, 2012.

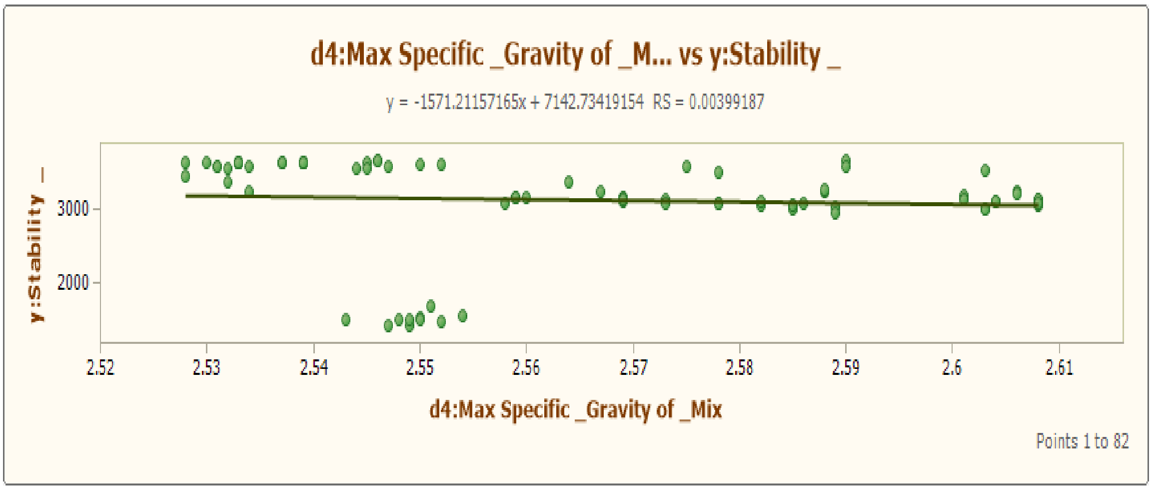
Appendix I: Input Variables vs Marshal Stability

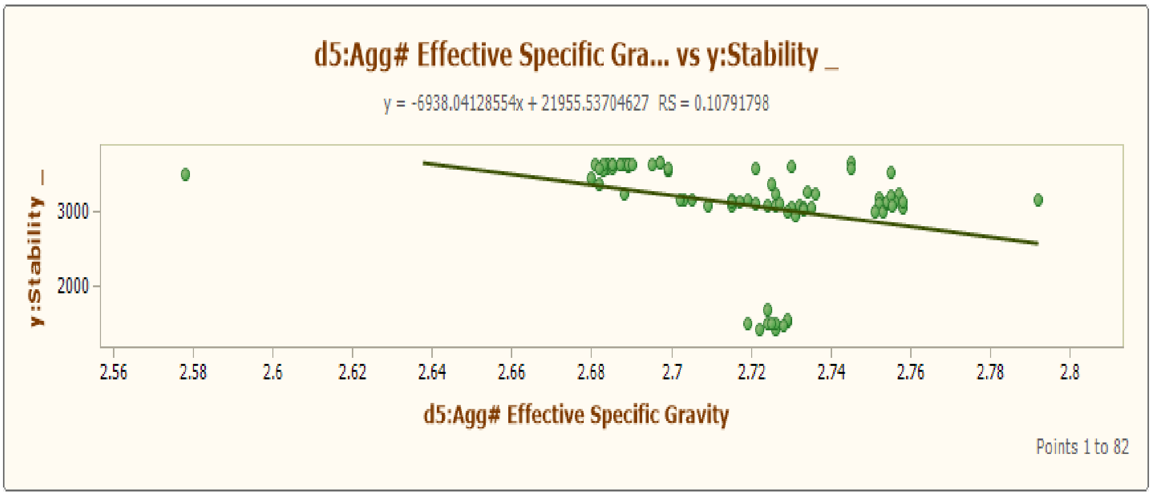


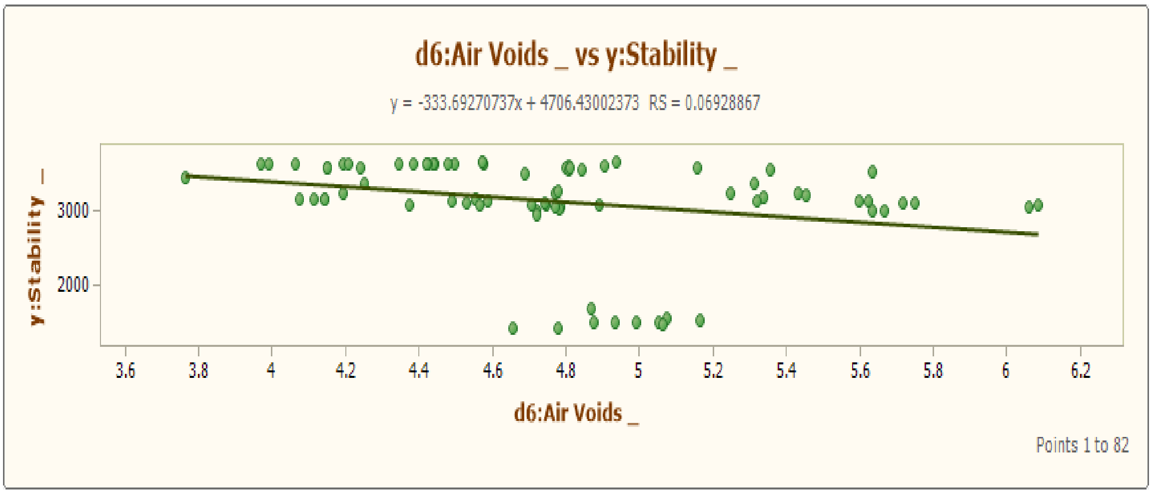




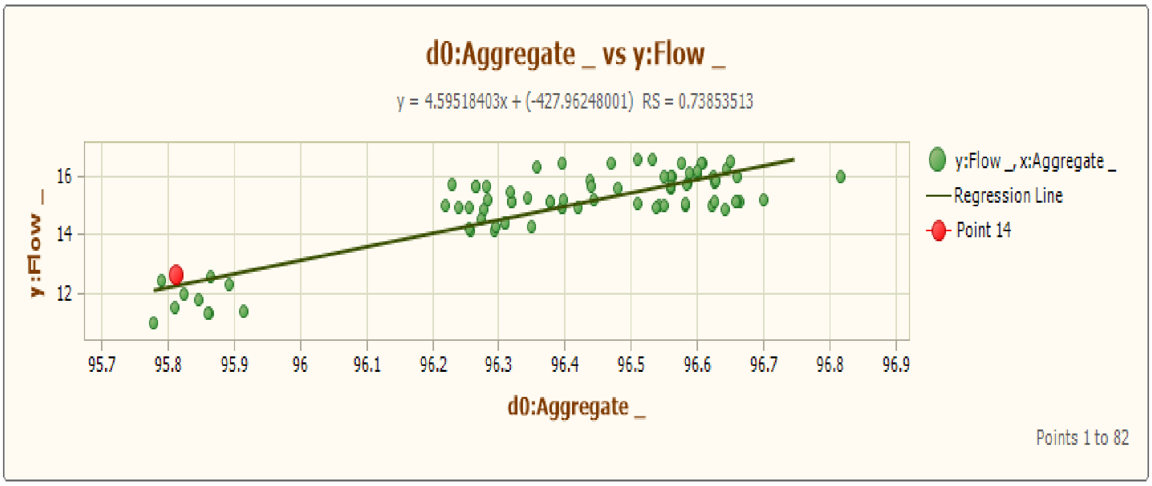

d3:Effective Asphalt Content vs y:Stability _
y = -715.8145451x + 5569.08302446 RS = 0.13226633
y:Stability _
3000
2000
2.8
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
5
d3:Effective Asphalt Content
Points 1 to 82

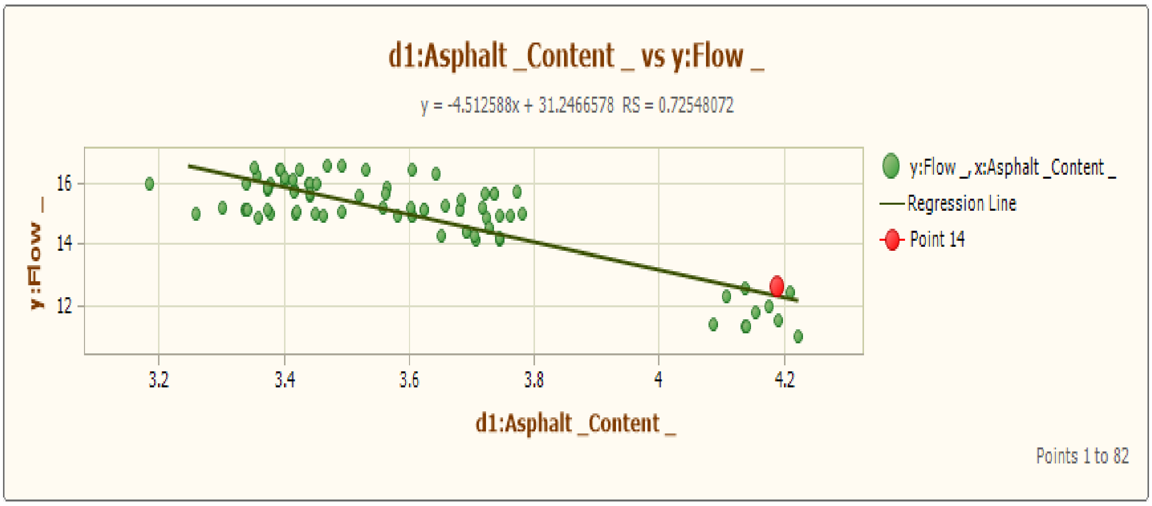

d4:Max Specific _Gravity of _M... vs y:Stability _
y = -1571.21157165x + 7142.73419154 RS = 0.00399187
y:Stability _
3000
2000
2.52
2.53
2.54
2.55
2.56
2.57
2.58
2.59
2.6
2.61
d4:Max Specific _Gravity of _Mix
Points 1 to 82

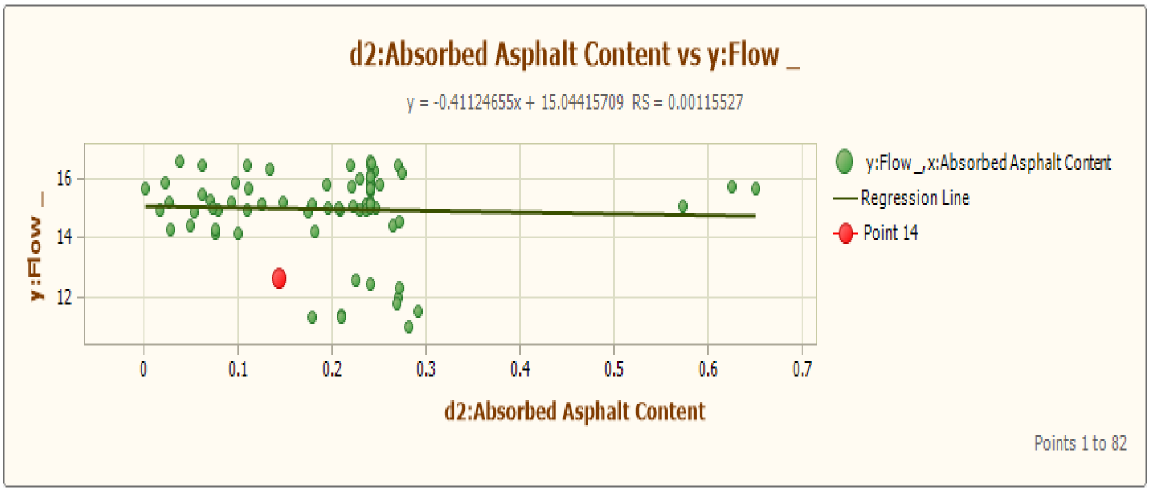

d5:Agg# Effective Specific Gra... vs y:Stability _
y = -6938.04128554x + 21955.53704627 RS = 0.10791798
y:Stability _
3000
2000
2.56
2.58
2.6
2.62
2.64
2.66
2.68
2.7
2.72
2.74
2.76
2.78
2.8
d5:Agg# Effective Specific Gravity
Points 1 to 82

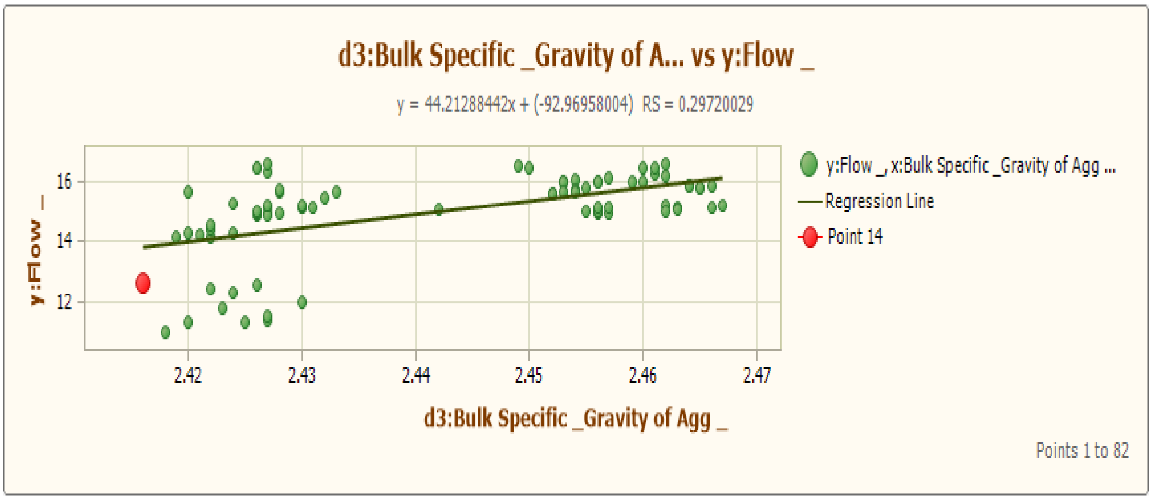


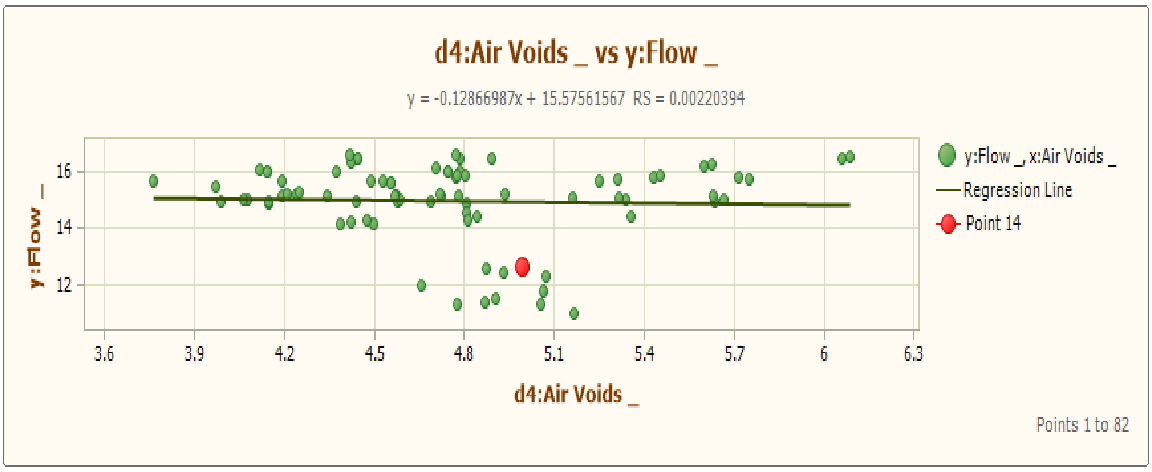
Appendix II: Input Variables vs Marshal Flow

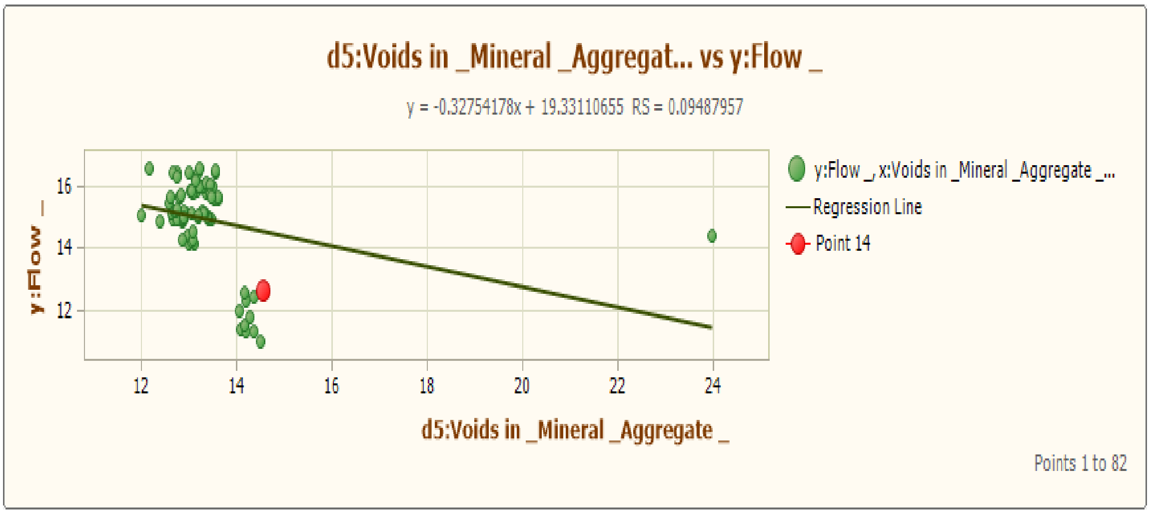


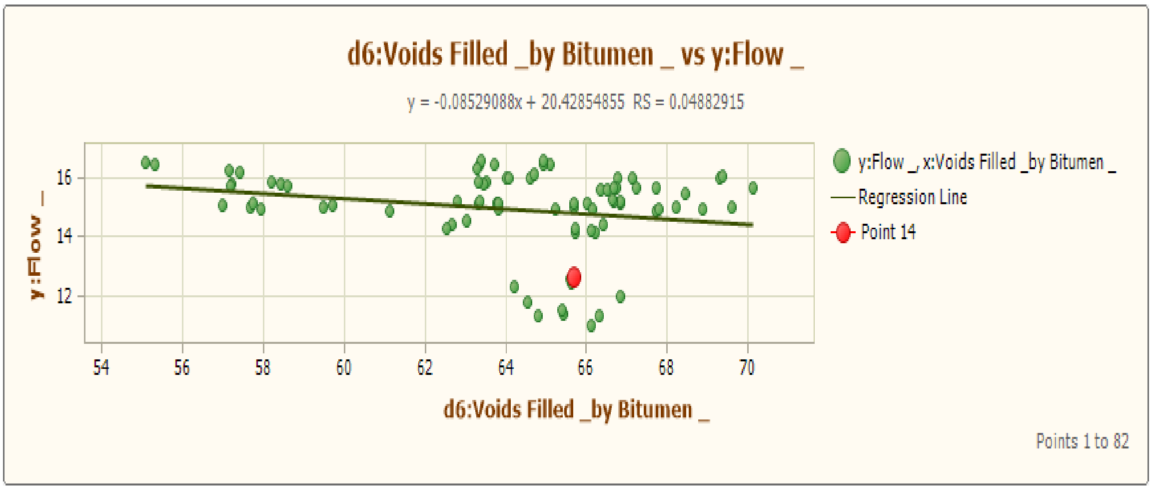



d2:Absorbed Asphalt Content vs y:Flow _
y = -0.41124655x + 15.04415709 RS = 0.00115527
y:Flow _
16
14
12
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
d2:Absorbed Asphalt Content
y:Flow _,x:Absorbed Asphalt Content
Regression Line
Point 14
Points 1 to 82


d3:Bulk Specific _Gravity of A... vs y:Flow _
y = 44.21288442x + (-92.96958004) RS = 0.29720029
y:Flow _
16
14
12
2.42
2.43
2.44
2.45
2.46
2.47
d3:Bulk Specific _Gravity of Agg _
y:Flow _, x:Bulk Specific _Gravity of Agg ...
Regression Line
Point 14
Points 1 to 82


d4:Air Voids _ vs y:Flow _
y = -0.12866987x + 15.57561567 RS = 0.00220394
y:Flow _
16
14
12
3.6
3.9
4.2
4.5
4.8
5.1
5.4
5.7
6
6.3
d4:Air Voids _
y:Flow _, x:Air Voids _
Regression Line
Point 14
Points 1 to 82


d5:Voids in _Mineral _Aggregat... vs y:Flow _
y = -0.32754178x + 19.33110655 RS = 0.09487957
y:Flow _
d5:Voids in _Mineral _Aggregate _
y:Flow _, x:Voids in _Mineral _Aggregate _...
Regression Line
Point 14
Points 1 to 82


d6:Voids Filled _by Bitumen _ vs y:Flow _
y = -0.08529088x + 20.42854855 RS = 0.04882915
y:Flow _
d6:Voids Filled _by Bitumen _
y:Flow _, x:Voids Filled _by Bitumen _
Regression Line
Point 14
Points 1 to 82

Appendix III: GEP performance in predicting Marshal Stability and Flow

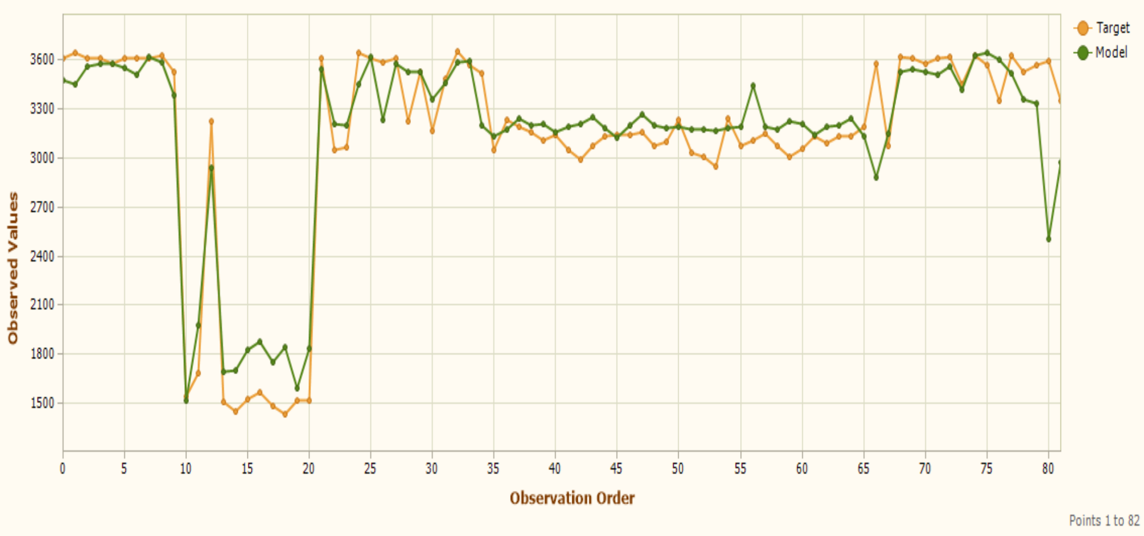


Figure 26: GEP performance while predicting Marshal Stability

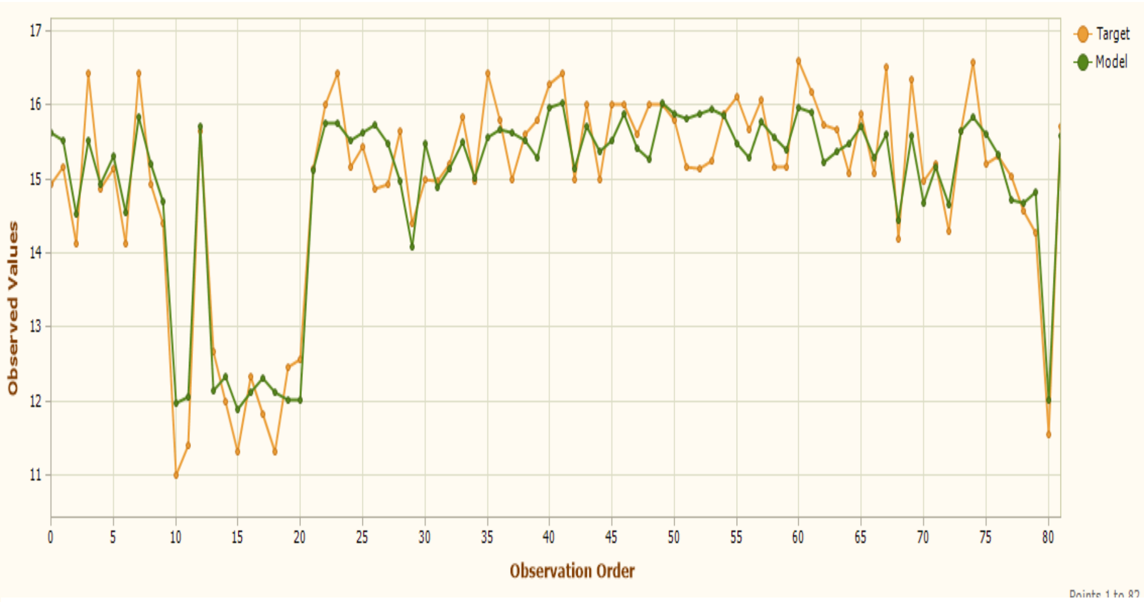


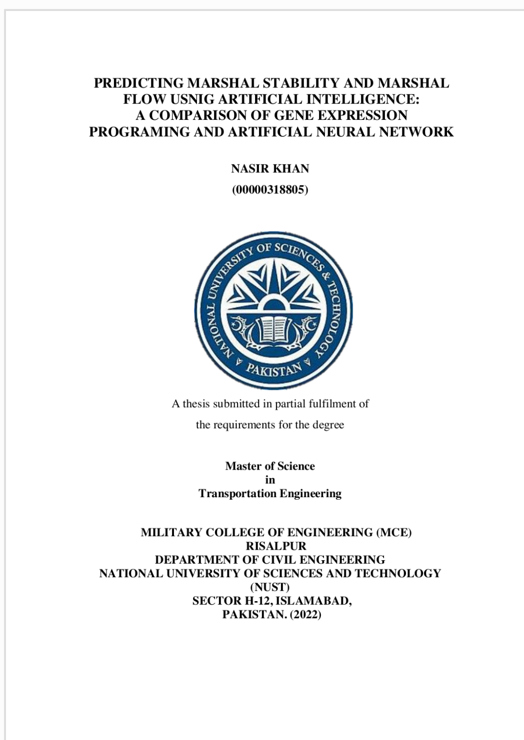
Figure 27: GEP performance while predicting Marshal Flow

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Nasir Khan
Assignment title:	06
Submission title:	Thesis
File name:	Thesis_Final.docx
File size:	1.72M
Page count:	72
Word count:	12,171
Character count:	63,423
Submission date:	21-Oct-2022 07:18AM (UTC-0700)
Submission ID:	1931551634


PREDICTING MARSHAL STABILITY AND MARSHAL FLOW USNIG ARTIFICIAL INTELLIGENCE: A COMPARISON OF GENE EXPRESSION PROGRAMING AND ARTIFICIAL NEURAL NETWORK

NASIR KHAN

(00000318805)

A thesis submitted in partial fulfilment of the requirements for the degree

Master of Science in Transportation Engineering

MILITARY COLLEGE OF ENGINEERING (MCE) RISALPUR

DEPARTMENT OF CIVIL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY (NUST)

SECTOR H-12, ISLAMABAD, PAKISTAN. (2022)

Thesis

ORIGINALITY REPORT

16%	11%	11%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	cyberleninka.org Internet Source	2%
2	Submitted to Higher Education Commission Pakistan Student Paper	1%
3	www.mdpi.com Internet Source	1%
4	Linyi Yao, Zhen Leng, Jiwang Jiang, Fujian Ni, Zili Zhao. "Nondestructive prediction of rutting resistance of in-service middle asphalt layer based on gene expression programing", Construction and Building Materials, 2021 Publication	1%
5	Submitted to Swinburne University of Technology Student Paper	<1%
6	Submitted to Laguna State Polytechnic University Student Paper	<1%

7	Behnood, Ali, Jan Olek, and Michal A. Glinicki. "Predicting modulus elasticity of recycled aggregate concrete using M5' model tree algorithm", Construction and Building Materials, 2015. Publication	<1 %
8	lib.dr.iastate.edu Internet Source	<1 %
9	Submitted to De Montfort University Student Paper	<1 %
10	dl.lib.uom.lk Internet Source	<1 %
11	www.biorxiv.org Internet Source	<1 %
12	www.udot.utah.gov Internet Source	<1 %
13	Handbook of Genetic Programming Applications, 2015. Publication	<1 %
14	Alireza Fallahpour, Kuan Yew Wong, Ezutah Udoncy Olugu, Siti Nurmaya Musa. "A Predictive Integrated Genetic-Based Model for Supplier Evaluation and Selection", International Journal of Fuzzy Systems, 2017 Publication	<1 %
	Submitted to Monash University	

15	Student Paper	<1 %
16	"Recent Developments in Sustainable Infrastructure (ICRDSI-2020)—GEO-TRA-ENV-WRM", Springer Science and Business Media LLC, 2022 Publication	<1 %
17	Weiguang Zhang, Adnan Khan, Ju Huyan, Jingtao Zhong, Tianyi Peng, Hanglin Cheng. "Predicting Marshall parameters of flexible pavement using support vector machine and genetic programming", Construction and Building Materials, 2021 Publication	<1 %
18	dergipark.ulakbim.gov.tr Internet Source	<1 %
19	www.researchgate.net Internet Source	<1 %
20	www.wispave.org Internet Source	<1 %
21	Hamed Majidifard, Behnam Jahangiri, Punyaslok Rath, Loreto Urra Contreras, William G. Buttlar, Amir H. Alavi. "Developing a prediction model for rutting depth of asphalt mixtures using gene expression programming", Construction and Building Materials, 2020	<1 %

Publication

22	dokumen.pub Internet Source	<1%
23	www.coursehero.com Internet Source	<1%
24	Alireza FALLAHPOUR, Atefeh AMINDOUST, Jurgita ANTUCHEVIČIENĖ, Morteza YAZDANI. "NONLINEAR GENETIC-BASED MODEL FOR SUPPLIER SELECTION: A COMPARATIVE STUDY", Technological and Economic Development of Economy, 2016 Publication	<1%
25	Submitted to The University of Manchester Student Paper	<1%
26	link.springer.com Internet Source	<1%
27	repository.tudelft.nl Internet Source	<1%
28	waberconference.com Internet Source	<1%
29	Muhammad Karami, Ainalem Nega, Ahdyeh Mosadegh, Hamid Nikraz. "Evaluation of Permanent Deformation of BRA Modified Asphalt Paving Mixtures Based on Dynamic Creep Test Analysis", Advanced Engineering Forum, 2016	<1%

Publication

30	ascelibrary.org Internet Source	<1 %
31	docs.neu.edu.tr Internet Source	<1 %
32	iopscience.iop.org Internet Source	<1 %
33	ir.unimas.my Internet Source	<1 %
34	www.emeraldinsight.com Internet Source	<1 %
35	recovery.preventionweb.net Internet Source	<1 %
36	Arash Pashazadeh, Mitra Javan. "Comparison of the gene expression programming, artificial neural network (ANN), and equivalent Muskingum inflow models in the flood routing of multiple branched rivers", Theoretical and Applied Climatology, 2019 Publication	<1 %
37	Submitted to Coventry University Student Paper	<1 %
38	Submitted to Jawaharlal Nehru Technological University Student Paper	<1 %

39 Kaval Chhabra, Divesh Agrawal, Saladi S. V. Subbarao. "chapter 12 Modeling of Polypropylene Modified Bitumen Mix Design Results Using Regression Analysis", IGI Global, 2017 <1%
Publication

40 biblioteca.versila.com <1%
Internet Source

41 coek.info <1%
Internet Source

42 www.igs.org.in:8080 <1%
Internet Source

43 H. Md. Azamathulla, Upaka Rathnayake, Ahmad Shatnawi. "Gene expression programming and artificial neural network to estimate atmospheric temperature in Tabuk, Saudi Arabia", Applied Water Science, 2018 <1%
Publication

44 Ozgan, E.. "Artificial neural network based modelling of the Marshall Stability of asphalt concrete", Expert Systems With Applications, 201105 <1%
Publication

45 transportproblems.polsl.pl <1%
Internet Source

46	Hamad Hassan Awan, Arshad Hussain, Muhammad Faisal Javed, Yanjun Qiu et al. "Predicting Marshall Flow and Marshall Stability of Asphalt Pavements Using Multi Expression Programming", Buildings, 2022 Publication	<1 %
47	Submitted to National University of Singapore Student Paper	<1 %
48	etheses.whiterose.ac.uk Internet Source	<1 %
49	upb.ro Internet Source	<1 %
50	Hassan, H.F.. "Potential uses of petroleum-contaminated soil in highway construction", Construction and Building Materials, 200510 Publication	<1 %
51	M. C. NATARAJA, M. A. JAYARAM, C. N. RAVIKUMAR. "KOHONEN'S FEATURE MAPS FOR FLY ASH CATEGORIZATION", International Journal of Neural Systems, 2011 Publication	<1 %
52	Submitted to University of Aberdeen Student Paper	<1 %
53	scholar.lib.vt.edu Internet Source	<1 %

text-id.123dok.com

54	Internet Source	<1 %
55	uwspace.uwaterloo.ca Internet Source	<1 %
56	www.ijeat.org Internet Source	<1 %
57	Mohammad Reza Mirzahosseini, Alireza Aghaeifar, Amir Hossein Alavi, Amir Hossein Gandomi, Reza Seyednour. "Permanent deformation analysis of asphalt mixtures using soft computing techniques", Expert Systems with Applications, 2011 Publication	<1 %
58	Sepehr Soleimani, Shabnam Rajaei, Pengcheng Jiao, Arash Sabz, Sina Soheilinia. "New prediction models for unconfined compressive strength of geopolymer stabilized soil using multi-gen genetic programming", Measurement, 2018 Publication	<1 %
59	ebin.pub Internet Source	<1 %
60	ir.uiowa.edu Internet Source	<1 %
61	kth.diva-portal.org Internet Source	<1 %

62	onlinepubs.trb.org Internet Source	<1 %
63	repository.nida.ac.th Internet Source	<1 %
64	scholarworks.utep.edu Internet Source	<1 %
65	ufdc.ufl.edu Internet Source	<1 %
66	www.geocities.com Internet Source	<1 %
67	www.graniteconstruction.com Internet Source	<1 %
68	"Applications of Cuckoo Search Algorithm and its Variants", Springer Science and Business Media LLC, 2021 Publication	<1 %
69	Ali Behnood, Dana Daneshvar. "A machine learning study of the dynamic modulus of asphalt concretes: An application of M5P model tree algorithm", Construction and Building Materials, 2020 Publication	<1 %
70	Mahmut Bilgehan. "A comparative study for the concrete compressive strength estimation using neural network and neuro-fuzzy	<1 %

modelling approaches", Nondestructive Testing and Evaluation, 2011
Publication

71	Mohammad Rezaei, Leila Hashemian, Alireza Bayat, Bryson Huculak. "Investigation of Rutting Resistance and Moisture Damage of Cold Asphalt Mixes", Journal of Materials in Civil Engineering, 2017 Publication	<1 %
72	Serkan Tapkın, Abdulkadir Çevik, Ün Uşar. "Prediction of Marshall test results for polypropylene modified dense bituminous mixtures using neural networks", Expert Systems with Applications, 2010 Publication	<1 %
73	dspace.nm-aist.ac.tz Internet Source	<1 %
74	eprints.usm.my Internet Source	<1 %
75	jestec.taylors.edu.my Internet Source	<1 %
76	techsofto.wordpress.com Internet Source	<1 %
77	"Advances in Data Mining. Applications and Theoretical Aspects", Springer Science and Business Media LLC, 2011 Publication	<1 %

78	Alavi, Amir H., Hassene Hasni, Nizar Lajnef, Karim Chatti, and Fred Faridazar. "An intelligent structural damage detection approach based on self-powered wireless sensor data", Automation in Construction, 2016. Publication	<1%
79	Ali Reza Ghanizadeh, Nasrin Heidarabadizadeh, Farhang Jalali. "Artificial neural network back-calculation of flexible pavements with sensitivity analysis using Garson's and connection weights algorithms", Innovative Infrastructure Solutions, 2020 Publication	<1%
80	Ankita Upadhya, M. S. Thakur, Parveen Sihag. "Predicting Marshall Stability of Carbon Fiber-Reinforced Asphalt Concrete Using Machine Learning Techniques", International Journal of Pavement Research and Technology, 2022 Publication	<1%
81	Benjamin J. Smith, Simon A. M. Hesp. "Crack Pinning in Asphalt Mastic and Concrete: Regular Fatigue Studies", Transportation Research Record: Journal of the Transportation Research Board, 2000 Publication	<1%
82	C.K. Lau, Kaushik Ghosh, M.A. Hussain, C.R. Che Hassan. "Fault diagnosis of Tennessee	<1%

Eastman process with multi-scale PCA and ANFIS", Chemometrics and Intelligent Laboratory Systems, 2013

Publication

---

83 Iftikhar Azim, Jian Yang, Muhammad Farjad Iqbal, Muhammad Faisal Javed, Sohaib Nazar, Feiliang Wang, Qing-feng Liu. "Semi-analytical model for compressive arch action capacity of RC frame structures", Structures, 2020 <1%

Publication

---

84 Mohamed Elkashef, R. Christopher Williams. "Improving fatigue and low temperature performance of 100% RAP mixtures using a soybean-derived rejuvenator", Construction and Building Materials, 2017 <1%

Publication

---

85 S.H.A. Kaboli, A. Fallahpour, N. Kazemi, J. Selvaraj, N.A. Rahim. "Electric energy consumption forecasting via expression-driven approach", 4th IET Clean Energy and Technology Conference (CEAT 2016), 2016 <1%

Publication

---

86 Singh, K.P.. "Artificial intelligence based modeling for predicting the disinfection by-products in water", Chemometrics and Intelligent Laboratory Systems, 20120515 <1%

Publication

87	Yasir Ali, Muhammad Irfan, Abdul Salam Buller, Hammad A. Khan, Hafiz M.F. Gul. "A binary logistic model for predicting the tertiary stage of permanent deformation of conventional asphalt concrete mixtures", Construction and Building Materials, 2019 Publication	<1 %
88	datospdf.com Internet Source	<1 %
89	digitalcommons.latech.edu Internet Source	<1 %
90	downloads.hindawi.com Internet Source	<1 %
91	epdf.pub Internet Source	<1 %
92	i-rep.emu.edu.tr:8080 Internet Source	<1 %
93	mafiadoc.com Internet Source	<1 %
94	mospace.umsystem.edu Internet Source	<1 %
95	ojs.imeti.org Internet Source	<1 %
96	ouci.dntb.gov.ua Internet Source	<1 %

97	trid.trb.org Internet Source	<1%
98	ucf.digital.flvc.org Internet Source	<1%
99	web.archive.org Internet Source	<1%
100	www.transport-links.org Internet Source	<1%
101	Fatai Anifowose, Jane Labadin, Abdulazeez Abdulraheem. "Ensemble model of Artificial Neural Networks with randomized number of hidden neurons", 2013 8th International Conference on Information Technology in Asia (CITA), 2013 Publication	<1%
102	Jian Liu, Fangyu Liu, Chuanfeng Zheng, Daodao Zhou, Linbing Wang. "Optimizing asphalt mix design through predicting effective asphalt content and absorbed asphalt content using machine learning", Construction and Building Materials, 2022 Publication	<1%
103	Jian Liu, Fangyu Liu, Hongren Gong, Ebenezer O. Fanijo, Linbing Wang. "Improving asphalt mix design by predicting alligator cracking and longitudinal cracking based on machine learning and dimensionality reduction	<1%

techniques", Construction and Building Materials, 2022
Publication

104 Mehran Mazari, Daniel D. Rodriguez. "Prediction of pavement roughness using a hybrid gene expression programming-neural network technique", Journal of Traffic and Transportation Engineering (English Edition), 2016 <1 %
Publication

105 eprints.uthm.edu.my <1 %
Internet Source

106 "Modeling the Effect of Filler Materials on Performance of Hot Mix Asphalt Using Genetic Programming", Airfield and Highway Pavements 2015, 2015. <1 %
Publication

107 "Proceedings of International Conference on Computer Vision and Image Processing", Springer Science and Business Media LLC, 2017 <1 %
Publication

108 Deepak Singh, Munendra Kumar. "Computation of energy dissipation across the type-A piano key weir by using gene expression programming technique", Water Supply, 2022 <1 %
Publication

109	Iftikhar Azim, Jian Yang, Muhammad Faisal Javed, Muhammad Farjad Iqbal, Zafar Mahmood, Feiliang Wang, Qing-feng Liu. "Prediction model for compressive arch action capacity of RC frame structures under column removal scenario using gene expression programming", Structures, 2020 Publication	<1%
110	Talal H. Fadhil, Rahel Khalid Ibrahim, Hardi Saadullah Fathullah. "The Influence of Curing Methods on Marshall Stability and Flow", IOP Conference Series: Materials Science and Engineering, 2020 Publication	<1%
111	Yinfei Du, Jiaqi Chen, Zheng Han, Weizheng Liu. "A review on solutions for improving rutting resistance of asphalt pavement and test methods", Construction and Building Materials, 2018 Publication	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		