

APPLICATION OF MACHINE LEARNING TECHNIQUES FOR ASPHALTIC MATERIAL MODELING

ABHARY ELEYEDATH



DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI

MARCH 2021

APPLICATION OF MACHINE LEARNING TECHNIQUES FOR ASPHALTIC MATERIAL MODELING

by

ABHARY ELEYEDATH

Department of Civil Engineering

submitted

In fulfilment of the requirements of the degree of Doctor of Philosophy

to the



Indian Institute of Technology Delhi

March 2021

DEDICATION

To my father, Velayudhan Eleyedath, for instigating the concept of dreams in me. To my mother, K. Sarojiny for nourishing me as her dream with her unwavering love and care, and to my brother, Azad Eleyedath for being the pillar of support and never allowing me to give up on my dreams. To my sister-in-law Sreerekha R for the support and understanding during the last lap of my research work. To my two-year-old niece Avani Azad for filling me with love and positivity whose presence was a blessing during the COVID 19 lockdown.

To my grandmother and other family members who helped me to inculcate discipline in my life and for encouraging me with their excitement.

Finally, to all my respected teachers for their valuable lessons, advises and blessings to make all my dreams come true.

CERTIFICATE

This is to certify that the thesis entitled “**Application of machine learning techniques for asphaltic material modeling**”, is being submitted by **Ms. Abhary Eleyedath** in partial fulfillment for the award of the degree of **Doctor of Philosophy** of the Indian Institute of Technology Delhi. This is a record of the research work and is entirely carried out by her under my supervision and guidance. The research report presented in this thesis has not been submitted for the award of any other degree or diploma.

Dr. Aravind Krishna Swamy
Associate Professor
Department of Civil Engineering
Indian Institute of Technology Delhi
New Delhi – 110016
India

ACKNOWLEDGEMENTS

The work presented in this thesis was carried out at the Department of Civil Engineering, Indian Institute of Technology Delhi under the supervision of Dr. Aravind Krishna Swamy. I am very much grateful to my Ph. D thesis supervisor Dr. Aravind Krishna Swamy for his invaluable suggestions, constant support and encouragement during the course of my research work. I feel proud working with him. I gained both field and laboratory knowledge during discussions with him. I like his attitude, honesty, and frankness. He always gave me freedom to do mistakes which apparently gave me confidence to grow as a researcher. Irrespective of the situation, he always greeted me with a smile which solved half of the problems. He was the source of constant motivation which helped me to carry out my research work in the right direction. I look up to him as he inspires me to become a better individual.

I am deeply grateful to my co-author Prof. G. V. Ramana, Head of the Department, Department of Civil Engineering, Indian Institute of Technology Delhi for his words of encouragement, and support throughout the research work and more than that for his friendly nature.

I would like to express my gratitude to the SRC members: Prof. A.K. Jain, Prof. Geetam Tiwari and Prof. Parag Singla for their valuable inputs at different stages of thesis work and for improving the quality of the work.

I pay my deep sense of gratitude to Prof. Kalaga Ramachandra Rao and Prof. Manoj M for their suggestions during progress presentations which helped me to improve my research work.

I also extend my gratitude to my co-author Dr. Siksha Swaroopa Kar, Senior Scientist, CSIR Central Road Research Institute, New Delhi.

I wish to acknowledge the support received from my senior colleagues and lab mates. I also owe my gratitude to my colleague Iqra Altaf Gillani for having wonderful discussions on machine learning. I also extend sincere thanks to my hostel mates, and well-wishers for their suggestions and help in completing the research. I further honestly thank many people who supported me directly/indirectly and whom I had encountered during this journey whose names are not mentioned here.

I would like to express gratitude to my family members, without whom my dream of completing a PhD would not have come true. Particularly thanks to my brother Azad Eleyedath, for providing me encouragement, advises and endless support throughout this journey. I am so proud of my family. Finally, I would like to thank God for making me believe in miracles.

ABHARY ELEYEDATH

ABSTRACT

It is well known fact that determination of properties of asphaltic materials in laboratory is laborious, expensive, time and resource consuming exercise. Hence, researchers have resorted to use of tools that provide a meaningful predictive model for material response. Within the domain of asphaltic materials, there are very few models available for predicting the material response. Owing to predictive capability, researchers/engineers have resorted to use of machine learning schemes to develop predictive models. However, most of the approaches/schemes are black box in nature. Under these circumstances, Gene Expression Programming (GEP) comes handy. GEP is a non-parametric technique that uses a symbolic regression approach to optimise the function itself. One of the main advantage of this technique is availability of equation that can be used as predictive model. The present work proposes GEP based models for predicting response of different asphaltic materials. These include (i) prediction of effective viscosity of binary blends of asphalt binders, (ii) prediction of density and viscosity of heavy crudes, (iii) prediction of foamed bitumen properties, and (iv) dynamic modulus of asphalt concrete. Several improvements have been suggested to improve prediction capability of GEP based approach. The same has been discussed in following paragraphs.

The first objective was to develop a viscosity mixing rule for binary blend of asphalt binders using GEP. This work presents a novel GEP based approach to obtain a viscosity-mixing rule. To develop these expressions, two distinct binary blends comprising of varying proportion of unmodified binder and polymer modified binder were prepared and tested for their resultant viscosity at different temperatures. The obtained data were used to (i) develop the GEP-based viscosity-mixing rule, and (ii) compare the resultant viscosity with viscosity-mixing rules reported in the literature (like the Arrhenius model). Statistical analysis indicated that the accuracy of GEP based mixing rule is superior to other viscosity-mixing rules reported in the literature.

The prediction of density and viscosity of heavy bitumen using GEP was chosen as the second objective. To evaluate the accuracy of proposed GEP based models, results reported by various researchers were utilized. Published literature on heavy bitumen including Athabasca, Cold lake and Gas free bitumen were used for this purpose. Using temperature, and pressure data as input, the density and viscosity of bitumen were predicted using GEP technique. The developed GEP based models were compared with the conventional empirical regression equations proposed by others. The statistical analysis indicated that GEP based models work better than currently used models for density and viscosity of bitumen.

The foamed bitumen parameters (i.e., *ER*, *HL*) are highly influenced by viscosity, foaming temperature and water content. With no predictive model that relates these variables with foaming parameters, development of predictive model was taken as third objective. The database consisting of *ER* and *HL*, and physico-chemical properties of six distinct binders was developed by research team at Central Road Research Institute (CRRI), New Delhi. All binders were tested for their physical properties and chemical composition, and the *ER* and *HL* were measured over the range of test temperatures (110°C to 200°C) and water contents (2% to 11%). The final database consisted of 166 observations. Expressions to predict *ER* and *HL* have been developed based on temperature, water content, viscosity, and chemical composition using this database through the GEP technique. The goodness of fit parameters indicated that models based on physical properties are able to predict *ER* and *HL* with reasonable accuracy and with the addition of chemical composition along with physical properties improved the accuracy of predictive models significantly. These expressions can be used to identify the feasible range of test conditions for the production of foamed bitumen before actual testing is commenced.

The fourth objective was to develop a novel hybrid clustering- GEP approach for predicting foamed bitumen properties. A database consisting of 190 observations was used. Essentially this database was obtained by addition of more binder data to database used with third objective. The Self-Organizing Map (SOM) based clustering of this database helped in obtaining homogeneous groups even with highly complex interaction. Further, C5.0 algorithm was used to decipher underlying patterns among clusters identified by SOM. GEP approach was used to develop four global models to predict *HL* and *ER*. Subsequently hybrid models were obtained through recalibration of these global models but using data from individual clusters. These models were different than those obtained while working on third objective. Major differences include (i) different functional form, (ii) additional model for *HL* in terms of *ER*, and (iii) increased accuracy with more information supplied. Statistical analysis indicated that hybrid models outperformed corresponding global models in all cases. Global sensitivity analysis indicated that among various parameters, water content had significant effect on *ER* prediction. This was followed by temperature, and viscosity. However, for predicting *HL*, this order was *ER* (if used), water content, temperature, and viscosity.

The fifth objective of research presents a novel hybrid Principal Component Analysis (PCA) - GEP approach to predict the $|E^*|$ of asphalt concrete. For this purpose, $|E^*|$ database developed during NCHRP 9-19 study was used. Using the information of all properties as input

(i.e., variables), the dimensionality was reduced using PCA. Such an exercise helped in removing the redundancy at input stage. The extracted principal components were used to develop first set of $|E^*|$ prediction models. The feature selection property of PCA was used to decide the parameters mostly contributing to the individual principal components (PC 's) and rank the same. Using the ranked variables as input, second set of $|E^*|$ prediction models were developed. Careful analysis of these two sets of models indicated that addition of parameters (or PC 's) beyond certain number did not contribute to accuracy of model. Comparison of models from these two sets indicated that predictive model obtained using variables as direct input resulted in improved accuracy. For better understanding, this finalized model was compared with other regression-based equations proposed by others. Comparison of goodness of fit indicators indicated that proposed hybrid model offers efficient and accurate alternative. The proposed model has flexibility to be used with any new database with recalibration.

The sixth (and final) objective of work evaluates the underlying patterns in prediction error that exists in these models. Again $|E^*|$ database developed during NCHRP 1-40D study was used for this purpose. Initially, the global dataset was compared against individual mixtures for similarity in terms of mean, range, and correlation. T-test showed that the individual mixture datasets are significantly different from the global dataset. Thus, calibration of predictive models was undertaken using global dataset and individual mixture dataset. Statistical indicators indicated improvement with mixture-wise calibration when compared to global calibration. To check the extent of prediction error, 'difference parameter' (DP) and 'ratio parameter' (RP) was introduced in this work. Q-Q plots and cumulative distribution plots constructed using DP and RP indicated highest and lowest error with Al-Khateeb and PCA-GEP models, respectively. For a detailed analysis, the entire range of measured $|E^*|$ was divided into subdivisions. In general, lower prediction error was observed in middle range of measured $|E^*|$. All the predictive models considered in this study overpredicted and underpredicted in lower and higher modulus range, respectively. Also, in the lower and higher modulus region, DP and RP showed skewed and flatter distribution when compared to normal distribution. Based on the numerical values of these distribution indicators, it was inferred that performance of all models are comparable in middle $|E^*|$ range. However, in extreme values of $|E^*|$, Hirsch and Al-Khateeb models performed poorly. In the same range, PCA-GEP, Original Wiczak and South Korean models performed well. Sensitivity analysis indicted that binder properties exhibited highest sensitivity followed by testing condition, aggregate gradation and volumetrics.

सार

यह सर्वविदित तथ्य है कि प्रयोगशाला में डामर सामग्री के गुणों का निर्धारण श्रमसाध्य, महंगा, समय और संसाधन खपत व्यायाम है। इसलिए, शोधकर्ताओं ने उन उपकरणों के उपयोग का सहारा लिया है जो भौतिक प्रतिक्रिया के लिए एक सार्थक भविष्य कहनेवाला मॉडल प्रदान करते हैं। डामर सामग्री के डोमेन के भीतर, सामग्री प्रतिक्रिया की भविष्यवाणी करने के लिए बहुत कम मॉडल उपलब्ध हैं। भविष्य कहनेवाला क्षमता के कारण, शोधकर्ताओं / इंजीनियरों ने भविष्य कहनेवाला मॉडल विकसित करने के लिए मशीन लर्निंग योजनाओं का उपयोग किया है। हालांकि, अधिकांश दृष्टिकोण / योजनाएं प्रकृति में ब्लैक बॉक्स हैं। इन परिस्थितियों में, जीन एक्सप्रेसन प्रोग्रामिंग (GEP) काम आता है। GEP एक गैर-पैरामीट्रिक तकनीक है जो फ़ंक्शन को अनुकूलित करने के लिए प्रतीकात्मक प्रतिगमन दृष्टिकोण का उपयोग करती है। इस तकनीक का एक मुख्य लाभ समीकरण की उपलब्धता है जिसका उपयोग भविष्य कहनेवाला मॉडल के रूप में किया जा सकता है। वर्तमान कार्य विभिन्न डामर पदार्थों की प्रतिक्रिया की भविष्यवाणी के लिए जीईपी आधारित मॉडल का प्रस्ताव करता है। इनमें (i) डामर बाँधने वालों की द्विआधारी मिश्रणों की प्रभावी चिपचिपाहट की भविष्यवाणी, (ii) घनत्व की घनत्व और भारी कूड़ों की चिपचिपाहट, (iii) फोमेड कोलतार गुणों की भविष्यवाणी, और (iv) डामर कंक्रीट के गतिशील छींटे शामिल हैं। जीईपी आधारित दृष्टिकोण की भविष्यवाणी क्षमता में सुधार के लिए कई सुधारों का सुझाव दिया गया है। निम्नलिखित पैराग्राफ में उसी पर चर्चा की गई है।

पहला उद्देश्य GEP का उपयोग करके डामर बाइंडरों के द्विआधारी मिश्रण के लिए एक चिपचिपापन मिश्रण नियम विकसित करना था। यह काम एक चिपचिपापन-मिश्रण शासन प्राप्त करने के लिए एक उपन्यास GEP आधारित दृष्टिकोण प्रस्तुत करता है। इन अभिव्यक्तियों को विकसित करने के लिए, दो अलग-अलग द्विआधारी मिश्रणों में अनमॉडिफाइड बाइंडर और पॉलीमर संशोधित बाइंडर के अलग-अलग अनुपात शामिल हैं, जिन्हें विभिन्न तापमानों पर उनके परिणामी चिपचिपाहट के लिए तैयार और परीक्षण किया गया था। प्राप्त डेटा का उपयोग (i) GEP- आधारित चिपचिपापन-मिश्रण नियम विकसित करने के लिए किया गया था, और (ii) साहित्य में बताए गए चिपचिपापन-मिश्रण नियमों (जैसे अरहेनियस मॉडल) के साथ परिणामी चिपचिपाहट की तुलना करते हैं। सांख्यिकीय विश्लेषण ने संकेत दिया कि GEP आधारित मिश्रण नियम की सटीकता साहित्य में बताए गए अन्य चिपचिपापन-मिश्रण नियमों से बेहतर है।

जीईपी का उपयोग करते हुए भारी कोलतार की घनत्व और चिपचिपाहट की भविष्यवाणी को दूसरे उद्देश्य के रूप में चुना गया था। प्रस्तावित GEP आधारित मॉडल की सटीकता का मूल्यांकन करने के लिए, विभिन्न शोधकर्ताओं द्वारा रिपोर्ट किए गए परिणामों का उपयोग किया गया था। अथाबास्का, कोल्ड लेक और गैस मुक्त कोलतार सहित भारी कोलतार पर प्रकाशित साहित्य का उपयोग इस उद्देश्य के लिए किया गया था। तापमान और इनपुट डेटा को इनपुट के रूप में उपयोग करते हुए, जीईपी तकनीक का उपयोग करके बिटुमेन की घनत्व और चिपचिपाहट की भविष्यवाणी की गई थी। विकसित GEP आधारित मॉडल की तुलना दूसरों द्वारा प्रस्तावित पारंपरिक अनुभवजन्य प्रतिगमन समीकरणों के साथ की गई थी। सांख्यिकीय विश्लेषण ने संकेत दिया कि जीईपी आधारित मॉडल बिटुमेन के घनत्व और चिपचिपाहट के लिए वर्तमान में उपयोग किए गए मॉडल की तुलना में बेहतर काम करते हैं।

फोमेड बिटुमेन पैरामीटर (यानी, ईआर, एचएल) चिपचिपापन, झाग तापमान और पानी की सामग्री से अत्यधिक प्रभावित होते हैं। कोई भविष्य कहनेवाला मॉडल नहीं है जो इन चरों को फोमिंग मापदंडों के साथ संबंधित करता है, पूर्वानुमानात्मक मॉडल का विकास तीसरे उद्देश्य के रूप में लिया गया था। ईआर और एचएल, और छह अलग-अलग बाइंडरों के भौतिक-रासायनिक गुणों वाले डेटाबेस को केंद्रीय सड़क अनुसंधान संस्थान (सीआरआरआई), नई दिल्ली में अनुसंधान टीम द्वारा विकसित किया गया था। सभी बाइंडरों को उनके भौतिक गुणों और रासायनिक संरचना के लिए परीक्षण किया गया था, और ईआर और एचएल को परीक्षण तापमान (110 डिग्री सेल्सियस से 200 डिग्री सेल्सियस) और पानी की सामग्री (2% से 11%) की सीमा पर मापा गया था। अंतिम डेटाबेस में 166 अवलोकन शामिल थे। जीईपी तकनीक के माध्यम से इस डेटाबेस का उपयोग करके तापमान, पानी की सामग्री, चिपचिपाहट और रासायनिक संरचना के आधार पर ईआर और एचएल की भविष्यवाणी करने की अभिव्यक्तियाँ विकसित की गई हैं। फिट मापदंडों की अच्छाई ने संकेत दिया कि भौतिक गुणों पर आधारित मॉडल उचित सटीकता के साथ ईआर और एचएल की भविष्यवाणी करने में सक्षम हैं और रासायनिक गुणों के साथ-साथ भौतिक गुणों के साथ-साथ अनुमानित मॉडल की सटीकता में काफी सुधार हुआ है। इन अभिव्यक्तियों का उपयोग वास्तविक परीक्षण शुरू होने से पहले फोमेड कोलतार के उत्पादन के लिए परीक्षण स्थितियों की व्यवहार्य श्रेणी की पहचान करने के लिए किया जा सकता है।

चौथा उद्देश्य एक उपन्यास हाइब्रिड क्लस्टरिंग-जीईपी दृष्टिकोण विकसित करना था जो कि फोमेड कोलतार गुणों की भविष्यवाणी के लिए था। 190 अवलोकनों वाले डेटाबेस का उपयोग किया गया था। मूल रूप

से यह डेटाबेस तीसरे उद्देश्य के साथ उपयोग किए जाने वाले डेटाबेस के लिए अधिक बाइंडर डेटा के अलावा प्राप्त किया गया था। इस डेटाबेस के स्व-आयोजन मानचित्र (SOM) आधारित क्लस्टरिंग ने अत्यधिक जटिल बातचीत के साथ सजातीय समूहों को भी प्राप्त करने में मदद की। इसके अलावा, C5.0 एल्गोरिथ्म SOM द्वारा पहचाने गए समूहों के बीच अंतर्निहित पैटर्न को समझने के लिए उपयोग किया गया था। जीएलपी दृष्टिकोण का उपयोग एचएल और ईआर की भविष्यवाणी करने के लिए चार वैश्विक मॉडल विकसित करने के लिए किया गया था। इसके बाद हाइब्रिड मॉडल इन वैश्विक मॉडल के पुनः स्थिरीकरण के माध्यम से प्राप्त किए गए थे, लेकिन व्यक्तिगत समूहों से डेटा का उपयोग करते हुए। ये मॉडल तीसरे उद्देश्य पर काम करते समय प्राप्त किए गए लोगों से अलग थे। प्रमुख अंतर में (i) विभिन्न कार्यात्मक रूप शामिल हैं, (ii) ER के संदर्भ में HL के लिए अतिरिक्त मॉडल, और (iii) आपूर्ति की गई अधिक जानकारी के साथ सटीकता में वृद्धि। सांख्यिकीय विश्लेषण ने संकेत दिया कि हाइब्रिड मॉडल सभी मामलों में संबंधित वैश्विक मॉडल से बेहतर प्रदर्शन करते हैं। वैश्विक संवेदनशीलता विश्लेषण ने संकेत दिया कि विभिन्न मापदंडों के बीच, ईआर भविष्यवाणी पर जल सामग्री का महत्वपूर्ण प्रभाव था। इसके बाद तापमान, और चिपचिपाहट थी। हालांकि, एचएल की भविष्यवाणी करने के लिए, यह आदेश ईआर (यदि इस्तेमाल किया गया), पानी की सामग्री, तापमान और चिपचिपाहट था।

शोध का पाँचवाँ उद्देश्य एक उपन्यास हाइब्रिड प्रिंसिपल कंपोनेंट एनालिसिस (PCA) प्रस्तुत करता है - GEP दृष्टिकोण की भविष्यवाणी करने के लिए $|E^*|$ डामर कंक्रीट की। इस प्रयोजन के लिए $|E^*|$ NCHRP 9-19 अध्ययन के दौरान विकसित डेटाबेस का उपयोग किया गया था। इनपुट (यानी, चर) के रूप में सभी गुणों की जानकारी का उपयोग करते हुए, पीसीए का उपयोग करके आयामीता कम हो गई थी। इस तरह के अभ्यास से इनपुट स्टेज पर अतिरेक को दूर करने में मदद मिली। निकाले गए प्रमुख घटकों का उपयोग पहले सेट के लिए किया जाता था। $|E^*|$ भविष्यवाणी मॉडल। पीसीए की सुविधा चयन संपत्ति का उपयोग ज्यादातर व्यक्तिगत प्रिंसिपल घटकों (पीसी) में योगदान करने वाले मापदंडों को तय करने और उसी को रैंक करने के लिए किया गया था। इनपुट के रूप में रैंक किए गए चर का उपयोग करना, का दूसरा सेट $|E^*|$ भविष्यवाणी मॉडल विकसित किए गए थे। मॉडल के इन दो सेटों के सावधानीपूर्वक विश्लेषण ने संकेत दिया कि निश्चित संख्या से परे मापदंडों (या पीसी) ने मॉडल की सटीकता में योगदान नहीं दिया। इन दो सेटों से मॉडलों की तुलना ने संकेत दिया कि प्रत्यक्ष इनपुट के रूप में चर का उपयोग करके प्राप्त होने वाले पूर्वानुमान मॉडल में सटीकता में सुधार हुआ। बेहतर समझ के लिए, इस अंतिम मॉडल की तुलना दूसरों द्वारा प्रस्तावित अन्य प्रतिगमन-आधारित समीकरणों के साथ की गई थी। फिट संकेतकों की अच्छाई की तुलना से संकेत

मिलता है कि प्रस्तावित हाइब्रिड मॉडल कुशल और सटीक विकल्प प्रदान करता है। प्रस्तावित मॉडल में लचीलेपन के साथ किसी भी नए डेटाबेस का उपयोग किया जाना है।

काम का छठा (और अंतिम) उद्देश्य भविष्यवाणी त्रुटि में अंतर्निहित पैटर्न का मूल्यांकन करता है जो इन मॉडलों में मौजूद है। फिर से $|E^*|$ NCHRP 1-40D अध्ययन के दौरान विकसित डेटाबेस का उपयोग इस उद्देश्य के लिए किया गया था। प्रारंभ में, वैश्विक डेटासेट की तुलना माध्य, श्रेणी और सहसंबंध के मामले में समानता के लिए व्यक्तिगत मिश्रण के खिलाफ की गई थी। टी-टेस्ट से पता चला कि व्यक्तिगत मिश्रण डेटासेट वैश्विक डेटासेट से काफी अलग हैं। इस प्रकार, वैश्विक डेटासेट और व्यक्तिगत मिश्रण डेटासेट का उपयोग करके पूर्वानुमान मॉडल का अंशांकन किया गया था। सांख्यिकीय संकेतकों ने वैश्विक अंशांकन की तुलना में मिश्रण-वार अंशांकन के साथ सुधार का संकेत दिया। भविष्यवाणी की त्रुटि की जांच करने के लिए, इस कार्य में 'डिफरेंस पैरामीटर' (DP) और 'अनुपात पैरामीटर' (RP) को पेश किया गया था। डीपी और आरपी का उपयोग करके निर्मित क्यू-क्यू भूखंड और संचयी वितरण भूखंडों ने क्रमशः अल-खतीब और पीसीए-जीईपी मॉडल के साथ उच्चतम और निम्नतम त्रुटि का संकेत दिया। एक विस्तृत विश्लेषण के लिए, मापा की पूरी श्रृंखला $|E^*|$ उपखंडों में विभाजित किया गया था। सामान्य तौर पर, मापा गया मध्यम श्रेणी में निम्न पूर्वानुमान त्रुटि देखी गई। $|E^*|$ इस अध्ययन में विचार किए गए सभी पूर्वानुमान मॉडल क्रमशः कम और उच्च मापांक रेंज में ओवररेटेड और अंडररेटेड हैं। इसके अलावा, निचले और उच्च मापांक क्षेत्र में, डीपी और आरपी ने सामान्य वितरण की तुलना में तिरछा और चापलूसी वितरण दिखाया। इन वितरण संकेतकों के संख्यात्मक मूल्यों के आधार पर, यह अनुमान लगाया गया था कि सभी मॉडलों का प्रदर्शन मध्य में तुलनीय है। $|E^*|$ सीमा। हालाँकि, चरम मूल्यों में $|E^*|$, हिर्श और अल-खतीब मॉडल ने खराब प्रदर्शन किया। उसी रेंज में, PCA-GEP, ओरिजिनल विटकैक और साउथ कोरियन मॉडल्स ने अच्छा प्रदर्शन किया। संवेदनशीलता विश्लेषण ने संकेत दिया कि बांधने वाली गुण परीक्षण स्थिति, समग्र उन्नयन और वॉल्यूमेट्रिक्स के बाद उच्चतम संवेदनशीलता प्रदर्शित करते हैं।

TABLE OF CONTENTS

CERTIFICATE	ii
ACKNOWLEDGEMENTS.....	iii
ABSTRACT	iv
सार	vii
TABLE OF CONTENTS.....	xi
List of Figures	xvi
List of Tables.....	xvii
List of notations	xviii
CHAPTER 1	1
INTRODUCTION.....	1
1.1 Background.....	1
1.2 Motivation	4
1.3 Objectives	5
1.4 Scope of the research.....	5
1.5 Research Methodology	6
1.6 Organization of Thesis.....	8
CHAPTER 2	9
BACKGROUND.....	9
2.1 Introduction	9
2.2 Mechanistic modeling of asphaltic materials	9
2.3 Regression based predictive models.....	10
2.4 Machine Learning Techniques	11
2.4.1 Artificial neural network (ANN)	12
2.4.2 Self-Organizing Map.....	14
2.4.3 Support vector machine (SVM).....	16

2.4.4 Genetic programming (GP)	18
2.5 Gene expression programming (GEP).....	19
2.6 Goodness of Fit indicators.....	23
2.7 Parametric study and sensitivity analysis	25
2.8 Concluding remarks.....	26
CHAPTER 3	27
VISCOSITY-MIXING RULE FOR ASPHALT BLENDS	27
3.1 Introduction	27
3.2 Background.....	28
3.3 GEP for Viscosity-Mixing Rule	29
3.4 Materials, Blend Preparation, and Testing	30
3.5 Results and discussions	32
3.5.1. Comparison of statistical performance of models	34
3.5.2 Parametric study and sensitivity analysis	35
3.6 Application	36
3.7 Concluding remarks.....	37
CHAPTER 4	38
DENSITY AND VISCOSITY OF BITUMEN.....	38
4.1 Introduction	38
4.2 Background.....	39
4.3 Data.....	41
4.4 Methodology Adopted.....	42
4.5 Predictive models developed using GEP.....	43
4.6 Performance analysis and comparison of models.....	43
4.7 Parametric study of developed models	47
4.8 Sensitivity analysis	50
4.9 Concluding remarks.....	50

CHAPTER 5	51
FOAMED BITUMEN CHARACTERIZATION	51
5.1 Introduction	51
5.2 Background.....	53
5.2.1. <i>Bitumen properties</i>	53
5.3 Raw Materials, Testing, and Properties.....	56
5.4 Development of predictive equations using GEP	61
5.5 Results and discussion	62
5.5.1. <i>Preliminary analysis</i>	62
5.5.2. <i>Predictive equations relating with physical properties</i>	73
5.5.3. <i>Predictive equations relating with SARA fractions</i>	73
5.5.4. <i>Predictive equations relating with FTIR bond indices</i>	74
5.5.5. <i>Multivariable Least Square Regression Model</i>	75
5.5.6. <i>Evaluation of accuracy of predictive equations</i>	76
5.6. Concluding remarks.....	79
CHAPTER 6	82
SOM-GEP MODEL FOR FOAMED BITUMEN CHARACTERISTICS.....	82
6.1 Introduction	82
6.2 Theoretical Background	87
6.2.1 <i>Decision Tree</i>	87
6.3 Asphalt Binder Properties and Testing.....	88
6.4 Methodology.....	90
6.5 Results and Discussion	95
6.5.1 <i>Viscosity-Temperature Relationship</i>	95
6.5.2 <i>Preliminary Correlation Analysis</i>	96
6.5.3 <i>SOM based Cluster Analysis</i>	98
6.5.4 <i>Development of Decision Tree</i>	99

6.5.5 GEP based Predictive Models	102
6.5.6 Multivariable Least Square Regression Model.....	104
6.5.7 Accuracy of Predictive Equations and Decision Trees.....	105
6.5.8 Global Sensitivity Analysis	108
6.6 Practical Application	117
6.7 Concluding remarks.....	118
CHAPTER 7	120
DYNAMIC MODULUS OF ASPHALT CONCRETE.....	120
7.1 Introduction	120
7.2 Theoretical background	130
7.2.1 Principal Component Analysis (PCA)	130
7.3 Database description.....	132
7.4 Methodology.....	134
7.5 Results and discussion	135
7.5.1 Preliminary analysis	135
7.5.2 Principal Component analysis results	137
7.5.3 Comparison of PCA-GEP model with other models.....	144
7.5.4 Parametric study of PCA-GEP model	147
7.5.5 Sensitivity analysis	150
7.6 Concluding remarks.....	150
CHAPTER 8	152
SYSTEMATIC ERROR IN DYNAMIC MODULUS MODELS.....	152
8.1 Introduction	152
8.2 Database Description.....	156
8.3 Methodology.....	157
8.4 Results and Discussion	161
8.4.1 Comparative analysis.....	163

8.4.2 Calibration of predictive models	169
8.4.3 Evaluation of extent of prediction error	173
8.4.4 Evaluation of underlying patterns in prediction error	180
8.4.5 Global sensitivity analysis	190
8.5 Concluding remarks.....	192
CHAPTER 9	195
CONCLUSIONS.....	195
9.1 Introduction	195
9.2 Contributions of research.....	200
9.3 Recommendations	200
Annexure I	202
References	205
Publications during Ph. D program.....	220
BIODATA	222

List of Figures

Figure 1.1: Energy consumption by various asphalt mixture production technologies	3
Figure 1.2: Methodology adopted in current research work	7
Figure 2.1: Process flow in artificial neural network	13
Figure 2.2: Multilayer feed forward artificial neural network	13
Figure 2.3: Self-organizing map clustering process	15
Figure 2.4: Support vector identification	16
Figure 2.5: Process flow in support vector machine	17
Figure 2.6: Expression Tree (ET) for Arrhenius equation	20
Figure 2.7: Process flow in gene expression programming	22
Figure 3.1: Comparison of predicted viscosity using various models	33
Figure 3.2: Comparison of predicted viscosity using various models	36
Figure 4.1: Comparison of predicted density against measured density	44
Figure 4.2: Comparison of predicted viscosity against measured viscosity	46
Figure 4.3 Parametric analysis of the GEP based density model.....	49
Figure 5.1: Foamed bitumen decay curve	53
Figure 5.2: Viscosity-temperature relationship of binders	58
Figure 5.3: Effect of viscosity and water content on expansion ratio and half-life	63
Figure 5.4: Effect of SARA fractions on foamed bitumen characteristics.....	68
Figure 5.5: Scatterplots relating bond indices with foamed bitumen characteristics.....	72
Figure 5.6: Comparison of measured with predicted foamed bitumen characteristics	78
Figure 6.1: Methodology adopted for the work	91
Figure 6.2: Effect of viscosity and temperature on Foamed bitumen characteristics	97
Figure 6.3: Correlation matrix plot of variables.....	98
Figure 6.4: Influence of variables in the clustering.....	99
Figure 6.5: Decision trees for finding clusters while predicting Expansion Ratio	100
Figure 6.6: Decision trees for finding clusters while predicting Half-Life	101
Figure 6.7: Comparison of predicted with measured foamed bitumen characteristics	107
Figure 6.8: Decision trees for finding clusters while predicting Expansion Ratio	113
Figure 6.9: Decision trees for finding clusters while predicting Half-Life	113
Figure 6.10: Comparison of predicted and measured foamed bitumen characteristics	117
Figure 7.1: Principal component analysis process flow	131
Figure 7.2: Methodology adopted in present study.....	134
Figure 7.3: Correlation matrix obtained with the database	136
Figure 7.4: Variance explained by principal components.....	138
Figure 7.5: Variation of fitness values	140
Figure 7.6: Representation of principal component loadings	141
Figure 7.7: Comparison of PCA-GEP based model with other models.....	145
Figure 7.8: Parametric study of PCA-GEP based model	150
Figure 8.1: Methodology adopted in present study.....	157
Figure 8.2: Comparison of properties with mixture-wise and global dataset	168
Figure 8.3: Comparison of predicted with measured $ E^* $ from various models	172
Figure 8.4: Q-Q plots constructed to check normality of difference parameter	177
Figure 8.5: Q-Q plots constructed to check normality of ratio parameter	180
Figure 8.6: Cumulative distribution curves of difference and ratio parameters.....	182
Figure 8.7: Underlying patterns in difference parameters.....	186
Figure 8.8: Underlying patterns in ratio parameters	190

List of Tables

Table 2.1: Chromosome representation of ET.....	20
Table 2.2: Summary of machine learning techniques	26
Table 3.1: Physical properties of binders used in present study.....	31
Table 3.2: Designation of binary blends and proportion of individual components.....	31
Table 3.3: Measured kinematic viscosity of binary blends, cP	32
Table 3.4: Ranking of various viscosity-mixing rules.....	34
Table 4.1: Summary statistics of datasets used in present study.....	41
Table 4.2: Summary of regression coefficients in finalized GEP based models.....	43
Table 4.3: Comparison of statistical parameters obtained using various models.....	44
Table 4.4 Sensitivity of variables in the developed GEP based models.....	50
Table 5.1: Physical properties of binder used in present study	57
Table 5.2: Summary of measured viscosity at various temperatures	58
Table 5.3: Chemical composition of binders determined through SARA analysis.....	59
Table 5.4: Bond index values computed using FTIR spectroscopy data.....	59
Table 5.5: Feasible conditions for production of foamed bitumen.....	60
Table 5.6: Statistical description of data on physical properties	60
Table 5.7: Summary of regression coefficients.....	75
Table 5.8: Summary of regression coefficients.....	76
Table 5.9: Summary of goodness of fit values	79
Table 6.1: Physical properties of binders used in present work	89
Table 6.2: Temperature range used for foamed bitumen production	90
Table 6.3: GEP model parameters.....	94
Table 6.4: Summary of statistics related to viscosity-temperature relationship.....	95
Table 6.5: Confusion matrix for the DT's	102
Table 6.6: Summary of coefficients in models for Expansion Ratio prediction	103
Table 6.7: Summary of coefficients in models for Half-Life prediction.....	104
Table 6.8: Statistical comparison of Global GEP models and MLSP models.....	105
Table 6.9: Summary of goodness of fit parameters.....	108
Table 6.10: Summary of global sensitivity analysis.....	109
Table 6.11: Summary of importance of variables from global sensitivity analysis	110
Table 6.12: Summary of coefficients in models for Expansion Ratio prediction	111
Table 6.13: Summary of coefficients in models for Half-Life prediction.....	111
Table 6.14: Summary of goodness of fit parameters.....	112
Table 6.15: Confusion matrix for performance evaluation of DT's.....	114
Table 6.16: Summary of global sensitivity analysis.....	115
Table 6.17: Summary of importance of variables from global sensitivity analysis	115
Table 7.1: Summary of testing protocols.....	122
Table 7.2: Summary of parameters included in dynamic modulus predictive models.....	129
Table 7.3: Summary of range and histogram of associated parameters	133
Table 7.4: Summary of principal component analysis output.....	138
Table 7.5: Variable loading on principal components.....	139
Table 7.6: Summary of parameters contributing to individual PC 's	142
Table 7.7: Statistical performance comparison of models	146
Table 7.8: Summary of sensitivity analysis.....	150
Table 8.1: Description of individual mixture data sets.....	158
Table 8.2: Statistical description of database	162
Table 8.3: Correlation matrix obtained using global dataset and mixture 111.....	164
Table 8.4: Summary of p-values obtained while conducting Welch t-test.....	169
Table 8.5: Goodness of fit indicators computed during calibration process	173
Table 8.6: Summary of statistics from Q-Q plots of difference parameter	174
Table 8.7: Summary of statistics from Q-Q plots of ratio parameter	178
Table 8.8: Distribution of ratio parameter and difference parameter	183
Table 8.9: Ranking of variables based on global sensitivity analysis	191

List of notations

$ E^* $	=	Dynamic modulus of asphalt mixture, 10^5 psi
$ E^* _m$	=	Dynamic modulus of asphalt mixture, psi
$ E^* _s$	=	Dynamic modulus of asphalt mixture, MPa
$ G^* $	=	Binder shear modulus, psi
$ G^* _b$	=	Binder shear modulus, Pa
$ G^* _g$	=	Glassy modulus, Pa = 10^9 Pa (145000psi)
δ_b	=	Phase shift angle, degrees
η	=	Asphalt binder viscosity, 10^6 psi
f	=	Loading frequency, Hz
V_a	=	Volume of air voids in the mix, %
V_{beff}	=	Volume of effective binder content, %
$p_{3/4}$	=	% weight of aggregates retained on the $\frac{3}{4}$ inch sieve
$p_{3/8}$	=	% weight of aggregates retained on the $\frac{3}{8}$ inch sieve
p_4	=	% weight of aggregates retained on the No.4 sieve
p_{200}	=	% weight of aggregates passing through the No.200 sieve
VMA	=	Voids in mineral aggregates in the mix, %
VFB	=	Voids filled with the binder in the mix, %
T	=	Test temperature, °C
AC	=	Asphalt content, %
σ_{amp}	=	Stress amplitude
ε_{amp}	=	Strain amplitude