CHARACTERISING ASPHALT MIXTURES USING IMPROVED MACHINE LEARNING APPROACHES BY OPTIMIZING MODEL PERFORMANCE AND RANK BASED MODEL VARIABLES

NISHIGANDHA RAJESHWAR JUKTE

DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI
JUNE 2025

CHARACTERISING ASPHALT MIXTURES USING IMPROVED MACHINE LEARNING APPROACHES BY OPTIMIZING MODEL PERFORMANCE AND RANK BASED MODEL VARIABLES

by

NISHIGANDHA RAJESHWAR JUKTE

Department of Civil Engineering

Submitted

In fulfilment of the Requirements of the Degree of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY
DELHI
JUNE 2025

DEDICATED TO MY PARENTS, MY FAMILY, MY GUIDE, AND MY FRIENDS. I AM FOREVER GRATEFUL FOR THEIR GUIDANCE AND BELIEF IN ME.

CERTIFICATE

This is to certify that the thesis titled "**Characterising Asphalt Mixtures Using Improved Machine Learning Approaches by Optimizing Model Performance and Rank Based Model Variables**", submitted by **Nishigandha Rajeshwar Jukte**, to the Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by her under my supervision and guidance, completed in February 2025. The thesis work, in my opinion, has reached the requisite standard, fulfilling the requirements for the degree of Doctor of Philosophy.

The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Aravind Krishna Swamy

(Professor)

Department of Civil Engineering

Indian Institute of Technology Delhi

New Delhi 110016

India

ACKNOWLEDGEMENTS

I want to thank everyone from the bottom of my heart for their support during my PhD journey. First and foremost, I want to express my gratitude to my supervisor, Prof. Aravind Krishna Swamy, for his time, helpful suggestions, and tireless encouragement. I express my gratitude to him for his invaluable advice, which has greatly improved my research. His timely input enabled me to strengthen my concepts and stay focused on goals at every stage of my work. His different perspectives and critical ways of thinking have always been inspiring for me. He also challenges his students to think critically, which improved my understanding of the subject. I am grateful for his understanding and kindness during the more challenging parts of my journey. His confidence in my abilities helped me to overcome difficulties. I consider myself privileged to have had him as my supervisor, and I sincerely appreciate his support.

I am also profoundly grateful to my student research committee members, Prof. Kumar Neeraj Jha, Prof. Arun Kumar from the Department of Civil Engineering and Prof. S. Pradyumna from the Department of Applied Mechanics, for their thoughtful suggestions and encouragement. I am also thankful to the other faculty members in Transportation Engineering for their critical comments and suggestions during my progress presentations each semester. Their thoughtful feedback pushed me to improve my work and deepened my understanding of my research topic. I appreciate the constructive discussions that often followed my presentations, which helped me refine my ideas and approach.

I would also like to thank the Transportation Engineering Laboratory staff members, particularly Mr. Amit Bundela, Mr. Kausik Pahari, Mr. Biswarup Banik, Mr. Ashok Kumar, and

Mr. Vishwas J. Though my path was largely independent, I appreciate the facilities and resources provided by them, which were essential for my work.

I would like to thank my seniors, Dr. Karanjeet Kaur and Dr. Anjali Balan, Manoj for their support and guidance throughout this process. I appreciate the time they took to share their knowledge and experiences with me. I am grateful to my friends and colleagues Ravi Raj, Akila, Prachi, Arpita, Hebah, Rohit, Sandeep, and Priti for their teamwork and keeping positive environment in the lab. Their encouragement, lab discussions, ideas and feedback have been invaluable both professionally and personally. Their positivity has helped me stay focused when it was much needed. The laughter, conversations, and memories created brought much-needed happiness and balance during challenging times. I am truly thankful to each of them for being an important part of my PhD journey.

I want to extend my heartfelt thanks to my close friend Neba for her support and encouragement throughout my PhD journey. I would be remiss in not mentioning Poonam, Vamsi, Amit, Yawar, and Debashish at IIT. I would also like to thank my friends from my earlier institute, Rubui, Vikram, Dr. Shubham, and Dr. Chintaman, for their continued support and motivation, which kept me grounded and inspired me to persevere.

I also want to thank my husband, Dr. Vivek Wagh, from the bottom of my heart. As he is also doing his PhD at IIT (BHU), the discussions with him and his advice was fruitful. The trust he has shown me in this journey is truly invaluable to me. I am fortunate that he has supported and understood me on this journey.

I want to thank my family, who have been essential to my life. I am genuinely grateful for everything they have done for me. Without their love and support, I would not be where I am today. I could not have walked this journey without my father, Mr. Rajeshwar Jukte, who always supported me in higher studies. He continuously inspired me in my journey. He raised my spirits when I was not motivated. He boosted my confidence every time by telling me I can do it. I would also like to thank my grandmother, grandfather, and Chhoti Aai for their love and support. I wish to thank my siblings, Dr. Snehal, Swapna, and Shreyas, for always supporting and being with me in every situation. I wish to thank my both brother in laws for their support and motivation. I would like to thank my source of happiness, i.e., the young ones in my family, Shivansh, Ramaa, and Mivaan. In this acknowledgement of my thesis, I extend my appreciation to all who have been part of this journey.

Finally, I want to thank my mother in heaven, Late Mrs. Renuka Jukte. I miss her every day. Even though she is not with me, I believe she guides me every time. She used to emphasize that I should go for higher studies. But along with that, she taught me to be a better person. What I am today is because of my parents.

Nishigandha Rajeshwar Jukte

ABSTRACT

The dynamic modulus ($|E^*|$) and phase angle (ϕ) of asphalt mixtures are two essential parameters for predicting pavement performance. These parameters are helpful in evaluating fatigue cracking and rutting in Mechanistic Empirical Pavement Design Guidelines (MEPDG). In addition to this, these parameters are also used to assess low temperature cracking, moisture damage, international roughness index, and reflective cracking. Therefore, predicting these parameters at a given temperature and for a different range of loading frequencies is essential. To evaluate the effectiveness of the proposed methodologies in first three objectives, dynamic modulus and phase angle measurements reported in national cooperative highway research program (NCHRP Project) 9-19 were used. Eight binders and five aggregate gradations originating from 201 distinct asphalt mixtures were identified. Dynamic modulus and phase angle mastercurves were constructed for individual mixtures using three mastercurve (sigmoidal, generalized sigmoidal, and CAM mastercurve) functions and two temperature shift factor determination approaches (free shifting approach and second order polynomial approach).

In addition, for improving the accuracy of phase angle prediction, the first objective of the study proposes a rational model to describe the shape of phase angle mastercurve. The proposed model is a combination of two sigmoidal functions having seven parameters. The proposed 7-parameter model was compared with popular sigmoidal, generalized sigmoidal, and CAM mastercurve models. Results indicate that the proposed model accurately describes the shape of phase angle mastercurve than other described models. Statistical analysis showed reasonable correlation among phase angle mastercurve characteristics (peak phase angle and its location) and other variables (i.e. mixture volumetrics, mixture gradation).

Mixturewise regression models were also proposed to predict phase angle mastercurve characteristics. Goodness of fit indicators verified that the proposed predictive models were able to predict phase angle mastercurve characteristics well.

Further, the study proposes an efficient approach to arrive at Gene Expression Programming (GEP) based models for $|E^*|$ and ϕ prediction. The proposed methodology encompasses (i) construction of $|E^*|$ and ϕ mastercurves, (ii) ranking of GEP based predictive expressions that uses reduced frequency, aggregate gradation, asphalt binder content, binder properties, and mixture volumetrics as input, and (iii) selection of predictive model at optimal tree depth, and (iv) recalibration. Thus, proposed methodology strives to attain right balance among selection of highly influencing variables, predictive model complexity, and predictive accuracy. To validate the proposed methodology, database reported in NCHRP 9-19 study was used. With application of proposed approach with increasing tree depths, model complexity, predictive accuracy and computational time increased. However, beyond certain tree depth the increment in model accuracy was marginal. This can be attributed to incorporation of highly influencing variables in initial stages and limitation on model forms during GEP module execution. The robustness of the proposed GEP based predictive expressions was checked with independent database and competitor models. The numerical values of goodness of fit indicators showed improved predictive accuracy through proper selection of input variables, and predictive model. The predictive accuracy was further improved through mixturewise recalibration. The proposed predictive expressions can be conveniently utilized for $|E^*|$ and ϕ prediction through recalibration and subsequent preliminary design.

Furthermore, the third objective proposed a deep learning approach to predict $|E^*|$ and ϕ. The input variables for deep learning model consisted of information regarding reduced frequency, binder properties, aggregate gradation, and mixture volumetrics. When compared to other input variables, reduced frequency and binder properties were highly correlated with dynamic modulus and phase angle. Further, recursive feature elimination was used to rank all input variables. Using these ranked input variables, deep learning based models for predicting dynamic modulus and phase angle were developed. The deep learning architecture (finalized through exhaustive optimization) dependent on parameter under consideration, and mastercurve construction approach was adopted. Detailed statistical analysis indicated dynamic modulus predictive models performed better when compared to phase angle predictive models. The numerical values of goodness of fit indicators showed that accuracy of deep learning based model was dependent on mastercurve construction approach adopted. Overall results indicated that the deep learning based models can predict dynamic modulus and phase angle with good accuracy. The output from the proposed deep learning models can be used as direct input into pavement design framework, which can result in accurate prediction of pavement performance.

The fourth objective of the study was to propose a machine learning based cost optimization of Marshall mix design. To develop the machine learning based cost optimization framework, different aggregate gradations in terms of D_i values from D_{90} to D_{10}, binder content, and bulk specific gravity of aggregate were used to predict Marshall test results and mix volumetrics. However, along with the aggregate gradation, coefficient of uniformity (C_u) and curvature (C_c) were used to develop the framework. To account for all shape factors at all the points on a gradation curve, new $C_u's$ and $C_c's$ were developed. The new $C_u's$ and $C_c's$ showed better correlation with mixture volumetrics. The four approaches were used for optimization with different machine learning predictive models using input independent variables as: (i) Approach 1: D_i values from D_{90} to D_{10} (ii) Approach 2: D_i values and traditional C_u C_c (iii) Approach 3: new developed C_u and C_c (iv) Approach 4: new developed $C_u's$ and

$C_c's$ developed and traditional C_u C_c. The cost optimization algorithm was developed for these four approaches by providing constraints on aggregate gradation and mixture volumetrics. The approaches with new $C_u's$ and $C_c's$ resulted in lower binder content when compared to approaches with D_i values. Also, the approaches with traditional C_u C_c provided similar aggregate gradation curves.

सार

डामर मिश्रण का गतिशील मापांक ($|E^*|$) और चरण कोण (ϕ) फुटपाथ प्रदर्शन की भविष्यवाणी के लिए दो आवश्यक पैरामीटर हैं। ये पैरामीटर एमईपीडीजी डिज़ाइन में थकान क्रैकिंग और रटिंग के आकलन में सहायक हैं। इसके अलावा, इन मापदंडों का उपयोग कम तापमान क्रैकिंग, नमी क्षति और सूचकांक और परावर्तक क्रैकिंग का आकलन करने के लिए भी किया जाता है। इसलिए, किसी दिए गए तापमान पर और लोडिंग आवृत्तियों की एक अलग श्रृंखला के लिए इन मापदंडों की भविष्यवाणी करना आवश्यक है। चरण कोण भविष्यवाणी की सटीकता में सुधार करने की दिशा में, अध्ययन का पहला उद्देश्य चरण कोण मास्टरवक्र के आकार का वर्णन करने के लिए तर्कसंगत मॉडल का प्रस्ताव करता है। प्रस्तावित मॉडल सात पैरामीटर वाले दो सिग्मोइडल फ़ंक्शन का संयोजन है। प्रस्तावित मॉडल की प्रभावशीलता का मूल्यांकन करने के लिए, एनसीएचआरपी प्रोजेक्ट 9-19 में रिपोर्ट किए गए चरण कोण माप का उपयोग किया गया था। 201 अलग-अलग डामर मिश्रणों से उत्पन्न आठ बाइंडरों और पांच समुच्चय ग्रेडेशन की पहचान की गई। प्रस्तावित 7-पैरामीटर मॉडल की तुलना प्रसिद्ध सिग्मॉइडल, सामान्यीकृत सिग्मॉइडल और सीएएम मास्टरकर्व मॉडल से की गई थी। परिणाम दर्शाते हैं कि प्रस्तावित मॉडल अन्य मॉडलों की तुलना में चरण कोण मास्टरवक्र के आकार का सटीक वर्णन करता है। सांख्यिकीय विश्लेषण ने चरण कोण मास्टरवक्र विशेषताओं (शिखर चरण कोण और उसका स्थान) और अन्य चर (यानी मिश्रण वॉल्यूमेट्रिक्स, मिश्रण ग्रेडेशन) के बीच उचित सहसंबंध दिखाया। चरण कोण मास्टरवक्र विशेषताओं की भविष्यवाणी करने के लिए मिश्रणवार प्रतिगमन मॉडल भी प्रस्तावित किए गए थे। फिट संकेतकों की

अच्छाई ने सत्यापित किया कि प्रस्तावित पूर्वानुमानित मॉडल चरण कोण मास्टरकर्व विशेषताओं की अच्छी तरह से भविष्यवाणी करने में सक्षम थे।

इसके अलावा, गतिशील मापांक ($|E^*|$) और चरण कोण (ϕ) भविष्यवाणी के लिए जीन एक्सप्रेशन प्रोग्रामिंग (जीईपी) आधारित मॉडल पर पहुंचने के लिए एक कुशल दृष्टिकोण। प्रस्तावित पद्धति में (i) $|E^*|$ का निर्माण शामिल है और ϕ मास्टरकर्व्स, (ii) जीईपी आधारित पूर्वानुमानित अभिव्यक्तियों की रैंकिंग जो इनपुट के रूप में कम आवृत्ति, समग्र ग्रेडेशन, डामर बाइंडर सामग्री, बाइंडर गुण और मिश्रण वॉल्यूमेट्रिक्स का उपयोग करती है, और (iii) इष्टतम पेड़ की गहराई पर पूर्वानुमानित मॉडल का चयन, और (iv)) पुनर्अंशांकन. इस प्रकार, प्रस्तावित पद्धति अत्यधिक प्रभावशाली चर, पूर्वानुमानित मॉडल जटिलता और पूर्वानुमानित सटीकता के चयन के बीच सही संतुलन प्राप्त करने का प्रयास करती है। प्रस्तावित पद्धति को मान्य करने के लिए, एनसीएचआरपी 9-19 अध्ययन में रिपोर्ट किए गए डेटाबेस का उपयोग किया गया था। पेड़ की गहराई बढ़ाने के साथ प्रस्तावित दृष्टिकोण के अनुप्रयोग के साथ, मॉडल जटिलता, पूर्वानुमान सटीकता और कम्प्यूटेशनल समय में वृद्धि हुई। हालाँकि, पेड़ की निश्चित गहराई से परे मॉडल सटीकता में वृद्धि मामूली थी। इसका श्रेय प्रारंभिक चरणों में अत्यधिक प्रभावशाली चर को शामिल करने और जीईपी मॉड्यूल निष्पादन के दौरान मॉडल रूपों पर सीमा को दिया जा सकता है। प्रस्तावित जीईपी आधारित भविष्य कहनेवाला अभिव्यक्तियों की मजबूती की जांच स्वतंत्र डेटाबेस और प्रतिस्पर्धी मॉडल के साथ की गई थी। फिट संकेतकों की अच्छाई के संख्यात्मक मूल्यों ने इनपुट चर और पूर्वानुमानित मॉडल के उचित चयन के माध्यम से पूर्वानुमानित सटीकता में सुधार दिखाया।

मिश्रणवार पुनर्अंशांकन के माध्यम से पूर्वानुमानित सटीकता में और सुधार किया गया। प्रस्तावित भविष्यसूचक अभिव्यक्तियों का उपयोग $|E^*|$ के लिए आसानी से किया जा सकता है और पुनर्अंशांकन और उसके बाद प्रारंभिक डिजाइन के माध्यम से ϕ भविष्यवाणी।

इसके अलावा, तीसरे उद्देश्य ने डामर मिश्रण के गतिशील मापांक और चरण कोण की भविष्यवाणी करने के लिए एक गहन शिक्षण दृष्टिकोण का प्रस्ताव दिया। प्रस्तावित दृष्टिकोण को मान्य करने के लिए राष्ट्रीय सहकारी राजमार्ग अनुसंधान कार्यक्रम 9-19 में रिपोर्ट किए गए गतिशील मापांक और चरण कोण डेटा का उपयोग किया गया था। इस डेटाबेस में 201 विशिष्ट डामर मिश्रणों का चयन किया गया। दो सिग्मॉइडल कार्यों और दो तापमान बदलाव कारक निर्धारण दृष्टिकोणों के संयोजन का उपयोग करके व्यक्तिगत मिश्रण के लिए गतिशील मापांक और चरण कोण मास्टरकर्व का निर्माण किया गया था। गहन शिक्षण मॉडल के लिए इनपुट चर में कम आवृत्ति, बाइंडर गुण, समग्र उन्नयन और मिश्रण वॉल्यूमेट्रिक्स के बारे में जानकारी शामिल थी। जब अन्य इनपुट चर की तुलना की जाती है, तो कम आवृत्ति और बाइंडर गुण गतिशील मापांक और चरण कोण के साथ अत्यधिक सहसंबद्ध होते हैं। इसके अलावा, सभी इनपुट चर को रैंक करने के लिए पुनरावर्ती सुविधा उन्मूलन का उपयोग किया गया था। इन रैंक किए गए इनपुट चर का उपयोग करके, गतिशील मापांक और चरण कोण की भविष्यवाणी के लिए गहन शिक्षण आधारित मॉडल विकसित किए गए थे। गहन शिक्षण वास्तुकला (संपूर्ण अनुकूलन के माध्यम से अंतिम रूप दिया गया) विचाराधीन पैरामीटर पर निर्भर है, और मास्टरकर्व निर्माण दृष्टिकोण अपनाया गया था। विस्तृत सांख्यिकीय विश्लेषण से संकेत मिलता है कि चरण कोण पूर्वानुमानित मॉडल

की तुलना में गतिशील मापांक पूर्वानुमानित मॉडल ने बेहतर प्रदर्शन किया। फिट संकेतकों की अच्छाई के संख्यात्मक मूल्यों से पता चला कि गहन शिक्षण आधारित मॉडल की सटीकता अपनाए गए मास्टरकर्व निर्माण दृष्टिकोण पर निर्भर थी। समग्र परिणामों से संकेत मिलता है कि गहन शिक्षण आधारित मॉडल अच्छी सटीकता के साथ गतिशील मापांक और चरण कोण की भविष्यवाणी कर सकते हैं। प्रस्तावित गहन शिक्षण मॉडल के आउटपुट का उपयोग फुटपाथ डिजाइन ढांचे में सीधे इनपुट के रूप में किया जा सकता है, जिसके परिणामस्वरूप फुटपाथ प्रदर्शन की सटीक भविष्यवाणी हो सकती है।

अध्ययन का चौथा उद्देश्य मार्शल मिक्स डिज़ाइन के मशीन लर्निंग आधारित लागत अनुकूलन का प्रस्ताव करना था। मशीन लर्निंग आधारित लागत अनुकूलन ढांचे को विकसित करने के लिए, मार्शल परीक्षण परिणामों की भविष्यवाणी करने और वॉल्यूमेट्रिक्स को मिश्रित करने के लिए D_{90} से D_{10} तक D_i मानों, बाइंडर सामग्री और समुच्चय के थोक विशिष्ट गुरुत्व के संदर्भ में विभिन्न समुच्चय ग्रेडेशन का उपयोग किया गया था। हालाँकि, समग्र उन्नयन के साथ-साथ रूपरेखा को विकसित करने के लिए एकरूपता और वक्रता के गुणांक का उपयोग किया गया था। ग्रेडेशन वक्र पर सभी बिंदुओं पर सभी आकार कारकों को ध्यान में रखने के लिए नए $C_u's$ और $C_c's$ विकसित किए गए। नए $C_u's$ और $C_c's$ ने मार्शल परीक्षण परिणाम मिश्रण वॉल्यूमेट्रिक्स के साथ बेहतर सहसंबंध दिखाया। चार दृष्टिकोणों का उपयोग इनपुट स्वतंत्र चर का उपयोग करके विभिन्न मशीन लर्निंग पूर्वानुमानित मॉडल के साथ अनुकूलन के लिए किया गया था: (i) दृष्टिकोण 1: D_{90} से D_{10} तक D_i मान (ii) दृष्टिकोण 2: D_i मान और

पारंपरिक C_u C_c (iii) दृष्टिकोण 3: नए विकसित $C_u's$ और $C_c's$ (iv) दृष्टिकोण 4: नए विकसित $C_u's$ और $C_c's$ के विकसित और पारंपरिक C_u C_c। कुल ग्रेडेशन और मिश्रण वॉल्यूमेट्रिक्स पर बाधाएं प्रदान करके इन चार दृष्टिकोणों के लिए लागत अनुकूलन एल्गोरिदम विकसित किया गया था। नए $C_u's$ और $C_c's$ के दृष्टिकोणों के परिणामस्वरूप D_i मानों वाले दृष्टिकोणों की तुलना में कम बाइंडर सामग्री प्राप्त हुई। इसके अलावा, पारंपरिक C_u C_cके दृष्टिकोण समान समग्र उन्नयन वक्र प्रदान करते हैं।

TABLE OF CONTENTS

List of Figures

List of Tables

List of Notations

$\lvert E^* \rvert$	Dynamic modulus of asphalt mixture, MPa
ϕ	Phase angle of asphalt mixture, °
T	Loading temperature, °C
f	Loading frequency, Hz
σ_o	Peak applied stress
ε_o	Peak resulting strain
$\lvert G^* \rvert$	Complex shear modulus of asphalt binder, MPa
δ_b	Binder phase angle of asphalt binder, °
$\rho_{3/4}$	% weight of aggregates retained on the ¾-inch sieve
$\rho_{3/8}$	% weight of aggregates retained on the 3/8 -inch sieve
ρ_4	% weight of aggregates retained on the No. 4 sieve
ρ_{200}	% weight of aggregates passing through the No. 200 sieve
AC	Asphalt binder content, %
V_a	Air voids in the mix, %
V_{beff}	Effective asphalt binder content, %
VMA	Voids in mineral aggregates, %
VFB	Voids filled with the bitumen, %
DV	Dependent variable
MV_i	Mixture indicator i (i = 1 through 5)
AGP_j	Aggregate gradation property indicator j (j = 1 through 4)
BP_k	Binder property indicator k (k = 1 through 2)
aT	Temperature shift factor
f_r	Reduced frequency, Hz
NoC	Number of coefficients
QoF	Quality of fit indicators
$SQFI$	Sum of quality of fit indicators
IVP	Independent variable proportion
R^2	Coefficient of determination

$Adj\ R^2$	Adjusted coefficient of determination
SSE	Sum of squared error
$RMSE$	Root mean squared error
C_u	Coefficient of uniformity
C_c	Coefficient of curvature
D_i	Aggregate particle size at every 10% interval from D_{90} to D_{10}, mm
V_a	Air voids, %
MAS	Maximum aggregate size, mm
$NMAS$	Nominal maximum aggregate size, mm
P_b	Binder content, %
C_i	Aggregate cost for bin i