Oklahoma Department of Transportation
200 NE 21st Street, Oklahoma City, OK 73105-3204

FINAL REPORT ~ FHWA-OK-26-01

INVESTIGATING THE AGING BEHAVIOR OF ASPHALT BINDERS AT DIFFERENT PRODUCTION STAGES AND DURING THE SERVICE LIFE OF THE PAVEMENT

Mohamed Elkashef, Ph.D., P.E.
Debakanta (Deb) Mishra, Ph.D., P.E.
Sina Mousavi Rad, Ph.D. Candidate
Mehdi Sadeghi, Ph.D. Student
Adeoluwa Gbolade, Ph.D. Candidate
Michael Olagunju, M.Sc.

School of Civil Engineering and Environmental Engineering
Oklahoma State University
Stillwater, Oklahoma

October 2025

The Oklahoma Department of Transportation (ODOT) ensures that no person or groups of persons shall, on the grounds of race, color, sex, religion, national origin, age, disability, retaliation or genetic information, be excluded from participation in, be denied the benefits of, or be otherwise subjected to discrimination under any and all programs, services, or activities administered by ODOT, its recipients, sub-recipients, and contractors. To request an accommodation please contact the ADA Coordinator at 405-521-4140 or the Oklahoma Relay Service at 1-800-722-0353. If you have any ADA or Title VI questions email ODOT-ada-titlevi@odot.org.

The Oklahoma State University herein certifies that, as of the date of submission to the Oklahoma Department of Transportation, the electronic version of this report meets all federal and state online accessibility requirements pertaining to:

> **United States Code Title 29, Section 794d vocational rehabilitation rights to access technology, and Oklahoma State Statutes Title 62, Section 34.20 duties relating to communications systems.**

and is ready for posting online.

INVESTIGATING THE AGING BEHAVIOR OF ASPHALT BINDERS AT DIFFERENT PRODUCTION STAGES AND DURING THE SERVICE LIFE OF THE PAVEMENT

FINAL REPORT ~ FHWA-OK-26-01
ODOT SPR ITEM NUMBER 2294

Submitted to:
Office of Research and Implementation
Oklahoma Transportation

Submitted by:
Mohamed Elkashef, Ph.D., P.E.
Debakanta (Deb) Mishra, Ph.D., P.E.
Sina Mousavi Rad, Ph.D. Candidate
Mehdi Sadeghi, Ph.D. Student
Adeoluwa Gbolade, Ph.D. Candidate
Michael Olagunju, M.Sc.

School of Civil and Environmental Engineering
Oklahoma State University

OKLAHOMA Transportation

October 2025

TECHNICAL REPORT DOCUMENTATION PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
FHWA-OK-26-01		

4. TITLE AND SUBTITLE Investigating the Aging Behavior of Asphalt Binders at Different Production Stages and During the Service Life of the Pavement	5. REPORT DATE Nov 2025 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Mohamed Elkashef Debakanta (Deb) Mishra Sina Mousavi Rad Michael Olagunju Mehdi Sadeghi Adeoluwa Gbolade	8. PERFORMING ORGANIZATION REPORT Click here to enter text.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Oklahoma State University; Stillwater Ok 74078	10. WORK UNIT NO. 11. CONTRACT OR GRANT NO. ODOT SPR Item Number 2294
12. SPONSORING AGENCY NAME AND ADDRESS Oklahoma Transportation Office of Research and Implementation 200 N.E. 21st Street, Rm. 3A7 Oklahoma City, OK 73105	13. TYPE OF REPORT AND PERIOD COVERED Final Report Aug 2022 - Nov 2025 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
Click here to enter text.

16. ABSTRACT

This study examines how asphalt mixtures change over time and how their performance is influenced by material selection and environmental factors. Thirty-six asphalt mixes produced in Oklahoma were collected and tested to understand their behavior in terms of cracking, rutting, moisture damage, and compaction. The work focuses on how aging affects these properties, since asphalt becomes stiffer and more brittle as it ages in the field. Cracking resistance was measured using the IDEAL-CT test, and both CT_{Index} and its underlying parameters were analyzed before and after lab-simulated long-term oven aging. Statistical and data analysis methods were used to interpret trends, and Machine Learning models were developed to predict CT_{Index} from mix variables. Rutting and moisture susceptibility were evaluated using the HWT and IDT-HT tests. Binder rheology was assessed through frequency–temperature sweep testing. Results showed that long-term aging significantly reduced cracking resistance and mix variables such as NMAS, RAP content, and binder grade significantly influenced CT_{Index}. The Corrected Rut Depth parameter provided a clearer indication of rutting than Total Rut Depth, and an IDT-HT strength of 240 kPa is suggested as a production acceptance threshold for rutting. Freeze–thaw conditioning affected rutting and stripping of most of the mixes. These findings support performance-based mix design and help guide mixture selection for durable pavements.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Aging, Cracking, Rutting, Moisture damage, Predictive Modeling	No restrictions. This publication is available from the Office of Research and Implementation, Oklahoma DOT.

19. SECURITY CLASSIF. (OF THIS REPORT)	20. SECURITY CLASSIF. (OF THIS PAGE)	21. NO. OF PAGES	22. PRICE
Unclassified	Unclassified	264	N/A

Form DOT F 1700.7 (08/72)

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Table of Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 Overview

The Balanced Mix Design (BMD) approach has gained widespread interest in recent years. Unlike traditional volumetric-based design, BMD incorporates performance testing to ensure mixtures can resist major pavement distresses such as cracking, rutting, and moisture damage and satisfy different performance criteria. Several states have already moved toward the implementation of BMD as part of their mix design specifications (Bennert et al., 2014; Diefenderfer & Bowers, 2019).

Among the major performance concerns in BMD, cracking has been identified as one of the most critical. So, a reliable cracking test should consider both crack initiation and crack propagation and should correlate with field performance. NCHRP project 9-57 investigated several cracking tests for BMD implementation based on the repeatability of the test results and ease of performance (F. Zhou et al., 2016). Several cracking tests, including the Semi-Circular Bending (SCB) (Mohammad and Kim, 2011), Illinois Flexibility Index Test (IFIT) (Ozer et al., 2016), and Indirect Tensile Asphalt Cracking Test (IDEAL-CT) (Romero-Zambrana, 2023), have been developed over the past years to address these needs and to evaluate the cracking resistance of mixtures at intermediate temperatures. Among these, the IDEAL-CT has been considered a rapid and simple test that can be implemented into practice, and is sensitive to key mixture variables such as binder grade, Asphalt Cement (AC) content, recycled material content, air voids, and aging level (F. Zhou et al., 2017). The IDEAL-CT is typically performed at a temperature of 25^{o}C and does not involve specimen cutting or notching. The load-deformation curve from the IDEAL-CT is used to calculate a Cracking Tolerance Index (CT_{Index}), which summarizes test outputs into a single indicator of cracking resistance.

While CT_{Index} is widely used, it has limitations. CT_{Index} is a single value that encompasses interrelated IDEAL-CT parameters, i.e. displacement at 75 percent post-peak load (l_{75}), slope at 75 percent post-peak load (m_{75}), and fracture energy. While informative, CT_{Index} alone might not provide a comprehensive picture of a mixture's cracking behavior. Mixtures with the same CT_{Index} values may behave differently under loading, i.e. different load-displacement curves. This highlights the importance of analyzing the abovementioned underlying IDEAL-CT parameters rather than relying on CT_{Index} alone. Advanced data interpretation methods such as Principal Component Analysis (PCA) can be applied to better explain the interaction among these parameters and their combined effect on cracking behavior. Similarly, predictive modeling techniques, including Machine Learning (ML), have been increasingly used in asphalt research studies to link materials properties and test outcomes, with purpose of predicting performance measures such as CT_{Index}. Moreover, advanced statistical analysis are valuable methods to interpret performance test results and help make meaningful inferences.

In addition to cracking, rutting remains another central focus of BMD. The Hamburg Wheel Tracking Test (HWTT) has been widely used by many state transportation agencies to evaluate both rutting resistance and moisture susceptibility. The Total Rut Depth (TRD) is used by many state transportation agencies to evaluate rutting resistance. However, the TRD does not separate between rutting caused by plastic deformation and that caused by stripping. An alternative parameter, i.e. Corrected Rut Depth (CRD), has been suggested to evaluate rutting that occurs solely due to plastic deformation without considering the effect of stripping (Yin et al., 2014).

Moisture susceptibility is also a key factor in BMD, as it significantly influences long-term pavement durability. Importantly, the interaction between aging and moisture resistance must be considered. Aging is one of the most influential factors in asphalt pavement performance, altering binder properties by increasing stiffness and brittleness, thereby reducing stress relaxation and deformability. Several studies have shown that aging can improve or worsen moisture resistance depending on the mixture characteristics (Albayati & Abduljabbar, 2019; Diab et al., 2019; Guo et al., 2014). Thus, evaluating the combined effect of aging and moisture conditioning is necessary to understand realistic field performance.

With several agencies moving towards the implementation of BMD, it is very important to understand how different mixture variables affect the mixture performance. On the other hand, aging, which is considered one of the most critical factors and causes several asphalt distresses, affects the overall asphalt pavement performance. Thus, the evaluation of the simultaneous effect of different mix variables and aging on the asphalt mixture performance is very significant.

The properties of the asphalt binder play a central role in all the performance of asphalt mixtures, including cracking, rutting, and moisture resistance. As a viscoelastic material, asphalt binder is sensitive to both temperature and loading time (Kamboozia et al., 2021). Aging increases binder stiffness, which improves rutting resistance at high temperatures but reduces cracking resistance at intermediate and low temperatures. To better characterize these effects, rheological testing such as frequency–temperature sweeps in the Dynamic Shear Rheometer (DSR) is commonly used. Indices, such as the Glover–Rowe (G-R) parameter, have been developed to link binder-level behavior to mixture performance indicators such as CT_{Index} (C. Chen et al., 2018a; Martin et al., 2024).

In summary, the BMD framework requires a comprehensive evaluation of cracking, rutting, and moisture resistance, while also accounting for durability and aging. This study contributes to that effort by examining mixture performance through IDEAL-CT, HWTT, and binder rheological testing, while also applying statistical and ML approaches to interpret results and identify relationships among key performance parameters.

1.2 Objectives

The main objectives of this study are to:

1) Assess the effect of lab-simulated long-term oven aging on the cracking resistance of plant mixes using the IDEAL-CT.
2) Investigate the impact of different mix variables on the CT_{Index} using statistical analyses.
3) Use ML to predict the cracking resistance of asphalt mixtures using different mix variables.
4) Evaluate the combined effect of moisture and aging on asphalt mixes and binders.
5) Analyze the effect of freeze-thaw on the rutting and stripping of asphalt mixes, using HWTT.

1.3 Literature Review

1.3.1 Aging

Aging is considered a primary factor that impacts the performance of asphalt mixtures. Asphalt binder oxidation and loss of volatiles, which contribute to asphalt aging, results in stiffening of the asphalt binder (Nikolaides, 2014). Temperature, light, and oxygen are the key factors influencing the binder's oxidative aging (Z. Chen et al., 2021; Pasandín et al., 2015). Aging is an irreversible process that occurs during asphalt production, referred to as Short-Term Aging (STA) (Grilli et al., 2017), and during the life time of the pavement, referred to as Long-Term Aging (LTA) (Maschauer et al., 2022). The performance characteristics of asphalt binders and mixtures are impacted by age-related phenomena which result in increases in viscosity and stiffness (Pasandín et al., 2015). The change in properties with aging can lead to pavement distresses and expensive maintenance and repairs (López-Montero et al., 2018). Standard procedures are available to simulate STA and LTA in the lab for asphalt binders and mixtures (Ziari et al., 2013).

1.3.2 Cracking

Cracking is considered one of the most common asphalt pavement distresses. Asphalt pavements deteriorate more quickly due to cracks, increasing maintenance costs. The cost of maintaining and repairing deteriorated pavements and roadways is significant each year. Research on the factors influencing the cracking behavior of asphalt mixtures is crucial if we are to reduce these expenses.

A variety of cracking tests, including the SCB (Mohammad & Kim, 2011), IFIT (Ozer et al., 2016), and IDEAL-CT (F. Zhou et al., 2017), were introduced in recent decades to evaluate the resistance of mixtures to cracking at intermediate temperatures. Among these cracking tests, the IDEAL-CT test, developed by researchers at Texas A&M University, has been considered a rapid and simple test that can be utilized for mix design, Quality Control (QC) and Quality Assurance (QA) of asphalt mixtures (F. Zhou et al., 2017).

The IDEAL-CT is typically performed at a temperature of 25°C, however agencies can select to vary the test temperature based on the binder's grade and climate. The test does not involve any specimen cutting or notching, and is considered as a simple alternative to the SCB test (Molenaar et al., 2002).

The IDEAL-CT uses an indirect tensile test configuration. The load-deformation curve from the IDEAL-CT is used to calculate a CT_{Index} parameter. The CT_{Index} is a measure of the cracking resistance, and it is calculated based on fracture energy, post-peak slope at 75% peak load (m_{75}), and displacement at 75% peak load (l_{75}). A high CT_{Index} denoting good cracking resistance is influenced by high fracture energy, high l_{75}, and a low m_{75}. The IDEAL-CT was shown to be sensitive to the Asphalt Cement (AC) content, binder grade, percentage of recycled materials, air void content, and aging level (F. Zhou et al., 2017). The results of the IDEAL-CT were validated using field sections in Texas, MnROAD, and the Federal Highway Administration (FHWA) accelerated load facility. The IDEAL-CT was also put through a ruggedness study to see how specimen thickness, loading rate, test temps, and air void content affected the results. The IDEAL-CT has the advantages of being easy to perform, using equipment that is easily accessible for standard indirect tensile strength testing, and being able to use the same size sample as the HWTT test (Molenaar et al., 2002).

1.3.3 Prediction Modeling

Overall, the evaluation of the effect of aging such as STA and LTA on the cracking resistance of asphalt mixes is critical for analyzing the pavement's performance. However, experimental methods for determining the cracking performance of asphalt pavements and aging conditioning impacts entail laborious, time-consuming, and expensive laboratory procedures. Therefore, ML can be utilized to estimate the cracking parameter values quickly and precisely under different aging conditions. One of the objectives of this study is to explore the use of ML to predict cracking resistance using IDEAL-CT.

ML techniques have become increasingly common in asphalt pavement engineering for predicting mixture performance, pavement condition, and binder properties. Compared with traditional regression methods, ML can capture complex relationships among material, environmental, and performance variables. This section highlights a few of the main studies conducted on the applications of ML in asphalt pavements and introduces some of the available ML models. Linear models such as Partial Least Squares (PLS) and multiple linear regression have been used to estimate parameters like pavement temperature (Chao & Jinxi, 2018), moisture content (Q. Li et al., 2022), and asphalt binder rheological properties (Xu et al., 2020). However, these methods are limited in handling nonlinear relationships often observed in asphalt materials.

Nonlinear algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Random Forests (RF), and Multivariate Adaptive Regression Splines (MARS) have shown higher predictive capability. ANN models, in particular, have been widely applied to predict fracture energy (Majidifard et al., 2019; Mirzaiyanrajeh et al., 2022), fatigue life (F. Xiao et al., 2009), rutting depth (Liu et al., 2022), and Marshall properties (Gul et al., 2022; Upadhya et al., 2022). Ensemble models like RF and Gradient Boosting (GB) have also gained attention due to their robustness and ability to manage large datasets (Nguyen et al., 2023). MARS offers interpretability through simple regression equations but may be less accurate for highly nonlinear behaviors (Ghanizadeh et al., 2020).

Recent studies highlight a growing interest in hybrid and explainable ML approaches that combine algorithms such as ANN with optimization or ensemble techniques (Majidifard et al., 2019). These methods improve model accuracy and help identify the key variables influencing performance indicators such as cracking resistance and rutting (Fathi et al., 2019). Overall, ML-based predictive tools are becoming essential for Balanced Mix Design (BMD) frameworks and pavement management, enabling more efficient and reliable performance evaluation of asphalt materials and mixtures.

Table 1 summarizes some of the previous research studies that used different ML models in asphalt materials/pavement applications. PLS, ANN, ENet, SVM, and MARS are some of the most common predictive models that have been used in the fields of asphalt materials and flexible pavement engineering.

Table 1. Previous Research on using ML to Predict Pavement Performance.

ML Method	Previous Studies	Prediction Variables	Findings
PLS (Linear)	(Chao & Jinxi, 2018)	Pavement Temperature	The model reliably and precisely estimated the pavement temperature.
PLS (Linear)	(Xu et al., 2020)	Binder Rheological Properties	PLS was reliable in forecasting the changes in the physical and rheological characteristics of SBS-treated asphalt under aging conditioning.
ANN (Non-Linear)	(Majidifard et al., 2019; Mirzaiyanrajeh et al., 2022)	Fracture Energy	The model used in both studies was suggested for pre-design purposes or to assess the fracture energy of asphalt mixes when testing is not practical.
ANN (Non-Linear)	(Fathi et al., 2019)	Alligator Deterioration Index (ADI)	The hybrid ML approach (ANN/RF) could make precise predictions about ADI to assess long-term pavement performance and degradation.
ANN (Non-Linear)	(Gul et al., 2022)	Marshall Stability, and Marshall Flow	For both marshal stability and flow, a database with 343 data points was created for modeling. The results of this study will help determine the best AI methods for rapidly and accurately estimating the Marshall Parameters.
ANN (Non-Linear)	(F. Xiao et al., 2009)	Fatigue Life	The study's findings indicated that ANN approaches are superior to conventional statistical-based prediction models for estimating the fatigue life of the modified mixes examined in this study.
ANN (Non-Linear)	(Upadhya et al., 2022)	Marshal Stability (MS)	It can be observed that in the testing phase with a small error band, the ANN outperformed other models in predicting the Marshal stability.
ANN (Non-Linear)	(M. Ling et al., 2017)	J-Integral of Top-Down Cracking (J-TDC)	A Finite Element Model produced results for 194,400 examples and its outcomes were used as the database. The ANN was an effective instrument with a remarkable level of prediction accuracy. The R^2 was greater than 0.99.
SVM (Non-Linear)	(Sandamal et al., 2023)	International Roughness Index (IRI)	Five ML models, including SVM, were used. The SVM model was not accurate in predicting the IRI when compared with other models.
SVM (Non-Linear)	(Gopalakrishnan & Kim, 2011)	Dynamic Modulus |E*|	A database, containing 7,400 data points, was used to develop the model, and the results indicated that the SVM model's prediction performance was comparable to the ANN and superior to multivariate regression-based models.
SVM (Non-Linear)	(Liu et al., 2022)	Rutting Depth	The database was taken from the United States Long-Term Pavement Performance (LTPP) program. SVM could predict the rut depth of asphalt mixes accurately.
MARS (Non-Linear)	(Ghanizadeh et al., 2020)	Flow Number	The model's R^2 for the training and testing sets were 0.96 and 0.97, respectively, demonstrating the model's high accuracy.

ML Method	Previous Studies	Prediction Variables	Findings
MARS (Non-Linear)	(Ikeagwuani, 2022)	Resilient Modulus (MR)	Dataset obtained from the LTPP program. The model predicted MR more accurately compared to previously developed AI technique models.

1.3.4 Principal Component Analysis

CT_{Index} condenses interrelated parameters (fracture energy, m_{75}, and l_{75}) into a single metric. While informative, this singular value might not fully capture the details of a mixture's cracking behavior. A more comprehensive understanding of cracking resistance requires examining the individual IDEAL-CT parameters. For example, it was shown that, in some cases, polymer-modification could lead to an increase in the fracture energy while increasing the post peak slope which could result in a similar or lower CT_{Index} (Yin et al., 2023). Hence, analyzing each IDEAL-CT parameter alongside CT_{Index} provides a richer understanding of the complex interplay between a mixture's properties and its susceptibility to cracking.

Principal Component Analysis (PCA), a statistical analysis method introduced by Pearson in 1901 (Pearson, 1901), excels at untangling inherent correlations between variables within a dataset. PCA identifies a new set of uncorrelated variables called Principal Components (PCs), which represent the most significant sources of variation within the data, effectively capturing the underlying structure despite the original variables' interdependencies. Each PC is a linear combination of the original variables, weighed to maximize the variance it explains. The first PC captures the direction of greatest variance, with subsequent PCs explaining progressively less variance while remaining uncorrelated to all preceding components. This process effectively reduces the dimensionality of the data while preserving the most critical information. By examining the influence of each original variable on the identified PCs, the underlying factors governing the data's behavior can be uncovered. This dimensionality reduction not only facilitates visualization of complex datasets but also aids in prioritizing the most relevant aspects for informed decision-making.

PCA has proven to be a valuable tool in asphalt pavement research. Wang et al. (Wang et al., 2018) employed PCA analysis of Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) data to differentiate fifteen bitumen samples from three distinct manufacturers. PCA effectively separated the samples into three clusters, with each cluster corresponding to a specific manufacturer. Zhao et al. used PCA and cluster analysis carried out on eight aging-resistance indexes to classify eighteen types of short-term aged asphalts into two groups according to the oil source (Zhao et al., 2021). Margaritis et al. utilized PCA and hierarchical cluster analysis on chemical and rheological properties to classify nineteen reclaimed asphalt binders into five distinct ageing clusters with statistically different properties reflecting their aging state (Margaritis et al., 2020). Tian et al. used PCA to assess the significance of indexes influencing the short-term aging behavior of asphalt binders (Tian et al., 2022).

PCA can be effectively combined with cluster analysis methods to detect groups and patterns in data (Granato et al., 2018; Margaritis et al., 2020; Zhao et al., 2021). By projecting data into a lower-dimensional space that captures the most significant variance, PCA helps to simplify the data structure. This simplification can lead to more reliable and distinct clusters when using cluster analysis. Classifying multiple samples based on a single parameter is often insufficient for a comprehensive study of sample types. K-means clustering, a popular unsupervised ML algorithm, addresses this by grouping data points into pre-defined k clusters based on the similarity of multiple parameters. It iteratively refines these clusters by strategically positioning k centroids and assigning data points to the nearest centroid based on Euclidean distance. Centroids are then recomputed as the mean of their assigned points. This process continues until the centroids stabilize, signifying convergence and the formation of tight clusters with minimized within-cluster variance. Further information regarding k-means clustering can be found elsewhere (Sinaga & Yang, 2020). K-means clustering is particularly well-suited for exploratory data analysis, uncovering hidden structures and patterns within datasets.

A comprehensive literature review revealed no prior applications of PCA and cluster analysis for asphalt mixture cracking assessment. This study addresses this gap by employing PCA and K-means clustering to analyze IDEAL-CT data and gain deeper insights into the complex relationship between individual parameters and overall cracking resistance.

1.3.5 Rutting

Asphalt pavement rutting remains a critical performance concern, and researchers are continually seeking efficient and reliable methods for its assessment. The HWTT is used by many states to evaluate both rutting and moisture resistance of asphalt mixes. A nationwide survey was conducted by the National Center of Asphalt Technologies (NCAT), revealing that fourteen states currently require HWTT in their specifications. Among these states, 6 states use it to evaluate rutting, while the remaining states employ it for both rutting and moisture damage evaluation (NCAT, 2020). The rutting specifications vary widely between states. For example, Texas Department of Transportation (TxDOT) specifies a minimum number of passes at 12.5 mm rut depth based on the high-temperature binder grade whereas Washington DOT (WSDOT) specifies the minimum number of passes at 10 mm rut depth based on the traffic level. The Oklahoma Department of Transportation (ODOT) currently uses HWTT to assess the rutting resistance of asphalt mixtures and TSR, as outlined in AASHTO T 283, to evaluate moisture resistance. However, there are concerns regarding TSR's effectiveness in accurately reflecting moisture performance in the field.

A HWTT curve undergoes three different stages, i.e. post-compaction, creep, and stripping denoting moisture damage. According to AASHTO T 324, the measured rut depth is used as an indication of rutting. Meanwhile, the Creep Slope (CS), Stripping Slope (SS), and Stripping Inflection Point (SIP) are utilized to assess moisture resistance. The measured rut depth will be referred to hereafter as TRD. From Figure 1, a typical view of the rut depth curve and the parameters calculated using AASHTO T 324 (AASHTO, 2011a) is shown. The NCAT carried out a recent study that suggested that the TRD should not be used to evaluate rutting since it includes the effect of stripping (Yin et al., 2014). It was also suggested that SIP is difficult to determine, which could lead to inaccuracies in the results. Alternative parameters, such as Stripping Number (SN), stripping life (LCST), and the CRD were proposed to evaluate the moisture damage and the rutting resistance, respectively. As shown in Figure 1, the proposed NCAT parameters isolate the effect of stripping and rutting. SN and LCST are two parameters to replace the subjectivity and inconsistency associated with SIP value. These parameters are obtained from fitting the entire HWTT curve using a visco-plastic model rather than drawing lines through selected portions of the data.

SN represents the wheel pass at which stripping begins. It is defined as the point where the curvature of the fitted HWTT curve changes from negative to positive. In practical terms, SN marks the first clear transition from rutting dominated by visco-plastic deformation to rutting dominated by moisture-induced stripping. Because SN is calculated directly from the mathematical curvature of the full curve, it avoids the subjectivity involved in fitting two straight lines to identify SIP. LCST represents the number of additional wheel passes the mixture can sustain after stripping has begun. It is defined as the difference between the SN and the wheel pass at which the rut depth reaches 12.5 mm. LCST therefore quantifies how quickly the mixture deteriorates once stripping initiates. A higher LCST indicates a mix that maintains structural integrity even after the onset of moisture damage. SIP requires the user to manually determine two "best-fit" straight lines for the creep and stripping phases. This process is subjective, varies between analysts, and is sensitive to the amount of data used. By contrast, SN and LCST rely on the full fitted curve and eliminate the need to visually pick transition points. This reduces analyst bias and improves consistency between laboratories (Yin et al., 2014). Field-performance studies demonstrated that SN and CRD correlate more strongly with measured pavement performance than SIP, making them more reliable for agency specification and acceptance decisions (Yin, Chen, et al., 2020).

The effect of rutting due to plastic deformation is captured by the CRD, whereas the effect of rutting due to stripping is given by the difference between the TRD and the CRD. Additionally, the LC12.5 mm is defined as the number of wheel passes at a TRD of 12.5 mm.

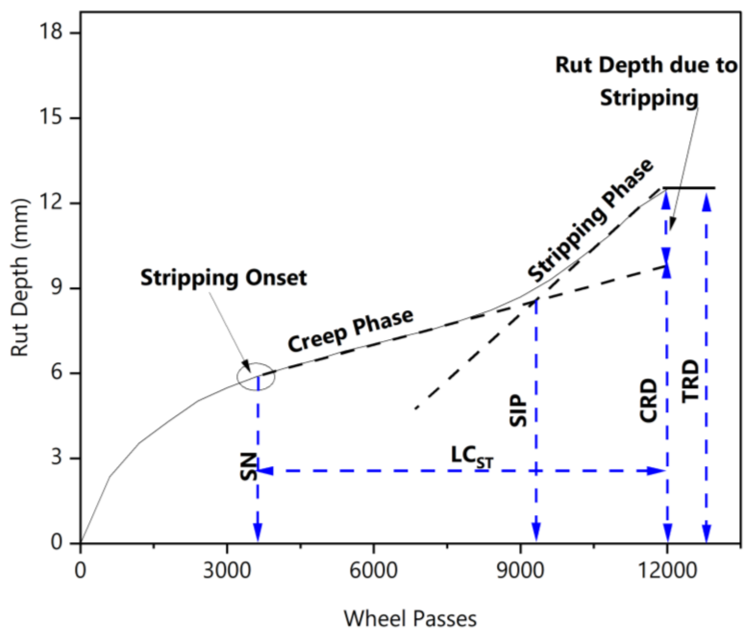


Figure 1. TRD, CRD, SN, and LCST Adapted from (Yin, Chen, et al., 2020).

A study was carried out to implement BMD in Wisconsin, and it was suggested to specify a maximum CRD at 20,000 passes (CRD_{20K}) which varies based on traffic level to evaluate rutting. A minimum SN of 2,000 was used as a criterion for moisture (West et al., 2021; Yin, Chen, et al., 2020).

The Indirect Tensile Test at High Temperature (IDT-HT) has emerged as a promising test due to its simplicity and ability to capture rutting resistance. One of the key advantages of the IDT-HT test is its shorter turnaround time of approximately six hours, including specimen preparation and testing, compared to the HWTT, which makes the IDT-HT test a promising candidate for practical implementation in BMD QC and QA, where timely results are crucial for facilitating production decisions without causing significant delays. Research has established a robust correlation between IDT-HT strength and rutting resistance determined using different rutting tests (Bennert et al., 2018; D. Christensen & Bonaquist, 2007; Yin, Taylor, et al., 2020). These findings suggest that IDT-HT can be used during mix design or as an acceptance tool during production. A study was conducted in Alabama to investigate the potential of using the IDT-HT as a surrogate rutting test for quality control and acceptance testing of BMD mixes, using a temperature of 50.2°C. A minimum IDT strength criterion of 20 psi was specified during the mix design stage (Yin & West, 2021). In a recent study, the potential of using the IDT-HT and other monotonic loading tests as screening tools for assessing the rutting potential of dense graded asphalt surface mixtures in Virginia was evaluated. A test temperature of 54.4°C was suggested to improve repeatability and provide a high degree of discrimination between mixes. A minimum IDT strength of 110 kPa (16 psi) and 133 kPa (19 psi) was proposed for mixes with low traffic and medium traffic, respectively (Boz et al., 2023). Table 2 shows a summary of the studies that proposed rutting performance criteria based on the IDT-HT. The test temperature and the proposed criteria varied widely between the studies. In Table 2, all studies used a loading rate of 50 mm/min unless otherwise noted.

Table 2. Summary of Previous Studies on IDT-HT.

Test Temperature (ºC)	Specimen Size (height × diameter)	Proposed Criteria based on IDT-HT Strength	Notes (Reference)
33ºC	100 mm × 150 mm	Excellent: > 440 kPa (64 psi) Good: > 320 to 440 kPa (46 to 64 psi) Fair: > 200 to 320 kPa (29 to 46 psi) Poor: ≤ 200 kPa (29 psi)	Test temperature was selected based on an approximate rheological equivalence approach considering the critical high pavement temperature and traffic loading conditions. A loading rate of 3.75 mm/min was used. (D. W. Christensen et al., 2000)
35ºC	30-50 mm × 150 mm	< 3 million ESALs: ≥ 32 kPa (5 psi) 3-30 million ESALs: ≥ 49 kPa (7 psi) > 30 million ESALs: ≥ 67 kPa (10 psi)	The proposed criteria were based on standard traffic speed (> 70 km/h) and IDT performed at 35ºC with a loading rate of 0.06 mm/min. (Pellinen & Xiao, 2006)
25ºC	100 mm × 150 mm	75 psi (11 kPa).	(Khosla & Harikrishnan, 2007)
40ºC	115 mm × 150 mm	<0.3 million ESALs: 50 to < 110 kPa (7 to < 16 psi) 0.3-3 million ESALs: 110 to < 170 kPa (16 to < 25 psi) 3-10 million ESALs: 170 to < 270 kPa (25 to < 39 psi) 10-30 million ESALs: 270 to < 430 kPa (39 to < 62 psi) 30-100 million ESALs: 430 to < 660 kPa (62 to < 96 psi) 100-300 million ESALs: ≥ 660 kPa (96 psi)	- Test temperature should be 9°C lower than the yearly, 7-day average, maximum pavement temperature. -Traffic level is at 70 km/h. - An equation for estimating the maximum allowable traffic based on IDT strength was proposed: $Traffic_{max} = 1.97 \times 10^{-5} \times IDT_{strength}^{2.549}$ (D. Christensen & Bonaquist, 2007)
44ºC	95 mm × 150 mm	High-performance thin overlay: ≥ 47 psi (324 kPa) Bituminous-rich intermediate course: ≥ 30 psi (207 kPa) High-RAP surface course: ≥ 47 psi (324 kPa) High-RAP intermediate/base course: ≥ 25 psi (172 kPa)	The temperature was set to 44°C, 10°C below the yearly 7-day average, maximum pavement temperature in New Jersey. (Bennert et al., 2018)
30ºC	63.5 mm x 101.6 mm	300 kPa (43 psi).	The test temperature was set to 30°C which was 20°C lower than the highest average pavement temperature. A loading rate of 2.5 and 3 mm/min was used. (Tran & Takahashi, 2018)

Test Temperature (ºC)	Specimen Size (height × diameter)	Proposed Criteria based on IDT-HT Strength	Notes (Reference)
Test Temperature (ºC)	**Specimen Size (height × diameter)**	**Proposed Criteria based on IDT-HT Strength**	**Notes (Reference)**
44ºC	95 mm × 150 mm	High-RAP surface course (extremely high traffic): ≥ 320 kPa (46 psi) High-RAP surface course (standard traffic): ≥ 155 kPa (22 psi) High-RAP intermediate/base course (extremely high traffic): ≥320kPa(46psi) High-RAP intermediate/base course (standard traffic): ≥ 155 kPa (22 psi) Bituminous-rich intermediate course: ≥ 200 kPa (29 psi) High-performance thin overlay: ≥ 320 kPa (46 psi)	The test temperature was set to 44°C, 10°C below the yearly 7-day average maximum pavement temperature, 20 mm below surface at 50% reliability in New Jersey. (Bennert et al., 2021)
50ºC	62 mm × 95 mm	138 KPa (20 psi)	(Yin & West, 2021)
49ºC	95 mm ×150 mm	150 kPa (22 psi) for less than 10 million ESALs.	The test temperature was set to 49°C, 10°C below the yearly 7-day average maximum pavement temperature, 20 mm below surface at 50% reliability in Virginia. (Meroni et al., 2021)
44ºC	n/a	207 KPa (30 psi).	The test temperature was set to 44°C, 10°C below the yearly 7-day average maximum pavement temperature, 20 mm below surface at 50% reliability in New Jersey. (Bennert et al., 2022)
54.4ºC	62 mm ×150 mm	≤299 AADTT: ≥105 kPa (15 psi) 300-999 AADTT: ≥120 kPa (17 psi)	(Seitllari et al., 2023)
50ºC	95 mm ×150 mm and 62 mm ×150 mm	135 kPa (20 psi).	Test was conducted at 50°C, 10°C lower than the maximum pavement temperature. (Vamsikrishna & Singh, 2023)
54.4ºC	62 mm ×150 mm	Mixture A (0-3 million ESALs): ≥110 kPa (16 psi) Mixture D (3-10 million ESALs): ≥133 kPa (19 psi)	Although 10°C below the yearly 7-day average maximum pavement temperature, 20 mm below surface at 50% reliability was 49°C, the test was conducted at 54.4°C (130°F) because it had improved

Test Temperature (°C)	Specimen Size (height × diameter)	Proposed Criteria based on IDT-HT Strength	Notes (Reference)
			repeatability and higher performance discrimination potential. (Boz et al., 2023)

1.3.6 Moisture Susceptibility

One of the most frequent types of asphalt pavement distress that can cause severe damage to pavements is moisture (Huang et al., 2010; R. Xiao et al., 2022). Stripping, raveling, fatigue cracking, and irreversible deformation are all possible outcomes of moisture damage (Caro et al., 2008) (Hamzah et al., 2015). Several research efforts has gone into researching the processes of moisture damage and creating test procedures that could be used to evaluate the moisture resistance of asphalt pavements (C. Ling et al., 2016; Moraes et al., 2017).

Several attempts have been made to investigate moisture resistance by analyzing asphalt mixes in a lab setting (Amelian et al., 2014; Lottman, 1978). Initially, simple methods like the boiling water test, static immersion test, and so on were used with loose asphalt mixes. The mixtures were then visually inspected to see if the asphalt layer had moved away from the aggregate surface (Weldegiorgis & Tarefder, 2015). Subsequently, an increasing number of experimental techniques utilizing compacted asphalt mixes were put forth, wherein the characteristics of compacted asphalt samples were evaluated before and subsequent to the action of water in order to determine their sensitivity to moisture (X. Chen & Huang, 2008).

The metric for evaluating moisture resistance in the commonly used modified Lottman test (AASHTO T 283) is the indirect Tensile Strength Ratio (TSR) (Lottman, 1978). The AASHTO T 283 test standard (AASHTO, 2014) adopted a modified Lottman technique employing partial vacuum saturation in order to obtain a level of saturation that ranges from 70 to 80 percent and prevent damage to the asphalt samples during conditioning.

Next, a freeze-thaw cycle is applied to the saturated specimen that has been vacuumed. AASHTO T 283 is used by the ODOT to assess the mixes' susceptibility to stripping in Oklahoma (ODOT, 2019) .

The impact of aging on the moisture resistance and performance of asphalt mixes is still unclear and barely evaluated (R. Xiao & Huang, 2022). One of the main variables influencing moisture damage is aging, which modifies the characteristics of the asphalt binder and reduces cohesion and adhesion, which could ultimately affect moisture damage. Therefore, when moisture damage is investigated, the impact of asphalt mixture aging should be taken into consideration to reflect the impact of this natural process on moisture damage (Arabani & Majd Rahimabadi, 2023).

Table 3 summarizes the previous research studies on the effect of different aging conditions on the moisture resistance of Hot Mix Asphalt (HMA) and Warm Mix Asphalt (WMA) asphalt mixes. The results show mixed findings, where some studies have showed that aging can increase moisture susceptibility, while other studies have shown aging to improve moisture resistance.

Table 3. Previous Research Studies on the Effect of Aging on Moisture Resistance.

Study	Aging Method	Test	Moisture Resistance (MR)	Description
(Diab et al., 2019)	160°C,16h oven aging	TSR	Aging affected the MR of HMA mixes with polymer modified binders, and had no effect on HMA mixes with neat binders	This study evaluated HMA mixes with and without polymer modification
(Albayati & Abduljabbar, 2019)	Different STA in terms of time and temperature (2, 4, and 6 hours at 120°C, 135°C, 150°C) for HMA & (2, 4, and 6 hours at 100°C, 115°C, 130°C) for WMA	TSR	Different STA conditions slightly increased the MR of different asphalt mixes	This study evaluated different asphalt mixes such as HMA and open-graded friction course as well as WMA mixes prepared with PG64-16
(Guo et al., 2014)	STA (16h at 130°C) & LTA (5 days at 85°C)	TSR	STA increased the MR of WMA-RAP mixes. LTA deceased the MR of WMA-RAP mixes	This study evaluated warm mix asphalt containing reclaimed asphalt pavement (WMA–RAP) materials (0% and 40% RAP)
(Ali et al., 2022)	LTA (120h at 85°C) based on AASHTO R30	TSR	LTA caused an increase in MR for HMA and a decrease in MR for WMA mixes	This study evaluated HMA and WMA asphalt mixes prepared with PG64-22
(B. Li et al., 2015)	STA (2, 4 and 8 h at 112°C, 129°C and 146°C)	TSR	All STA conditioning at different temperatures and times increased the MR of WMA	This study evaluated Sasobit- WMA asphalt mixes
(Punith et al., 2012)	LTA (120h at 85°C) based on AASHTO R30	HWTT	LTA improved MR of WMA mixtures	This study evaluated foamed WMA mixes containing moist aggregate

1.3.7 Compactability

Achieving optimal compaction in HMA pavement layers is paramount to preventing various pavement distresses including rutting, moisture damage, and cracking. Numerous factors influence a contractor's ability to achieve the target density for HMA mixtures such as weather conditions, the underlying layers' support capacity, layer thickness, the type of compaction equipment used, and the inherent characteristics of the HMA mixture itself (Leiva & West, 2008). Researchers have proposed employing laboratory-measured parameters of the mixture and its components as indicators of both HMA compactability and resistance to permanent deformation (Leiva & West, 2008).

Understanding the compactability of asphalt mixtures is essential for both designing optimal mixes and achieving successful field construction (Hu et al., 2017). It is critical to understand factors that may impact compactability of asphalt mixtures such as binder content and type, compaction temperature and effort, and aggregate gradation and texture (Pouranian & Haddock, 2020).

The Superpave Gyratory Compaction (SGC), a widely accepted method in mix design, allows for the evaluation of asphalt mixture compactability through various parameters derived from the compaction curve. One such measure is the Compaction Energy Index (CEI) (Faheem et al., 2005; Peterson et al., 2004). CEI reflects the variation in a mixture's shear resistance during compaction (Leiva & West, 2008). The concept of the Locking Point (LP) can be used as a threshold for over compaction in the SGC. As compaction progresses, the aggregate particles move closer together, reducing voids and ultimately locking into place, forming a stable internal structure (Vavrik, 2000).

The slope of the densification curve or Compaction Slope (CS), calculated from the percentage of maximum density (%G_{mm}) versus the log number of gyrations (starting at gyration 10), is another potential indicator of compactability. Leiva and West demonstrated that these laboratory-derived parameters (CEI, LP, and CS) are effective in describing the compactability of asphalt mixtures (Leiva & West, 2008). In this study, the effect of aging on the cracking performance and compactibility for various plant-produced mixes will be evaluated.

1.3.8 Effect of Aging on Binder's Rheology

Asphalt binder is the viscoelastic glue in asphalt mixtures that holds aggregates together and greatly influences pavement performance. Binder rheological properties, how the binder responds to stresses or deforms under various temperatures and loading rates, are critical to the pavement's resistance to distresses like rutting and cracking. These properties are time- and temperature-dependent; for example, at high temperatures or slow loading, binders are less stiff and more viscous, whereas at low temperatures or fast loading they exhibit higher stiffness and more elasticity. Aging of the binder over time further alters these rheological characteristics, generally making the binder stiffer and more brittle (R. M. Anderson et al., 2011a).

Asphalt binders exhibit viscoelastic behavior, having both elastic and viscous responses. Two fundamental rheological properties measured in standard tests are the complex shear modulus ($|G^*|$) and the phase angle (δ). The complex modulus $|G^*|$ represents the binder's overall stiffness under oscillatory loading, while the phase angle δ (0° to 90°) indicates the balance of elastic and viscous response ($\delta = 0°$ being purely elastic solid and $\delta = 90°$ a purely viscous liquid). These properties are typically measured using a DSR (D. W. Christensen, 2023). The DSR applies a cyclic torsional strain on a binder sample (sandwiched between parallel plates) and measures the resulting stress. From this, $|G^*|$ and δ are obtained at the test temperature and loading frequency. Under standard Superpave Performance Grade (PG) specifications according to the AASHTO T315 (AASHTO, 2012, p. 315), DSR tests are performed at high service temperatures on unaged or short-term aged binder to control rutting (using the parameter $|G^*|/\sin\delta$), and at intermediate service temperatures on aged binder to address fatigue cracking (R. M. Anderson et al., 2011a).

Beyond the standard one-frequency PG tests, a common approach to fully characterize binder rheology is to conduct a frequency–temperature sweep in the DSR (H. Chen et al., 2025). In this test, the binder's $|G^*|$ and δ are measured over a range of loading frequencies (e.g. 0.1 rad/s to 100 rad/s) and multiple temperatures, typically within the linear viscoelastic range. The resulting data can be shifted using time–temperature superposition to construct a master curve of binder stiffness and phase angle over a broad range of frequencies at a reference temperature (Rowe, 2011). The frequency–temperature sweep is especially useful for evaluating binder properties related to cracking, since many cracking phenomena occur at low loading rates or intermediate temperatures. Researchers use master curve data to derive specialized parameters for cracking performance, such as the Glover-Rowe parameter as an index to (R. Zhang et al., 2022).

Anderson et al. (2011) conducted a landmark study evaluating aged binders and found that as binder oxidative age increases, its ductility at intermediate temperatures drops sharply, corresponding with an increased tendency for cracking (R. M. Anderson et al., 2011b). They introduced the G-R parameter as an indicator of when an asphalt binder becomes so brittle that it cannot relax thermal stresses fast enough to prevent cracking. G-R parameter quantifies age-related embrittlement using DSR measurements: higher G-R values mean a stiffer, more brittle binder that is more crack-susceptible (R. M. Anderson et al., 2011b). Multiple studies confirm that binders with high G-R correspond to higher observed cracking. For instance, Yin et al. (2017) and Kaseer et al. (2018) showed that as binders age, their G-R parameter consistently increases, and those values correlated with measured cracking in mixtures and pavements (Kaseer et al., 2018; Yin et al., 2017). Rodezno et al. (2021) reported the same trend for RAP containing binders had elevated G-R and tended to fall above empirically derived cracking thresholds (Rodezno et al., 2021).

It is useful to consider some threshold values that researchers have proposed. Using Black space diagrams (plots of |G*| vs δ, a tool to visualize binder brittleness), there are two tentative G-R limits for unmodified binders: roughly 180 kPa as a threshold for the onset of block cracking and 600 kPa for severe cracking visible on the pavement surface (Ishaq & Giustozzi, 2022; Martin et al., 2024).

Unlike neat binders, modified binders tend to show a rubbery plateau behavior in their phase angle master curves at the intermediate frequencies/temperatures. This region has been observed for different modified binders such as polymer-modified ones (Lin et al., 2019; H. Zhang et al., 2019)). The study by Asgharzadeh et al. (2013) focused on developing an empirical model to describe master curves of modified asphalt binders. Traditional models often capture |G*| master curves well but sometimes fail to represent the irregularities observed in the phase angle master curves of polymer-modified binders. They proposed a Double-Logistic (DL) function that can model modified binders using parameters with physical significance, such as plateau phase angle, glassy modulus, rheological index, and crossover frequency. DL model fits diverse master curve patterns, properly predicts plateau regions, and provides reliable rheological parameters. Overall, the work demonstrates that phase angle master curves provide deeper insights than |G*| alone, and the DL model is a practical tool for capturing the viscoelastic behavior of neat and modified binders (Asgharzadeh et al., 2013).

2 METHODS AND MATERIALS

2.1 Plant Mixes

A total of thirty-six plant-produced mixes were sampled from different asphalt plants in Oklahoma. Table 4 provides a list of all sampled mixes showing mix code, mix type, RAP content, binder's PG, virgin AC, total AC content, and Recycled Binder Ratio (RBR). RBR is the ratio of the recycled binder to total binder content. It should be noted that, according to ODOT standard specifications, mixes that contain RAP are not allowed to be used as a surface course.

Mixes marked with an asterisk are WMA. A large percentage of the sampled mixes were HMA with only eight of the 36 mixes being WMA. Two of the sampled mixes were Stone Matrix Asphalt (SMA) mixes. The RAP contents of the mixes ranged from 0 to 25 percent by weight of the mix. Figure 2 illustrates the number of mixes corresponding to different mix factors such as mix type, mix production, PG grade, and RAP content.

The mixes included different binder grades such as PG58-28, PG64-22, PG70-28, and PG76-28. The PG58-28 is not commonly used in Oklahoma whereas the PG64-22 is the most used binder grade. NMAS included four different sizes: 4.75, 9.5, 12.5, and 19 mm. The 4.75 mm NMAS is a very fine mix that is not typically used except for specialized applications such as bituminous curbs. The 9.5 mix is a fine mix that is used for leveling and thin lifts. The 12.5 and 19 mm NMAS mixes are generally used for surface and base courses, respectively (Moraes et al., 2023).

Table 4. List of Asphalt Mixtures.

Mix ID	Mix Code	Mix Type	RAP (%)	Virgin Binder's PG	Virgin AC (%)	Total AC (%)	RBR
M1	S4qc0131791400	S4	0	PG76-28	4.9	4.9	0
M2	S3pv0262100600	S3	25	PG64-22	3.1	4.3	0.28
M3	S4pv0262100500	S4	25	PG64-22	3.9	4.9	0.20
M4	S4qc0131901500	S4	25	PG64-22	4.0	4.9	0.18
M5	M2qc0132101000	SMA	0	PG76-28	6.1	6.1	0
M6	S6qc0131702201	S6	25	PG64-22	4.7	6.0	0.22
M7*	WS5qc0101881400	S5	0	PG64-22	6.0	6.0	0
M8	S5qc0131681300	S5	25	PG64-22	3.9	5.3	0.26
M9*	WS3qc0102100600	S3	25	PG64-22	3.0	4.3	0.30
M10	S3qc0522100100	S3	20	PG64-22	3.9	4.9	0.20
M11*	WS5qc0101992400	S5	0	PG70-28	5.4	5.4	0
M12	S5c00932200100	S5	0	PG70-28	5.4	5.4	0
M13	S4pv0442100100	S4	0	PG70-28	5.4	5.4	0
M14	S3qc0062000100	S3	20	PG64-22	3.3	4.3	0.23
M15*	WS4qc0102091500	S4	0	PG76-28	4.9	4.9	0
M16	S4qc0062000900	S4	20	PG64-22	4.3	5.3	0.19
M17*	WS4qc0102192500	S4	0	PG70-28	4.9	4.9	0
M18*	WS4qc0102000910	S4	0	PG70-28	4.9	4.9	0
M19	S4qc0062000200	S4	20	PG64-22	4.3	5.3	0.19
M20	S4qc0061800100	S4	0	PG64-22	5.3	5.3	0
M21	S3qc0062000501	S3	20	PG64-22	3.3	4.3	0.23
M22	S4pv0262201300	S4	0	PG70-28	4.9	4.9	0
M23	S5pv0262101600	S5	15	PG58-28	4.8	5.5	0.13
M24	S4pv0262200300	S4	25	PG64-22	3.5	4.9	0.29
M25*	WS3qc0102101900	S3	25	PG64-22	3.5	4.8	0.27
M26*	WS3pv0262001600	S3	25	PG64-22	3.5	4.5	0.22
M27	S4qc0061800800	S4	0	PG70-28	5.3	5.3	0
M28	S3qc0522190100	S3	20	PG64-22	3.5	4.9	0.29

Mix ID	Mix Code	Mix Type	RAP (%)	Virgin Binder's PG	Virgin AC (%)	Total AC (%)	RBR
M29	S570-2665Gyro	S5	0	PG70-28	6.1	6.1	0
M30	M2c00932200101	SMA	0	PG76-28	7.5	7.5	0
M31	S5pv0262101000	S5	0	PG64-22	5.6	5.6	0
M32	S3pv0262200300	S3	0	PG64-22	4.8	4.8	0
M33	S4qc0062002100	S4	0	PG70-28	5.3	5.3	0
M34	S4pv0442100100	S4	0	PG70-28	5.4	5.4	0
M35	B4pv0442301700	S4	25	PG58-28	4.3	5.6	0.23
M36	S4c00932301600	S4	0	PG70-28	4.9	4.9	0

*WMA mixes

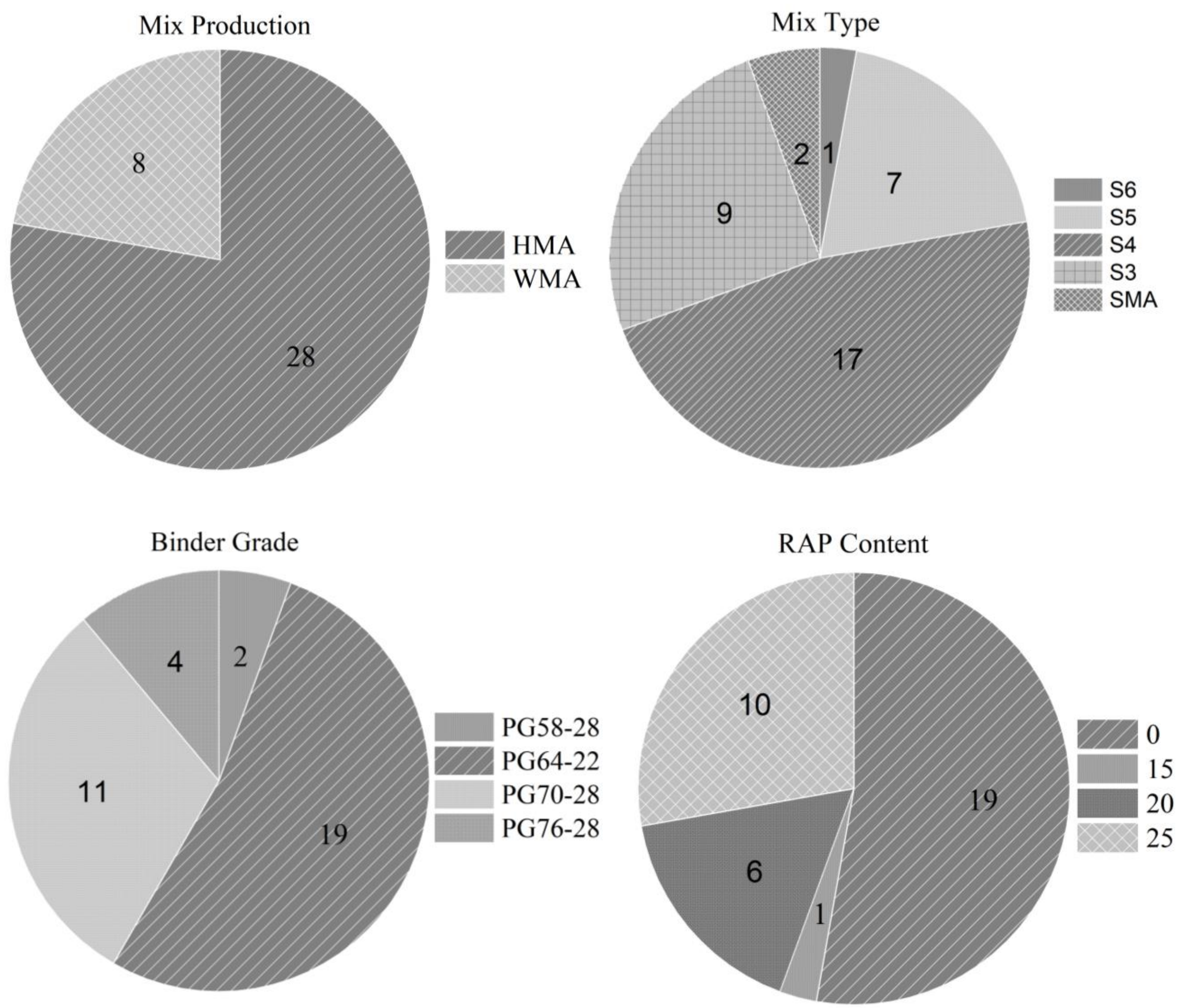


Figure 2. Number of Mixes Corresponding to Different Mix Factors.

All sampled mixes were designed based on the ODOT standard specifications (ODOT, 2019). The mix designs were provided by the Oklahoma DOT. The mix code, shown in Table 4, is a unique code that identifies the mix in the Oklahoma DOT database. The mixes had different gradations, AC content, and RAP content as shown in Table 4. The mix type indicates the NMAS, and the type of gradation as defined in the Oklahoma DOT specifications. The Oklahoma DOT specifies mixes based on their NMAS and gradation where S3, S4, S5, and S6 mixes represent mixes with 3/4" (19 mm), 1/2" (12.5 mm), 3/8" (9.5 mm) and #4 sieve (4.75 mm), respectively. S3, S4, S5, and S6 mixes are designed according to the Superpave method as specified in the ODOT specifications. SMA is a separate category of mixes that typically contain high AC content and use a polymer-modified PG76-28 binder. The SMA mixes sampled in this study had an NMAS of 1/2". The sampled mixes were categorized as either HMA or WMA. The WMA mixes were produced using either a WMA mix additive, or a foaming agent as given by the mix designs. All the WMA mixes used a foaming process called Terex except for mix 26 which used a chemical additive called Evotherm P25. The total AC percentage includes virgin binder and recycled binder from RAP. For virgin mixes with no RAP, the virgin AC percentage is equal to the total AC percentage. Also, those mixes with S4 (12.5 mm) and S5 (9.5 mm) mix type without any RAP and having insoluble residue higher than 30 can be considered as surface courses. SMA mixes are also surface courses. S3 mixes with NMAS of 19 mm are intermediate courses.

Figure 3 provides a summary of all the sampled mixes according to their mix type, binder's PG, and RAP content. The mixes included 19 virgin mixes with no RAP and 17 mixes with RAP. All the mixes containing RAP used either PG58-28 or PG64-22 binder. Polymer-modified binders, i.e., PG70-28 and PG76-28, were only used with virgin mixes containing no RAP. The addition of RAP usually results in increased mix stiffness, hence soft virgin binders such as PG64-22 or PG58-28 are used to mitigate the effect of RAP. All mixes containing 20 percent RAP used PG64-22 binder. The PG58-28 binder was used with only two mixes containing 15 percent and 25 percent RAP. It is worth noting that the current Oklahoma DOT standard specifications do not allow the use of RAP in surface layers. Hence all mixes containing RAP are intended to be used in base courses, shoulders, driveways, or in temporary detours. The list of mixes contained SMA mixes, both were prepared using a PG76-28 binder and one S6 mix containing 25 percent RAP. S6 mixes are very fine mixes that are not commonly used. According to the ODOT statistics for mix usage in 2022, the S6 mixes constitute only 1percent of the total mix tonnage in Oklahoma.

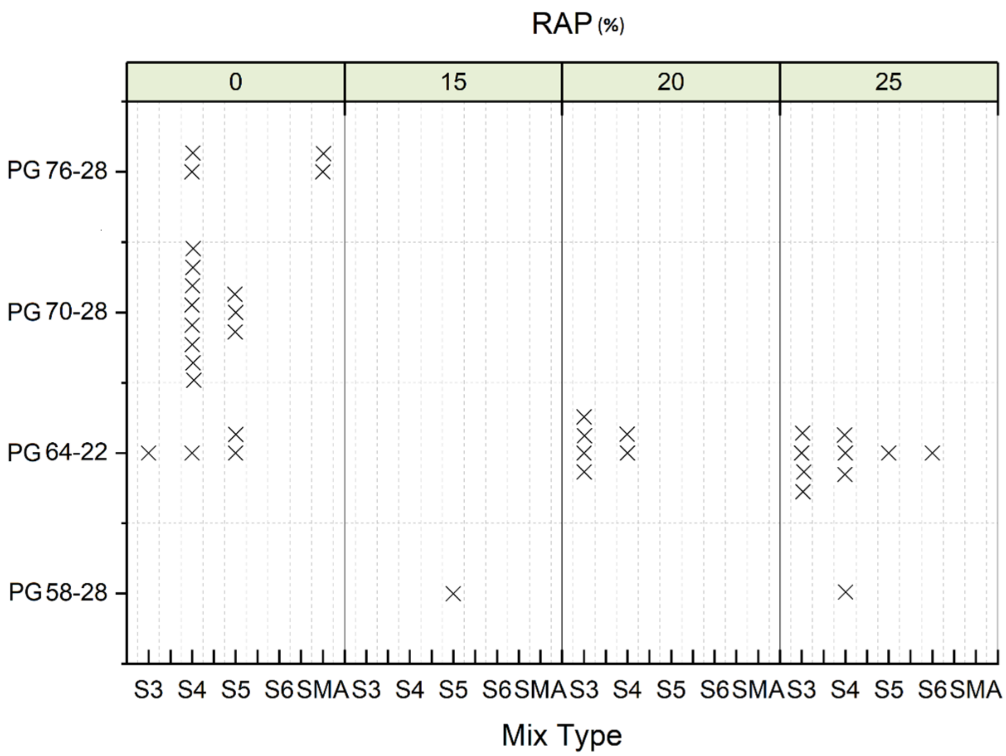


Figure 3. Number of Mixtures based on Mix Type, Binder's PG, and RAP Content.

Figure 4 shows the AC content of all mixes sampled in this study based on their mix type (i.e., S3, S4, S5, S6, and SMA). It can be noted that mixes with lower NMAS have higher AC content. The relationship between the NMAS and AC content is governed by the minimum VMA requirement for different mix types. As per the ODOT specifications, the minimum VMA requirement is higher for mixes with lower NMAS, which usually results in more AC content. The ODOT specifications also specify a minimum AC content based on the mix type, where mixes with finer gradation, i.e., lower NMAS, have a higher minimum AC. The correlation between NMAS, VMA, and AC content will be discussed further in the following sections. SMA mixes showed the highest AC content among all mixes due to their gap gradation and high minimum VMA requirement.

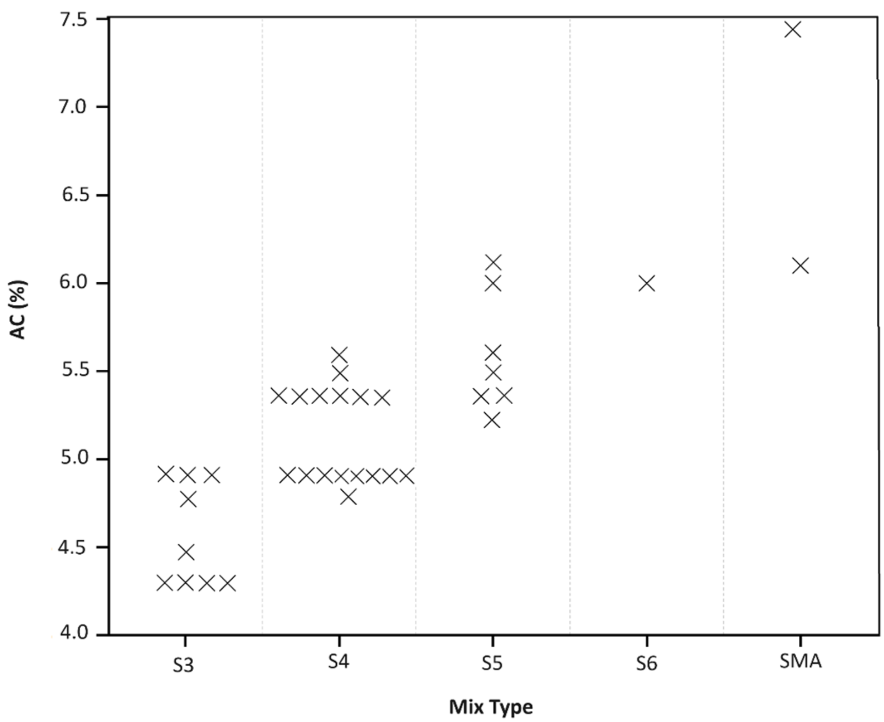


Figure 4. Number of Mixes based on Mix Type, and AC Content.

2.2 Methodology

To achieve the objectives of this study, different plant-produced asphalt mixes were sampled from different asphalt plants across Oklahoma. A total of thirty-six mixes were sampled, representing a wide range of materials and mix designs approved by the ODOT.

Plant-produced mixes were sampled in 5-gallon metal buckets and delivered to the lab. The buckets were heated for 1.5 to 2 hours at 150°C to split the mix into single-size samples using a sample splitter. A mechanical splitter was then used to split the mix into 2600g samples according to AASHTO R 47 (AASHTO, 2019). Each bucket resulted in about 7 to 9 specimens.

The mixes were reheated in the lab and compacted to produce Plant-Mixed Lab-Compacted (PMLC) specimens. The PMLC specimens were tested using the IDEAL-CT at 25^{o}C as part of the benchmarking study. A detailed discussion of benchmarking efforts is provided in the benchmarking study report. In the current study, the mixes were further evaluated to assess the impact of LTA conditioning on the cracking resistance using IDEAL-CT. A long-term oven aging protocol was used in the lab to simulate in-service field aging. Additional testing was also conducted on PMLC specimens using HWTT to evaluate their rutting and moisture resistance. Moreover, combine effect of aging and moisture were evaluated on the PMLC performance.

Prior to producing the specimens, the theoretical maximum specific gravity (G_{mm}) provided in the mix design was verified in the lab, as part of benchmarking study. Trial IDEAL-CT specimens then were compacted in this study to determine the required compaction weight to achieve the target air void level.

2.2.1 IDEAL-CT Testing

The aging study involved subjecting the loose mixes to long-term oven aging and performing IDEAL-CT testing on LTA conditioned specimens. Plant-produced mixes were placed in pans and long-term aged in the oven for 8 hours at 135^{o}C based on AASHTO R 121-24 (AASHTO, 2024). This LTA procedure was used to simulate 3-5 years of in-service aging based on Alabama climate.

After aging, PMLC samples were heated at compaction temperatures (based on HMA or WMA mixes) for 0.5-1 hour by checking the temperature constantly. A total of five IDEAL-CT specimens were compacted at the target air void level of 7 ± 0.5 percent. The air void measurements were conducted according to AASHTO T 269 (AASHTO, 2022).

The specimens were compacted to a height of 62 mm and a diameter of 150 mm. The IDEAL-CT was conducted as per ASTM D8225 (ASTM, 2019a) at a temperature of 25°C.

The IDEAL-CT specimens were sealed in plastic bags and conditioned in a water bath at a temperature of 25°C for 2 hours prior to testing. The IDEAL-CT was conducted using an IDEAL-Plus device. A total of five specimens were tested and the average CT_{Index} was reported. The IDEAL-CT test results of the LTA PMLC specimens were compared with those of the PMLC specimens tested in benchmarking study to assess the impact of aging.

2.2.2 ML Modeling and PCA Analysis

An analysis framework based on ML methods was presented to predict the cracking resistance of asphalt mixtures under different aging conditions using easy-to-establish mix volumetric properties. The research framework of ML modeling is illustrated in Figure 5.

Various parametric and non-parametric ML models were developed to predict the CT_{Index} of Oklahoma asphalt mixes under LTA conditions, using mix volumetric parameters and binder PG. The continuous variables included the n parameter (gradation parameter from the 0.45 power curve), VMA, VFA, DBR, total Asphalt Cement (AC) content, G_{mm}, and RAP content. The categorical variables included NMAS, binder PG, and mix type. All features were obtained from ODOT mix designs.

Data preprocessing involved checking skewness, identifying highly correlated parameters, and applying centering and scaling. The models were trained and evaluated, and their accuracy was compared using performance metrics. Detailed process can be found in section 3.1.5.

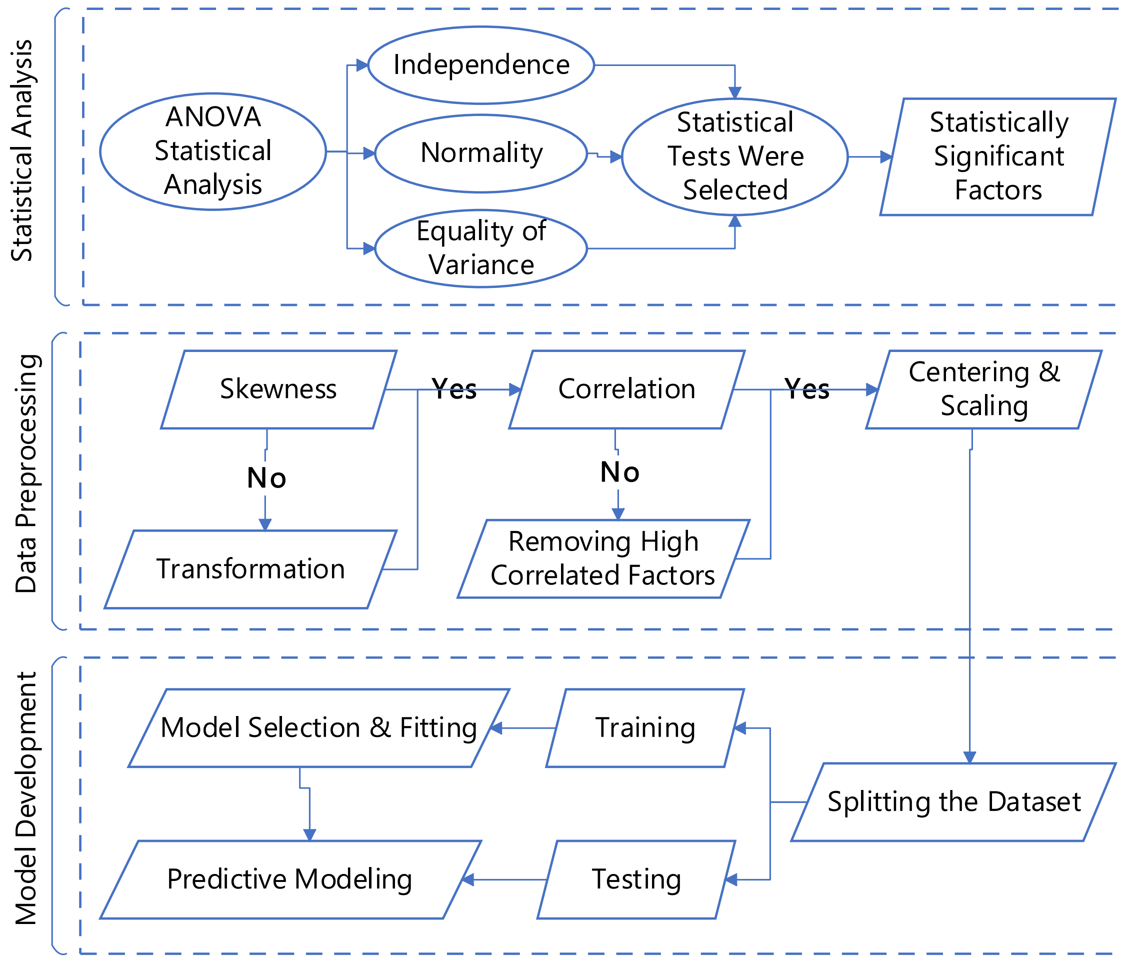


Figure 5. Prediction Modeling Framework.

Additional analysis such as PCA was employed on IDEAL-CT data to analyze and describe cracking performance of mixes under LTA conditioning. PCA was employed to explore the underlying structure within the data and identify potential relationships between the IDEAL-CT parameters and PCs. K-means clustering was utilized to explore the formation of distinct mixture groups based on these parameters. The PCA framework was illustrated in Figure 6.

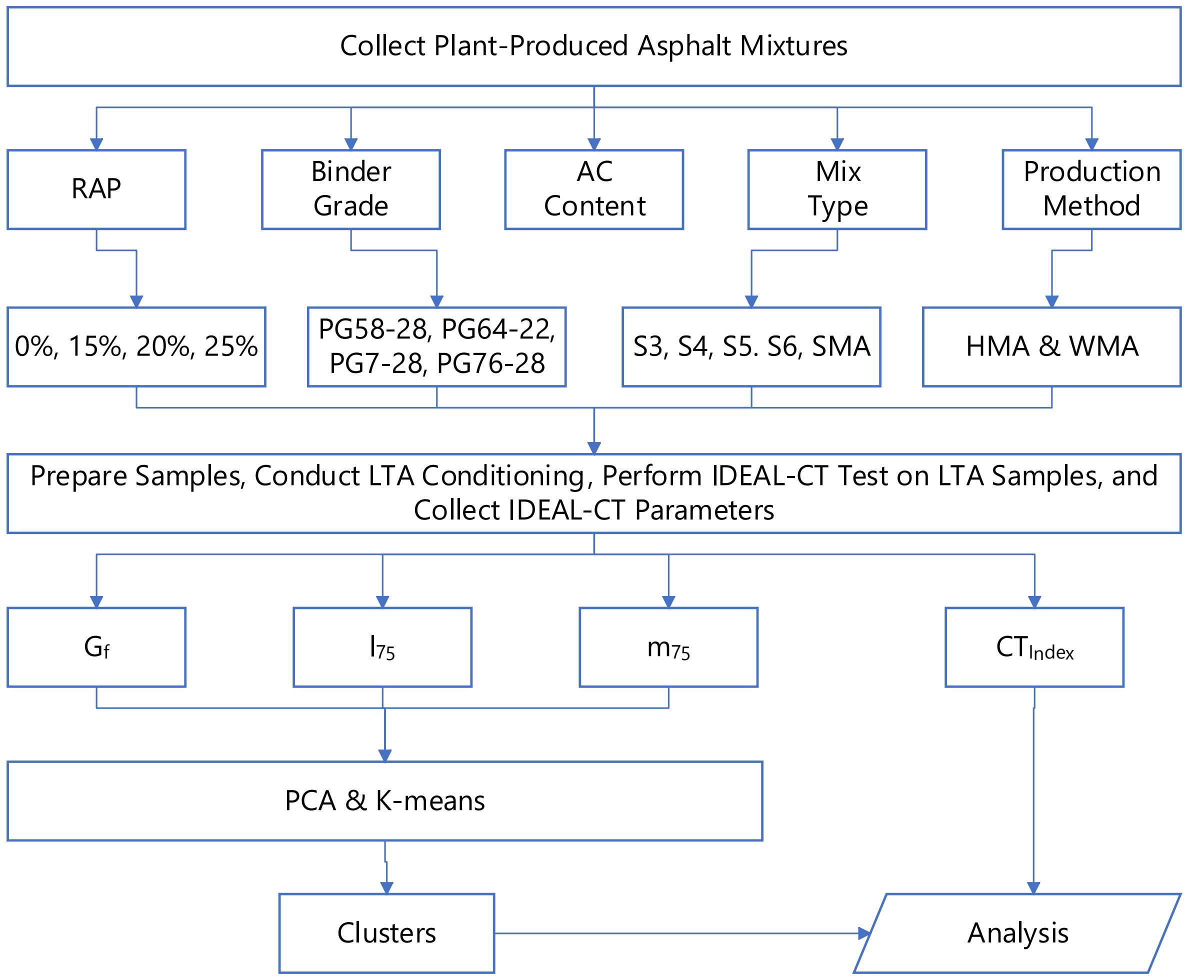


Figure 6. PCA Study Framework.

2.2.3 HWT Testing

To conduct HWTT on all PMLC specimens, additional PMLC specimens were produced using the same procedure adopted in the benchmarking study, which involves splitting the buckets, reheating the samples to the compaction temperature for 2 hours, and compacting them to the desired air void level. The specimens were tested using HWTT to evaluate the mix rutting and moisture resistance. A total of four HWTT samples were prepared. The HWTT specimens were not subjected to LTA. The specimens were compacted at an air void level of 7 ± 0.5 percent using a Superpave gyratory compactor. The specimens were cut to fit the HWTT molds. The HWTT was conducted at 50°C according to AASHTO T 324 (AASHTO, 2011b) and OHDL 55 (OHDL, 2014) using specimens measuring 150 mm in diameter and 62 mm in height. The samples were conditioned in water for 30 minutes at the test temperature (50ºC) before starting the test.

The HWTT was allowed to run up to a maximum of 20,000 passes or until a maximum rut depth of 12.5 mm is reached. As per OHD L 55, an initial run is conducted using two specimens to determine the rut depth at the specified number of wheel passes. The specified number of wheel passes is determined based on the binder's PG as per the ODOT specifications. If the average rut depth from the initial run exceeded 8 mm, another run was conducted using the remaining two specimens, and the average rut depth from the two runs was calculated. OHD L 55 also specifies that the results should not be reported if the rut depth at the specified number of wheel passes exceeds 18 mm. ODOT passing criteria for the HWTT is based on a minimum rut depth of 12.5 mm at 10,000, 15,000, and 20,000 passes for PG64, PG70, and PG76 binders, respectively.

Testing was conducted using two different devices; an Asphalt Pavement Analyzer Junior (APA Jr.) and an HWTT device called SmartTracker from InstroTek. More information about these two devices is presented in Section 2.3.3.

New rutting and moisture parameters (CRD and SN) as mentioned in NCAT report were also conducted on HWTT data of asphalt mixtures in this study to provide a comprehensive analysis of TRD due to rutting not because of stripping. Detailed process can be found in sections 3.2.1 and 3.2.2.

Moreover, the potential of using IDT-HT was examined as a surrogate QC/QA test during production. To conduct IDT-HT to assess its potential for acceptance during production, additional specimens were conditioned in the incubator at a temperature of 50°C for 2 hours before testing. The IDT-HT testing was conducted using an IDEAL-Plus device. Figure 7 illustrates the rutting study framework.

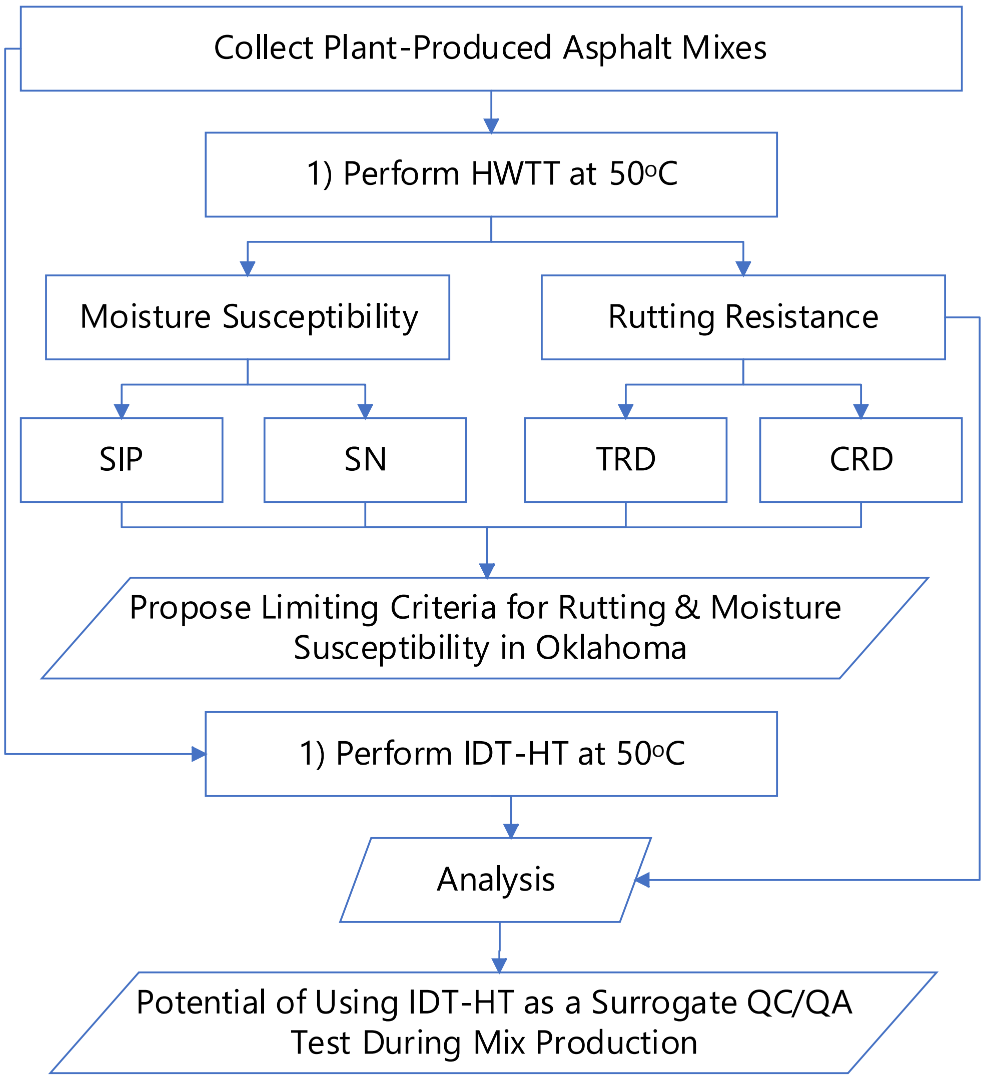


Figure 7. Rutting Study Framework.

To evaluate the influence of freeze-thaw moisture conditioning on the high-temperature mixture performance, four additional samples of nine selected PMLC mixes, based on materials availability, varying in NMAS were prepared and conditioned. Fo each NMAS (19mm, 12.5mm and 9.5mm), three different mixes were selected. The specimens were subjected to one freeze-thaw cycle following AASHTO T283 procedures, which included vacuum saturation, freezing at −18 °C for at least 16 hours, and thawing in a water bath at 60°C for 24 hours. Then, conditioned samples were tested by HWTT and

IDT-HT (2 replicates per mix for each test), same procedures as conducted before, to evaluate the effect of moisture conditioning on the rutting and stripping parameters.

2.2.4 Combined Aging and Moisture

For evaluating the combined effect of aging and moisture, STA and LTA conditioned specimens for seven selected mixtures based on materials availability, all with the same NMAS of 12.5mm, were prepared. Six samples per mix for each aging condition were compacted. Then, moisture conditioning according to AASHTO T 283 was applied on the three samples of each mix. For each aging condition, three samples were control (dry) and other three samples conducted moisture conditioning (wet). The specimens measuring 150 mm in diameter and 62 mm in height were compacted at an air void level of 7 ± 0.5 percent.

Based on the listed mixtures in Table 5, there are two different binder types (PG64-22 and PG70-28). Two mixtures with PG64-22 contain 20% RAP. The mix type of all mixtures is S4.

Table 5. List of Selected Mixes for the Moisture Susceptibility Assessment.

Mix ID	Mix Composition	Virgin AC %	Total AC %	Insoluble Residue
M13	S4-0RAP-PG70-28	5.4	5.4	44.5
M16	S4-20RAP-PG64-22	4.3	5.3	5.1
M19	S4-20RAP-PG64-22	4.3	5.3	5.1
M20	S4-0RAP-PG64-22	5.3	5.3	38.7
M22	S4-0RAP-PG70-28	4.9	4.9	46.8
M27	S4-0RAP-PG70-28	5.3	5.3	40.5
M33	S4-0RAP-PG70-28	5.3	5.3	40.6

The results of tensile strength and TSR values were compared before and after aging conditioning to identify changes. Additionally, the effect of aging conditioning on the cracking parameters of the mixtures were evaluated before and after moisturedamaged procedures to determine the influence of both aging and moisture on the cracking susceptibility.

The tested samples were collected and kept in plastic back after test. Their binders were extracted from the mixtures using an automated asphalt analyzer following ASTM D8159 (ASTM, 2019b). This procedure ensured complete separation of binder from the aggregates and fillers. The extracted solutions were then processed with a rotary evaporator and recovery device according to ASTM D5404/5404M (ASTM, 2021) to obtain the recovered binders. Then, the recovered binder samples from each conditioning stage (STA-Dry, STA-Wet, and LTA-Dry) were carefully sealed in small metal cans and stored at ambient laboratory temperature until further testing.

For binder evaluation, binder testing was performed on the extracted and recovered binders to evaluate the combined effect of aging and moisture on intermediate-temperature cracking behavior of binders. The rheological properties of the recovered binders were characterized by frequency-temperature sweep test using DSR. Frequency-temperature sweep test was conducted in strain-controlled mode at an intermediate range of temperatures (5, 15, and 25^{o}C) and frequencies (0.1–100 rad/s). The test geometry was selected based on the intermediate-temperature range and binder stiffness, 8 mm parallel plate with 2 mm gap for both STA and LTA samples to avoid instrument compliance issues. From the sweep data, the G-R parameter was computed as a combined indicator of stiffness and relaxation capability of the binders at intermediate temperatures. Then, the obtained G-R parameter conditioned samples were compared to evaluate the combined effect of aging and moisture on the binder cracking performance.

2.2.5 Compactability

The compaction raw data of both STA specimens (prepared for HWTT) and LTA specimens (prepared for IDEAL-CT) were obtained from the SGC device. Different compaction parameters such as CEI, LP, CS were calculated from the raw data obtained from SGC and then the effect of aging conditioning and mix parameters on these compaction variables were evaluated. Figure 8 illustrates the research framework of this study.

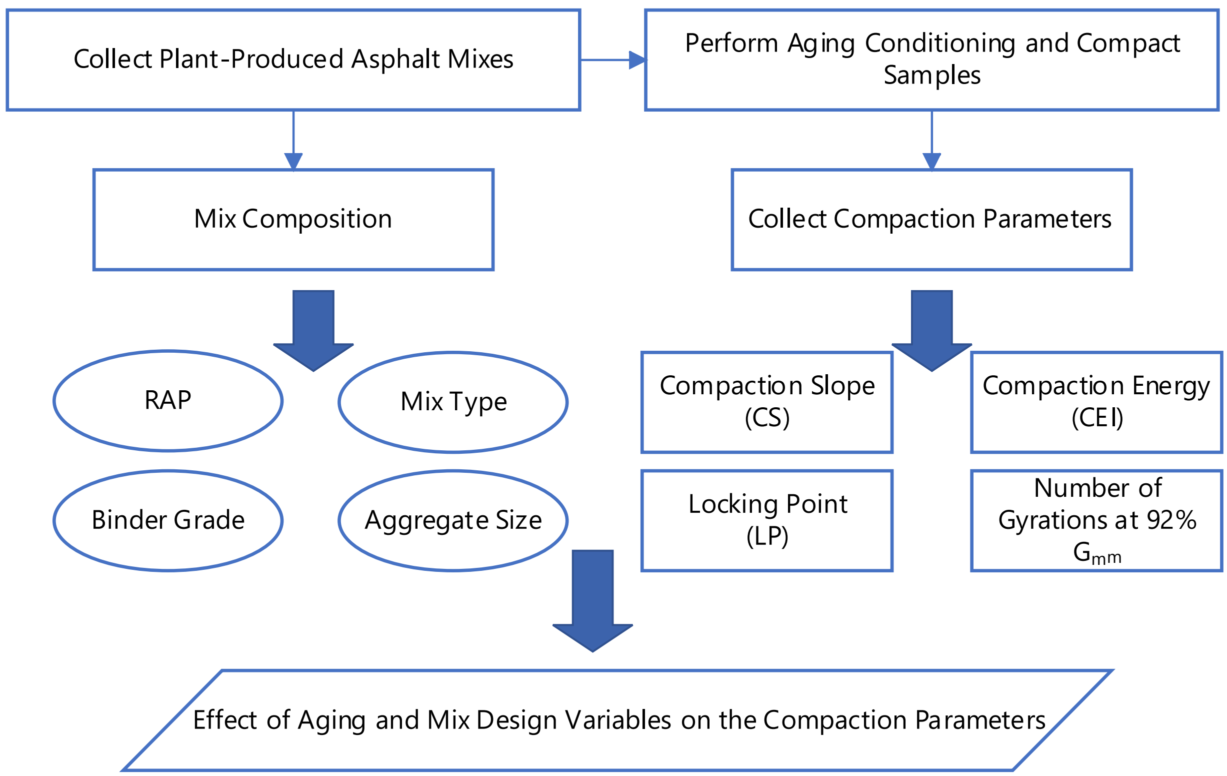


Figure 8. Compactability Study Framework.

CEI for each mix was defined as the area under the compaction curve between the 8th gyration and 92% of G_{mm} (Bahia et al., 1998), as shown in Figure 9. CEI quantifies the overall compaction effort required to densify a mixture throughout the compaction process. LP also defined as the first instance of two consecutive gyrations resulting in the same sample height (Leiva & West, 2008).

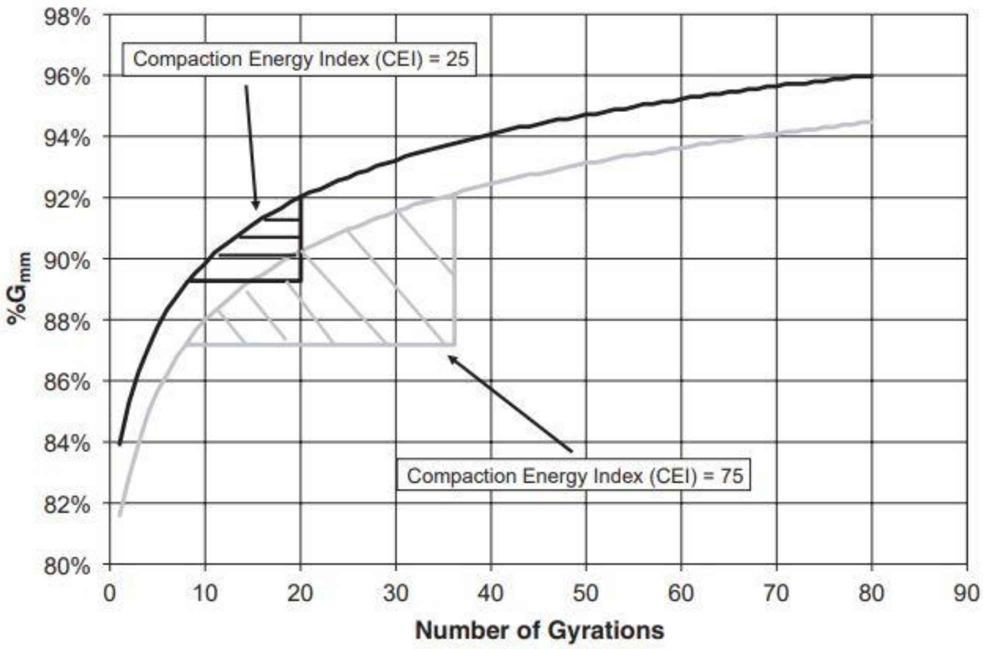


Figure 9. Illustration of CEI (Leiva & West, 2008).

According to ODOT specifications (ODOT, 2019), the design number of gyrations (N_{des}) and initial number of gyrations (N_{ini}), which represent the compaction effort applied during mix design, are specified based on binder grade, as shown in Table 6.

Table 6. Number of SGC Gyrations for Different PG in Oklahoma (ODOT, 2019).

Binder PG	N_{ini}	N_{des}
64-22	6	50
70-28	7	65
76-28	8	80

The CS parameter is calculated by using Equation 1. This parameter represents the rate of densification during the linear compaction phase, reflecting how quickly the mix compacts between the initial and design gyrations (Leiva & West, 2008).

$$CS = 100 * \frac{\%G_{mm}\ at\ N_{des} - \%G_{mm}\ at\ N_{ini}}{\log(N_{des}) - \log(N_{ini})} \quad (1)$$

Where *CS* is compaction slope, G_{mm} *at* N_{des} and G_{mm} *at* N_{ini} are the levels (%) of compaction (density) achieved at the design number of gyrations (N_{des}) and N_{ini}, respectively.

2.3 Test Methods

This section provides an overview of the cracking test, i.e., IDEAL-CT, rutting and moisture test, i.e., HWTT, and moisture test, i.e., modified Lottman that were used in this study.

2.3.1 Indirect Tensile Asphalt Cracking Test (IDEAL-CT)

The IDEAL-CT was developed by researchers at Texas A&M (F. Zhou et al., 2017) (Zhou et al., 2017). The IDEAL-CT uses an indirect tensile test configuration as shown in Figure 10. The test is conducted at 25°C and is considered a simple and practical alternative to the SCB test configuration since it does not involve any specimen cutting or notching. Several studies have shown that IDEAL-CT provides a good correlation with the IFIT test (Elkashef et al., 2022; F. Zhou et al., 2017) .

Figure 10. IDEAL-CT Setup.

A typical IDEAL-CT load-displacement curve is shown in Figure 11. The IDEAL-CT results are used to calculate a CT_{Index}, according to ASTM D8225. The CT_{Index} is based on the fracture energy (G_f), m_{75}, and l_{75}, as shown in Figure 11. A high CT_{Index} is an indication of better cracking resistance. A high fracture energy, high l_{75} a low m_{75} contribute to a high CT_{Index} and better-cracking resistance.

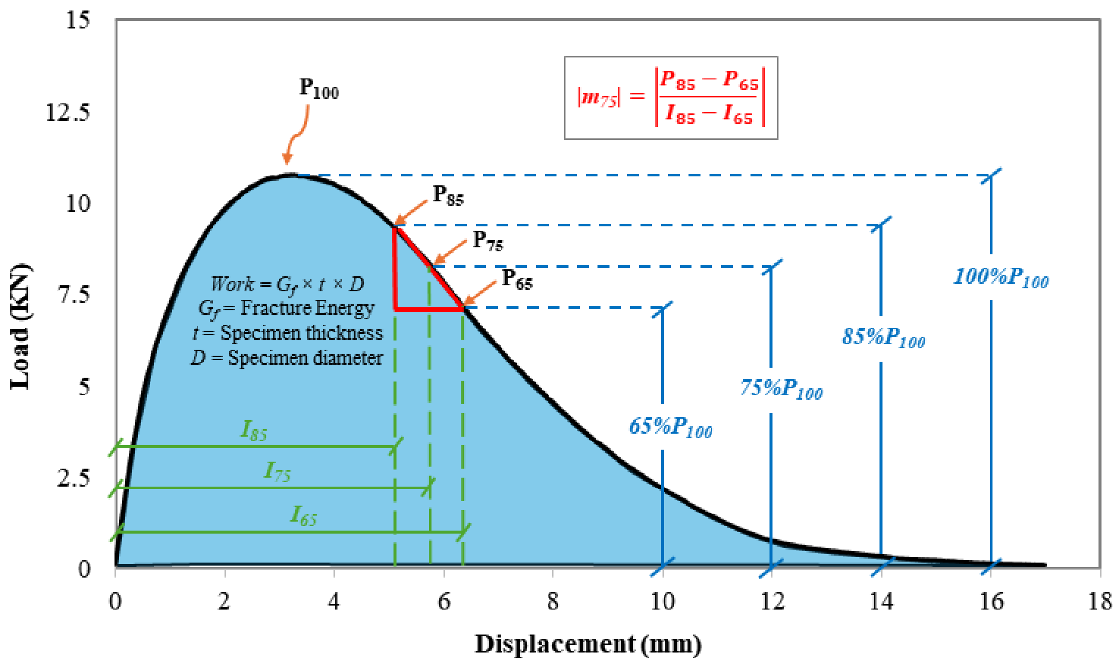
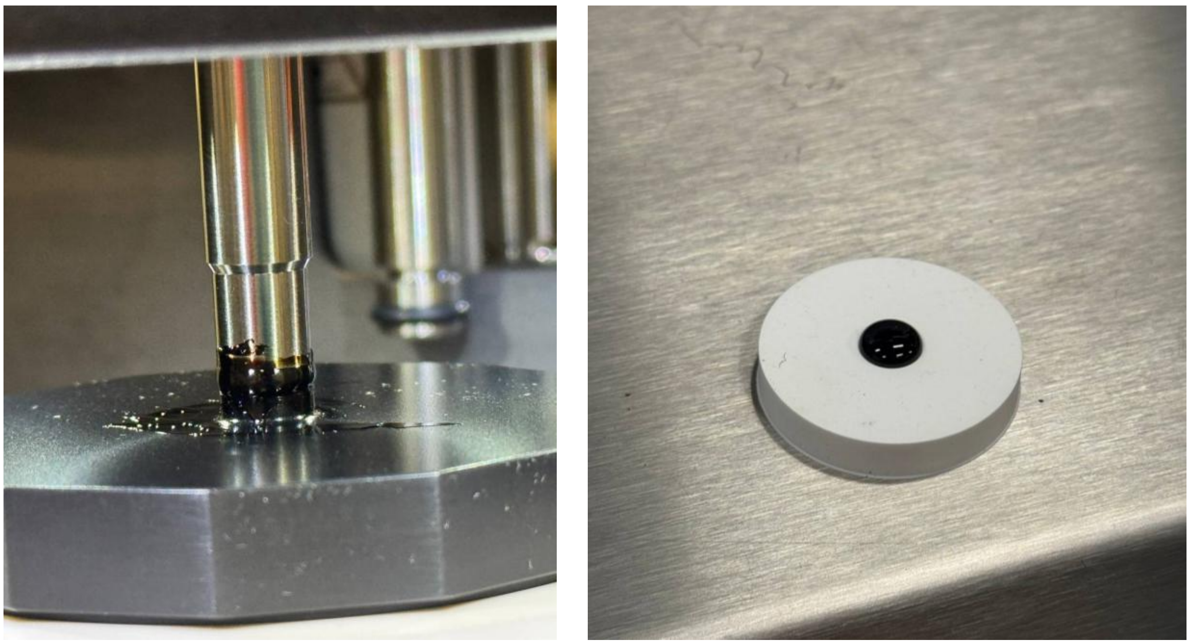


Figure 11. Typical Load-Displacement Curve for IDEAL-CT.

Several studies have suggested that the CT_{Index} from the IDEAL-CT and FI from the IFIT have shown a good correlation with field cracking including fatigue, thermal, and reflective cracking (Kriz et al., 2014; Ozer et al., 2018). The IDEAL-CT was validated using test sections at the FHWA and a limited number of test sections in Texas (C. Chen, 2020; Hawraa & Baaj, 2018). The work conducted at NCAT has also shown that the Illinois SCB test and the IDEAL-CT test were able to identify mixes with good and bad cracking performance using seven short-term aged lab-mix lab-compacted mixtures (C. Chen, 2020).

2.3.2 Dynamic Shear Rheometer (DSR) and Frequency-Temperature Sweep

The frequency sweep in this study was conducted in frequency ranges between 0.1 and 100 rad/s at intermediate temperatures of 5°C, 15°C, and 25°C and shear strain of 0.1% using a DSR instrument as shown in Figure 12(a). The plate's gap and diameter utilized in this study were 2 mm and 8 mm, respectively as presented in Figure 12(b) (Z. Zhou et al., 2020).



(a) (b)

Figure 12. DSR Test Setup; (a) Trimmed Sample between Plates, (b) 8-mm Binder Sample.

The objective of this test is to specify |G*| and δ of binders at varying frequencies, as well as to construct complex shear modulus and phase angle master curves. The resulting data at temperatures of 5°C and 25°C was shifted horizontally using time–temperature superposition to their counterparts at the reference temperature of 15°C to construct a master curve of binder stiffness and phase angle over the range of frequencies (D. W. Christensen, 2023). Master curves succinctly represent binder rheological behavior; for instance, the low-frequency region of the curve corresponds to cold-temperature or long-duration loads, where |G*| tends to be high and δ tends to drop, whereas the high-frequency region corresponds to warm, rapid loads.

The complex shear modulus master curves were constructed using the Christensen-Anderson-Marasteanu (CAM) model (D. W. Christensen & Anderson, 1992) for all the binders. The same model was applied for construction of phase angle master curves of neat and modified binders. If CAM model cannot fit the model for modified binders, the DL model (Asgharzadeh et al., 2013) should be used for construction of phase angle master curves.

Complex modulus master curves were constructed using CAM model presented in Equation 2 (Wu et al., 2021; Yusoff et al., 2013).

$$|\boldsymbol{G}^*| = \frac{|G_g^*|}{\left[1+\left(\frac{f_c}{f'}\right)^k\right]^{m/k}} \tag{2}$$

Where $|G_g^*|$ is glassy shear modulus considered 10^9 Pa, in this study, K and m are fitting constants, f′ is reduced frequency (Hz), and f_c is cross-over frequency (Hz).

Figure 13 shows the complex modulus master curve construction. As can be seen, first, the measured complex modulus values of the binder at temperatures of 5°C and 25°C were shifted horizontally to the reference temperature of 15°C. Then, the CAM model was fitted to the complex modulus values versus reduced frequency data. In this study, the same approach was adopted to construct complex modulus master curves for all binder types.

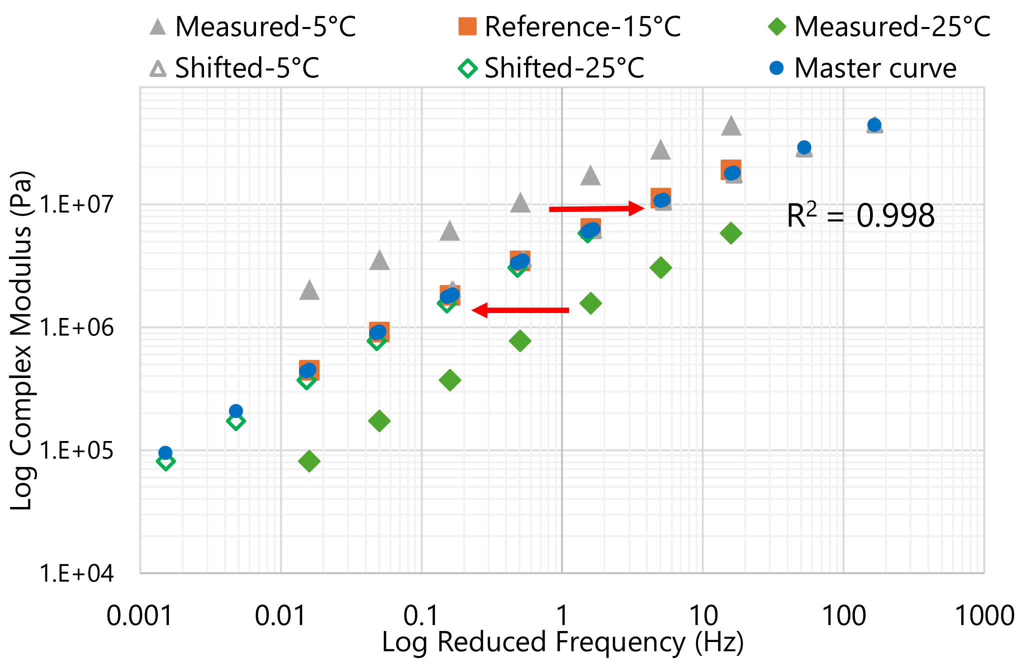


Figure 13. Complex Modulus Master Curve Construction.

Figure 14 schematically shows the CAM model parameters. From these curves, crossover frequency (f_c), crossover modulus (G_c*) and rheological index (R) were extracted and used to valuate rheological properties of asphalt binders. Crossover frequency is the frequency at which the storage and loss modulus are equal at the reference temperature (15°C in this study) corresponding to a phase angle of 45° (δ = 45°). The crossover frequency is binder specific and is a hardness parameter which indicates the general consistency of a binder. A smaller f_c value indicates that the material is exhibiting elastic behavior over a larger range of frequencies. The complex shear modulus of the binder at f_c is defined as the crossover modulus. R index is the difference between the log of the glassy modulus (10^9 Pa in this study) and the log of |G*| at the crossover frequency ($\log(G_g^*/G_c^*)$) (D. A. Anderson et al., 1994a). A larger value for R denotes more pronounced elastic behavior relative to the viscous behavior at the intermediate range of frequencies and temperatures (Farrokhzade et al., 2022). R correlates with important binder properties, including fatigue resistance, chemical composition and degree of oxidative hardening (D. W. Christensen et al., 2017).

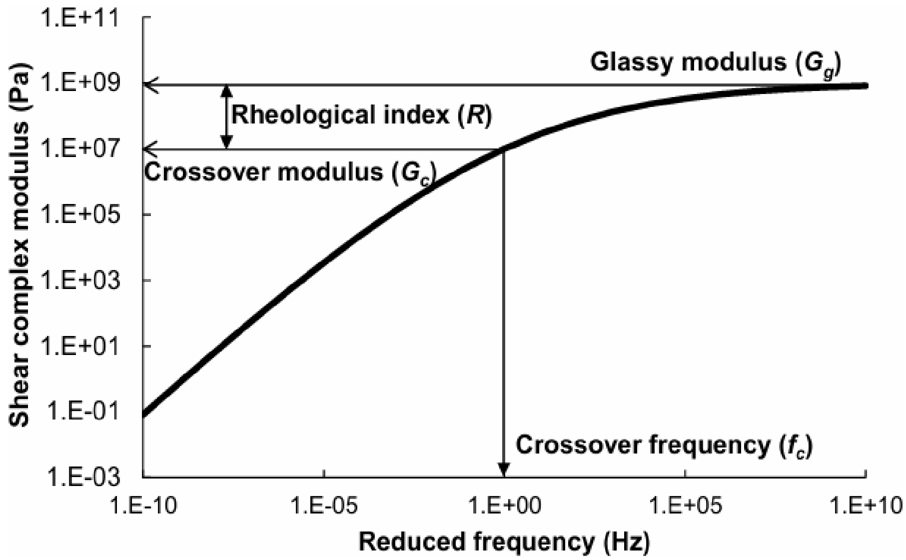


Figure 14. CAM Model Parameters (Farrokhzade et al., 2022).

After constructing complex modulus master curves of binders, the phase angle master curves were also constructed at the reference temperature of 15°C using the CAM model presented in Equation 3 (Safaei et al., 2016).

$$\delta = \frac{90 \times m}{\left[1+\frac{f'}{f_c}\right]^k} \tag{3}$$

Where K and m are fitting constants, f' is reduced frequency (Hz), and f_c is cross-over frequency (Hz).

The stages mentioned for constructing the complex modulus master curves are also true for phase angle master curves construction, as shown in Figure 15.

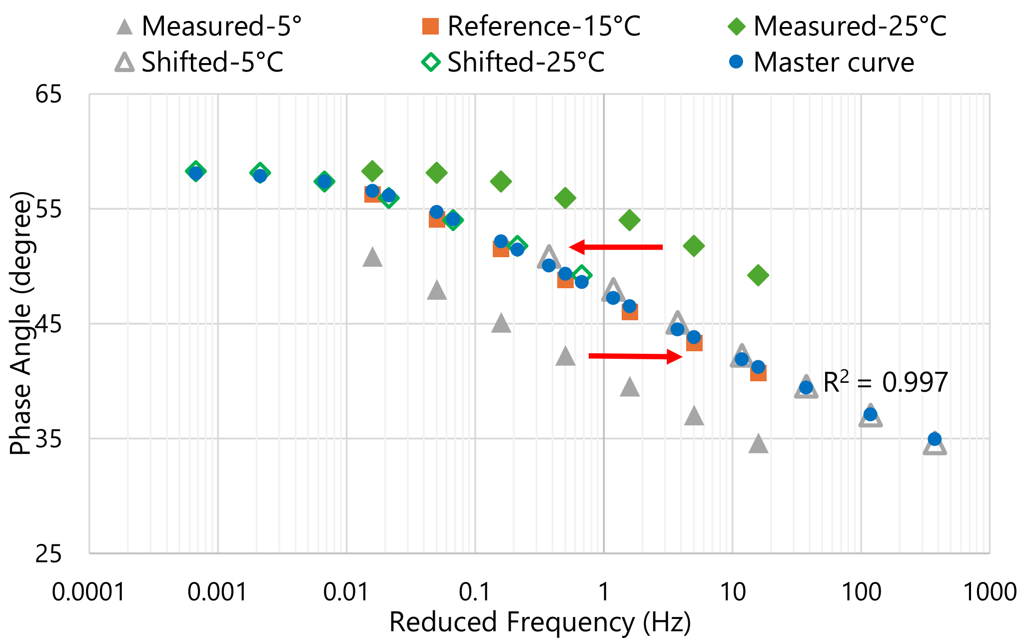


Figure 15. Phase Angle Master Curve Construction.

Glover-Rowe (G-R) Parameter is widely used as an indicator of long-term binder embrittlement at intermediate service temperatures. It was derived from the earlier Glover parameter. Glover et al., (2005) found that the reciprocal of the ductility of aged binders at 15°C correlated well with G′/(η′/G′), a DSR function relating storage modulus G′ and dynamic viscosity η′) (Glover et al., 2005). Rowe simplified this into the G-R parameter (Equation 4), defined purely in terms of |G*| and δ at a low loading frequency (Rowe, 2011):

$$G - R = |G^*|\frac{(\cos\delta)^2}{\sin\delta} \tag{4}$$

G-R typically evaluated at 15°C and 0.005 rad/s (very low frequency) (Mensching et al., 2015). Intuitively, G-R combines binder stiffness with a penalty for low phase angle (since a brittle binder with low δ will increase G-R). A high G-R means the binder is both stiff and unable to relax (high stiffness and low δ), signaling high cracking potential (Martin et al., 2024). As noted NCHRP research report 927, the G-R thresholds are about 180 kPa as the warning (crack onset) for LTA and 600 kPa as the limit (extensive cracking) for STA for binders (Ishaq & Giustozzi, 2022; Kaseer et al., 2018). Asphalt binders with higher G-R values are expected to have experienced a greater level of aging than those with lower G-R (C. Chen et al., 2018b).

2.3.3 Hamburg Wheel Tracking Test (HWTT)

The HWTT is conducted by applying a load that simulates moving traffic using two steel wheels with 158 lbs. each. The wheels travel across the specimens at a speed of 53 ± 2 passes per minute. The specimens are submerged in water at a given test temperature. The test temperature is specified based on the binder grade, and climatic conditions. The average rut depth is measured and is used as an indication of the rutting resistance of the mix. Figure 16 shows the HWTT setup and the observed rutting in the specimens after the test (Rahman et al., 2011). A typical rutting curve is shown in Figure 17. An initial post-compaction stage is observed, followed by a creep stage with a constant slope. As the Number of Wheel Passes (NWP) increases, stripping could occur indicating the moisture sensitivity of the mix. A SIP is an indication of the onset of stripping failure.

(a) (b)

Figure 16. HWTT Setup Showing Specimens; (a) Before, and (b) After Testing.

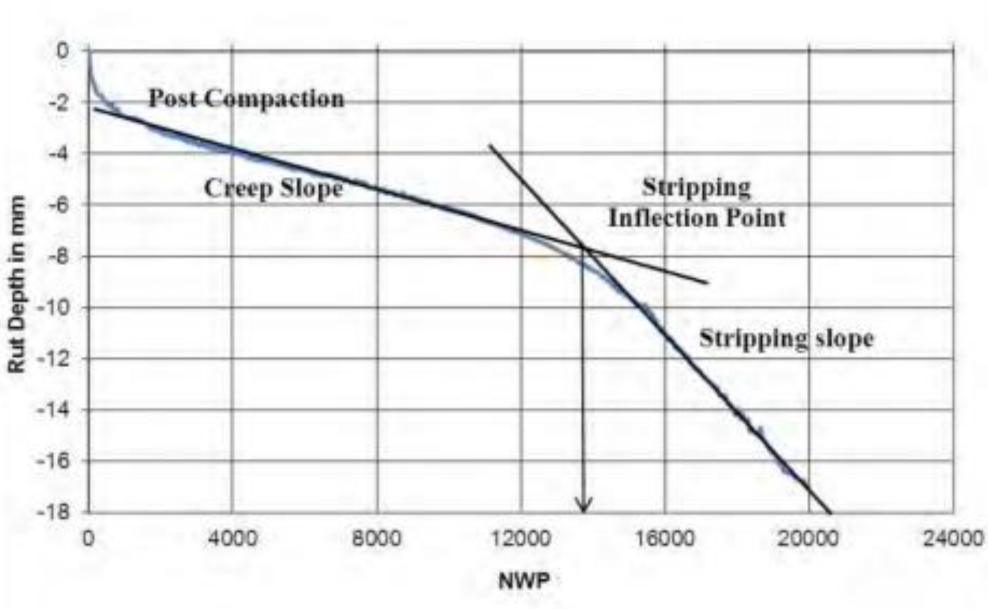


Figure 17. Typical Rut-Depth Curve for HWTT.

In this study, The HWTT was conducted at the Ingevity facility in Tulsa, OK using a device called SmatTracker from InstroTek. However, the APA Jr. started to show errors in the middle of testing, so the research team decided to continue the HWTT at the Ingevity facility in Tulsa, OK using an HWTT device called SmatTracker from InstroTek. The test procedure is outlined below:

- Following the preparation of the samples, two samples from each mix were placed in each of the right and left moulds. The moulds containing the specimens were placed in the confining plate and fixed in place.
- The filling/preheating option was then selected to condition the samples for 30 minutes. The specimen was then slid down the guide rails into the water bath, and the retaining nut was tightened by hand to hold the specimen in place. The device lid was then closed, and the test setup began.
- The water mode is selected with the wheel pass limit set to 20,000, and with the rut depth feature enabled and set to 12.5 mm. The speed of the wheels was set at 52 passes per minute and combined mode is selected to ensure both the right and left wheels work together.

- The “rut depth” feature was then selected so that the device stops when the rut depth is equal to or greater than 12.5 mm. The water mode was checked again, and the air mode was unchecked, the device began to heat the water to the set temperature of 50°C.
- The machine was checked to see if it was ready by checking if there were any active alarms. A USB drive was inserted into the device to save the result of the test. The test was then allowed to start.

2.3.4 Modified Lottman test (AASHTO T 283)

For the AASTHO T 283 test, a total of six specimens measuring 150 mm in diameter and 62 mm in height were prepared. The specimens were divided into two sets of three specimens each. One set representing the control was tested in an unconditioned state (dry), while a second set was conditioned (wet) according to AASTHO T 283. The mixes were short-term aged for 2 hours at 150^{o}C prior to compaction according to the ODOT specifications. Compaction was done using the SGC at an air void level of 7 ± 0.5 percent.

The AASHTO T 283 moisture conditioning protocol involves vacuum saturation to bring the specimens to a saturation level between 70 and 80 percent. The vacuum saturated specimens are sealed with a plastic film, placed in a plastic bag with 10 ml of water, and stored in a freezer for a minimum of 16 hours. Figure 18 shows the specimen after freezing. Following the freezing cycle, the specimens are taken out of the plastic bag and the plastic film removed. The specimens are then immersed in a hot water bath at 140^{o}C for 24 hours. The AASTHO T 283 is meant to simulate freeze-thaw action in the field. All specimens were conditioned in a water bath at 25^{o}C for 2 hours before testing. The dry and conditioned specimens were tested using the indirect-tensile strength test in the IDEAL-Plus device. TSR is determined as the ratio between the strength of the conditioned and dry specimens.

(a)

(b)

(c)

(d)

Figure 18. Process of Moisture Conditioning.

2.3.5 Indirect Tensile Test – High Temperature (IDT-HT)

The IDT-HT was conducted as per ASTM D8225 at a temperature of 50°C. However, the test specimens were conditioned at 50°C to match the temperature at which the HWTT specimens were conditioned. From the test results, strength and stiffness are calculated. For conditioning the samples prior to testing, the compacted specimens were placed in a conditioning chamber for 2 hours ± 10 minutes at 50°C ± 1°C. To conduct the test, the IDEAL-CT device was used. Figure 19 shows the conditioning of samples in the environmental chamber and IDT-HT test setup. A typical IDEAL-CT load-displacement curve is shown in Figure 20. The IDT results are used to calculate strength and stiffness, according to ASTM D8225.

(a) (b)

Figure 19. IDT-HT; (a) Samples Conditioning at 50°C, (b) Test Setup.

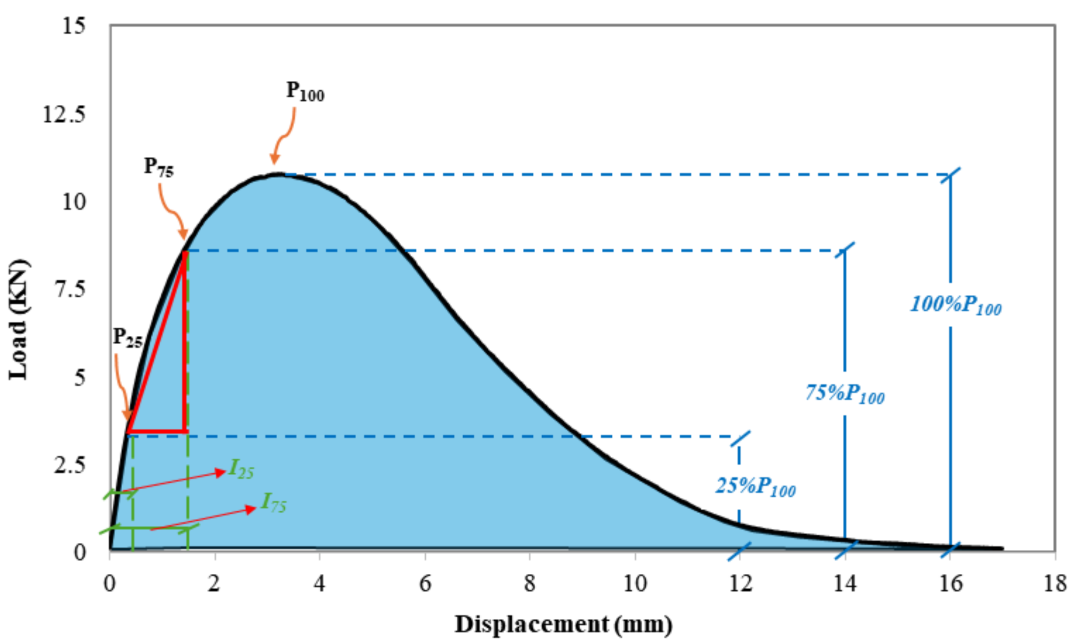


Figure 20. Typical Load-Displacement Curve for IDT-HT.

The IDT strength is commonly calculated to assess the rutting potential of the asphalt mixture. As shown in Equation 5, the IDT-HT strength is calculated based on the peak load obtained during the test and the specimen dimensions. A higher IDT strength value signifies better rutting resistance in the asphalt mixture.

$$S_t = \frac{2000P}{\pi t D} \quad (5)$$

Where, P = peak load (N), t = specimen thickness (mm), and D = specimen diameter (mm).

$$Stiffness = \left(\frac{P_{75} - P_{25}}{l_{75} - l_{25}}\right) \quad (6)$$

Where, P_{75} = 75% of the peak load (N) before the post-peak stage, P_{25} = 25% of the peak load (N) before the post-peak stage, l_{75} = displacement (mm), corresponding to 75% of the peak load before the post-peak stage, and l_{25} = displacement (mm), corresponding to 25% of the peak load before the post-peak stage.

For the IDT-HT stiffness, the calculation is the same as the m_{75}, but this time we calculate the slope (stiffness), as mentioned in Equation 6, of the line between the l_{75} and l_{25}. Note that the displacements mentioned above occur prior to reaching the peak load.

3 RESULTS AND ANALYSIS

3.1 Cracking

3.1.1 CT_{Index} Results

Figure 21 provides a plot of the frequency and cumulative distribution of the CT_{Index} values of the STA and LTA specimens. It can be shown that 50% of the sampled mixes had an average CT_{Index} of 75 under STA condition and 45 under LTA condition.

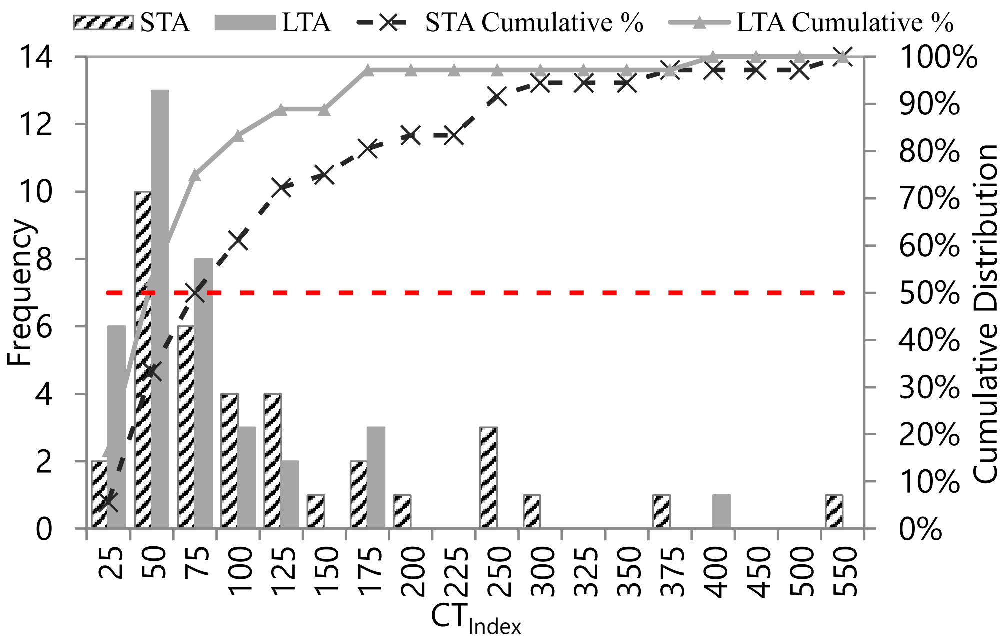


Figure 21. Frequency and Distribution CT_{Index} at STA and LTA Conditions for all Mixes.

Figure 22, Figure 23, Figure 24, and Figure 25 show the frequency distribution and cumulative plots of the CT_{Index} values for S3, S4, S5, and SMA mixes, respectively. As previously mentioned, S4 mixes are generally used for surface courses, whereas S3 mixes are used for base courses. The current ODOT BMD provisional specifications specify a CT_{Index} of 60 for S3 mixes, and a CT_{Index} of 100 for S4 mixes at short-term aged conditions. Based on the reported results, only three S3 mixes and nine S4 mixes meet the BMD provisional specifications.

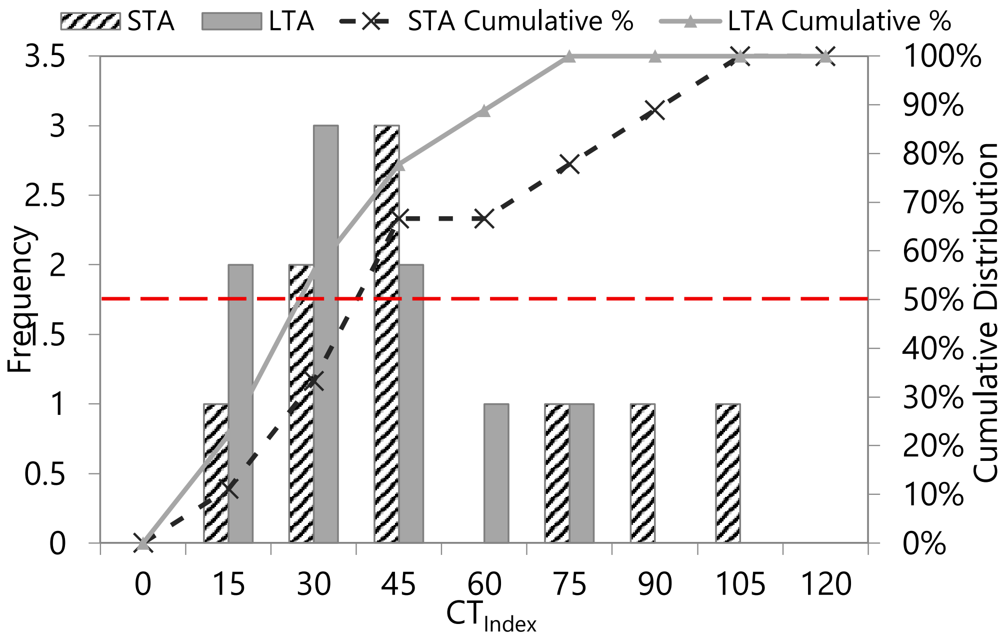


Figure 22. Frequency and Distribution of CT_{Index} at STA and LTA Conditions for S3 Mixes.

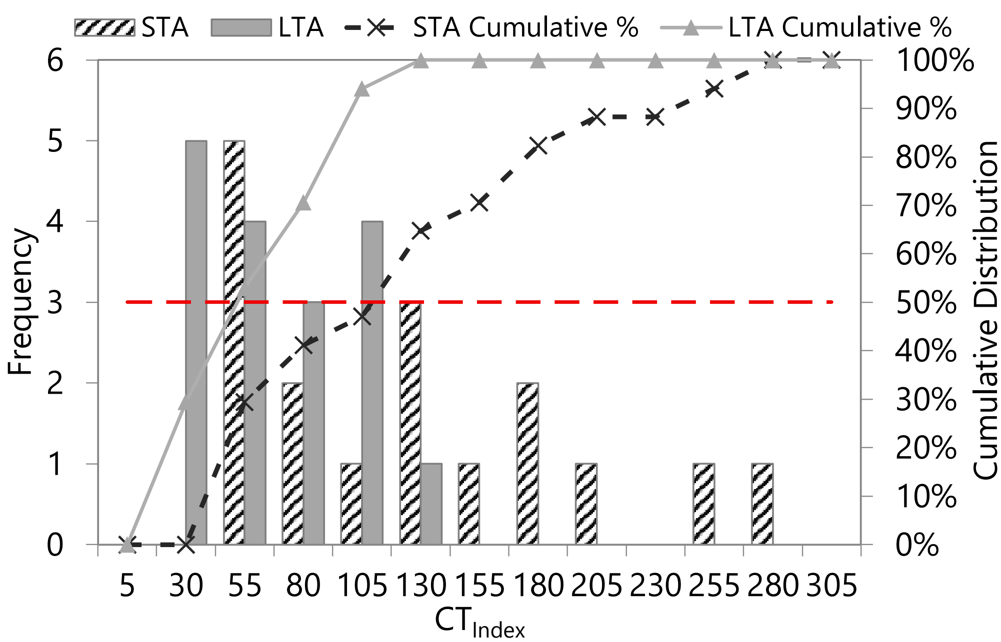


Figure 23. Frequency and Distribution of CT_{Index} at STA and LTA Conditions for S4 Mixes.

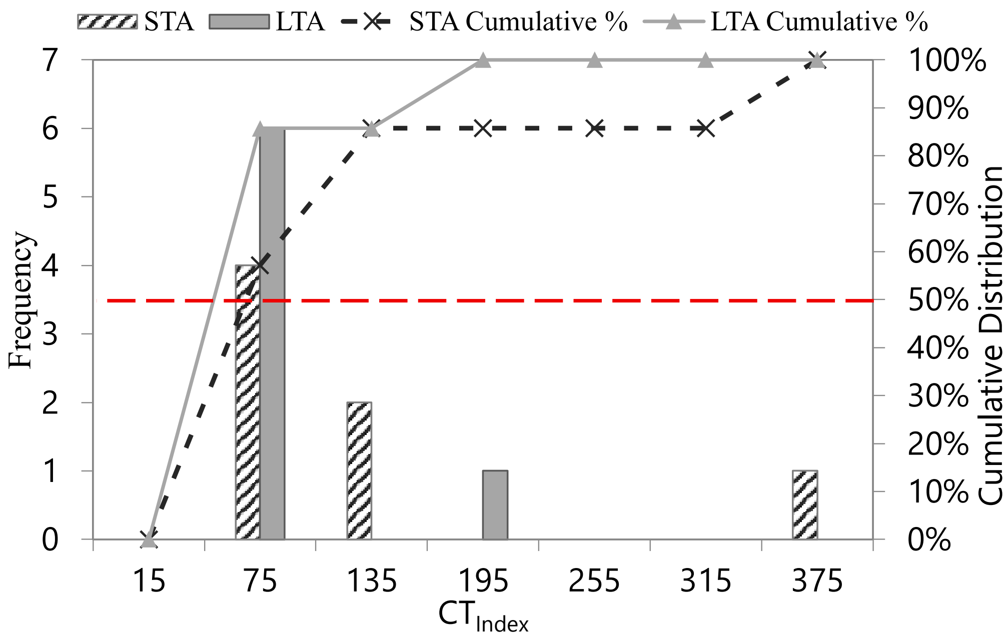


Figure 24. Frequency and Distribution of CT_{Index} at STA and LTA Conditions for S5 Mixes.

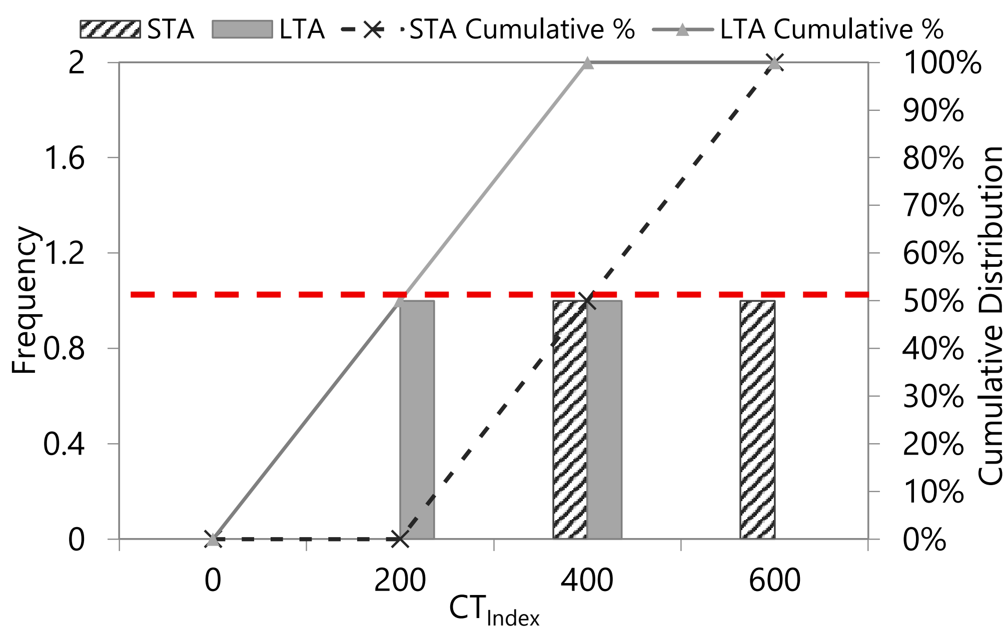


Figure 25. Frequency and Distribution of CT_{Index} at STA and LTA Conditions for SMA Mixes.

3.1.2 Effect of Aging on CT_{Index}

The results of CT_{Index} at both the STA and LTA conditions are presented in Figure 26. As shown in Figure 26, all data points fall below the equality line, indicating that LTA resulted in a reduction in the CT_{Index}. On average, it was noted that the CT_{Index} values of the long-term aged samples were 57 percent of those of the short-term aged samples. Previous research studies have also shown that the CT_{Index} is sensitive to aging (Radeef et al., 2021, 2022).

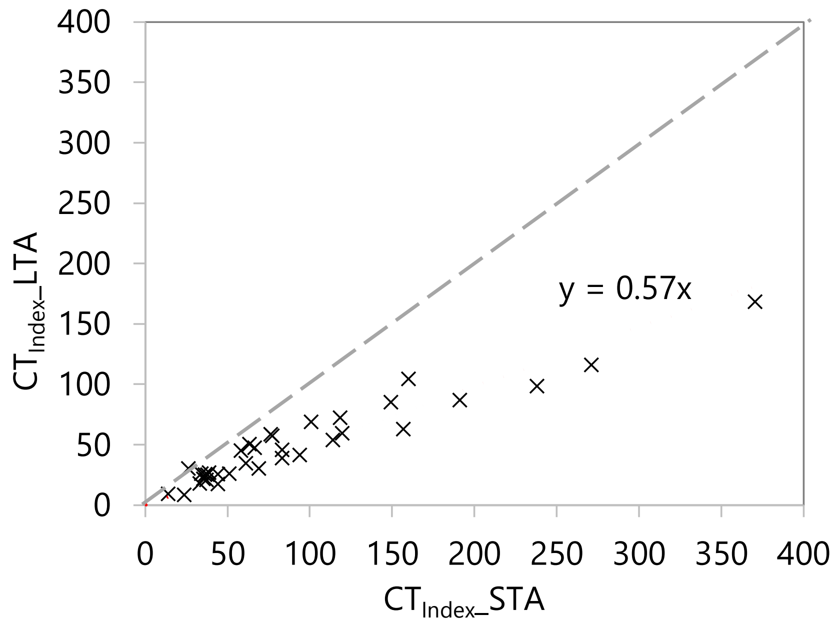


Figure 26. Comparison between CT_{Index} at STA and LTA Conditions.

The ratio between CT_{Index}_LTA and CT_{Index}_STA was calculated and plotted in Figure 27 versus the total AC content and RAP content. The data does not show a specific trend, and there is no clear correlation between the total AC content and the rate of aging for the different mixes containing different RAP content. One of the mixes shows a CT_{Index} ratio that is higher than one. This mix is clearly an outlier.

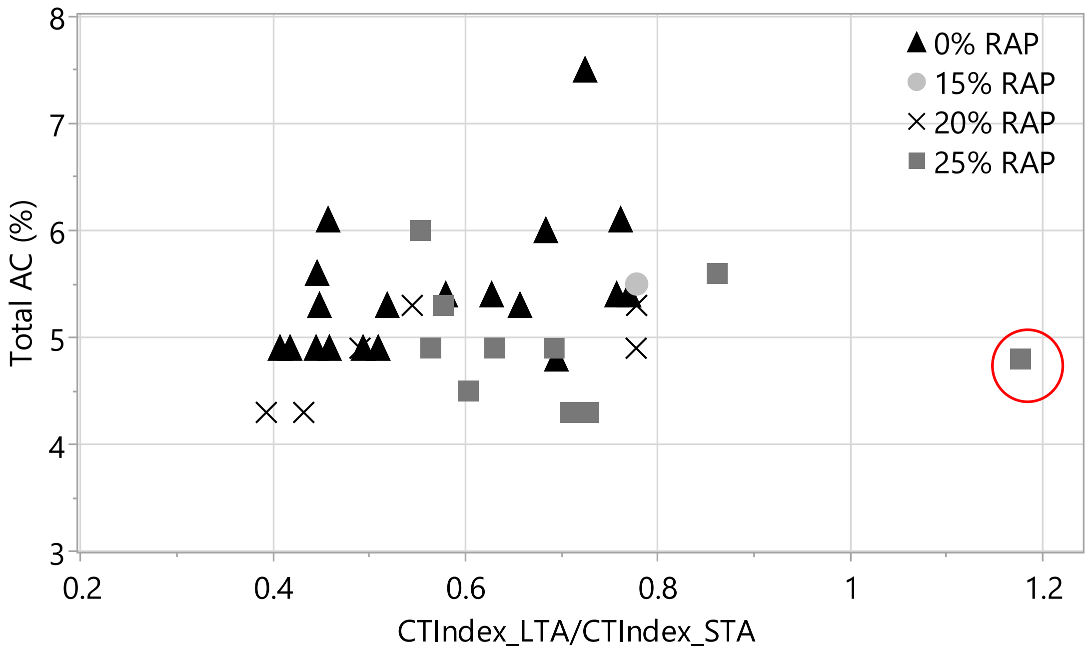


Figure 27. The Ratio of CT_{Index}_LTA to CT_{Index}_STA at Different AC and RAP Content.

The ratio between CT_{Index}_LTA and CT_{Index}_STA was calculated and plotted in Figure 28 versus the RBR. There appears to be no direct correlation between the RBR and the aging rate. One mix shows a CT_{Index} ratio of higher than one and is identified as an outlier.

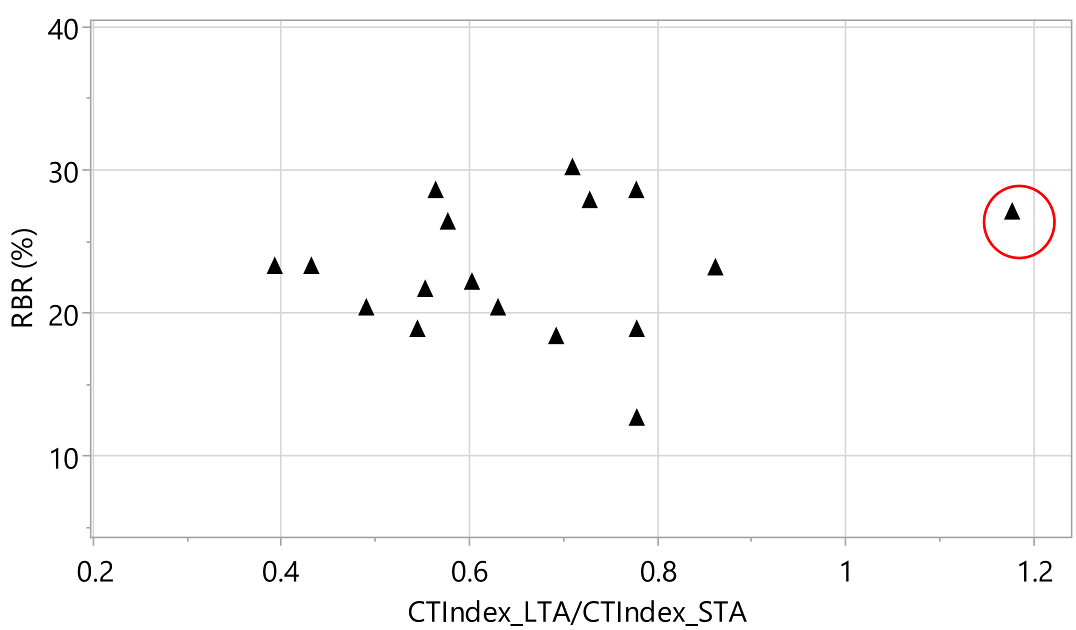


Figure 28. The Ratio of CT_{Index}_LTA to CT_{Index}_STA versus RBR

The ratio between the CT_{Index}_LTA and CT_{Index}_STA was plotted against the binder's PG as shown in Figure 29. The data was presented in box plot format. The mix showing a ratio of higher than one was marked as an outlier. The list of mixes included one mix with a PG58-28 binder as indicated previously. The figure shows a clear trend, with mixes having higher binder grades showing a larger reduction in their CT_{Index} with aging.

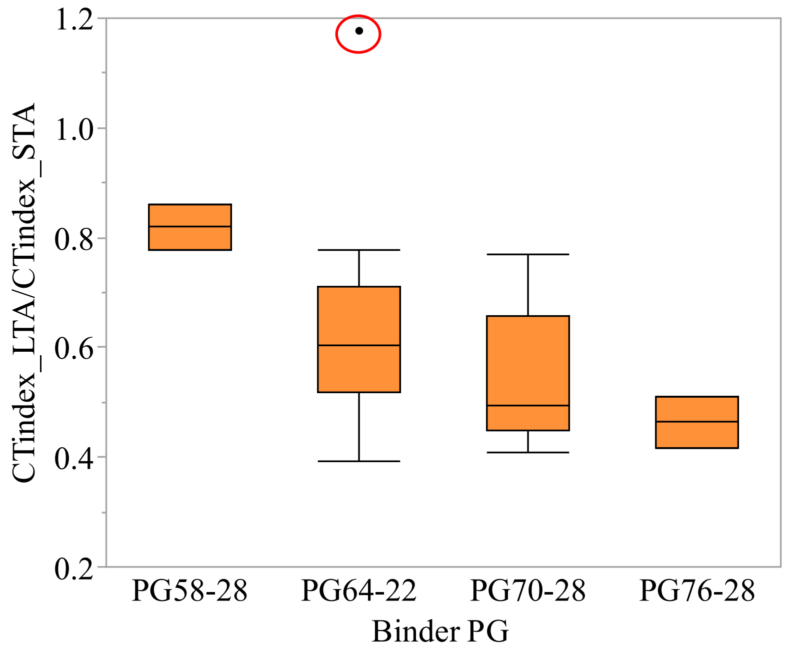


Figure 29. The Ratio of CT_{Index}_LTA to CT_{Index}_STA for Mixes with Different PG.

The relationship between the mix type and the CT_{Index}_LTA to CT_{Index}_STA ratio is shown in Figure 30. It is evident that there is no clear correlation between the mix type and the rate of aging.

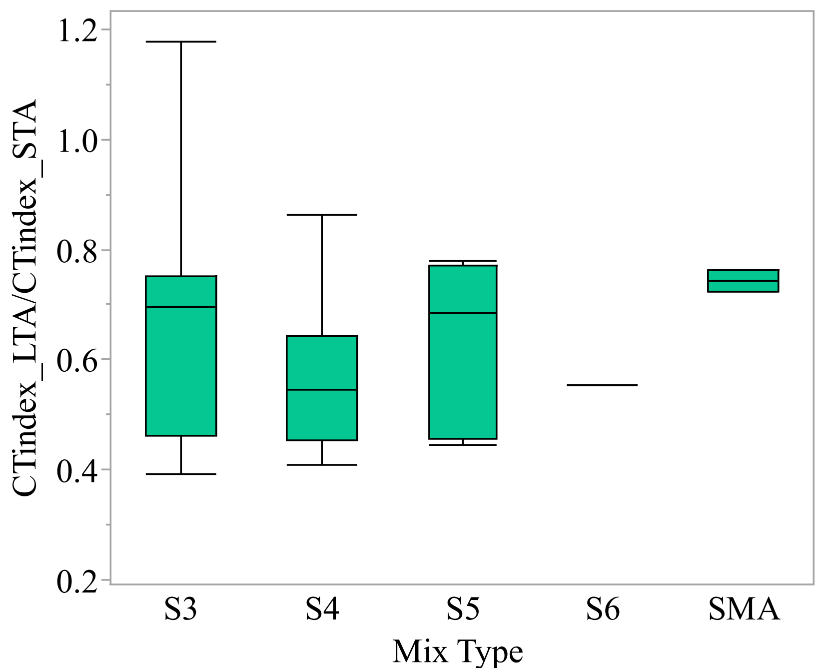


Figure 30. The Ratio of CT_{Index} (LTA) to CT_{Index} (STA) for Mixes with Different Mix Types.

Based on the above results, it can be shown that no clear correlation was noted between the aging rate and any of the mixture variables. A statistical analysis using ANOVA is presented in the next section to provide a more robust comparison between the mixes.

3.1.3 Effect of Aging on other Cracking Parameters

The calculation of the CT_{Index} involves different parameters such as fracture energy, m_{75}, and l_{75}. An increase in fracture energy and l_{75} results in an increase in CT_{Index} whereas an increase in m_{75} results in a decrease in CT_{Index}. The change in fracture energy with LTA for all mixes is shown in Figure 31. The results indicate that on average the fracture energy did not vary significantly with LTA.

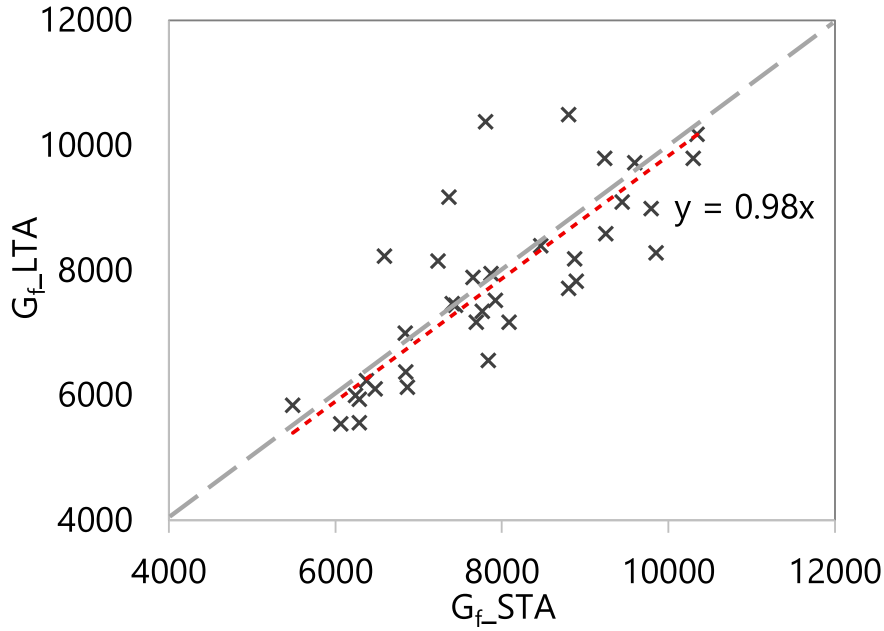


Figure 31. Fracture Energy Results at STA and LTA Conditions.

The I_{75} was plotted for both short-term aged and long-term aged samples as shown in Figure 32. It is shown that LTA conditioning results in an average decrease of I_{75} by about 20 percent.

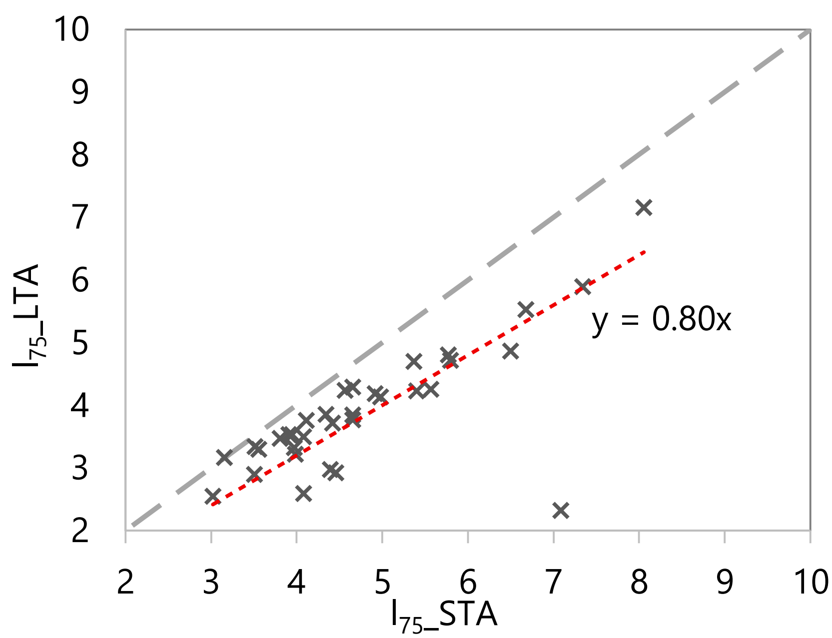


Figure 32. I_{75} Results at STA and LTA Conditions.

Figure 33 shows the relationship between the m_{75} at both STA and LTA conditions. It can be noted that the post-peak slope increased for all mixes because of LTA. Following LTA, the m_{75} increased by an average of 41 percent for all mixes.

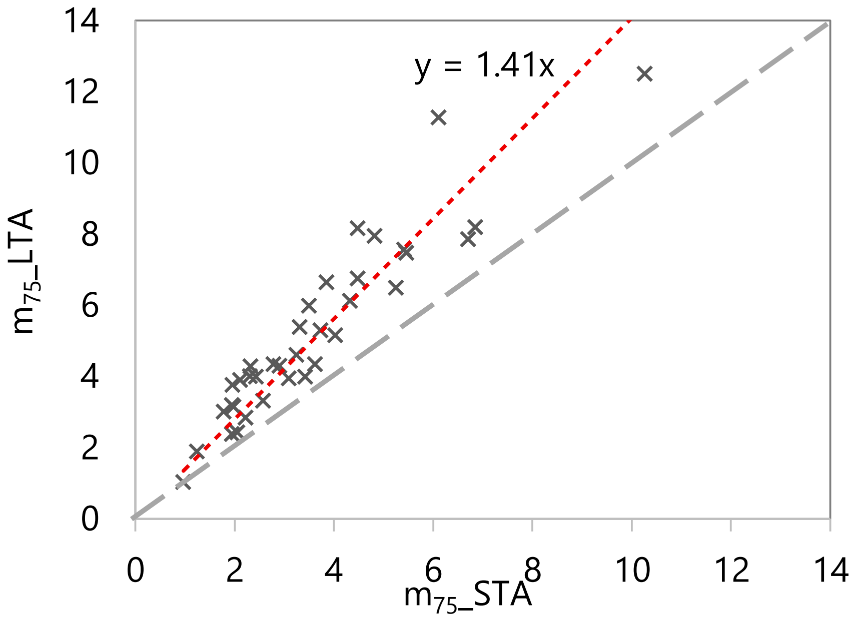


Figure 33. m_{75} Results at STA and LTA Conditions.

The peak load from the deformation curve was plotted for both short-term and long-term aged conditions as shown in Figure 34. The peak load is not used in the calculation of the CT_{Index}, however, it was plotted here to illustrate the effect of LTA conditioning on the mix strength. It is shown that LTA results in an average increase in peak load of about 17 percent. A statistical analysis based on ANOVA will be presented in the following sections to further analyze the effect of aging on the different parameters from the IDEAL-CT test.

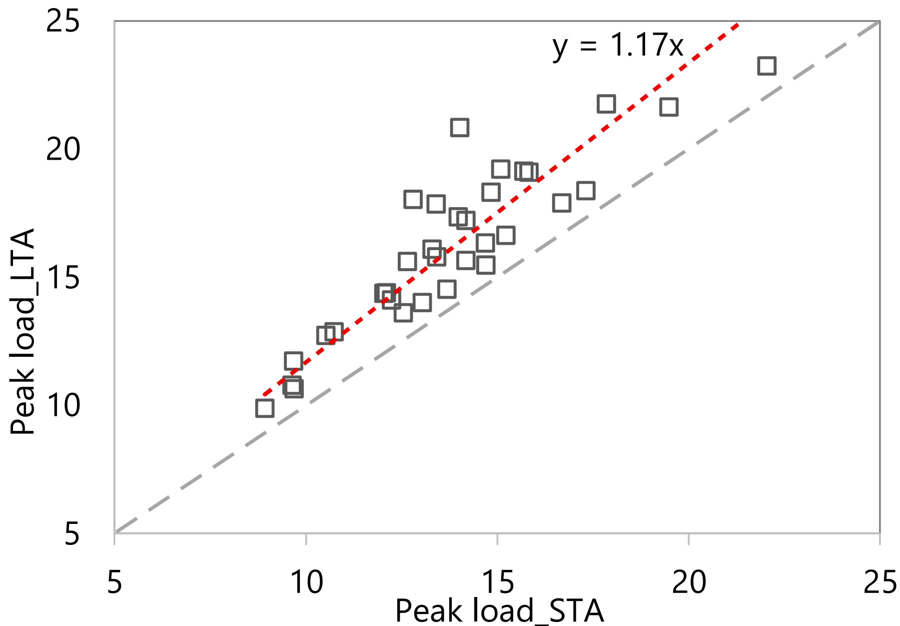


Figure 34. Peak Load Results at STA and LTA Conditions.

Table 7 presents the results of the t-test on the above-mentioned different IDEAL-CT parameters. The results indicated that since the p-values of all parameters are less than alpha (0.05) except for fracture energy, we can reject the null hypothesis of the test. That means we have sufficient evidence to say that the mean of peak load, m_{75}, and l_{75} between STA and LTA is different thus showing the effect of LTA on the above-mentioned parameters.

Table 7. t-Test Results on Different IDEAL-CT Parameters.

Parameters	P-value	Significant?
Fracture Energy	0.847	Not Significant
m_{75}	<0.001	Significant
l_{75}	<0.001	Significant
Peak load	<0.001	Significant

Table 8 provides a summary of all the IDEAL-CT parameters such as CT_{Index}, fracture energy, tensile strength, m_{75}, and l_{75} for all mixtures.

Table 8. IDEAL-CT Summarized Results.

Mix ID	CT_{Index}	CT_{Index}	G_f (J/m^2)	G_f (J/m^2)	Tensile Strength (psi)	Tensile Strength (psi)	l_{75} (mm)	l_{75} (mm)	m_{75} (kN/mm)	m_{75} (kN/mm)
Mix ID	STA	LTA	STA	LTA	STA	LTA	STA	LTA	STA	LTA
M1	51	26	6866	6127	141	148	3.7	5.3	4.0	3.3
M2	14	9	6846	6368	209	204	10.3	12.5	3.0	2.5
M3	35	22	7918	7519	169	187	5.4	7.5	3.6	3.3
M4	39	27	6286.8	5931	133	141	4.1	5.1	3.9	3.5
M5	243	158	10345	10171	129	140	1.9	2. 4	7.1	2.3
M6	83	46	7763	7340	127	145	3.2	4.6	5.6	4.3
M7	38	25	7868	7942	167	190	5.4	7.6	3.8	3.5
M8	37	21	6241	5997	139	155	4.3	6.1	4.0	3.2
M9	36	25	9790	8985	209	221	6.8	8. 2	4.5	2.9
M10	83	39	8800	7707	148	177	3.3	5.4	4.7	3.8
M11	76	59	5488	5836	97	112	2.2	2.8	4.6	4.2
M12	77	57	6287	5558	111	129	2.6	3.3	4.9	4.2
M13	149	85	9236	9793	129	163	2.4	3.9	5.8	4.7
M14	24	8	6065	5542	152	179	6.1	11.3	4.1	2.6
M15	238	98	8804.6	10484	130	169	1.9	3.7	6.5	4.9
M16	34	26	7403	7468	163	175	5.2	6.5	3.5	3.3
M17	271	116	10301	9790	121	148	1.8	3.0	6.7	5.5
M18	114	54	6474	6103	102	117	1.9	3.1	5.4	4.2
M19	33	18	6373	6231	143	170	4.5	6.7	3.5	2.9
M20	120	59	6834	6996	102	128	1.9	3.2	5.0	4.1
M21	44	17	7838	6551	158	175	4.8	7.9	4.4	3.0
M22	157	63	9249	8584	128	155	2.3	4.3	5.8	4.8
M23	58	45	7442	7437	140	147	3.6	4.3	4.3	3.8
M24	61	34	7651	7883	144	171	3.5	5.9	4.1	3.5
M25	26	30	7801	10378	181	204	6.7	7.8	3.2	3.2
M26	44	26	7230	8148	150	189	4.5	8.2	3.9	3.5
M27	160	105	8891	7822	122	124	2.1	2.4	5.4	4.7
M28	66	47	8088	7163	146	143	3.4	3.9	4.1	3.8
M29	370	168	8875	8183	94	111	1.2	1.9	7.3	5.9
M30	531	377	9854	8273	92	86	0.9	1.0	8.1	7.2
M31	69	30	8468	8392	149	190	3.8	6.6	4.4	3.7
M32	101	69	9442	9088	145	157	3.1	3.9	4.7	4.3
M33	94	42	7688	7168	126	152	2.8	4.3	4.7	3.8
M34	118	72	7364	9167	112	161	2.3	4.0	5.4	4.6
M35	63	50	6588	8225	120	158	2.9	4.3	4.1	3.8
M36	191	87	9595	9723	123	162	2.1	3.9	6.3	5.2

The following discussion provides specific examples to illustrate the effect of LTA on the different parameters used in the calculation of the CT_{Index}. Mixes M31, M7, M20 were used for this discussion.

Figure 35 provides a plot of the STA and LTA load-displacement curves for mix 31. M31 is an S5 mix with a PG64-22 binder and no RAP. According to Table 8, there was no significant difference in fracture energy for this mix, which represents the area under the curve. The STA specimen showed more ductility as given by the l75 parameter. m75 decreased significantly following LTA.

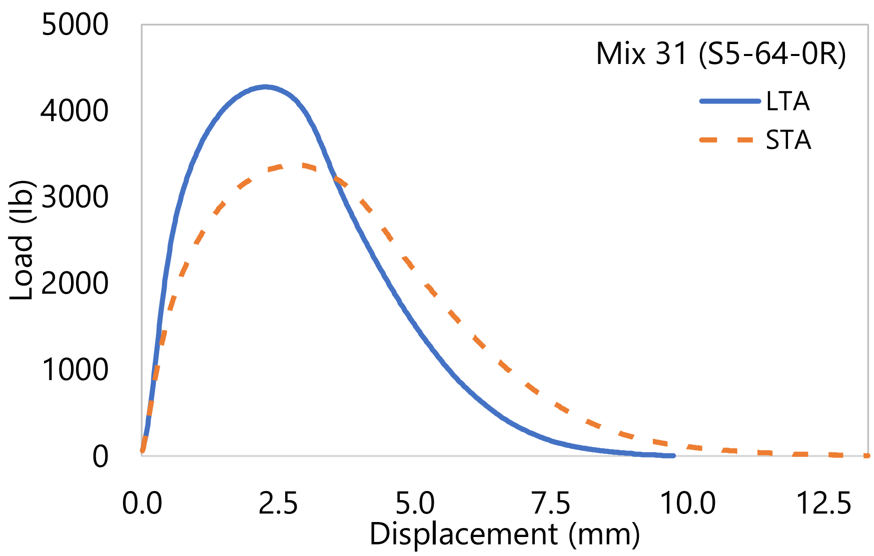


Figure 35. Load-Displacement Curves of Mix M31 at STA and LTA Conditions.

Figure 36 shows a plot of the load-deformation curves for Mix 7. M7 is an S5 mix with a PG64-22 binder and no RAP, using WMA technology. It is shown that there was a slight decrease in fracture energy. The m_{75} increased significantly with LTA, indicating a steeper slope and a more abrupt failure. The l_{75} parameter decreased with LTA.

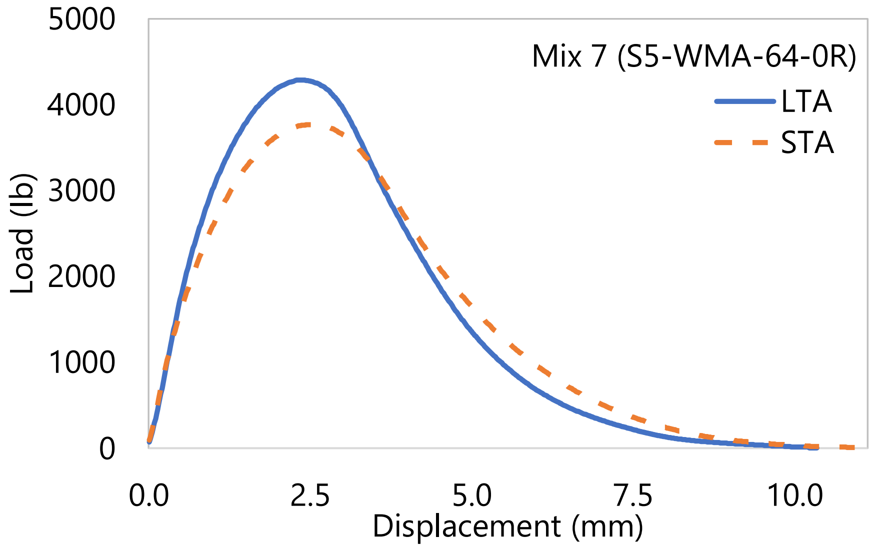


Figure 36. Load-Displacement Curves of Mix M7 at STA and LTA Conditions.

Figure 37 shows a plot of the load-deformation curves for Mix 20. M20 is an S4 mix with a PG64-22 binder and no RAP. It is shown that there was no significant difference in fracture energy. The m_{75} increased significantly with LTA, indicating a steeper slope and a more abrupt failure. The l_{75} parameter decreased with LTA.

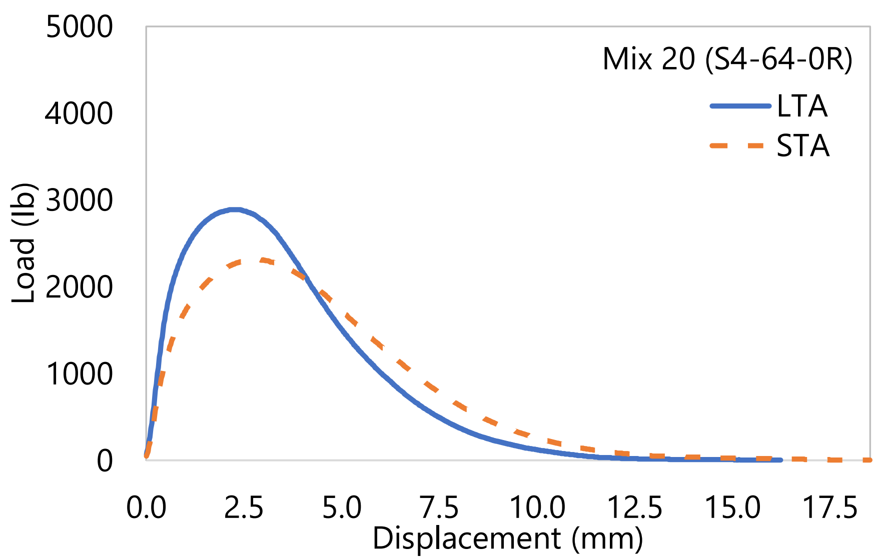


Figure 37. Load-Displacement Curves of Mix M20 at STA and LTA Conditions.

3.1.4 Statistical Analysis Using ANOVA

The first statistical approach selected for this study was the Analysis of Variance (ANOVA). Before conducting the ANOVA, it is important to test for its inherent assumptions to ensure none of the underlying assumptions is violated during the analysis process. The main assumptions for ANOVA are Independence and Random Sampling, Normality, and Homogeneity of Variance (Homoscedasticity) (Fisher, 1925; Statistics Solutions, 2013; Zack, 2019). Due to the nature of the laboratory testing matrix design, the assumptions of independence and random sampling are automatically satisfied for this dataset. The next assumption to check was normality. This was done by performing the Shapiro-Wilk test on the CT_{Index} data (both STA and LTA data combined) (Shapiro & Wilk, 1965; Zack, 2019). The null hypothesis for the Shapiro-Wilk Test is that the data is normally distributed. The p-value for the given dataset was < 0.001, as shown in Figure 38 for the combined CT_{Index} values at the STA and LTA conditions. A low p-value for the Shaprio-Wilk test indicates that the data is not normally distributed. This is visually evident by looking at the CT_{Index} histogram in Figure 38 which does not appear as normally distributed. The next step was to repeat this step for STA and LTA CT_{Index} data separately to see if the normality assumption was better met when the data was divided based on aging condition, but the individual subsets were still not normally distributed from the Shapiro-Wilk test as shown in Figure 39 and Figure 40 for the STA and LTA results, respectively.

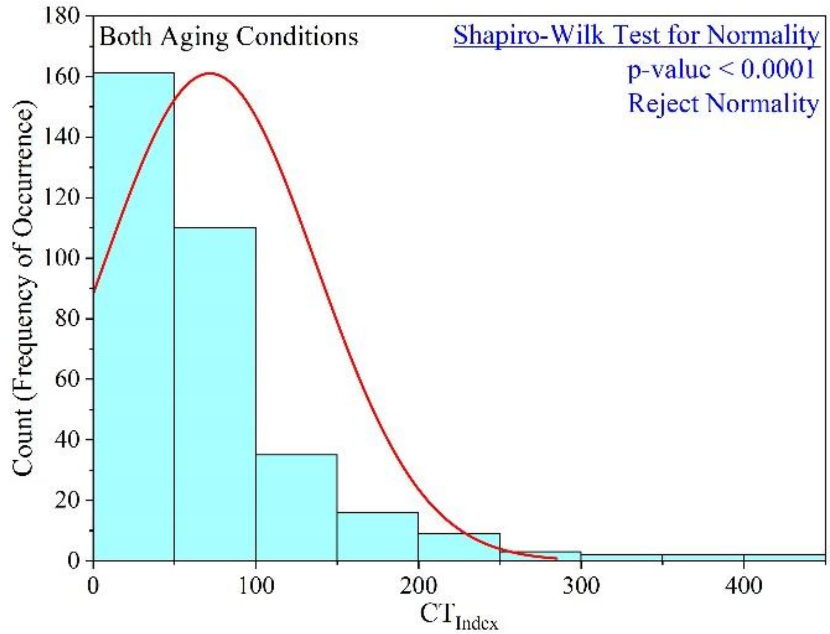


Figure 38. Plots showing histograms and normality test Results for combined STA and LTA $CT_{Index.}$

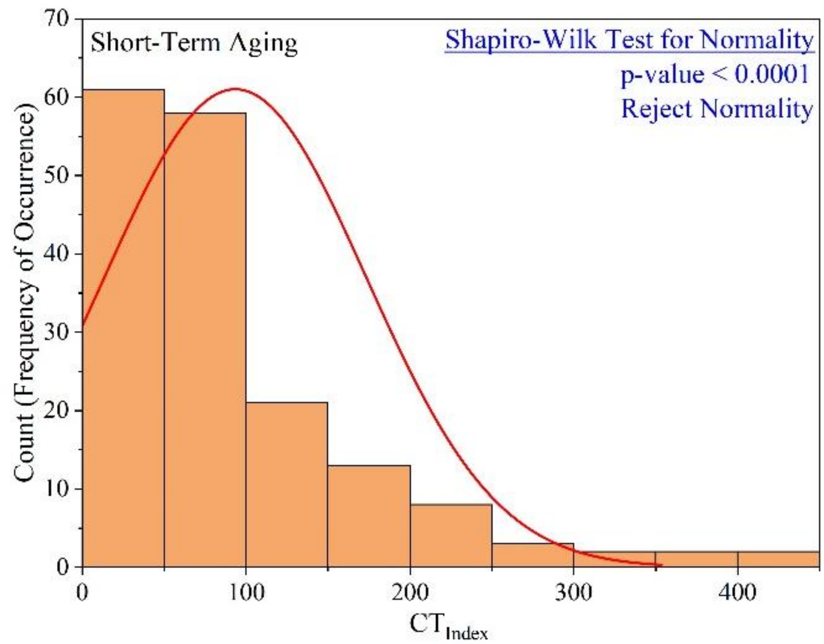


Figure 39. Plot showing histograms and normality test results for STA $CT_{Index.}$

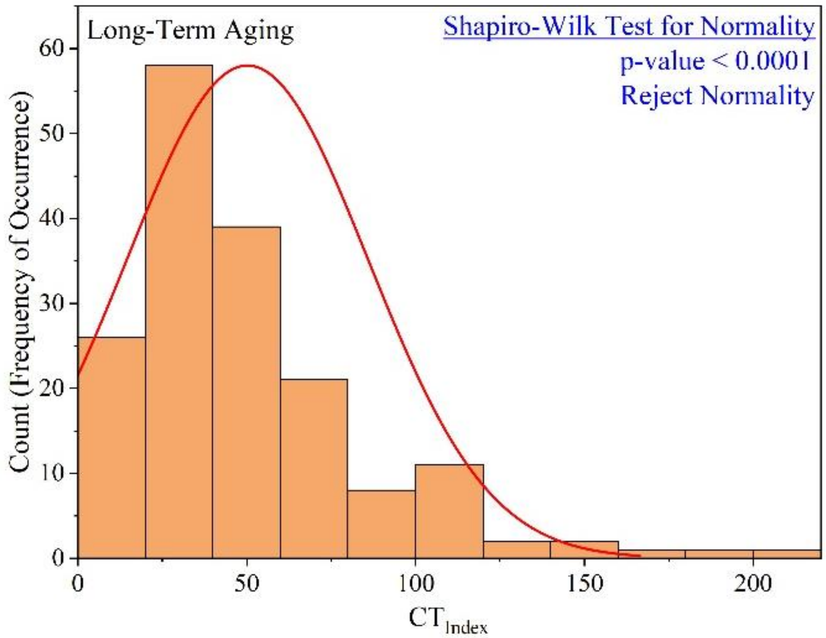


Figure 40. Plot showing Histograms and Normality Test Results for LTA CT_{Index}.

Once the non-normality of the CT_{Index} data was established, the natural next step was to consider transformation approaches that may result in a normal distribution. Several different transformation options are available to consider, such as reciprocal transformation, square-root transformation, log (natural) transformation, etc (Heidel, 2023; Onofri, 2019; Sharanappa, 2022). Among the ones considered, the log (natural) transformation was most effective in terms of addressing the non-normality of the distribution. Figure 41 shows the distribution of $\ln(CT_{Index})$ data, and the result of the Shapiro-Wilk test which had a p-value of 0.6107; therefore, normality cannot be rejected.

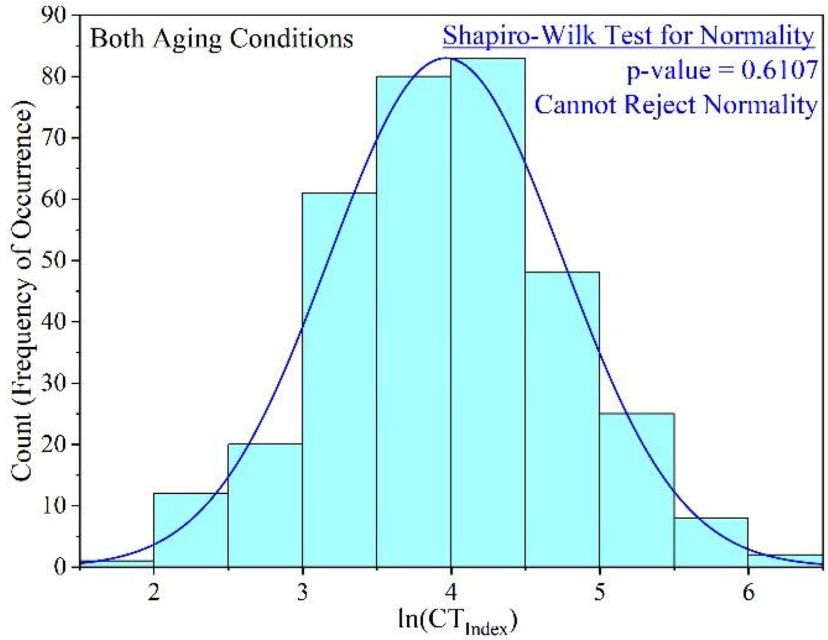


Figure 41. Plots showing Histograms and Normality Test Results for combined STA and LTA; ln(CT_{Index}).

Once the normality assumption was addressed, the next step was to check for equality of variance using Levene's test (Carroll & Schneider, 1985; Levene, 1960; Statistical Methods, 2023). One of the factors that was used to perform the Levene tests is the NMAS, which had four levels. The result of the test with NMAS is presented in Figure 42. From the figure, the null hypothesis associated with Levene's test (equality of variance) is rejected at a 5% confidence level (p-value = 0.0001), which means the dataset violates the condition of equality of variance. The inequality of the variances can be visualized by the difference in the spread of the boxes in Figure 42 (Statistical Methods, 2023) . It is, therefore, necessary to select a test that is "robust" against the violation of the equality of variance assumption. One such test is Welch's ANOVA Test (Welch, 1938). It is important to note that although Welch's test is robust against unequal group variances, it still requires the data to be normally distributed (Frost, 2018; Stephanie, 2023). Therefore, for the purpose of this study, Welch's tests were conducted on the transformed CT_{Index} data [ln(CT_{Index})]. It is also important to note that Welch's ANOVA test is a One-Way parametric test. So, to determine how the factors affect the CT_{Index} under different aging conditions, we split the CT_{Index} data based on their aging condition (STA or LTA), and therefore the subsequent Welch's ANOVA tests were performed separately on the two subsets.

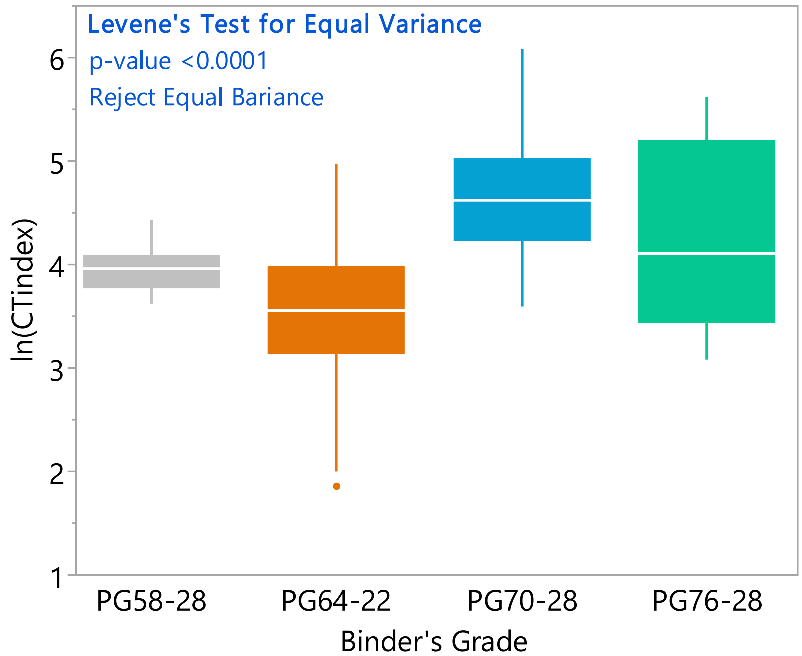


Figure 42. Box Plot showing the Levene Test of ln(CT_{Index}) Values Corresponding to Different NMAS.

The factors (and levels) considered using Welch's ANOVA were: NMAS (4.75, 9.5, 12.5, 19), course type (surface course and intermediate course), mix type (HMA and WMA), RAP content (0%, 15%, 20%, and 25%), binder grade (PG58-28, PG64-22, PG70-28, and PG76-28), and binder type (unmodified and polymer-modified). The significance levels (p-values) obtained from these analyses are listed in Table 9. From the result of Welch's ANOVA tests, considering the log-transformed CT_{Index} values after both STA and LTA conditions, all factors except mix type were found to be significant. In other words, there was no significant difference between the mean ln(CT_{Index}) values for HMA or WMA after STA and LTA. As a reminder, if the p-value for a parameter is less than 0.05, taking a 95% confidence level into account, it indicates that it has a significant effect on the results acquired (in this case the mean ln(CT_{Index}) value), otherwise, it is considered insignificant.

Table 9. Significance Levels of Different Factors on CT_{Index} values.

Parameter	ln (CT_{Index})	ln (CT_{Index})	ln (CT_{Index})	ln (CT_{Index})	ln (CT_{Index})	ln (CT_{Index})
Condition	STA	STA	STA	LTA	LTA	LTA
Test	Welch's ANOVA	Games-Howell	Games-Howell	Welch's ANOVA	Games-Howell	Games-Howell
Factor	p-value	Pairs	p-value	p-value	Pairs	p-value
Mix Type	0.669	HMA vs WMA	0.669	0.279	HMA vs WMA	0.279
NMAS	<0.001	4.75 vs 9.5 4.75 vs 12.5 4.75 vs 19 9.5 vs 12.5 9.5 vs 19 12.5 vs 19	0.834 0.815 <0.001 0.513 0.002 <0.001	<0.001	4.75 vs 9.5 4.75 vs 12.5 4.75 vs 19 9.5 vs 12.5 9.5 vs 19 12.5 vs 19	1.000 0.838 <0.001 0.928 <0.001 <0.001
Course Type	<0.001	Surface Course vs Intermediate Course	<0.001	<0.001	Surface Course vs Intermediate Course	<0.001
Binder Grade	<0.001	58-28 vs 64-22 58-28 vs 70-28 58-28 vs 76-28 64-22 vs 70-28 64-22 vs 76-28 70-28 vs 76-28	0.015 <0.001 0.173 <0.001 0.031 0.772	<0.001	58-28 vs 64-22 58-28 vs 70-28 58-28 vs 76-28 64-22 vs 70-28 64-22 vs 76-28 70-28 vs 76-28	<0.001 <0.001 0.995 <0.001 0.082 0.335
Binder Type	<0.001	Unmodified vs Polymer-modified	<0.001	<0.001	Unmodified vs Polymer-modified	<0.001
RAP Content	<0.001	0 vs 15 0 vs 20 0 vs 25 15 vs 20 15 vs 25 20 vs 25	<0.001 <0.001 <0.001 0.020 <0.001 0.897	<0.001	0 vs 15 0 vs 20 0 vs 25 15 vs 20 15 vs 25 20 vs 25	0.020 <0.001 <0.001 <0.001 <0.001 0.664

After Welch's ANOVA test, the next step was to carry out a pairwise comparison post hoc test to identify significant differences between specific groups. The Games-Howell test is a suitable multiple comparison method for our type of data as it can deal with unequal variances across groups (Games & Howell, 1976). The result of the Games-Howell tests is also presented in Table 9. Considering the NMAS factor, samples with 19 mm NMAS appeared to have significantly different mean ln(CT_{Index}) from samples with lower NMAS after STA and LTA. Samples in other NMAS groups do not have significantly different mean ln(CT_{Index}) values from one another after either aging condition. For the binder grade factor, PG58-28 vs PG76-28, and PG70-28 vs PG76-28 were found to not have significantly different mean ln(CT_{Index}) values after short-term aging. These two pairs as well as the 64-22 vs 76-28 pair also have non-significant differences in their mean log-transformed values after LTA. Samples with high RAP content (20% and 25%) appear to not have a significant difference in their mean ln(CT_{Index}) values after either STA or LTA conditioning.

An ANOVA was also conducted to investigate the effect of different mix factors on the ratio of the CT_{Index}_LTA to CT_{Index}_STA (log-transformed). However, the Shapiro-Wilk test revealed that the data was not normally distributed (p-value = 0.0057). To address this, a natural log transformation was applied, resulting in a normally distributed dataset (p-value = 0.2341). Levene's test was then used to check for homogeneity of variances. The null hypothesis of equal variances was not rejected for mix type, NMAS, binder PG, or RAP (p-values = 0.239, 0.466, 0.515, and 0.497, respectively). The aforementioned indicated that the assumptions of equal variance and normality are both true. Thus, a one-way ANOVA and Tukey's test were used to analyze the ratio of CT_{Index}_LTA to CT_{Index}_STA ratio. The results are summarized in Table 10. It can be inferred from the results that that RAP content (0% vs 25%) and binder grade (58-28 vs 76-28), as well as binder type, were significant factors influencing the CT_{Index} ratio.

Table 10. Significance Levels of Different Factors on the CT_{Index}_LTA/CT_{Index}_STA Ratio.

Test	One-Way ANOVA	One-Way ANOVA	Tukey's HSD post-hoc	Tukey's HSD post-hoc
Factors	**p-value**	**Significance?**	**Pairs**	**p-value**
Mix Type	0.387	Not Significant	HMA vs WMA	0.387

NMAS	0.526	Not Significant	4.75 vs 9.5 4.75 vs 12.5 4.75 vs 19 9.5 vs 12.5 9.5 vs 19 12.5 vs 19	0.973 1.000 0.960 0.708 0999 0.555
Course Type	0.205	Not Significant	Surface Course vs Intermediate Course	0.205
Binder Grade	0.042	Significant	58-28 vs 64-22 58-28 vs 70-28 58-28 vs 76-28 64-22 vs 70-28 64-22 vs 76-28 70-28 vs 76-28	0.365 0.129 0.049 0.540 0.208 0.585
Binder Type	0.041	Significant	Unmodified vs Polymer-modified	0.041
RAP Content	0.042	Significant	0R vs 15R 0R vs 20 0R vs 25R 15R vs 20 15R vs 25 20R vs 25R	0.429 0.999 0.050 0.535 0.964 0.261

3.1.5 Predictive Modeling

The next step in this study was to apply ML-based models that can predict the CT_{Index} values of asphalt mixtures, given mix volumetrics-related parameters. This section presents details about data pre-processing and predictive model implementation.

Data preprocessing is a critical first step in the science of predictive modeling. Different models have different sensitivities to the type of predictors in them, and their performance can be improved by transforming the data to reduce the impact of skewness or outliers. Data pre-processing techniques may consist of the addition (incorporation of dummy variables), deletion (near-zero variance estimators and multi-collinearity), or transformation (Box-Cox, Yeo-Johnson, and Spatial Sign) of a set of data. Data transformations are used to improve the numerical stability of some calculations during model development. As most ML models work with a normal distribution, a check for skewness was required to be carried out on the predictors. This is essential because the skewness of distributions can lead to the drawing of wrong inferences from the model (Das & Imon, 2016). A distribution is said to be un-skewed when it is roughly symmetric, and the probability of falling on either side of the distribution's mean is nearly equal (Kuhn & Johnson, 2013). The general rule of thumb is that distribution with a skewness statistic between -0.5 to +0.5 is nearly normal, and the closer to zero the skewness statistic is, the better it is (Gawali, 2021; Menon, 2023). Any predictor that has a skewness value not within this range requires transformation to improve its skewness (Kuhn & Johnson, 2013). Common transformation approaches include replacing the data with its square, cube, inverse, etc. The choice of method of transformation to be used depends on the extent of skewness of the distribution.

For this study, the Box-Cox transformation was considered for transforming all continuous predictors to get their skewness statistics as close to zero as possible (Kuhn & Johnson, 2013). The Box-Cox transformation was developed for the transformation of outcome data, but it has also been found to be effective in transforming continuous predictors (Box & Cox, 1964). It estimates a distinct parameter, lambda (λ), which is used to index each value in accordance with Equation 7 below, where x and x* are the values before and after transformation, respectively.

$$x^{*} = \begin{cases} \frac{x^{\lambda}-1}{\lambda}, & if\ \lambda \neq 0 \\ \log(x), & if\ \lambda = 0 \end{cases} \tag{7}$$

Table 11 shows the lambda values and the skewness statistics of the predictors before and after transformations. The Box-Cox transformation adequately reduced the skewness statistics of most of the predictors. The skewness statistic for DBR (Dust to Binder Ratio) was nearly zero, and therefore, did not change after the Box-Cox transformation. This is evident by the lambda value being 1.1 (close to 1) for DBR. From Equation 6, a lambda value equal to 1 will result in the subtraction of 1 from all values of the predictor, and this will not cause a change in its distribution/skewness. Two predictors were not transformed because they contained non-positive values: RAP (which has some values that are zero), and LT-PG (Low-Temperature Performance Grade; which has all its values as negative). Box-Cox transformation is only applicable to positive values (Bhalla, 2017). For the sake of thoroughness, other forms of transformations (Yeo-Johnson and Spatial Sign) were also considered in this study as shown in Table 11 ((Rousseeuw, 1985; Yeo & Johnson, 2000).

After the application of these two transformation methods, one predictor still had skewness statistics outside the range of -0.5 to 0.5, for each Yeo-Johnson and Spatial Sign as opposed to none for Box-Cox transformation (Gawali, 2021; Menon, 2023). Therefore, the Box-Cox transformation was most suitable for transforming the predictors and was adopted for the predictive modeling study.

It should be mentioned that the "n" parameter used in the predictive modeling is known as the shape factor. The shape factor is very important in asphalt mix design as it refers to the form or angularity of the particles used in the mix. This parameter gives information about the particle shape and surface texture of the aggregate. Now, a lower shape factor generally indicates a more angular or irregular-shaped particle, while a higher shape factor suggests a more rounded particle.

Table 11. Skewness Statistics for Different Predictor Variables before and after Transformations.

	NMAS	n	RAP	G_{mm}	VMA	VFA	DBR	HT-PG	LT-PG	AC Total
Skewness before transformation	0.31	0.25	0.09	0.31	0.17	0.43	0.03	0.44	-0.24	0.15
Skewness after Box-Cox	~0.00	0.01	0.09	0.16	~0.00	0.05	0.03	0.02	-0.24	~0.00
Lambda for Box-Cox	0.62	0.16	_	-2.00	-0.6	-2.00	0.94	-1.36	_	0.31
Skewness after Yeo-Johnson	~0.00	~0.00	0.01	0.61	0.01	-0.02	-0.01	~0.00	-0.24	~0.00
Skewness after Spatial Sign	0.74	0.32	0.13	0.05	0.01	-0.07	0.33	0.31	-0.26	0.01

note: gradation parameter derived by fitting the aggregate gradation to 0.45 power curve; VMA: Voids in Mineral Aggregates; VFA: Voids Filled with Asphalt; DBR: Dust to Binder Ratio; AC Total: Total Binder Content in Mix

After the transformation of the predictors, a between-predictor correlation check was carried out, to determine predictors that are substantially correlated and pose a threat of redundancy and instability to the models (Kuhn & Johnson, 2013). Figure 43 shows the correlation matrix for the different predictors used in this modeling effort. The cutoff value for the correlation coefficient was set to 0.9, in accordance with common practice in the literature. As a result, NMAS and VMA was removed for having a correlation value higher than 0.9 (Kuhn, 2015; Kuhn & Johnson, 2013). Centering and scaling of the predictors were also carried out as required by several of the models considered in this study. Centering was carried out by subtracting the average value from every value of a predictor; this would make the predictor have a mean of zero. Scaling, on the other hand, was performed by dividing each value by its standard deviation, and this resulted in the standard deviation of all values of the predictor becoming 1. It should be noted that although centering and scaling improves the numerical stability of the calculations made in ML, they cause the loss of interpretability of individual values.

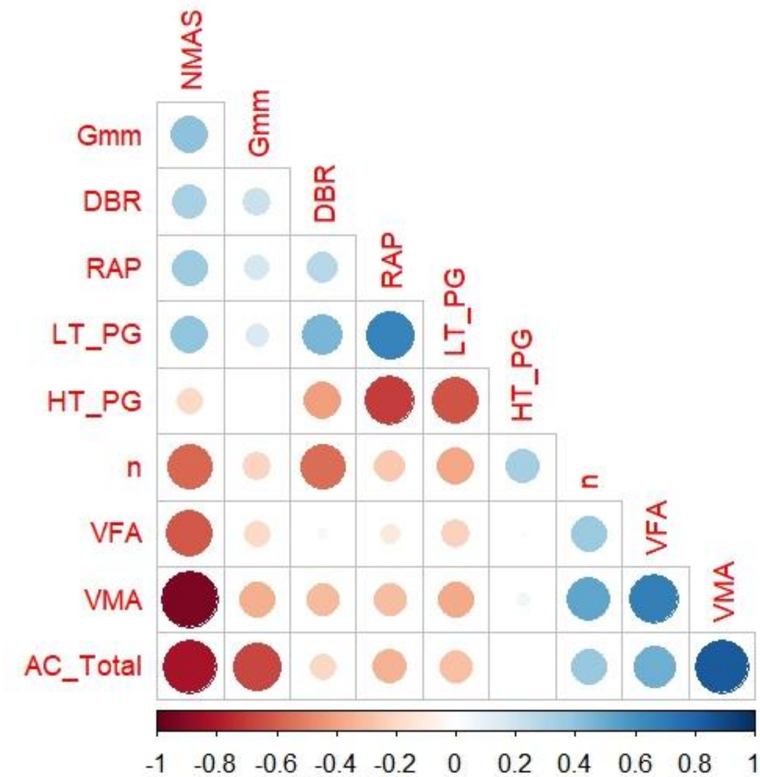


Figure 43. Correlation Matrix showing Between-Predictor Correlation Coefficients.

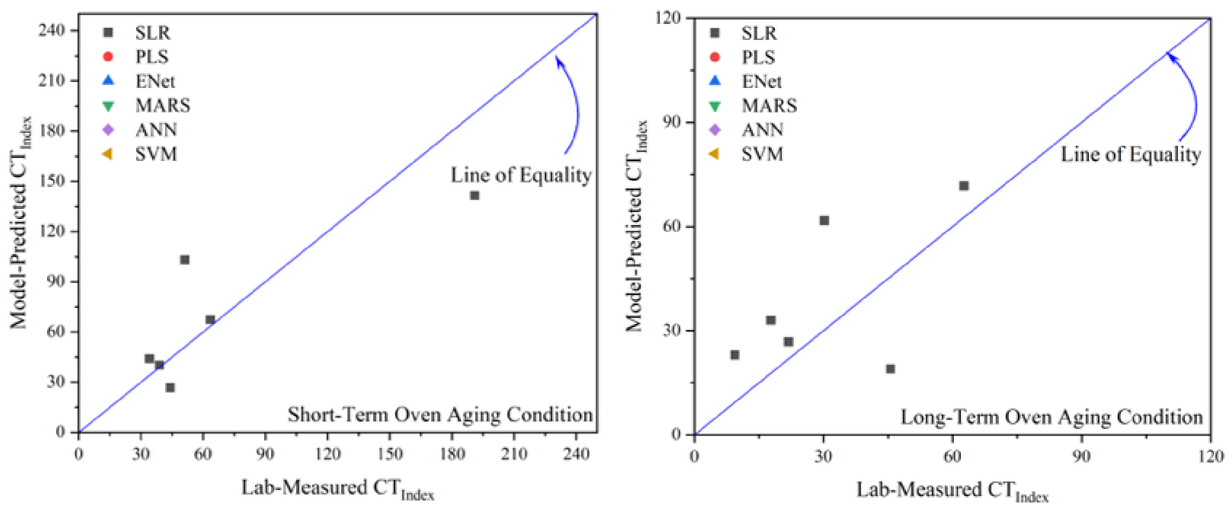
An exploratory analysis was also carried out on the response variable to see the degree of skewness of its distribution. The CT_{Index} had a skewness of 2.21, and using natural logarithmic transformation proved productive because it changed the skewness of the distribution to 0.09. This step is important because it provides an idea of what is to be done during the modeling process – the natural log of CT_{Index} should be modeled instead of just the CT_{Index}. While a Box-Cox transformation could have been used on the response variable, the use of the natural log to transform the CT_{Index} values allows for easier interpretation of the predictions of the model.

Two classes of models were considered: parametric and nonparametric. Parametric methods assume the functional form of the model (Brownlee, 2016). Once we know the model form, we can estimate the coefficients and standard error of each coefficient which allows us to make inferences on the model coefficients. These parametric models have low flexibility but are highly interpretable. Non-parametric methods, on the other hand, do not make explicit assumptions about the functional form of the model (Brownlee, 2016; Gurudath, 2020). Hence, they are highly flexible, which allows for high predictability but with a loss of interpretability. The more flexible a model is, the less interpretable it is (James et al., 2021). Putting both classes of models into consideration allows us to evaluate the tradeoff between prediction accuracy and model interpretability. The models selected for this study range from low flexibility to high flexibility, and low interpretability to high interpretability. Models that require tuning were trained using k-fold cross-validation to calculate the optimal hyperparameters necessary to train the model.

As the first step of the modeling stage, the available data was split into a proportion of 80:20 for training and testing, and k-fold cross-validation was set to be 5 folds. K-fold validation is an important component of ML and is done by splitting the data into roughly equal sets. Samples in k-1 sets are used for the training of the model, and samples of the remaining set are used for validation (Kuhn & Johnson, 2013). The training and validation will be done k number of times, and this helps the model improve its performance (Kim, 2009; Molinaro, 2005). The “set.seed” function was applied inside R to ensure we are using the same training/testing datasets across the different models being compared and to aid in reproducibility and repeatability when additional data is incorporated (Kuhn & Johnson, 2013).

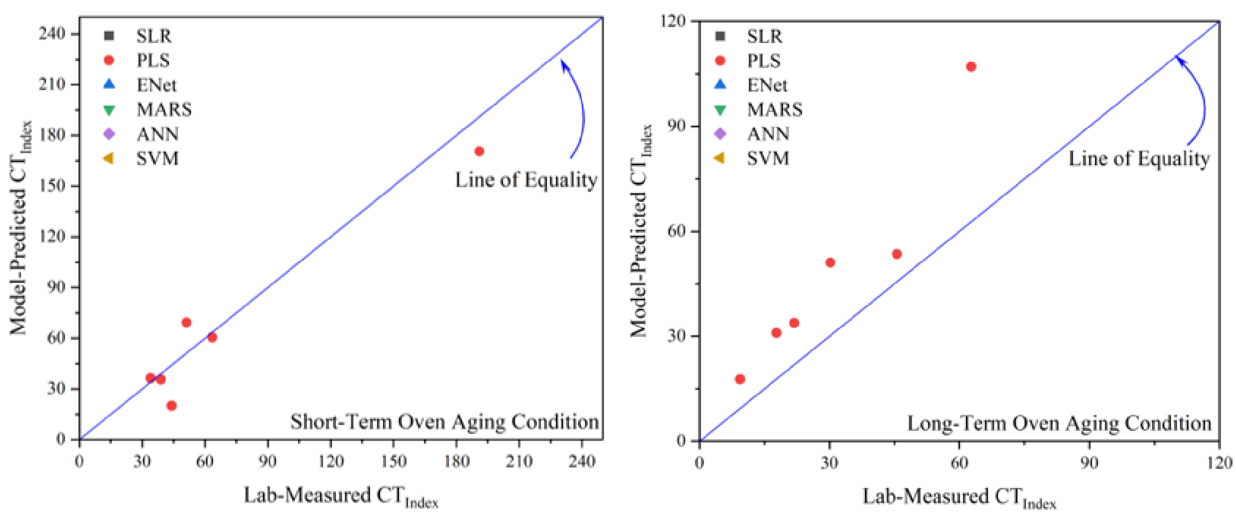
The linear models considered in this study were: Simple Linear Regression (SLR), Partial Least Squares (PLS), and the Penalized regression elastic net (ENet). SLR is one of the simplest and most used models in ML, and it functions by establishing linear relationships between dependent and independent variables (Maulud & Abdulazeez, 2020). PLS is a statistical method that is covariance-based and was designed to deal with multiple regression when the data has small samples or multicollinearity (Pirouz, 2006). Vinzi and Russolillo (2013) defined PLS as a set of iterative algorithms that are based on least squares which are used to implement a broad spectrum of both explanatory and exploratory multivariate techniques (Esposito Vinzi & Russolillo, 2013). An advantage of PLS is that it allows the reduction of dimensionality of correlated variables and models the underlying, shared information of the variables (Korstanje, 2021). The non-linear models considered were: multivariate adaptive regression splines (MARS), artificial neural network (ANN), and support vector machine (SVM). An overview of the different predictive models used in this study is provided in the Appendix.

Figure 44, Figure 45, Figure 46, Figure 47, Figure 48, and Figure 49 present the relationship between the lab-measured CT_{Index} values and predicted CT_{Index} values for all the models used in this study. The closeness of the data points to the equality line depicts how accurate the prediction is for that instance. For the LTA condition, a lot of the predicted values fall above the equality line which indicates that the CT_{Index} of the mixtures are predicted to be more than what they really are. For both conditions, the predictions by ANN seem to be the most accurate, because most of ANN's data points are very close to the equality line. The coefficient of determination (R^2), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values of the different models on the testing set are presented in Figure 50. R^2 gives a deep insight into how well the model predicts the variance of the response variable using the predicted variables, while RMSE tells us how far away the predicted values are from the observed values, hence it is an indication of whether the model is error-prone or not. MAE measures the average of the residuals (Chugh, 2020; Safjan, 2023; Zack, 2021). Lower values of RMSE and MAE indicate higher accuracy of the model, while a higher value of R^2 is better. From Figure 50, the ANN and PLS model have the highest R^2 values (0.96 and 0.85, respectively) and lowest RMSE and MAE values relative to other models considered. This implies that it is most capable of predicting the variance in the response variable, and it is also relatively least prone to having errors in its predictions. ENet showed the worst performance in predicting the response variable.



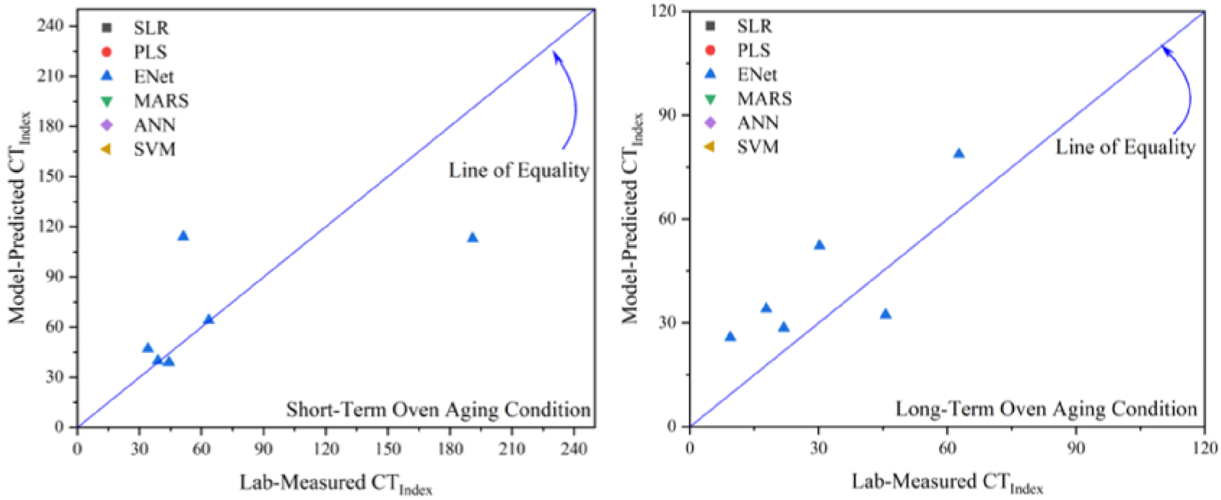
(a) (b)

Figure 44. Comparing the Lab-Measured and SLR Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.



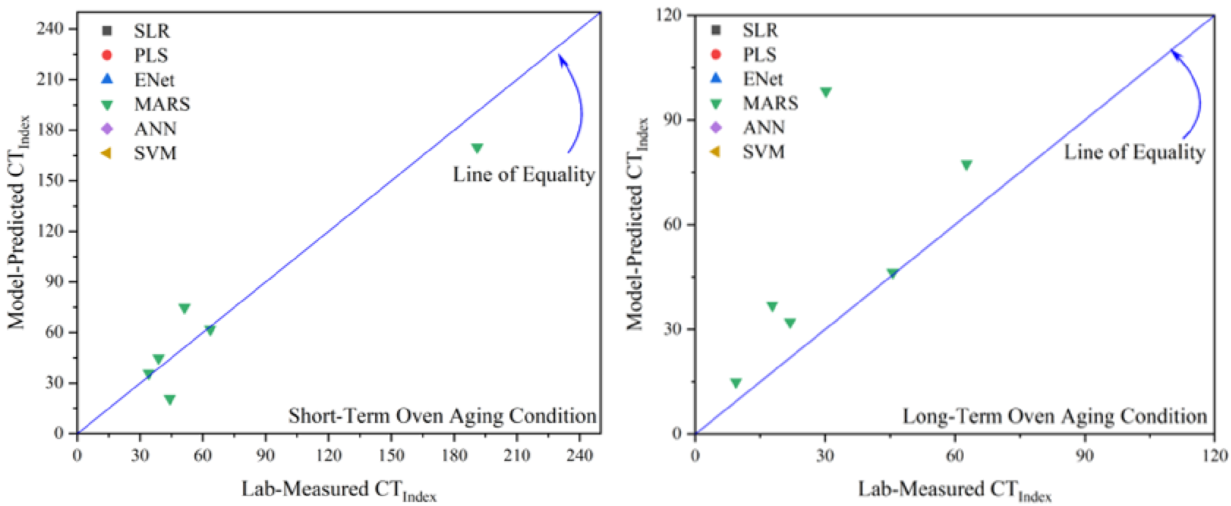
(a) (b)

Figure 45. Comparing the Lab-Measured and PLS Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.



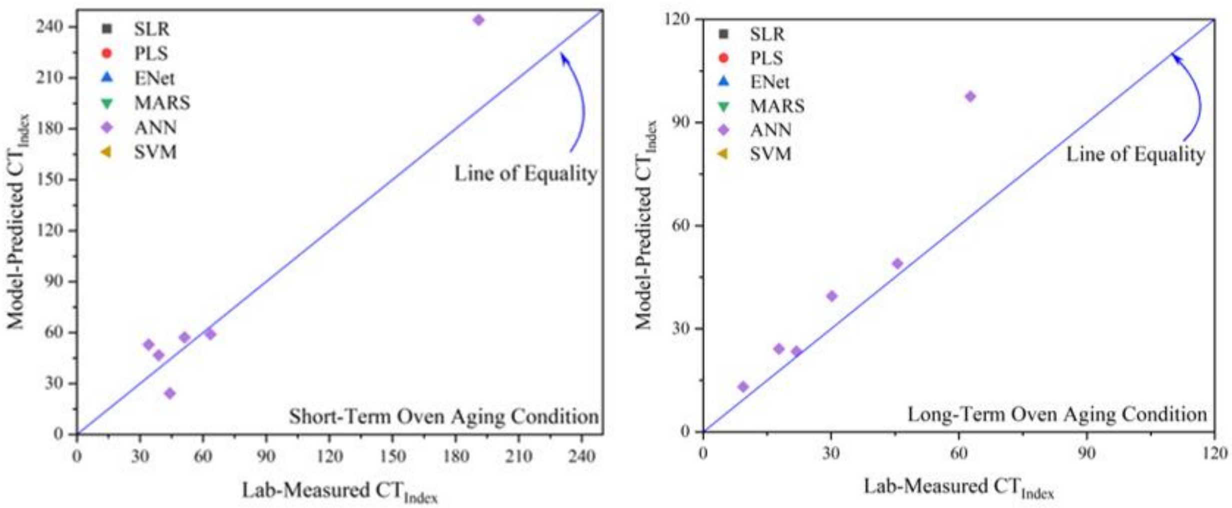
(a) (b)

Figure 46. Comparing the Lab-Measured and ENet Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.



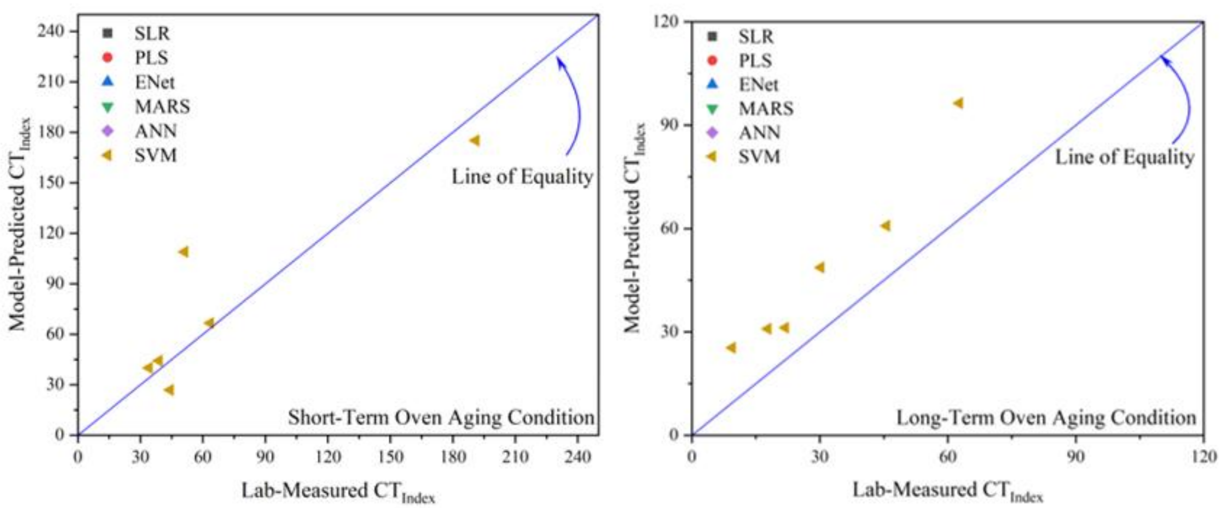
(a) (b)

Figure 47. Comparing the Lab-Measured and MARS Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.



(a) (b)

Figure 48. Comparing the Lab-Measured and ANN Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.



(a) (b)

Figure 49. Comparing the Lab-Measured and SVM Model-Predicted CT_{Index} Values under; (a) STA, and (b) LTA Conditions.

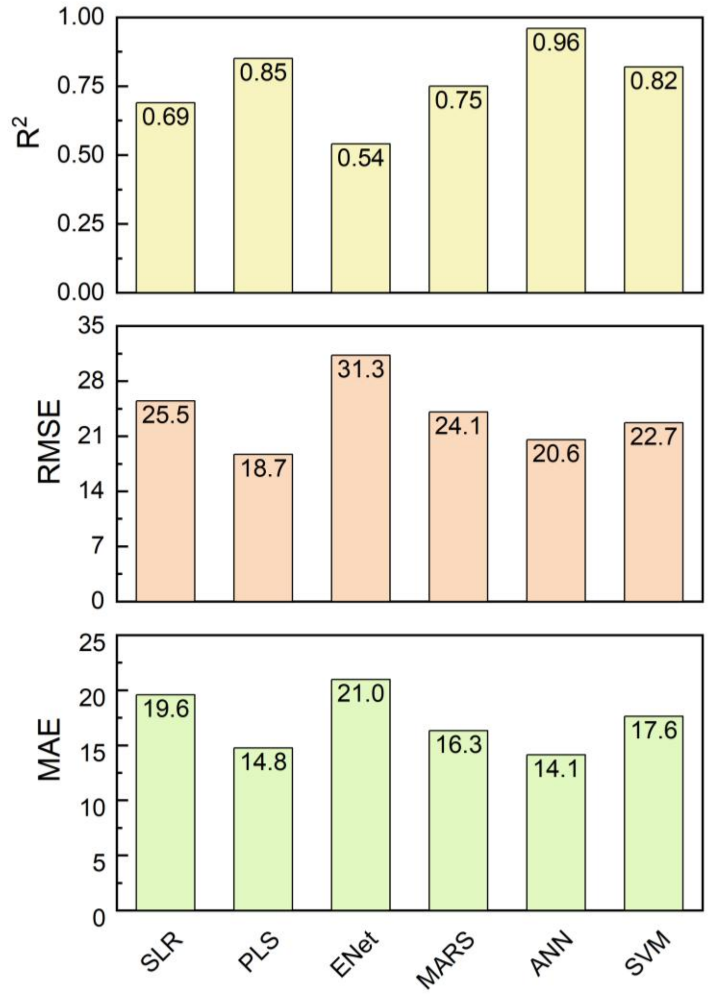


Figure 50. Comparing the Effectiveness of Different Models Considered.

3.1.6 Principal Component Analysis

PCA was employed to identify underlying patterns and reduce dimensionality in the dataset containing the IDEAL-CT parameters, i.e. m_{75}, l_{75}, and fracture energy. As explained previously, PCA accomplishes this by creating a new set of uncorrelated variables, called principal components (PCs), that capture the maximum variance in the original data.

The data was standardized by subtracting the mean of each IDEAL-CT parameter from its corresponding values and then dividing by the standard deviation. This ensured all variables were on the same scale and prevented features with larger scales from dominating the analysis. The covariance matrix was computed to capture the linear relationships between the IDEAL-CT parameters. Eigenvalue decomposition was then performed on the covariance matrix to identify eigenvalues and eigenvectors. Eigenvalues quantify the variance captured by each PC. A larger eigenvalue indicates that the corresponding PC explains a greater portion of the total data variance. The sum of all eigenvalues corresponds to the total variance in the data. Eigenvectors are unit vectors that identify the directions within the data that exhibit the most variation. Eigenvectors can be thought of as arrows pointing towards where the data spreads the most. Each PC is associated with a unique eigenvector, and these eigenvectors collectively define the new axes (PCs) onto which the data is projected. The loading matrix was calculated by multiplying the eigenvectors by the square root of the eigenvalues. It indicates the weight or contribution of each parameter to the formation of a particular PC. The loadings range from -1 to 1, with a higher value (positive or negative) indicating a stronger correlation with a particular PC.

In real-world applications, selecting all principal components is often impractical due to computational limitations. Instead, a smaller subset is chosen based on the amount of variance explained by PCs. Eigenvalues were ranked in descending order, with the first PC capturing the most variance. This study employed a common criterion, retaining only the PCs that cumulatively explain at least 90% of the total variance in the data. Based on this criterion, the first two principal components were retained. The data was presented on the PCA score plot which is a scatter plot of the data projected onto the retained PCs. Score plot can be used to visualize the underlying structure and potential clusters in the data. The results of this will be detailed in the results and discussion section.

3.1.6.1 K-Means Clustering

Following the PCA, K-means clustering was employed to identify distinct groups within the asphalt mixture data, using m_{75}, l_{75}, and fracture energy. K-means clustering yielded a set of clusters, each potentially representing a distinct group within the data. Data points within a cluster share similar characteristics based on the selected parameters, i.e. m_{75}, l_{75}, and fracture energy. While the number of clusters (k-value) can influence how data points are grouped, a k-value of 6 was considered for this study based on trial and error. This selection ensured a clear and meaningful separation of clusters in the two-dimensional space (PC1 vs. PC2) of the PCA score plot. A clear separation allows for a more robust interpretation of the results. Specific clusters can be linked to potential variations or behaviors in the asphalt mixture data based on the underlying relationships captured by the PCs. The resulting clusters were mapped onto the PCA score plot. This visualization served as a valuable tool for identifying groups within the asphalt mixture data that exhibit similar properties and potentially share analogous load-displacement behavior.

All data analysis including PCA and K-means clustering analyses were conducted using JMP® Pro 16 software. This report presents its findings in three different sections. The first section analyzes the correlation between IDEAL-CT parameters. The second section presents the results of PCA and explores the relationship between principal components and IDEAL-CT parameters. Finally, the third section describes the results obtained using k-means clustering.

3.1.6.2 Correlation Between IDEAL-CT Parameters

As explained above, PCA is used to identify correlations between parameters and develop new uncorrelated parameters, i.e. principal components. The correlations between the IDEAL-CT parameters were examined by plotting the relationship between these parameters for all thirty-three mixes, as shown in Figure 51. The correlation coefficients between the parameters were determined as shown in Table 12. The higher the absolute value of the correlation coefficient, the stronger the correlation between the parameters. A positive correlation coefficient indicates positive correlation whereas a negative correlation coefficient indicates negative correlation. Based on the data presented in Figure 51 and Table 12, it is shown that there is a strong negative correlation between m_{75} and l_{75}. The data also suggests that there is a positive correlation between the fracture energy and l_{75}, and a weak negative correlation between fracture energy and m_{75}.

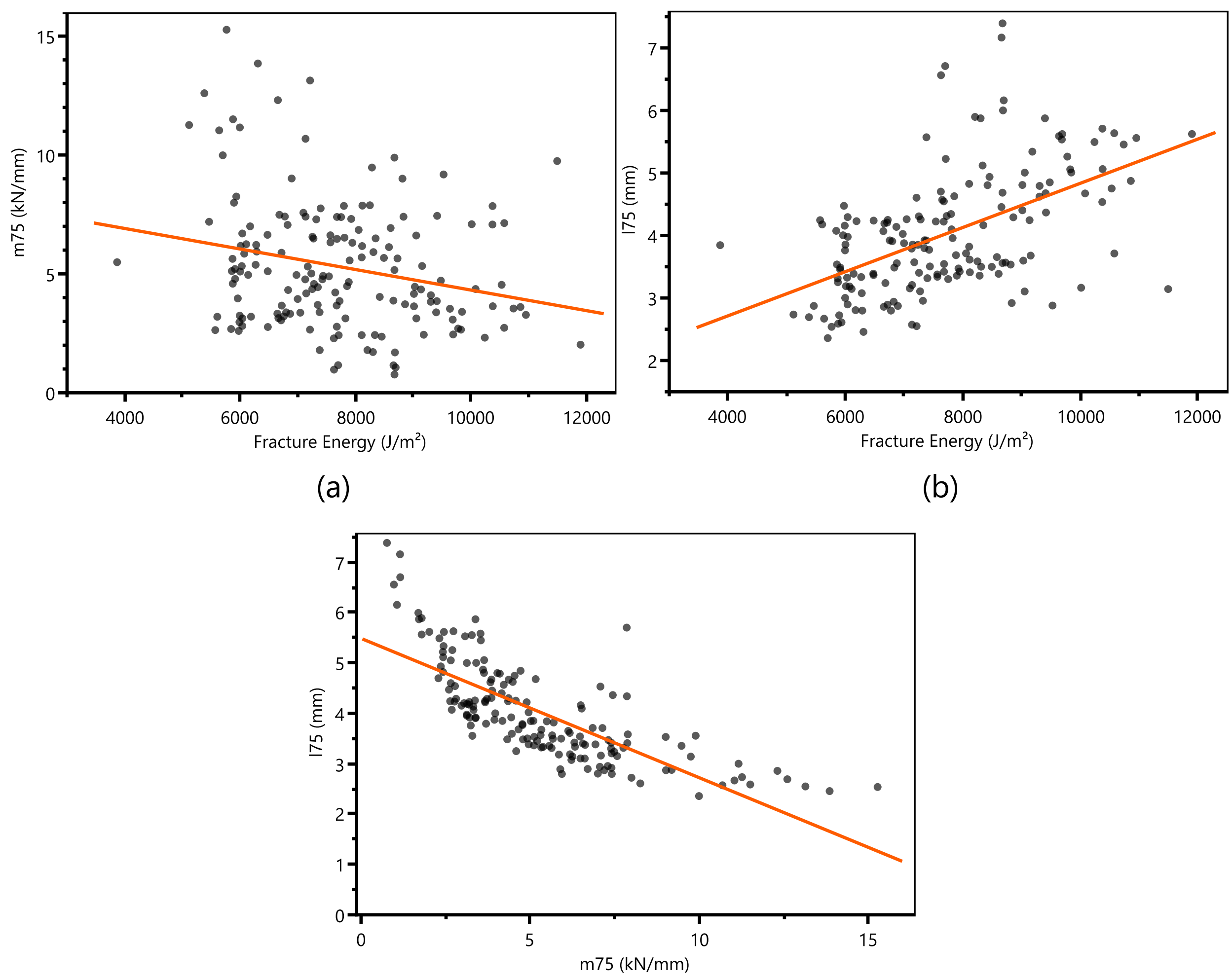


Figure 51. Relationship between IDEAL-CT Parameters; (a) m_{75} vs. Fracture Energy, (b) l_{75} vs. Fracture Energy, and (c) l_{75} vs. m_{75}.

Table 12. Correlation Matrix of m_{75}, l_{75}, and Fracture Energy.

Variables	Fracture Energy	l_{75}	m_{75}
Fracture Energy	1.00	0.53	-0.24
l_{75}	0.53	1.00	-0.76
m_{75}	-0.24	-0.76	1.00

The strong negative correlation between l_{75} and m_{75} is demonstrated in Figure 52 using the load-displacement curves for mixes M2 and M30 as an example. Mix M2 exhibits a high m_{75} and a low l_{75} whereas M30 exhibits a low m_{75} and a high l_{75}. This strong correlation shows the interdependency between the two parameters and the need to use PCA to untangle these correlations.

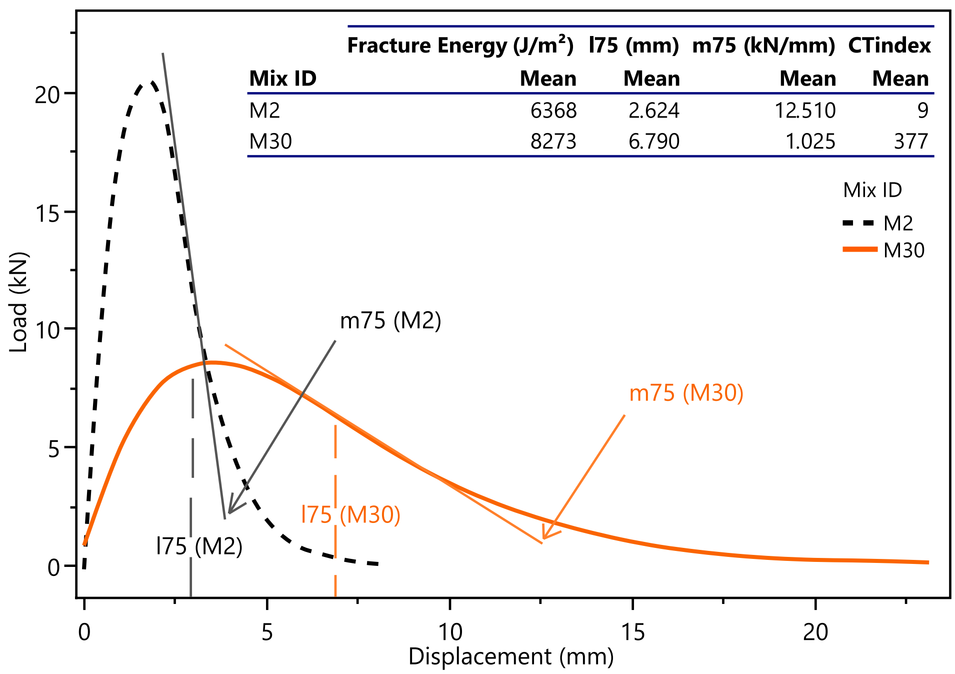


Figure 52. Average Load-Displacement for M2 and M30.

An example of the weak correlation between fracture energy and m_{75} is given by Figure 53 using the load-displacement curves of mixes M28 and M33. While M33 shows a steeper slope, i.e. higher m_{75}, compared to M28, the areas under the curves for both mixes, i.e. related to fracture energy, are approximately the same.

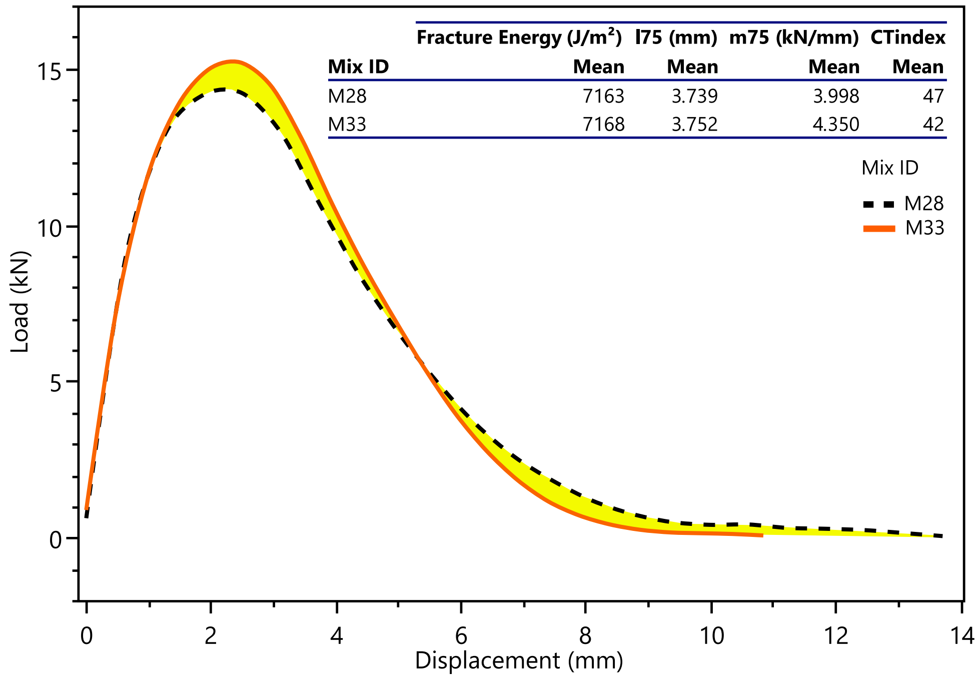


Figure 53. Average Load-Displacement for M28 and M33.

The above observations highlight the inherent interrelationships between the IDEAL-CT parameters. PCA offers a valuable tool in this context, where it not only allows for the reduction of the number of parameters but also leads to the creation of new components that are uncorrelated with each other.

3.1.6.3 Principal Components

Table 13 shows the eigenvalues resulted from PCA. The first principal component, i.e. PC1, with an eigenvalue of 2.05 captures a much larger portion of the variance compared to PC2 (0.78) and PC3 (0.17). Together, PC1 and PC2 capture 94% of the total variance in the data.

Eigenvector coefficients represent the weights (or importance) of the original variables, (i.e. l_{75}, fracture energy, and m_{75}) in that principal component (Table 13). For example, in PC1, l_{75} has a higher weight (0.66) compared to fracture energy (0.47), indicating it contributes more to the variance captured by PC1. The eigenvector coefficients act as a recipe for constructing the direction of the eigenvector in the space defined by the original variables.

Loading matrix is calculated by multiplying the eigenvectors by the square root of the eigenvalues (Table 13). For instance, the I_{75} loading for PC1 (0.95) is obtained by multiplying the eigenvector (0.66) by the square root of its eigenvalue (2.05). A high I_{75} loading for PC1 suggests a strong positive correlation between I_{75} and the direction of maximum variance captured by PC1. As the loading matrix suggests, PC1 is primarily influenced by I_{75} and m_{75}, with less impact from the fracture energy. An increase in PC1 indicates an increase in both I_{75} and fracture energy, and a decrease in m_{75}. On the other hand, PC2 shows a strong association with fracture energy with less impact from m_{75} and almost no impact from I_{75}. An increase in PC2 indicates a decrease in both the fracture energy and m_{75}.

Table 13. PCA Results; Eigenvalues, Eigenvectors and Loading Matrix

PCA Results	**Principal Component PC1**	**Principal Component PC2**	**Principal Component PC3**
Eigenvalues	2.05	0.78	0.17
Percent variance explained	68	26	6
Cumulative percent variance explained	68	94	100
Eigenvectors			
Fracture energy	0.47	-0.83	-0.31
I_{75}	0.66	0.10	0.74
m_{75}	-0.59	-0.55	0.59
Loading matrix			
Fracture Energy	0.67	-0.73	-0.13
I_{75}	0.95	0.09	0.31
m_{75}	-0.84	-0.49	0.25

Figure 54 presents a PCA score plot where each data point represents a mixture (mix ID) and is labeled with its corresponding mean CT_{Index}. Each axis in the score plot corresponds to a principal component. This two-dimensional graph serves as a reliable approximation of the original three-dimensional space defined by the initial parameters, i.e. m_{75}, l_{75}, and fracture energy. Further discussion regarding the variation of the principal components with the different IDEAL-CT parameters will be presented later.

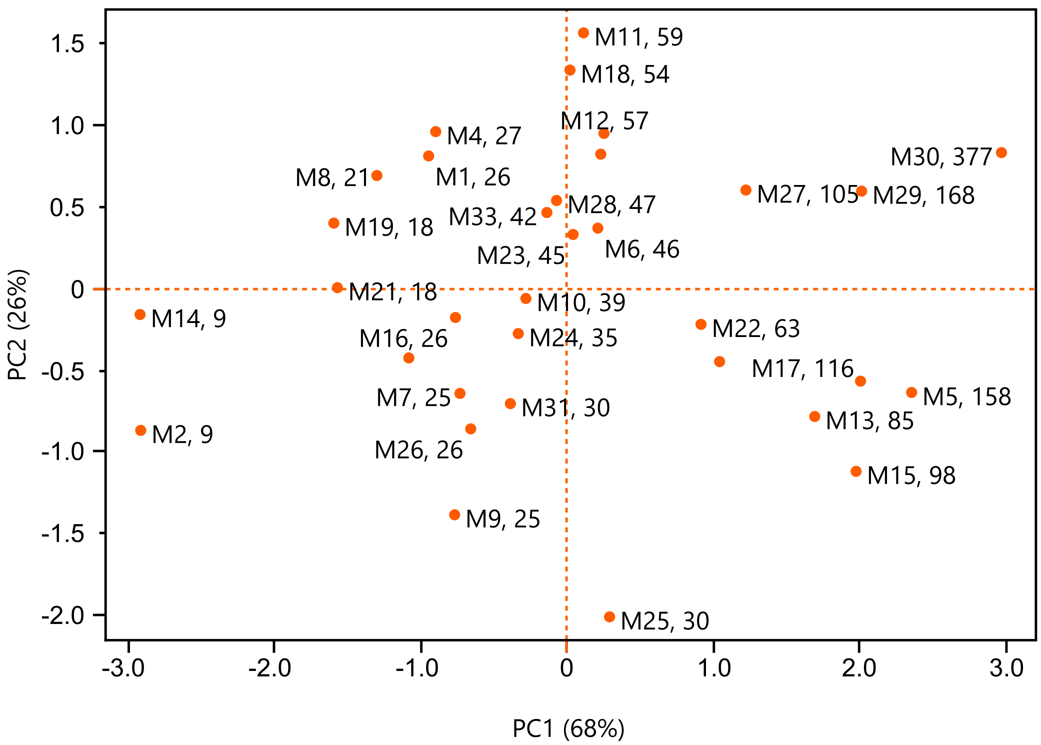


Figure 54. PCA Score Plot (showing Mix ID and Mean CT_{Index} as Labels).

3.1.6.4 Relation Between Principal Components and IDEAL-CT Parameters

To further investigate the significance of PC1 and PC2 and how they relate to the IDEAL-CT parameters and the shape of the load-displacement curve, four mixtures were selected with different PC1 and PC2 values as shown in Figure 55. The figure shows the mean PCs and standard deviation error bars for each mix. Figure 56 shows the load-displacement curves for these four mixtures.

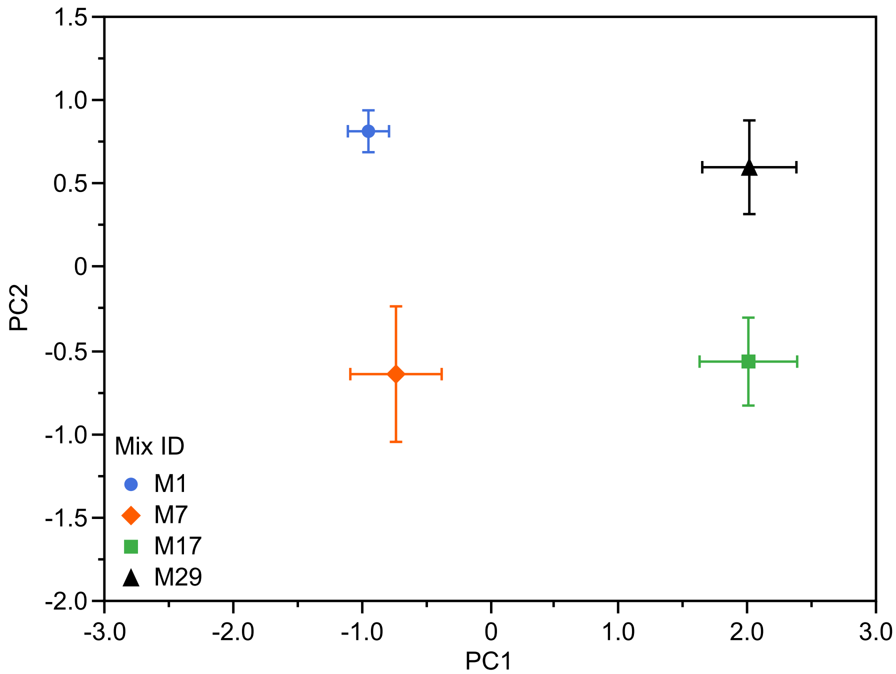


Figure 55. Score Plot for M1, M7, M17, and M29.

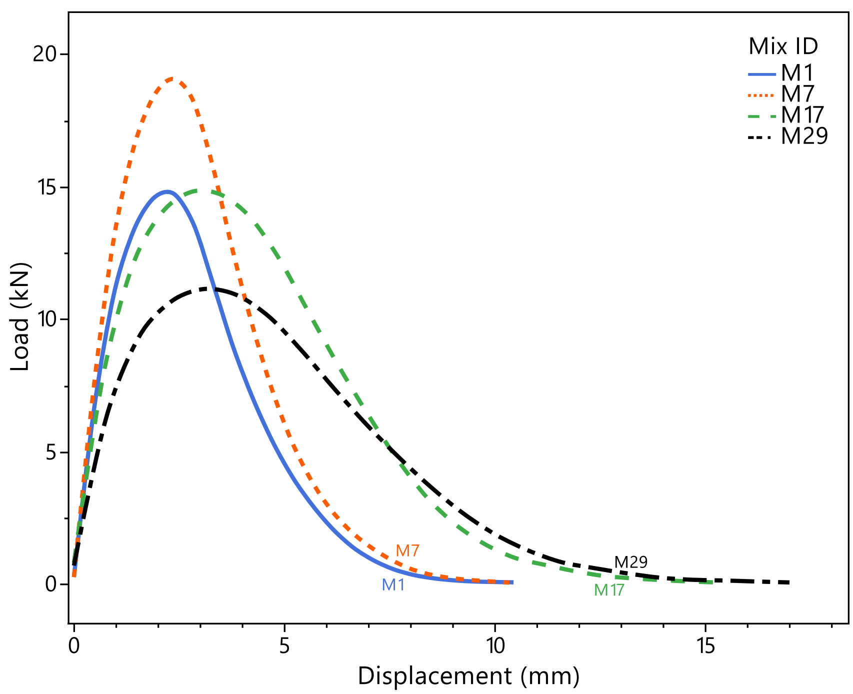


Figure 56. Average Load-Displacement Curve for M1, M7, M17, and M29.

The mean IDEAL-CT parameters for the four mixtures, i.e. M1, M7, M17 and M29, are shown in Table 14.

Table 14. Mean IDEAL-CT Parameters for M1, M7, M17, and M29

Mix ID	CT_{Index}	l_{75} (mm)	m_{75} (kN/mm)	Fracture energy (J/m^2)
M1	26	3.34	5.29	6127
M7	25	3.54	7.56	7942
M17	116	5.28	3.01	9790
M29	168	5.69	1.89	8183

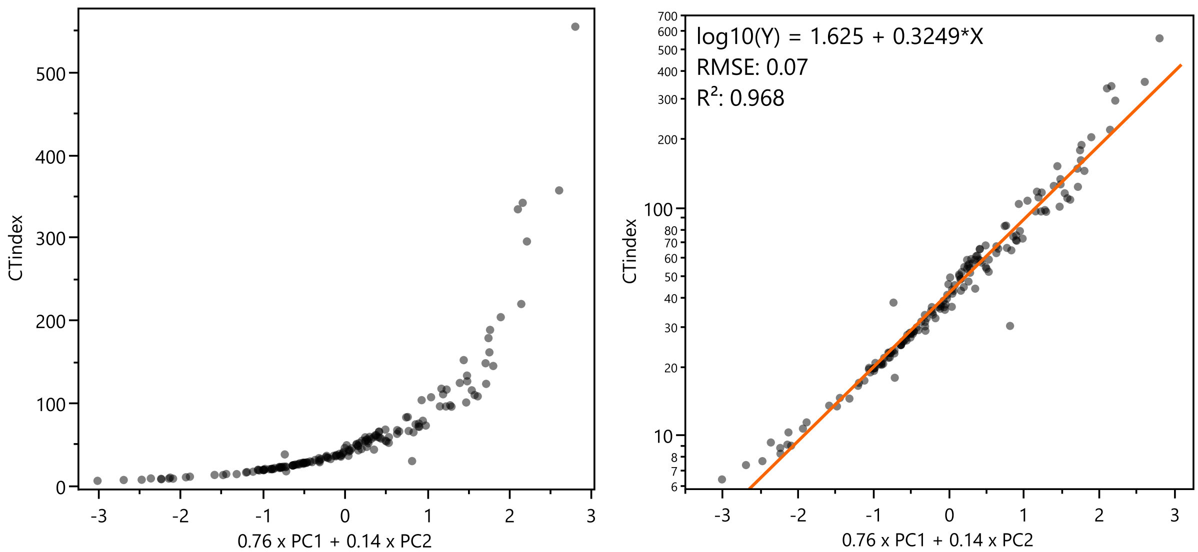
A comparison between mixes M7 vs. M17 and M1 vs. M29, i.e. mixes with varying PC1 and constant PC2, reveals the following:

- An increase in l_{75} and a decrease in m_{75} is noted between the mixes as PC1 increases. The fracture energy also appears to increase with the increase of PC1.
- As PC1 increases, the load-displacement curve exhibits a base widening, signifying enhanced ductility and fracture energy alongside reduced stiffness and peak load (Figure 56).

A comparison between the M1 vs. M7 and M17 vs. M29, i.e. mixes with varying PC2 and constant PC1, reveals the following:

- A decrease in both fracture energy and m_{75} is noted between the mixes as PC2 increases. There also seems to be minimal impact on l_{75} with the change on PC2.
- As PC2 increases, the load-displacement curve does not exhibit a significant base narrowing (indicating constant l_{75}). However, a decrease in fracture energy is observed, accompanied by a decrease in the m_{75} (i.e. a less steep slope).
- Mixes M1 and M7 show constant CT_{Index} despite having different load displacement curves as captured by the change in PC2.

The relationship between the principal components and the CT_{Index} was further investigated. As shown in Figure 57, a weighted summation of PC1 and PC2 was shown to correlate with the CT_{Index}. Using partial least square (PLS) regression, it was shown that CT_{Index} correlates exponentially with 0.76×PC1 + 0.14×PC2 as shown in Figure 57(a). Figure 57(b) shows the same plot using log(CT_{Index}) versus the weighted summation of PC1 and PC2, indicating a linear relationship.



(a) (b)

Figure 57. (a) CT_{Index} and (b) log of CT_{Index} vs. Weighted Summation of PC1 and PC2.

Figure 58 shows the variation of the different IDEAL-CT parameters with the principal components using contour plots. The CT_{Index} increases with the increase of PC1 and PC2 as shown by the correlation presented previously. The m_{75} increases in the same direction as CT_{Index}, while l_{75} trends in the opposite direction to m_{75}. An increase in PC1 and PC2 is associated with a decrease in m_{75}, which results in higher CT_{Index}. It is also noted that l_{75} does not show notable variation with PC2.

Furthermore, the direction of change for fracture energy appears to be relatively perpendicular to the direction of change observed in m_{75}, l_{75} and CT_{Index}. This implies that fracture energy does not strongly correlates with the CT_{Index} where mixes with similar fracture energy can exhibit significantly different CT_{Index} values. As previously discussed, fracture energy primarily influences the direction defined by PC2, while CT_{Index} is more heavily influenced by variations along PC1 which is dominated by l_{75} and m_{75}, as shown in Figure 58. Therefore, changes in fracture energy along PC2 may not significantly impact CT_{Index}.

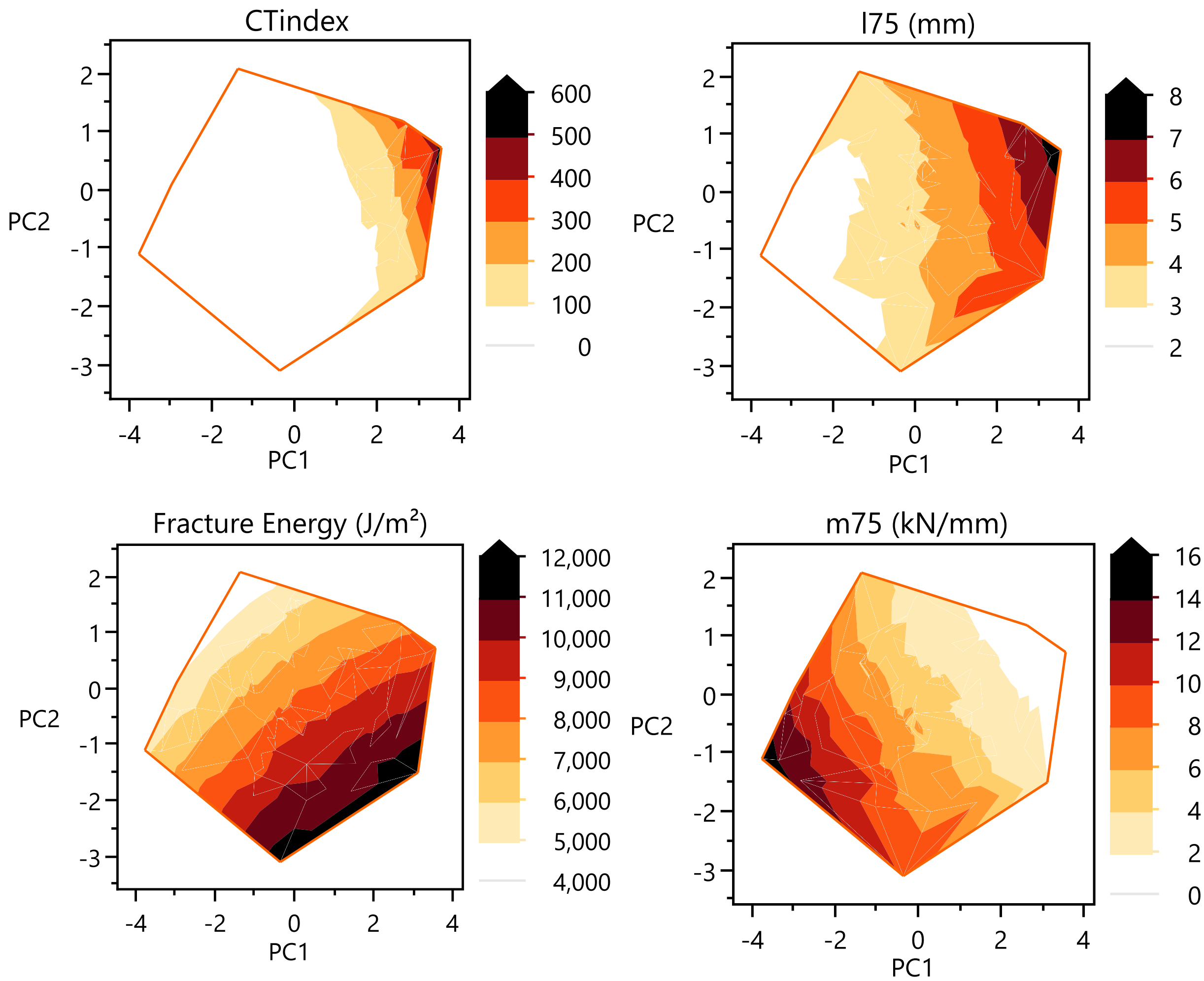


Figure 58. Contour Plots of different IDEAL-CT parameters such as $CT_{Index,}$, l_{75}, m_{75}, and Fracture Energy vs. PCA Score Plot.

3.1.6.5 Utilizing K-means Clustering to Detect Mixture Groups

K-means clustering was employed to identify subgroups within the mixture data. This method utilized m_{75}, l_{75}, and fracture energy as input variables. The resulting clusters are visualized on the PCA score plot as shown in Figure 59. A total of six clusters were identified.

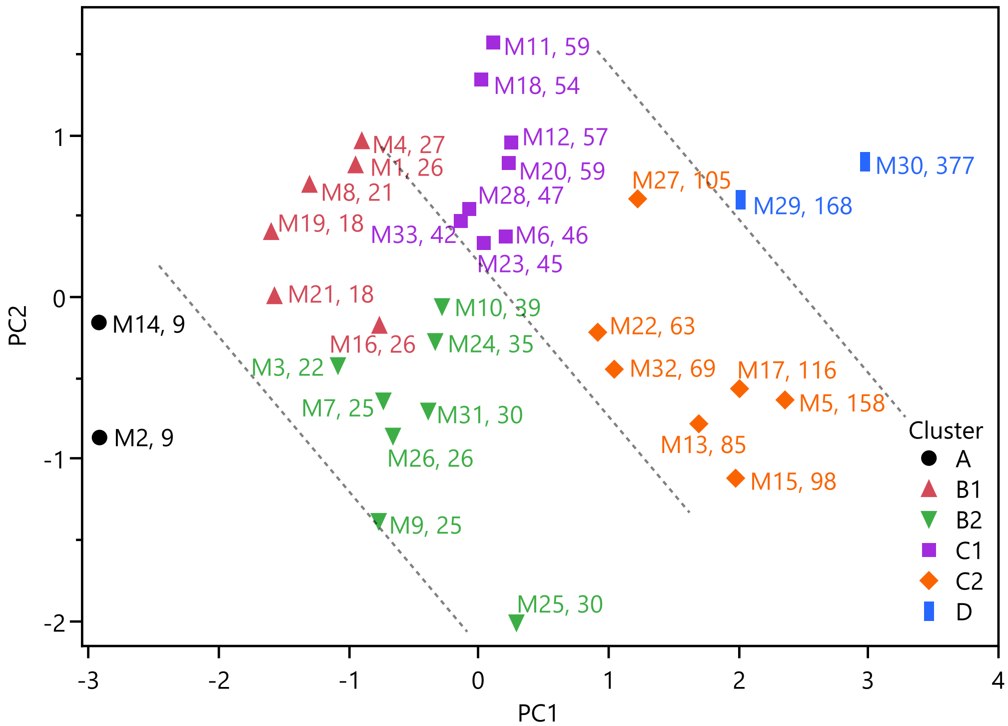


Figure 59. Clusters shown on the PCA Score Plot (Labels are Mix ID and Mean CT_{Index}).

While selecting a lower or higher k-value could potentially merge or separate some clusters, a k-value of 6 was chosen for this analysis based on trial and error to ensure clear distinction between the clusters along both the x and y directions of the PCA score plot. The labels shown in Figure 59 next to each mix indicate the mix ID and mean CT_{Index}. The mean IDEAL-CT parameters for the clusters are shown in Table 15. As indicated previously, the CT_{Index} increases with the increase of PC1 and PC2. Cluster A on the lower left corner of the score plot has the lowest CT_{Index} whereas cluster D on the upper right corner of the plot has the highest CT_{Index}. Clusters B1 and B2 lie in a direction perpendicular to the change in CT_{Index} and hence have comparable average CT_{Index}. The same is also true for clusters C1 and C2. Accordingly, cluster B was defined to combine both clusters B1 and B2 and cluster C was defined to combine both clusters C1 and C2.

Table 15. Mean IDEAL-CT Parameters for Six Clusters

Cluster	CT_{Index}	l_{75} (mm)	m_{75} (kN/mm)	Fracture energy (J/m^2)
A	9	2.637	12.070	6072
B1	22	3.158	6.353	6316
B2	29	3.524	7.046	8519
C1	53	4.082	3.683	6892
C2	97	5.028	3.659	9628
D	272	6.248	1.454	8165

Figure 60 depicts the average load-displacement curves for the main four clusters, i.e. A, B, C, and D. The shape of the load-displacement curve shifts to the right with l_{75} increasing, m_{75} decreasing, and fracture energy increasing, resulting in an increase in CT_{Index}.

Figure 61 shows the average load-displacement curves for the subclusters, i.e. B1, B2, C1 and C2. It is shown that the load-displacement curves change notably between B1 to B2 and between C1 to C2 even though there is relatively small change in the CT_{Index}. The change in the load-displacement curve is mainly due to an increase in energy fracture, combined with an increase in m_{75}, which act against each other to cancel out or minimize the effect on the CT_{Index}.

Based on the clustering results and the corresponding load–displacement curves, it can be observed that clusters exhibiting favorable cracking performance generally show peak loads in the range of 10–15 kN. Curves within this peak-load range tend to provide a balanced combination of fracture energy, $l_{75,}$ and m_{75}, leading to optimal IDEAL-CT parameters and CT_{Index} values.

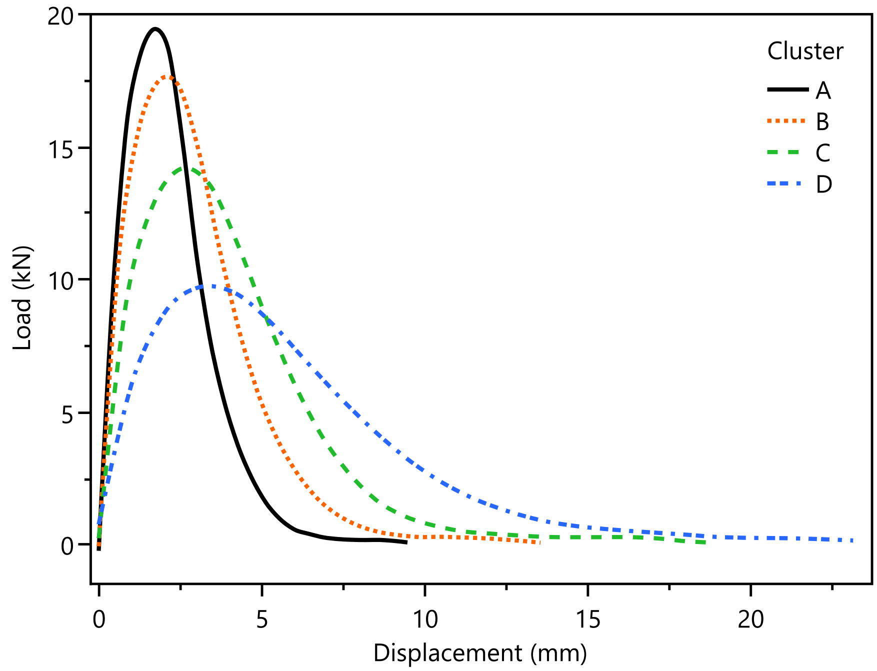


Figure 60. Average Load-Displacement Curve for Clusters A, B, C, and D.

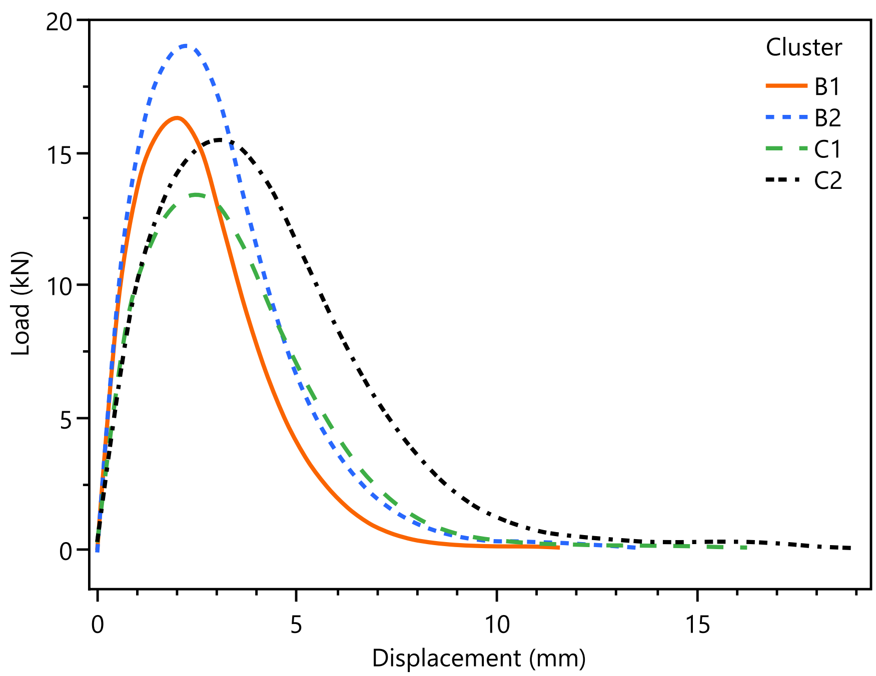


Figure 61. Average Load-Displacement Curve for Subclusters B1, B2, C1, and C2.

The six clusters, representing different modes of behavior, and their mean load-displacement curves, are shown on the PCA score plot in Figure 62. There are four regions in the score plot that exhibit a diagonal relationship to each other. The lower left and upper right regions together determine the directional changes in m_{75}, I_{75}, and CT_{Index}. Conversely, the upper left and lower right regions are primarily associated with the directional change in fracture energy.

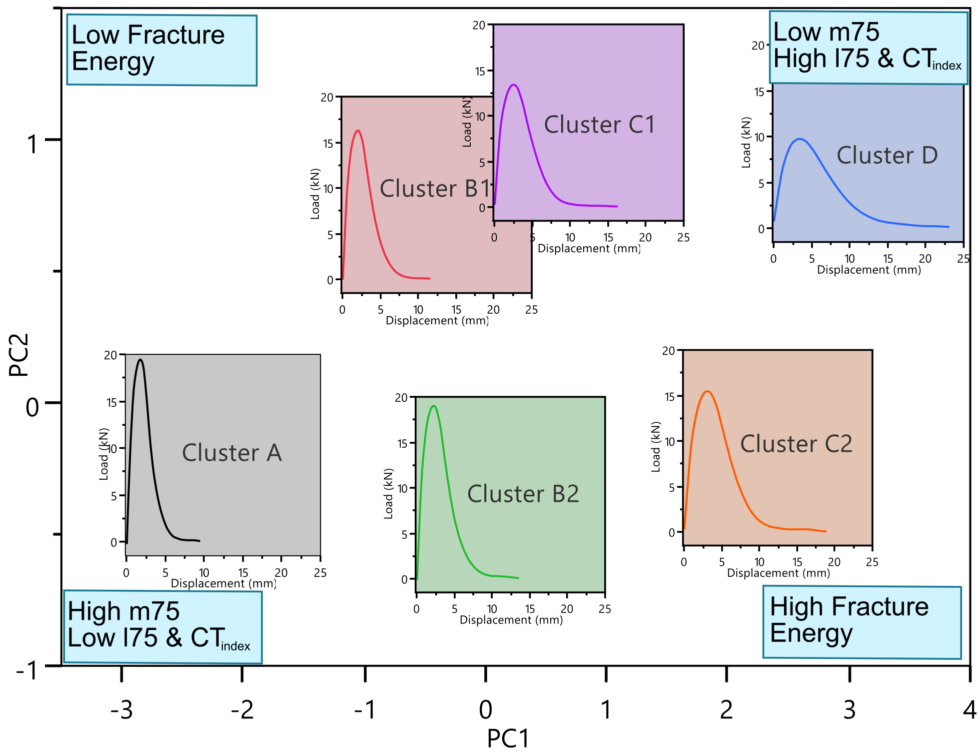


Figure 62. Clusters and their Mean Load-Displacement Curves on the PCA Score Plot.

3.1.6.6 Prevalent Mixture Variables Among Clusters

To further examine the mixture compositions within each cluster, the prevalent mixture variables were identified as shown in Table 16. To determine these prevalent characteristics, the mode was calculated for categorical variables such as binder modification (neat, i.e. PG58-28 and PG64-22 or modified, i.e. PG70-28 and PG76-28), RAP, and mix type. The mode represents the most frequently occurring value within each cluster for these variables. For the continuous variable total AC, the median was calculated. The median is less susceptible to outliers compared to the mean and can be a more robust measure for skewed data.

The analysis of median AC content across clusters reveals a trend along the PC1 and PC2 axes of the score plot. Cluster A, positioned in the low PC1 and PC2 region, exhibits the lowest median AC. This observation aligns with the fact that an increase in AC content typically results in an increase in cracking resistance as measured by CT_{Index}.

Table 16. Mode of Mixture Variables (n/total) and Median Total AC% for each Cluster

Cluster	Binder Modification*	RAP*	Mix type*	Total AC (%)**
A	Neat (2/2)	RAP (2/2)	S3 (2/2)	4.3
B	Neat (13/14)	RAP (11/14)	S4 (6/14)	4.9
C	Modified (10/15)	No RAP (12/15)	S4 (8/15)	5.3
D	Modified (2/2)	No RAP (2/2)	SMA (1/2)	6.8

**Mode of variable in each cluster;*
***Median of variable in each cluster.*

Figure 63 demonstrates a clear association between the location of clusters within the PCA score plot and the composition of the mixtures. Cluster A is comprised entirely of mixtures with neat binder, while cluster D contains only mixtures with modified binder. This observation suggests a gradual increase in the mixtures containing modified binder relative to neat binder moving from cluster A to D in the score plot, which suggests that modified binders generally performed better in terms of cracking resistance compared to neat binders. Similarly, cluster A is characterized by the mixtures containing RAP, while mixtures in cluster D have no RAP. This trend indicates a transition from mixtures containing RAP (cluster A) to those with no RAP (cluster D) moving across the PCA score plot, which suggests that the CT_{Index} decreased as more RAP was added. Mixtures assigned to clusters A, located in the low PC1 and PC2 region of the PCA score plot and exhibiting low to medium CT_{Index} values, are exclusively S3 mixtures. Mixtures within clusters B and C, positioned in the central region of the plot and associated with medium to high CT_{Index} values, are mostly composed of S4 mixtures. Finally, the two mixtures in cluster D, situated in the upper right corner with high CT_{Index} values, are SMA and S5 mixtures.

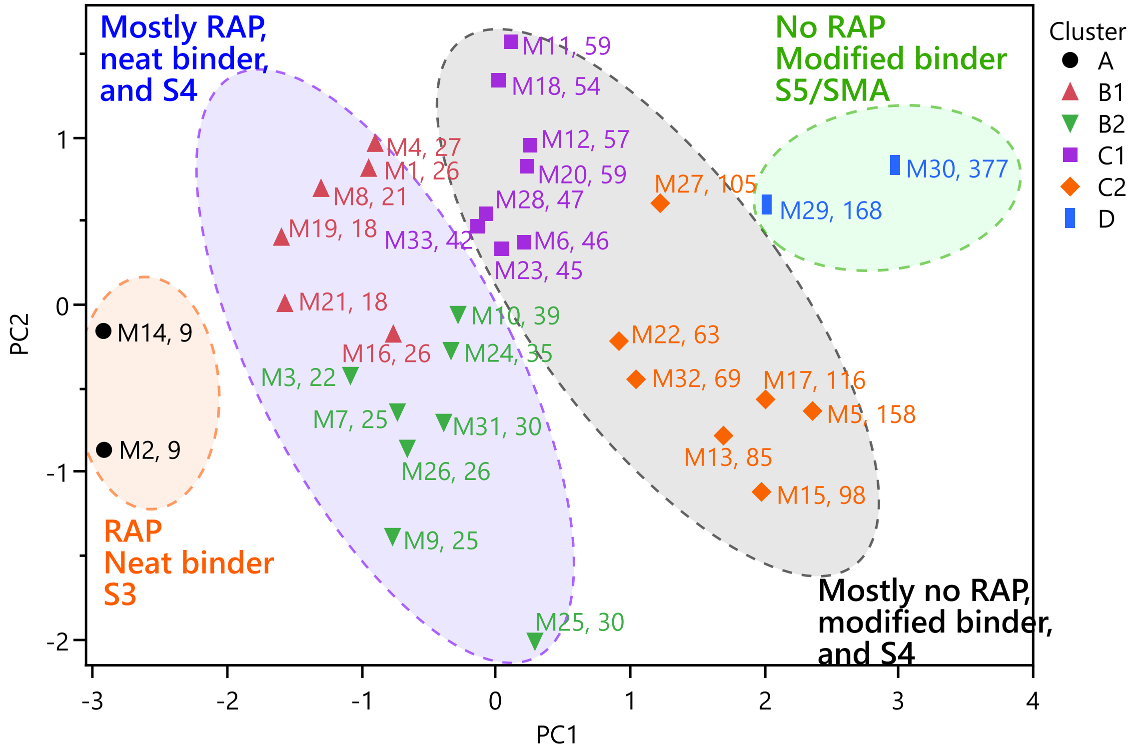


Figure 63. Prevalent Binder and RAP Content in Clusters (Labels are Mix ID and CT_{Index}).

3.1.6.7 PCA Summary

In this study, PCA was used to analyze the IDEAL-CT results of thirty-three plant-produced mixes subjected to LTA condition. Additionally, k-means clustering was employed to explore the underlying structure within the data and identify distinct mixture groups. In summary, this study demonstrates the effectiveness of using PCA and k-means clustering to analyze the complex relationships between IDEAL-CT parameters, CT_{Index}, and mixture composition. These techniques provided valuable insights that simple data analysis methods might have missed. The findings suggest that a combination of factors, including l_{75}, m_{75}, fracture energy, and mixture composition, influences cracking resistance in asphalt mixtures. The ability to identify these relationships can aid in developing and optimizing asphalt mixtures for improved performance and durability.

3.2 Rutting and Moisture Sensitivity

A total of thirty-three plant produced mixes were tested for rutting and moisture susceptibility using the HWTT. The results of the HWTT were analyzed to determine the TRD to assess rutting and the CS, SS, and SIP to assess moisture susceptibility as per AASHTO T324. Additionally, the results were analyzed using the NCAT method to determine the CRD for rutting, and the SN, LC_{ST}, and $LC_{12.5}$ for moisture susceptibility.

A rut depth curve undergoes different stages as shown in Figure 64(a). An initial post-compaction stage is followed by a creep phase with a constant slope. If stripping occurs, a stripping phase is observed after the creep phase. The post-compaction phase typically occurs during the first 1,000 wheel passes, in which the specimens undergo consolidation under load. The creep phase is represented by the first steady-state portion of the curve, in which the specimens show a relatively constant increase in rut depth with wheel passes.

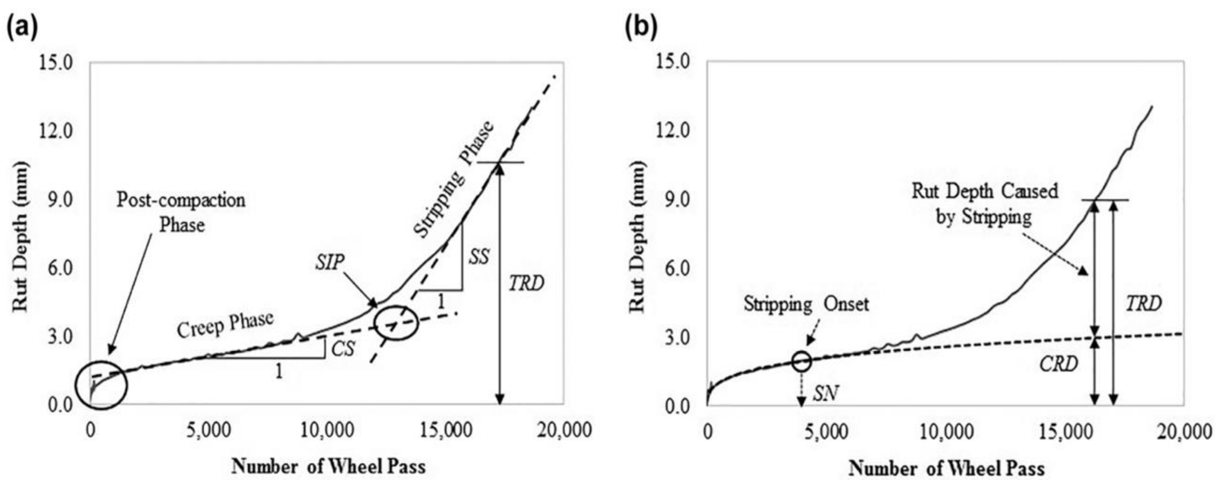


Figure 64. Rutting and Moisture Parameters using; (a) AASHTO T 324 Method, and (b) NCAT Method (Yin, Chen, et al., 2020).

The stripping phase, if it occurs, is represented by a second steady-state portion of the curve with a steeper slope than the creep phase. The accumulated rut depth in the stripping phase is mainly caused by the stripping of asphalt binder from the aggregates rather than the permanent deformation of the mixture. Traditional HWTT parameters specified in AASHTO T 324 include the TRD at a certain number of wheel passes, CS, SS, and SIP. SIP is defined as the number of wheel passes at the onset of the stripping phase and is graphically represented as the intersection of the CS and the SS. Asphalt mixtures with lower TRD, CS, and SS, and higher SIP values are expected to exhibit good performance in relation to rutting resistance and moisture resistance (Yin, Chen, et al., 2020). The AASHTO T 324 method however does not isolate the effect of stripping and rutting beyond the SIP. Once stripping occurs, the TRD is caused by both rutting and stripping. To address this limitation an alternative method was proposed by researchers at NCAT as shown Figure 64(b). Based on the NCAT method, parameters such as LC_{ST} and SN are used to evaluate moisture, and the CRD is used to evaluate rutting (Yin, Chen, et al., 2020).

In this study, the HWTT results will be analyzed using both the AASHTO T 324 method and the NCAT method, to evaluate the performance of the plant-produced mixes.

3.2.1 HWTT Moisture Results

3.2.1.1 AASHTO Parameters & TSR

Using the HWTT rut depth curve, the moisture parameters were calculated for all 33 mixes. OriginPro, a graphical and statistical software, was used to plot the curves and determine the CS, SS, and SIP. By applying the smooth curving feature in OriginPro, any noise in the measurement was reduced and a smooth curve was obtained. The CS was obtained by applying the line fitting tool to the region where creep begins and ends. To get the SS, if any, the line fitting tool is used on the region where the stripping begins and ends. To obtain the SIP, the SS and CS are extended to determine the point of intersection that represents the SIP. The calculation of the CS, SS, and the SIP were all done in accordance with AASHTO T324, as presented in Table 17.

Table 17. AASHTO Parameters for Moisture Susceptibility.

Mix ID	CS	SS	SIP
M1	2.00E-04	No SS	No SIP
M2	1.10E-04	No SS	No SIP
M3	1.90E-04	6.40E-04	15,512
M4	1.70E-04	3.40E-03	15,943
M5	9.30E-04	No SS	No SIP
M6	4.00E-04	2.50E-03	9,554
M7	8.30E-05	No SS	No SIP
M8	2.10E-05	3.8E-03	11,336
M9	3.00E-04	No SS	No SIP
M10	8.50E-05	No SS	No SIP
M11	2.40E-04	7.57E-03	9,242
M12	1.10E-04	1.21E-03	17,821
M13	5.20E-05	No SS	No SIP
M14	1.50E-04	1.16E-03	16,164
M15	6.00E-05	No SS	No SIP
M16	4.29E-04	No SS	No SIP
M17	2.10E-04	1.62E-03	15,438
M18	4.90E-04	4.40E-03	11,115
M19	4.60E-05	3.50E-03	10,943
M20	2.50E-04	2.23E-03	9,534
M21	1.20E-04	2.2E-04	17,916
M22	1.58E-04	2.55E-03	11,569
M23	5.20E-04	1.41E-03	6,610
M24	5.14E-04	1.65E-03	8,066
M25	6.82E-05	No SS	No SIP
M26	3.86E-04	1.79E-03	9,090
M27	4.39E-05	No SS	No SIP
M28	9.70E-05	No SS	No SIP
M29	1.99E-04	No SS	No SIP
M30	3.90E-04	5.6E-04	15,211
M31	2.23E-04	2.02E-03	9,264
M32	2.14E-04	1.12E-03	12,876
M33	9.30E-05	No SS	No SIP

The distribution of the SIP values for the 33 mixes is shown in Figure 65. Fourteen out of the thirty-three mixes showed no stripping, whereas seven mixes showed significant stripping with a SIP less than 10,000 passes. Currently, the ODOT does not use SIP to evaluate moisture. Several states, however, including Georgia, Massachusetts, Washington, and Maine, have adopted a minimum SIP of 15,000 passes as the moisture susceptibility criteria. By applying this criterion to the 33 PMLC Oklahoma mixes, 12 mixes making up 36% of the 33 mixes will fail the criteria and be out of the specification.

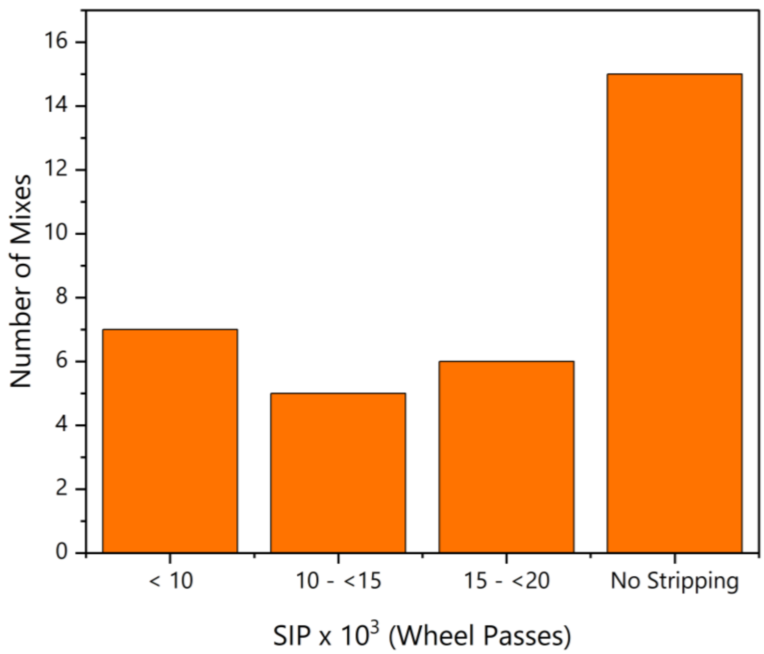


Figure 65. The SIP Distribution of the PMLC mixes.

The ODOT uses the AASHTO T 283, i.e., modified Lottman test, to assess moisture, and a TSR value of 0.8 and 0.75 is required during mix design and field acceptance, respectively.

The TSR from the ODOT mix design was plotted against the SIP as shown in Figure 66. It can be noted that there was no correlation between the TSR and SIP. All mixes are ODOT-approved mixes hence all TSR values were above the ODOT minimum criterion of 0.8. The TSR and the SIP results are summarized, in Table 18.

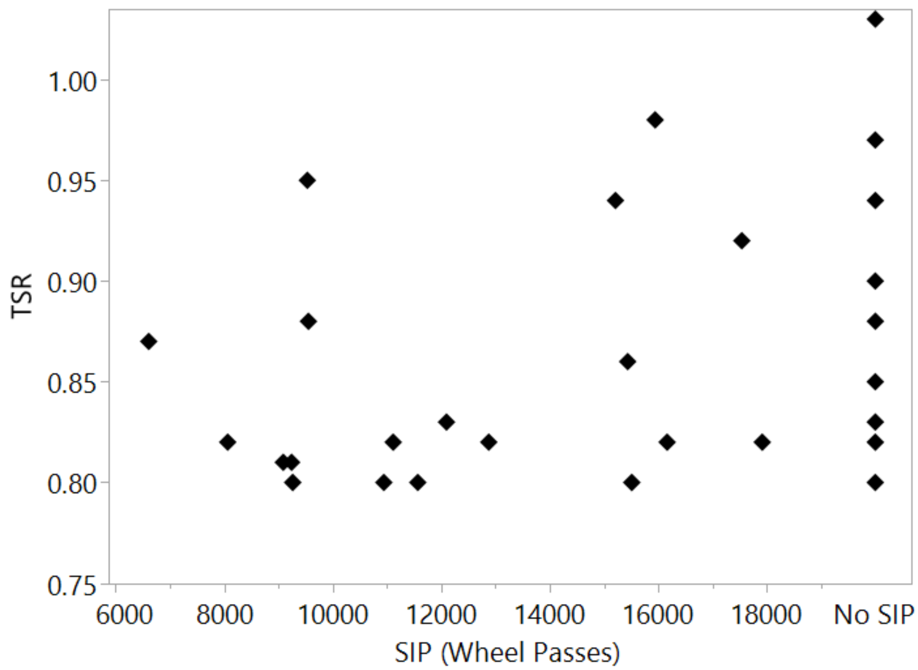


Figure 66. Relationship between TSR and SIP.

Table 18. TSR and SIP for all 33 PMLC Mixes

Mix ID	TSR	SIP
M1	0.83	No SIP
M2	0.82	No SIP
M3	0.80	15,512
M4	0.98	15,943
M5	0.94	No SIP
M6	0.88	9,554
M7	1.03	No SIP
M8	0.83	11,336
M9	0.80	No SIP
M10	0.90	No SIP
M11	0.81	9,242
M12	0.92	17,821
M13	0.97	No SIP
M14	0.82	16,164
M15	0.85	No SIP
M16	0.80	No SIP
M17	0.86	15,438
M18	0.82	11,115
M19	0.80	10,943
M20	0.95	9,534
M21	0.82	17,916
M22	0.80	11,569
M23	0.87	6,610
M24	0.82	8,066
M25	0.88	No SIP
M26	0.81	9,090
M27	0.83	No SIP
M28	0.90	No SIP
M29	0.82	No SIP
M30	0.94	15,211
M31	0.80	9,264
M32	0.82	12,876
M33	0.82	No SIP

3.2.1.2 NCAT Parameter Calculation

NCAT procedure was followed for calculating CRD, SN, and LCST. All calculations were performed using a workflow in JMP Pro software, as follows:

- JMP Pro was used to fit the full HWTT curve using the NCAT visco-plastic deformation model, which separates rutting-induced deformation from stripping-induced deformation. This method is described in the NCAT studies (Yin, Chen, et al., 2020; Yin et al., 2014).
- Using the fitted model parameters, CRD, SN, and LCST were calculated.
- SN was determined as the load level where the fitted curve transitions from negative to positive curvature (onset of stripping).
- CRD was computed by projecting the visco-plastic portion of the curve beyond SN and isolating the permanent deformation.
- LCST was calculated as the number of additional wheel passes required for stripping-related deformation to reach 12.5 mm.

More details about the NCAT parameters calculation can be found in APPENDIX E.

3.2.2 Relationship Between the SN, SIP, and SS

The SN parameter was calculated based on the NCAT method and plotted against the SIP parameter as shown in Figure 67. A power function was fitted to these data points, resulting in an R^2 value of 0.90, indicating a strong correlation between them. This analysis is consistent with the findings reported by (Yin, Chen, et al., 2020).

For mixtures that did not exhibit measurable SN or SIP values, a value corresponding to 20,000-wheel passes was assigned only for plotting purposes and to illustrate the correlation between SN and SIP. These assigned values were not used for interpretation beyond visualization.

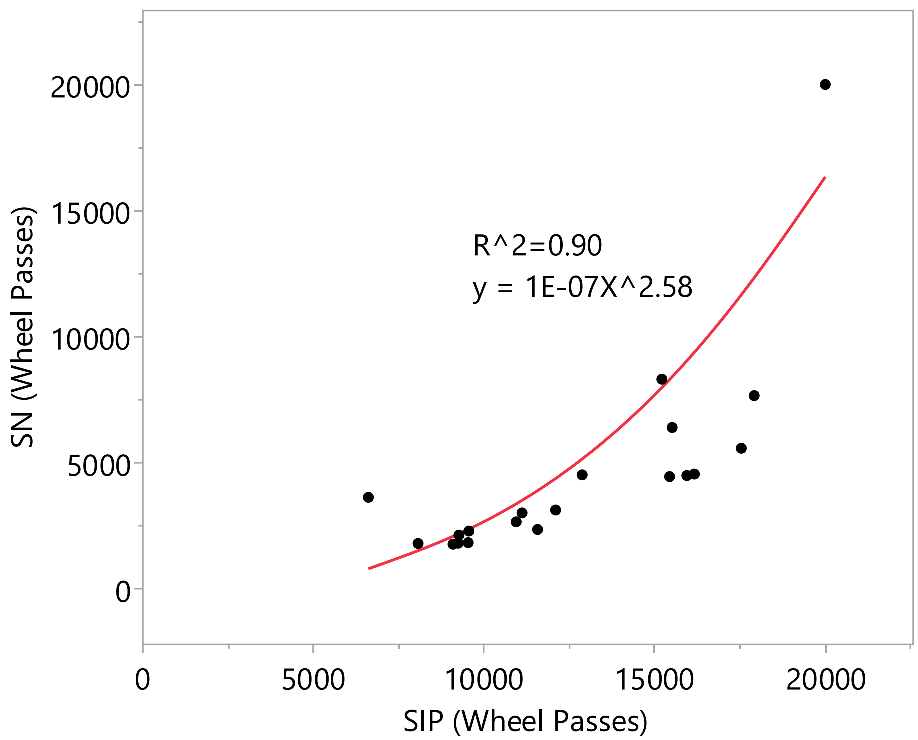


Figure 67. Correlation between SN and SIP.

The SN and SIP results are also presented in Table 19. It was noted that, for those mixes that exhibited stripping, SN occurred at a lower number of wheel passes compared to the SIP. This is because the SN occurs at the onset of stripping, where the curvature changes from negative to positive, while SIP occurs at the onset of significant stripping at the intersection of the CS and SS.

Table 19. SN and SIP for all 33 PMLC Mixes

Mix ID	SIP	SN
M1	No SIP	No SN
M2	No SIP	No SN
M3	15,512	6,403
M4	15,943	4,499
M5	No SIP	No SN
M6	9,554	2,716
M7	No SIP	No SN
M8	11,336	3,133
M9	No SIP	No SN
M10	No SIP	No SN
M11	9,242	1,817
M12	17,821	5,581
M13	No SIP	No SN
M14	16,164	4,558
M15	No SIP	No SN
M16	No SIP	No SN
M17	15,438	4,457
M18	11,115	3,018
M19	10,943	2,664
M20	9,534	1,839
M21	17,916	7,667
M22	11,569	2,361
M23	6,610	3,635
M24	8,066	1,806
M25	No SIP	No SN
M26	9,090	1,777
M27	No SIP	No SN
M28	No SIP	No SN
M29	No SIP	No SN
M30	15,211	8,318
M31	9,264	2,036
M32	12,876	4,527
M33	No SIP	No SN

In Figure 68, the SS was plotted against the SN. A power function was fitted to these parameters, resulting in an R^2 value of 0.61, indicative of a reasonable correlation between the two parameters. This outcome is consistent with the conclusions drawn by (Yin, Chen, et al., 2020). Figure 68 represents only those mixes have SN and SS values.

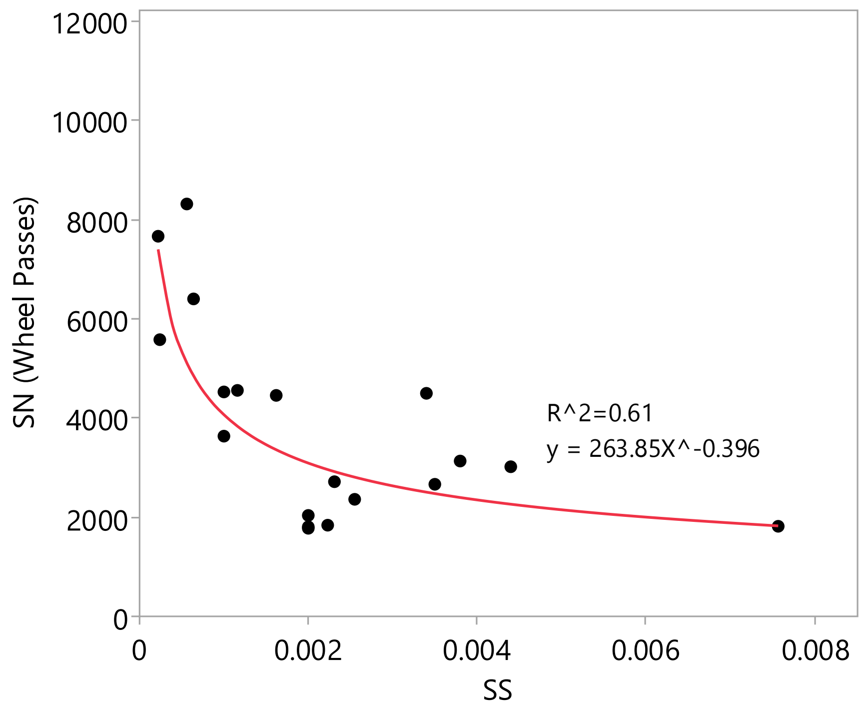


Figure 68. Correlation between SN and SS.

A minimum SN of 2,000 wheel passes was recommended by (Yin, Chen, et al., 2020) and was proposed for BMD implementation by (West et al., 2021). A minimum SN of 2,000 wheel passes is proposed for all Oklahoma asphalt mixes (Olagunju et al., 2025). This criterion was applied to all 33 PMLC mixes, as shown in Figure 69 showing that 87% of the mixes meet this requirement. Additionally, based on the relationship between SN and SIP, a minimum SN of 2,000 passes coincides with a minimum SIP of 10,000 passes.

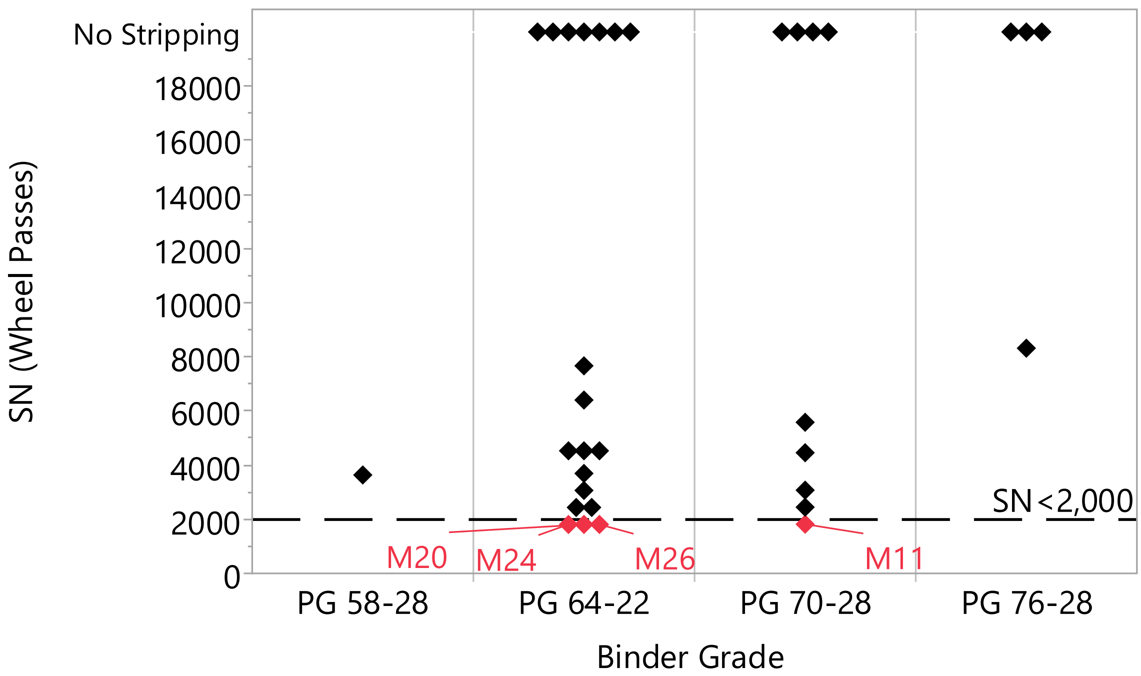


Figure 69. SN versus Binder Grade.

3.2.3 HWTT Rutting Results

3.2.3.1 TRD versus CT_{Index}

The TRD at 10,000 passes was determined for all 33 PMLC mixes and plotted against the CT_{Index} at the short-term aged condition as shown in Figure 70, to examine the balanced performance of the mixes. These results are also given in Table 20.

From the data presented in Figure 70, there was no clear correlation between the STA-CT_{Index} and TRD. Several mixes showed comparable CT_{Index} values with significantly different rut depth. For example, mixes M6, M10, M11, and M12 showed a CT_{Index} of about 80 ±5, whereas their TRD varied significantly ranging between 2.9 mm for M12 and 9.3 mm for M11. Other mixes showed similar TRD with significantly different CTIndex. For example, mixes M1, M5, M10, M17 and M29 showed a TRD at 10,000 passes of about 3.8 to 3.95 mm, but their CT_{Index} values varied significantly ranging between 51 for M1 and 370 for M29.

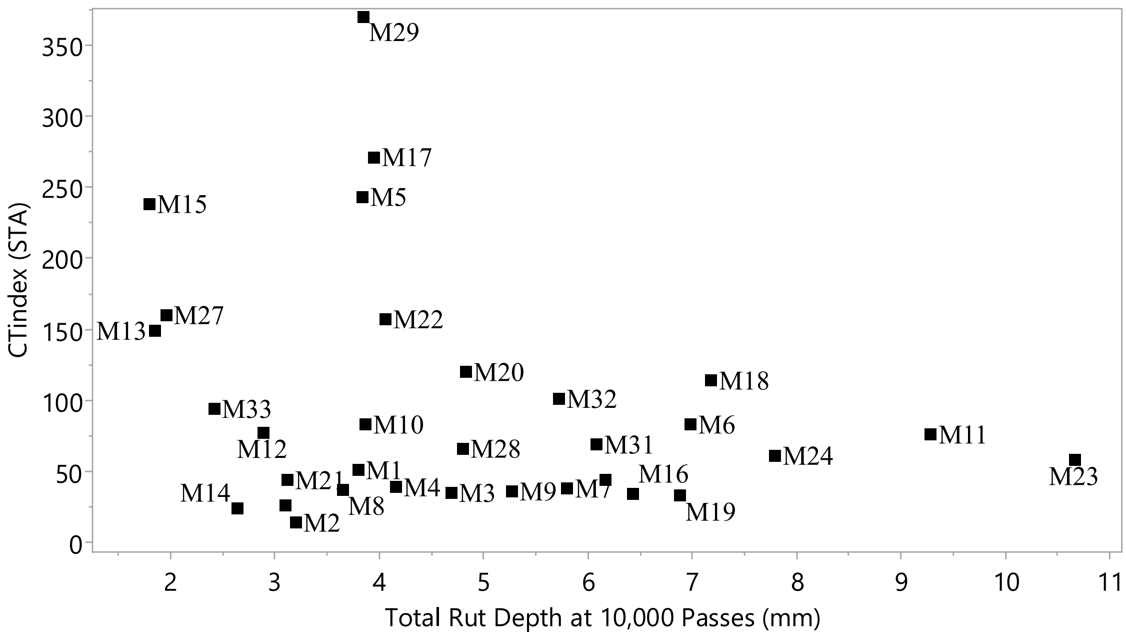


Figure 70. Rut Depth at 10,000 Passes from HWTT versus CT_{Index} from IDEAL-CT at STA.

Table 20. CT_{Index} at STA and LTA Conditions, and TRD at 10,000 Passes.

Mix ID	CT_{Index} (STA)	TRD at 10,000 Passes (mm)
M1	51	3.80
M2	14	3.20
M3	35	4.69
M4	39	4.16
M5	243	3.84
M6	83	6.98
M7	38	5.80
M8	37	3.65
M9	36	5.27
M10	83	3.87
M11	76	9.28
M12	77	2.89
M13	149	1.85
M14	24	2.64
M15	238	1.80
M16	34	6.43
M17	271	3.95
M18	114	7.18
M19	33	6.88
M20	120	4.83
M21	44	3.12
M22	157	4.06
M23	58	10.66
M24	61	7.79
M25	26	3.10
M26	44	6.17
M27	160	1.96
M28	66	4.80
M29	370	3.85
M30	531	6.12
M31	69	6.08
M32	101	5.72
M33	94	2.42

3.2.3.2 Relationship between TRD and Binders' PG

The TRD at 10,000 passes was plotted against the mix type as shown in Figure 71 for mixes using neat binders, i.e. PG58-28 and PG64-22, Figure 72 for mixes using PG70-28 binder, and Figure 73 for mixes using PG76-28 binder. It was noted that the mean rut depth decreased with binder stiffness, where mixes using PG58-28 and PG64-22 had a mean rut depth of 5.3 mm, mixes with PG70-28 had a mean rut depth of 4.16 mm, and mixes with PG76-28 had a mean rut depth of 3.89 mm. No clear trend was noted between the mix type and TRD.

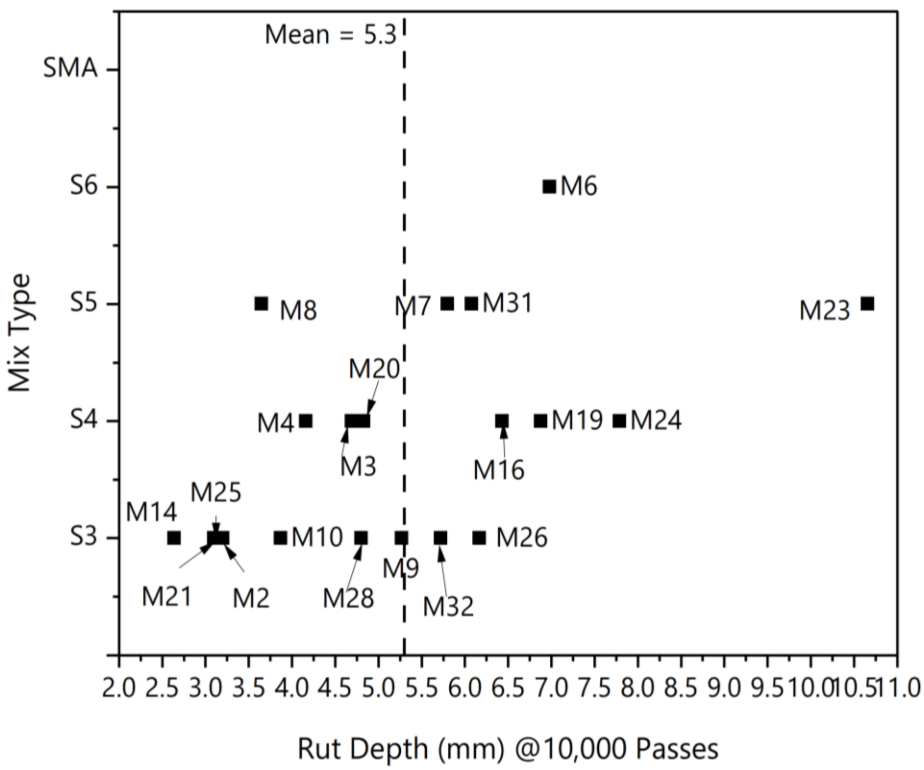


Figure 71. TRD at 10,000 Passes from HWTT versus Mix Type for PG58-28 and PG64-22.

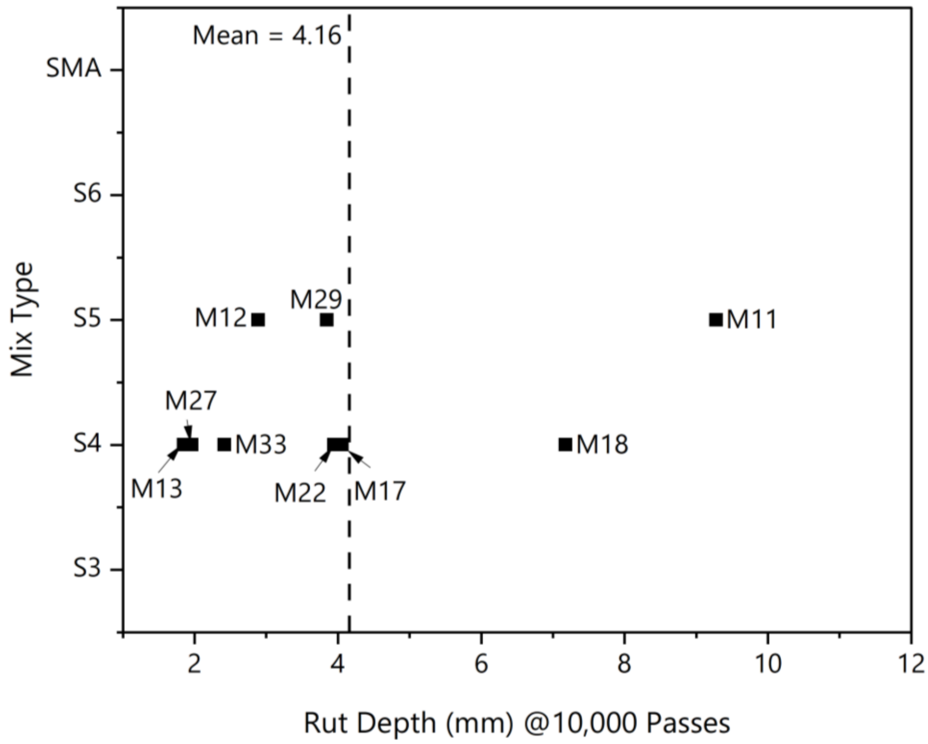


Figure 72. Rut Depth at 10,000 passes from HWTT versus Mix Type for PG70-28.

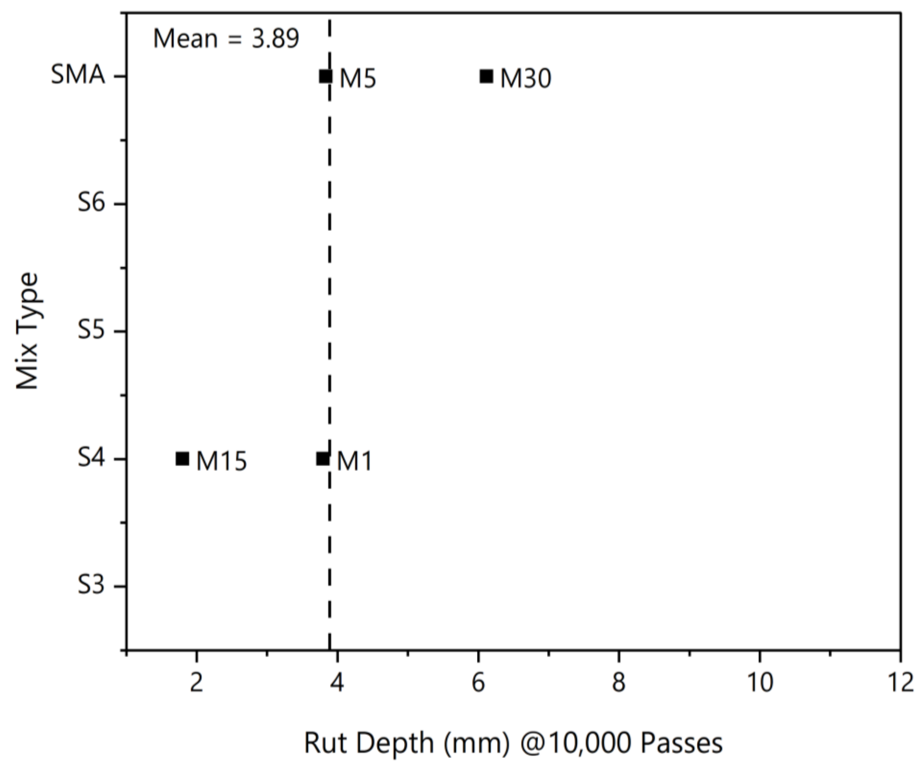


Figure 73. Rut Depth at 10,000 passes from HWTT versus Mix Type for PG76-28.

3.2.3.3 Comparison between TRD for LMLC and PMLC

The TRD results for the PMLC and LMLC were compared as shown in Table 21. The PMLC specimens were tested at the Ingevity lab in Tulsa, Oklahoma, using a SmarTracker device from InstroTek, while the LMLC specimens were tested by ODOT during the mix design process using a WheelTracker device from Troxler. The TRD values presented in Table 21 were determined at 10,000 passes for mixes containing PG58-28 and PG64-22 binders, 15,000 passes for mixes containing PG70-28 binder, and 20,000 passes for mixes containing PG76-28 binder, in accordance with ODOT specifications. As shown in Table 21, all mixes met the ODOT rut criterion of a maximum TRD of 12.5 mm at the specified number of wheel passes during production, except for mixes M11 and M18. Notably, both M11 and M18 are WMA mixes utilizing foamed asphalt technology, produced at the same asphalt plant. These were the only two mixes, out of the 33 sampled mixes, that were produced at this plant. WMA mixes that were produced at other asphalt plants such as mixes M7, M9, M15, M17, M25 and M26 satisfied the ODOT rut-depth criterion during production. Additionally, the LMLC TRD for mix M29 was determined using the APA and therefore was excluded from further analysis.

Table 21. Summary of the LMLC and PLMC TRD Results.

Mix ID	Specification's Wheel Passes	LMLC TRD (mm)	PMLC TRD (mm)
M1	20,000	9.81	4.82
M2	10,000	1.88	3.2
M3	10,000	1.48	4.69
M4	10,000	4.27	4.16
M5	20,000	1.57	4.74
M6	10,000	7.68	6.98
M7	10,000	3.36	6.28
M8	10,000	5.97	3.65
M9	10,000	0.26	5.27
M10	10,000	1.34	3.87
M11	15,000	1.62	16.85**
M12	15,000	4.85	3.18
M13	15,000	1.93	2.09
M14	10,000	5.97	2.64
M15	20,000	2.75	2.32
M16	10,000	7.71	6.43
M17	15,000	2.16	5.55
M18	15,000	2.71	15.74**
M19	10,000	12.15	6.88
M20	10,000	3.64	4.83
M21	10,000	3.95	3.08
M22	15,000	8.49	12.24
M23	10,000	3.25	10.53
M24	10,000	1.37	7.79
M25	10,000	1.59	3.1
M26	10,000	2.43	6.16
M27	15,000	1.94	2.19
M28	10,000	1.34	4.8
M29	15,000	1.24*	4.89
M30	20,000	8.1	10.26
M31	10,000	2.35	6.08
M32	10,000	8.88	5.72
M33	15,000	5.43	2.93

**Rut depth was determined using the APA.*

***Mixes failed ODOT rut-depth criterion during production.*

The PMLC TRD was plotted against the LMLC TRD as shown in Figure 74 for mixes containing PG58-28 and PG64-22 binders, Figure 75 for mixes containing PG70-28 binder, and Figure 76 for mixes containing PG76-28 binder. No clear trend was observed between the PMLC TRD and LMLC TRD across the various mixes. In some cases, the PMLC TRD was higher than the LMLC TRD, indicating a reduction in rutting resistance during production. whereas in other instances the LMLC TRD was higher than the PMLC TRD, indicating an increase in rutting resistance during production.

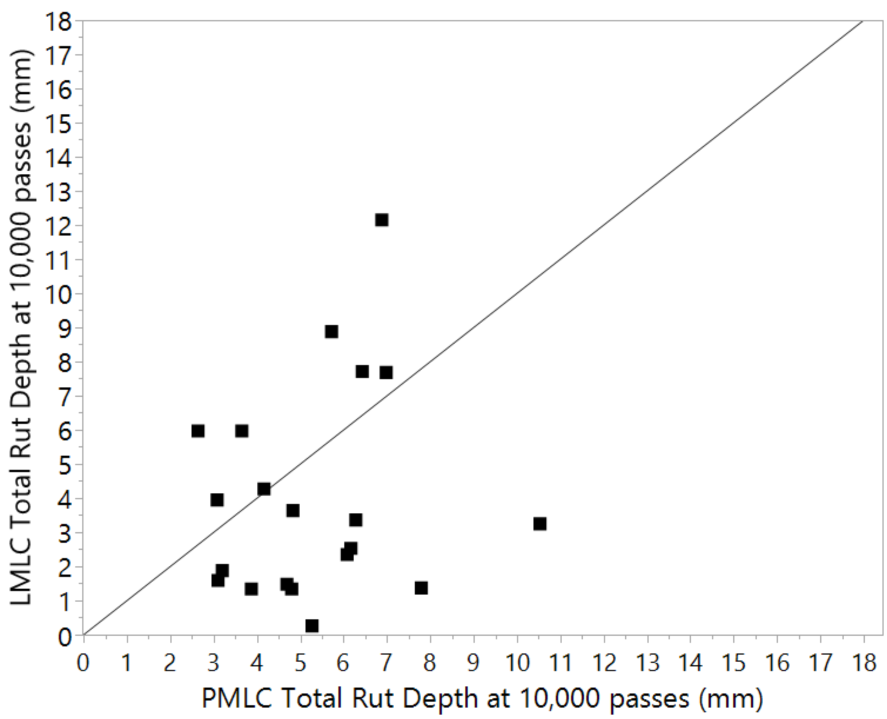


Figure 74. PMLC vs. LMLC TRD at 10,000 Passes for Mixes with PG64-22 and PG58-28.

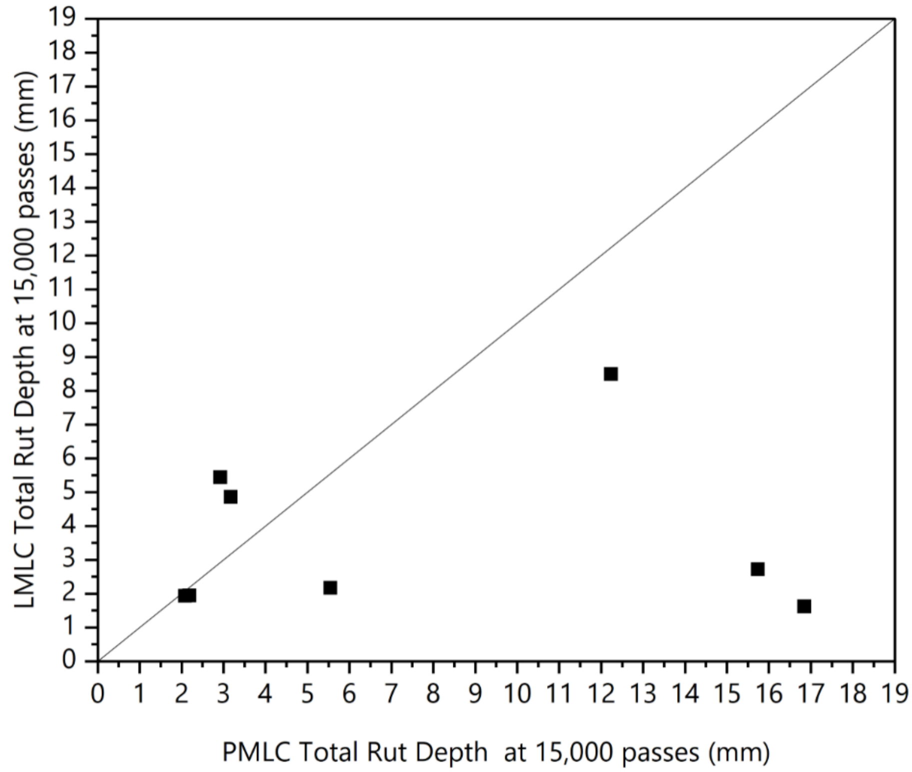


Figure 75. PMLC vs. LMLC TRD at 15,000 Passes for Mixes with PG70-28.

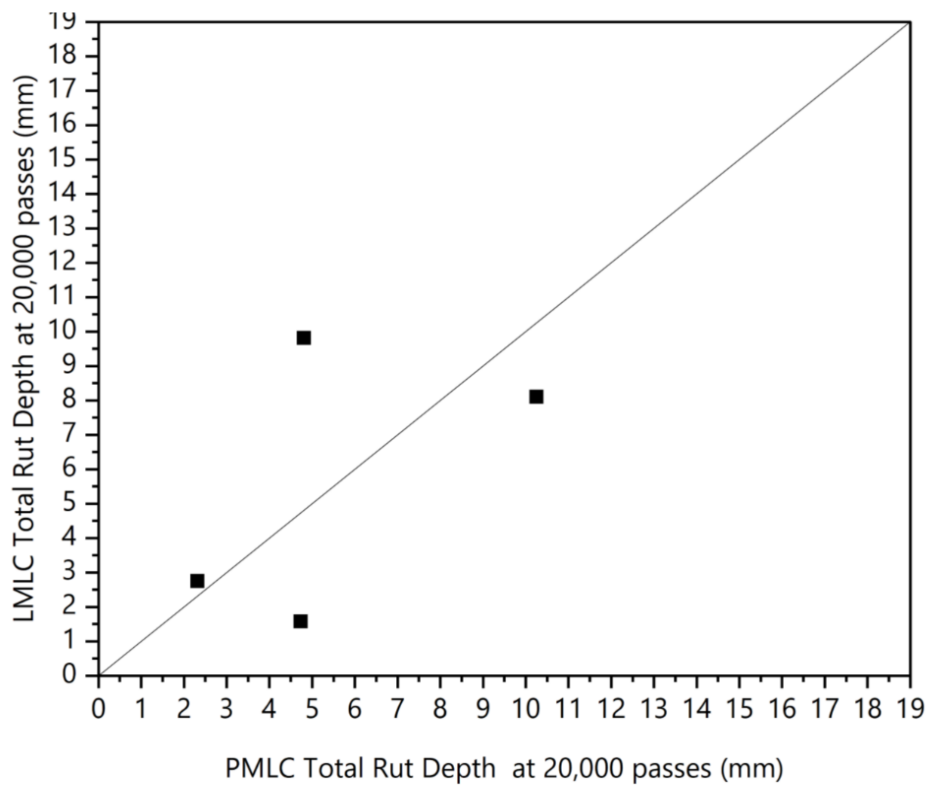


Figure 76. PMLC vs LMLC TRD at 20,000 Passes for Mixes with PG76-28.

3.2.3.4 Corrected Rut Depth (CRD)

The CRD was calculated using the NCAT method at different numbers of wheel passes. The analysis in this section is limited to PMLC only since the rut depth curves for the LMLC were not available, so it was not possible to calculate their CRD. Table 22 shows a summary of TRD and CRD for mixes containing PG58-28 and PG64-24, whereas Table 23 and Table 24 show TRD and CRD for mixes containing PG70-28 and PG76-28, respectively. For all mixes, CRD_{10K}, CRD_{15K}, and CRD_{20K}, representing CRD at 10,000, 15,000, and 20,000 passes, respectively were calculated. TRD_{10K}, TRD_{15K}, and TRD_{20K}, denotes the TRD at 10,000, 15,000, and 20,000 passes, respectively. TRD was shown at 10,000 passes for mixes containing PG58-28 and PG64-22, at 10,000 and 15,000 passes for mixes containing PG70-28, and at 10,000, 15,000, and 20,000 passes for mixes with PG76-28.

Table 22. CRD and TRD for Mixes Containing PG58-28 and PG64-22 Binders.

Mix ID	TRD_{10K} (mm)	CRD_{10K} (mm)	CRD_{15K} (mm)	CRD_{20K} (mm)	% TRD due to Rutting [(CRD_{10K}/TRD_{10K}) *100]
M2	3.20	3.20	3.70	4.23	100
M3	4.70	4.70	6.22	6.56	100
M4	4.16	4.07	4.52	4.85	98
M6	6.98	5.16	5.93	6.52	74
M7	5.80	5.80	6.47	6.91	100
M8	3.65	3.12	3.52	3.82	85
M9	5.27	5.27	6.81	8.43	100
M10	3.87	3.87	4.32	4.83	100
M14	2.80	2.13	2.32	2.46	76
M16	6.43	6.43	8.61	11.56	100
M19	6.88	4.28	5.01	5.56	62
M20	5.30	2.70	3.03	3.28	51
M21	3.22	3.15	3.62	3.99	98
M23	10.66	7.85	9.36	10.55	74
M24	7.79	2.64	2.97	3.21	34
M25	3.10	3.10	3.46	3.63	100
M26	6.17	2.83	3.30	3.67	46
M28	4.80	4.80	5.30	5.87	100
M31	6.08	2.51	2.60	2.65	41
M32	5.72	4.96	5.38	5.66	87

Table 23. CRD and TRD for Mixes Containing PG70-28 Binder.

Mix ID	TRD_{10K} (mm)	TRD_{15K} (mm)	CRD_{10K} (mm)	CRD_{15K} (mm)	CRD_{20K} (mm)	% TRD due to Rutting [(CRD_{15K}/TRD_{15K}) *100]
M11	9.28	47.1	2.46	2.73	2.93	6
M12	2.89	3.25	2.85	3.13	3.49	96
M13	1.85	2.09	1.85	2.09	2.27	100
M17	3.95	5.55	2.95	3.00	3.03	54
M18	7.18	23.34	5.88	6.86	7.62	29
M22	4.06	12.24	2.52	2.79	2.98	23
M27	1.96	2.19	1.96	2.19	2.38	100
M29	3.85	4.89	3.85	4.89	6.15	100
M33	2.42	2.93	2.42	2.93	3.52	100

Table 24. CRD and TRD for Mixes Containing PG76-28 Binder.

Mix ID	TRD_{10K} (mm)	TRD_{15K} (mm)	TRD_{20K} (mm)	CRD_{10K} (mm)	CRD_{15K} (mm)	CRD_{20K} (mm)	% TRD due to Rutting [(CRD_{20K}/TRD_{20K}) *100]
M1	3.80	4.55	4.82	3.80	4.55	4.82	100
M5	3.84	4.33	4.74	3.84	4.33	4.74	100
M15	1.80	2.08	2.32	1.80	2.08	2.32	100
M30	6.12	8.05	10.68	5.77	6.89	7.76	73

The CRD isolates the effect of rutting and stripping. The impact of stripping is reflected in the difference between TRD and CRD at the same number of wheel passes, which is influenced by the SIP and SS. The ratio of CRD to TRD, at a given number of wheel passes, provides a measure of the rutting versus stripping. The CRD/TRD was calculated for all mixes as shown in Table 22, Table 23, and Table 24. A CRD/TRD percentage of 100 indicates that the entire measured rut depth was due to rutting, denoting no effect of stripping on rutting. A lower CRD/TRD percentage indicates that the measured rut depth is primarily affected by stripping. For example, mix M11 showed a CRD/TRD of 6% indicating that 94% of the measured rut depth was due to stripping.

Figure 77, Figure 78, and Figure 79 show a plot of the data presented in Table 22, Table 23, and Table 24, respectively. All mixes that fall on the equality line exhibited no stripping at the specified number of passes. However, several mixes displayed stripping, as indicated by CRD being lower than TRD. Notably, none of the mixes containing PG76-

28 binder showed any signs of stripping at 20,000 passes, except for mix M30. For mixes M11 and M18, the test stopped before reaching 15,000 passes due to measured rut depth exceeding 12.5 mm. To determine the TRD_{15K} for mixes M11 and M18, the rut depth was extrapolated by extending the CS of the rut depth curve to estimate the rut depth at the desired number of passes. This extrapolation involved projecting the trend of the rut depth curve to predict the rut depth point at the specified number of wheel passes. For instance, one can use Equation 8 to determine the rut depth at 15,000 passes with the CS and rut depth values at test completion available.

$$RD_{15K} = [(15000 - LC_{TC}) * SS] + RD_{TC} \qquad (8)$$

Where RD_{15K} = Rut depth at 15,000 passes (mm), LC_{TC} = Load cycles at test completion, SS = Stripping Slope, and RD_{TC} = Rut depth at test completion (mm).

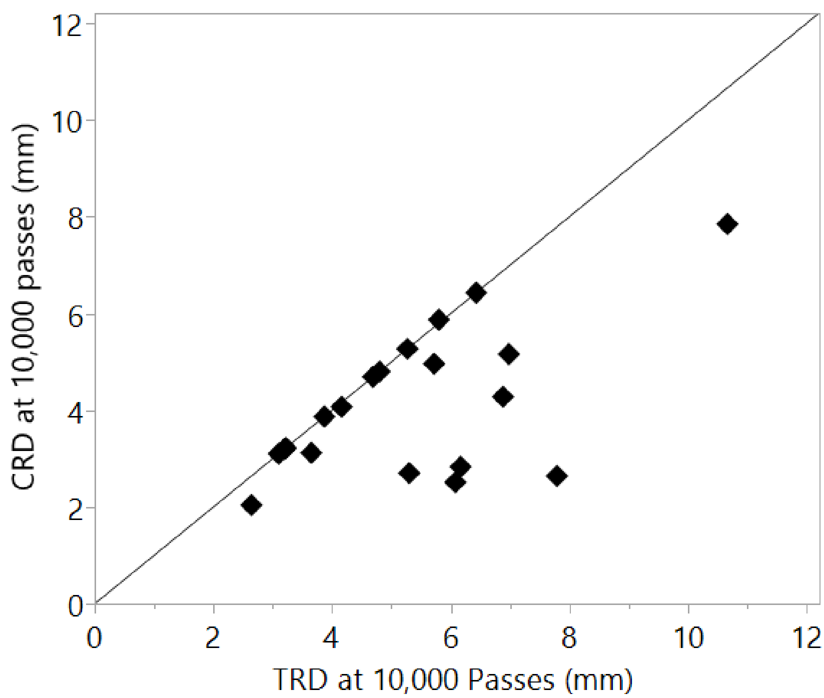


Figure 77. TRD and CRD at 10,000 Passes for Mixes with PG58-28 and PG64-22.

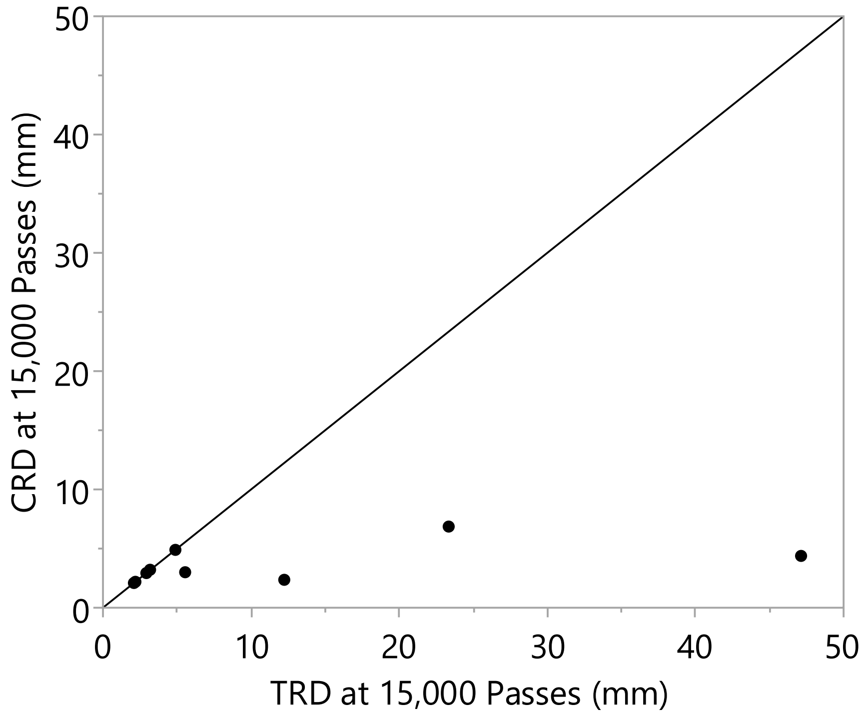


Figure 78. TRD and CRD at 15,000 Passes for Mixes with PG70-28.

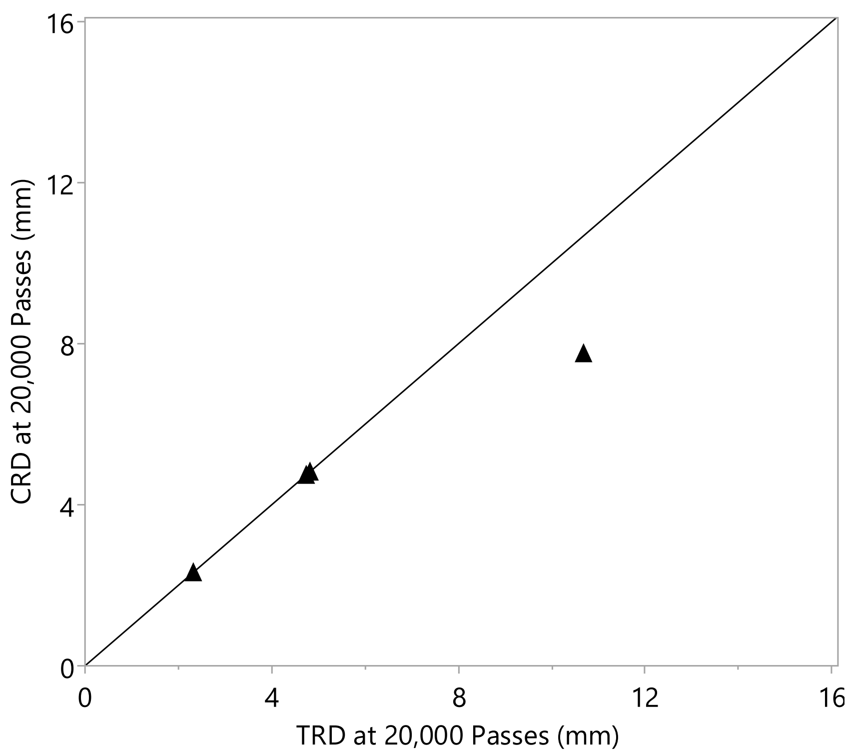


Figure 79. TRD and CRD at 20,000 Passes for Mixes with PG76-28.

3.2.3.5 Effect of Insoluble Residue on Rutting Performance

The insoluble residue in asphalt mixes measures the quantity of aggregate material that remains undissolved in hydrochloric acid. A high percentage of insoluble residue indicates the presence of higher amounts of siliceous minerals, which are considered more polish-resistant than carbonate materials and provide improved skid resistance. The test is conducted according to OHD L-25 and is used as an indicator of the appropriateness of the aggregate to be used in surface mixes. According to ODOT specifications, coarse aggregate used in surface mixes must meet a minimum insoluble residue requirement, which varies based on the binder's PG. For mixes containing PG64-22 binder, a minimum insoluble residue of 30% is required, while mixes with PG70-28 or PG76-28 binders must have at least 40% insoluble residue. The insoluble residue is calculated for the combined coarse aggregate in the mix and reported in the mix design.

Figure 80 presents a plot of the insoluble residue for all mixes. The data indicates that all mixes containing PG70-28, and PG76-28 binders have an insoluble residue greater than 40%, meeting the ODOT criteria for surface mixes. In contrast, most mixes with PG64-22 binder exhibit an insoluble residue below 30%.

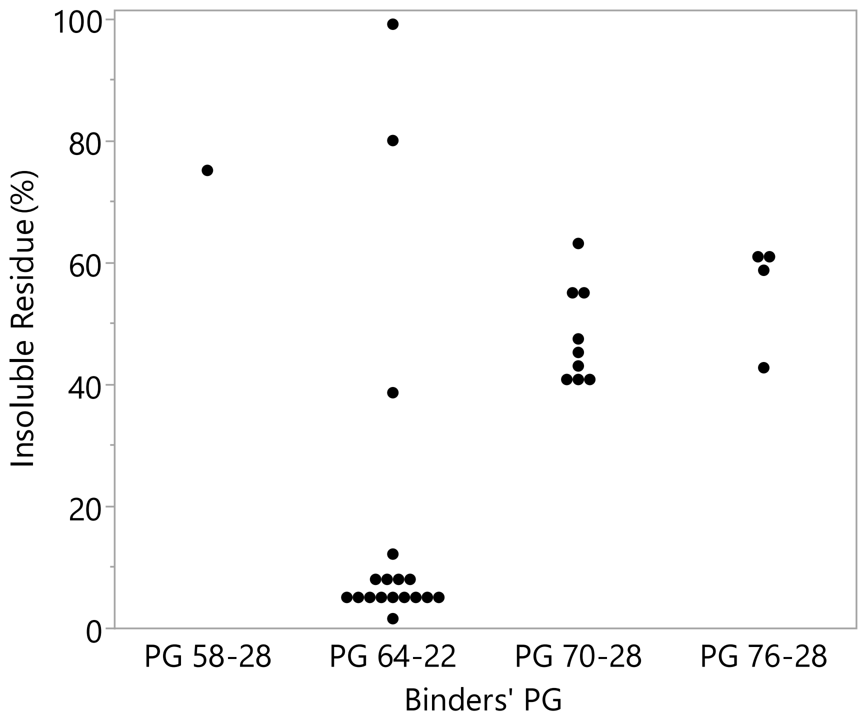


Figure 80. Binder's PG versus Insoluble Residue.

A low insoluble residue may be an indication of poor resistance to polishing and a weak aggregate structure. To evaluate the impact of insoluble residue on the CRD, the CRD_{10k} was plotted against the insoluble residue as shown in Figure 81. The results show that the insoluble residue percentage did not have an impact on rutting resistance as measured by the CRD_{10k}.

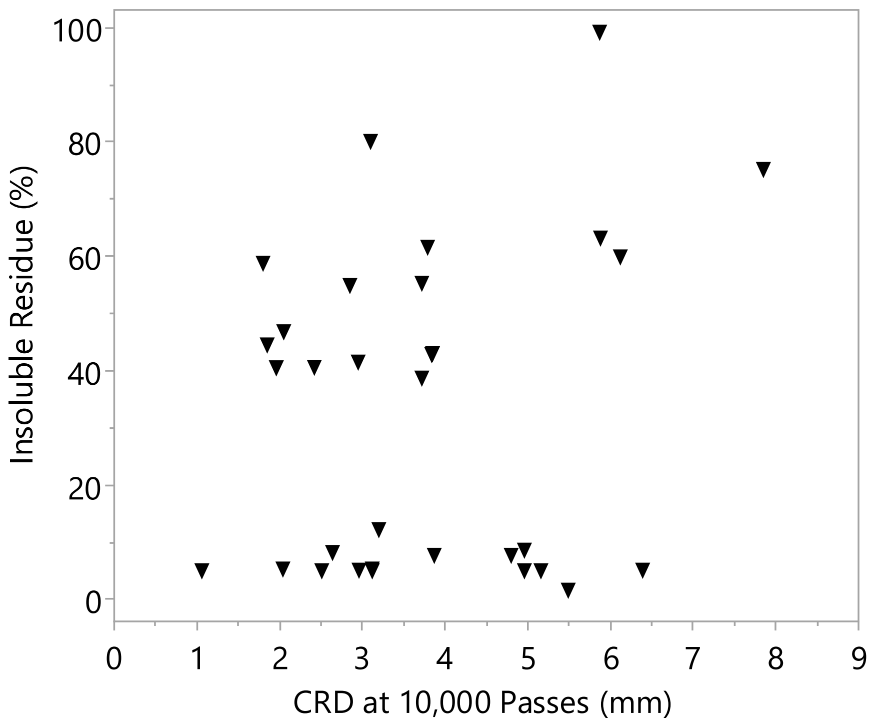


Figure 81. Insoluble Residue versus $CRD_{10K.}$

It is important to note that the ODOT specification assumes that the RAP aggregate has zero insoluble residue. As a result, for mixes that contain RAP, the insoluble residue of the combined aggregate is typically lower than that of virgin mixes due to the assumption of zero insoluble residue for RAP. Figure 82, illustrates the effect of RAP on the insoluble residue where most of the mixes with 20 and 25% RAP exhibit very low insoluble residue due to the assumption that RAP has zero insoluble residue.

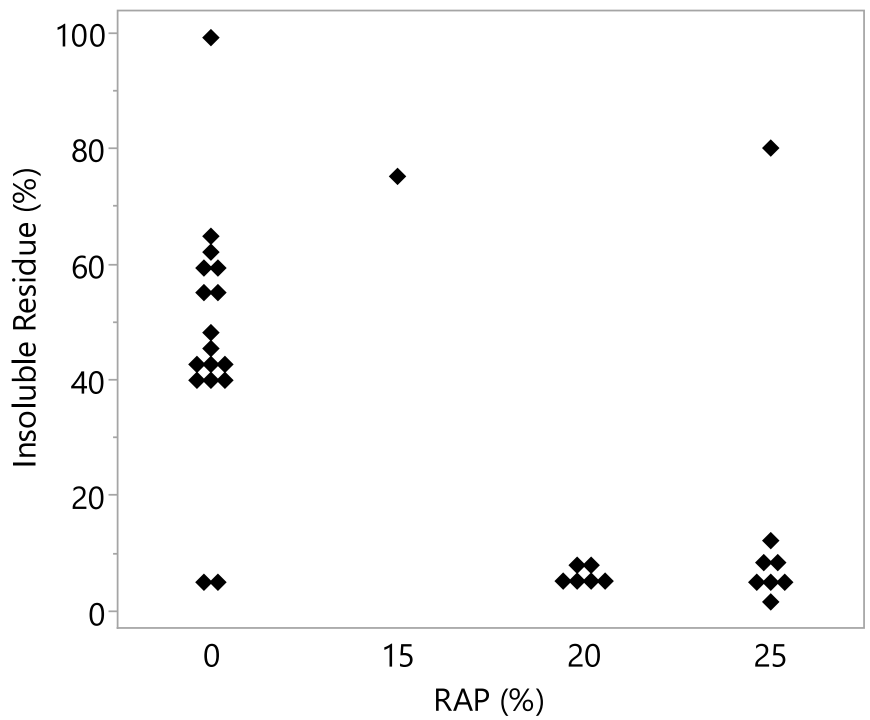


Figure 82. RAP Content versus Insoluble Residue.

From the above, it can be concluded that the insoluble residue cannot be relied upon as an indication of rutting resistance. In most cases, RAP leads to an increased rutting resistance and a reduction in the insoluble residue, due to the assumption of zero insoluble residue for RAP.

3.2.3.6 Proposed Rutting Criteria Using CRD

Based on the above discussion, the CRD parameter seems to offer a more accurate indication of rutting compared to TRD, as it is unaffected by stripping. A recent study conducted by NCAT using WisDOT mixes recommended maximum thresholds for CRD_{20k} based on mix type and traffic level (West et al., 2021). The recommended maximum CRD_{20K} is 8mm for low traffic mixes, 7mm for medium traffic mixes, and 6mm for high traffic and SMA mixes.

Considering that Oklahoma uses PG64-22, PG70-28, and PG76-28, binders for low-traffic, medium-traffic, and high-traffic mixes, respectively, the WisDOT recommended CRD thresholds were used to evaluate the Oklahoma mixes based on the binder's PG according to the following: $CRD_{20K} \leq$ 8mm for mixes with PG58-28 and PG64-22, $CRD_{20K} \leq$ 7mm for mixes with PG70-28, and $CRD_{20K} \leq$ 6mm for mixes with PG76-28 mixes, as given in Table 25.

Table 25. Proposed Rut Criteria Based on Maximum CRD at 20,000 Passes.

Binder's PG	**Maximum CRD at 20,000 passes (mm)**
PG58-28	8
PG64-22	8
PG70-28	7
PG76-28	6

The mixes were evaluated using the proposed criterion as shown in Figure 83. The mixes marked in red did not meet the proposed criterion. It is noted that three mixes (i.e. M9, M16, and M23) failed to meet the proposed criterion for PG58-28 and PG64-22 of $CRD_{20K} \leq$ 8mm. Mix M18 did not meet the PG70-28 criterion of $CRD_{20K} \leq$ 7mm and mix M30 did not meet the PG76-28 criterion of $CRD_{20K} \leq$ 6mm.

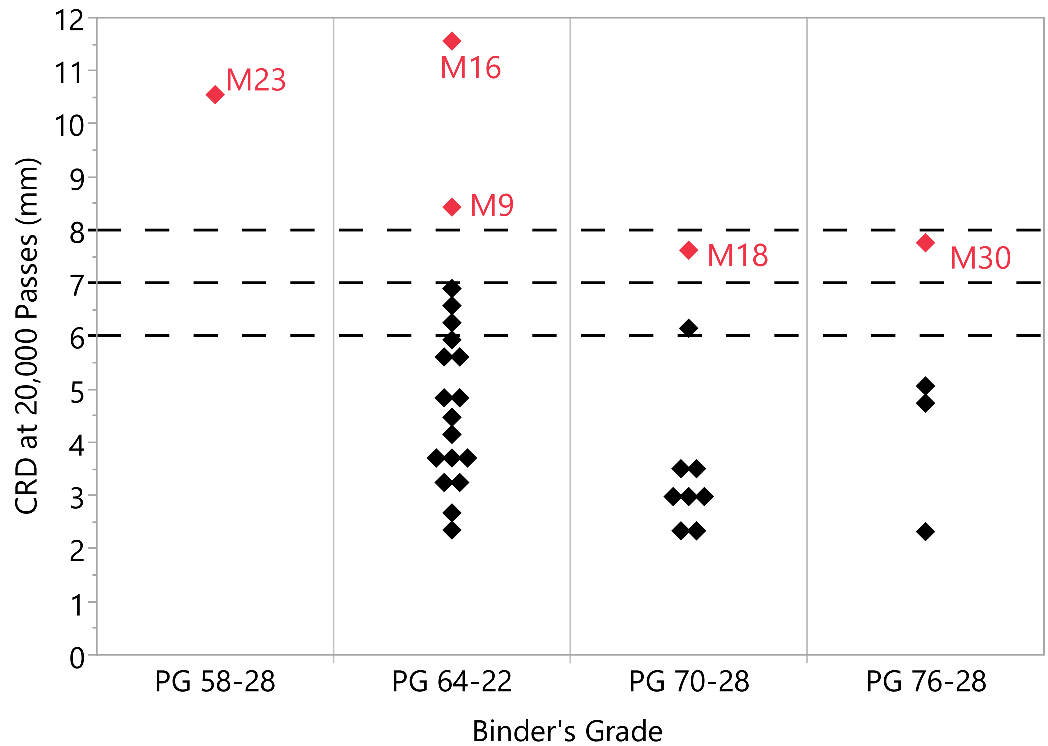


Figure 83. CRD_{20K} Rutting Criteria.

An alternative criterion can be established based on selecting a maximum number of wheel passes at a constant CRD depth. The criterion was developed using 10,000 passes for PG58-28 and PG64-22, 15,000 passes for PG70-28, and 20,000 passes for PG76-28. To determine the alternative criterion, the relationship between the CRD_{20k} and CRD_{10k} was plotted in Figure 84 for mixes with PG64-22, and the relationship between the CRD_{20k} and CRD_{15k} was plotted in Figure 85 for mixes with PG70-28.

Using Figure 84, a CRD_{20k} of 8mm is shown to be equivalent to a CRD_{10k} of 6mm, and using Figure 85, a CRD_{20k} of 7mm is shown to be equivalent to a CRD_{15k} of 6mm. Therefore, a maximum CRD of 6mm at 10,000, 15,000, and 20,000 wheel passes, for mixes with PG64-22, PG70-28, and PG76-28, respectively, was proposed. A summary of the alternative criterion is shown in Table 26.

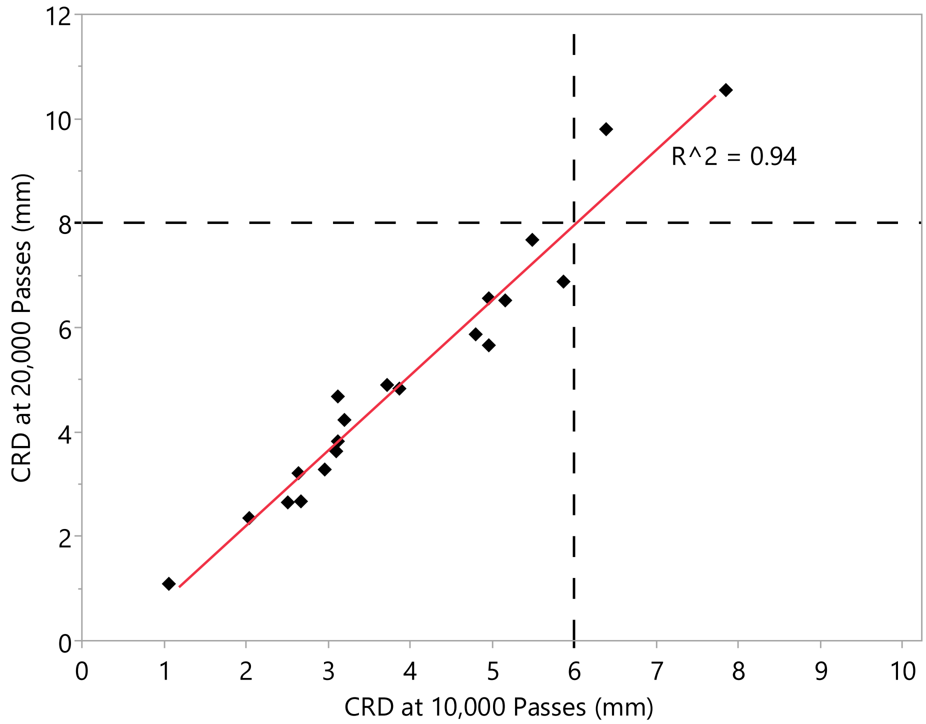


Figure 84. Correlation between the CRD20K and CRD10K for Mixes with PG58-28 and PG64-22.

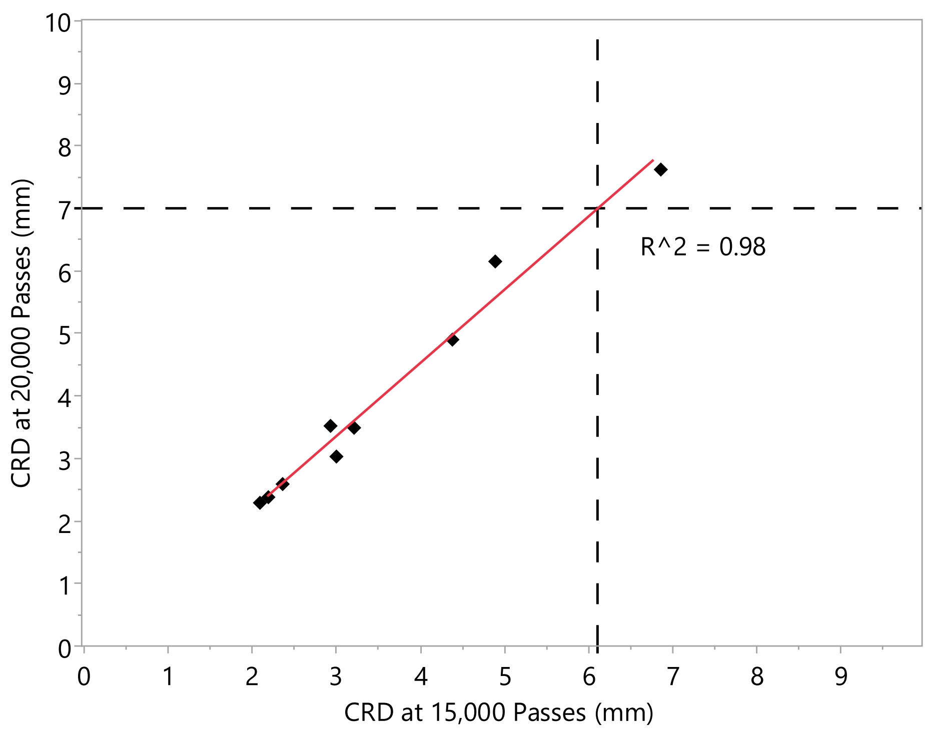


Figure 85. Correlation between the CRD_{20K} and CRD_{15K} for Mixes with PG70-28.

Table 26. Proposed Rut Criteria Based on Minimum Wheel Passes at a Maximum CRD of 6mm.

Binder's PG	Minimum Wheel Passes at a Maximum CRD of 6mm
PG58-28	10,000
PG64-22	10,000
PG70-28	15,000
PG76-28	20,000

3.2.4 Summary of HWTT Rutting and Moisture

Rutting and moisture susceptibility are significant performance criteria that can be effectively analyzed with the HWTT results. Figure 86 integrates both the SN criteria (i.e. minimum SN of 2,000-wheel passes) and the CRD criteria (i.e. CRD_{20K} ≤ 6mm for PG76-28 binders, CRD_{20K} ≤7mm for PG70-28 binders, and CRD_{20K} ≤8mm for PG58-28 and 64-22 binders). Notably, the data reveals that mixes M11, M20, M24, and M26 did not meet the SN criteria but showed strong rutting performance. On the other hand, mixes M9, M16, M18, M23, and M30 satisfied moisture susceptibility criterion but fell short in rutting performance. This highlights the need for separate performance criteria to accurately assess both rutting and moisture susceptibility. To improve moisture resistance in mixes that perform well in rutting but fail moisture tests, anti-stripping agents could be used. Therefore, mixes M11, M20, M24 and M26, which demonstrated good rutting performance, would be good candidates for use of an anti-stripping agent.

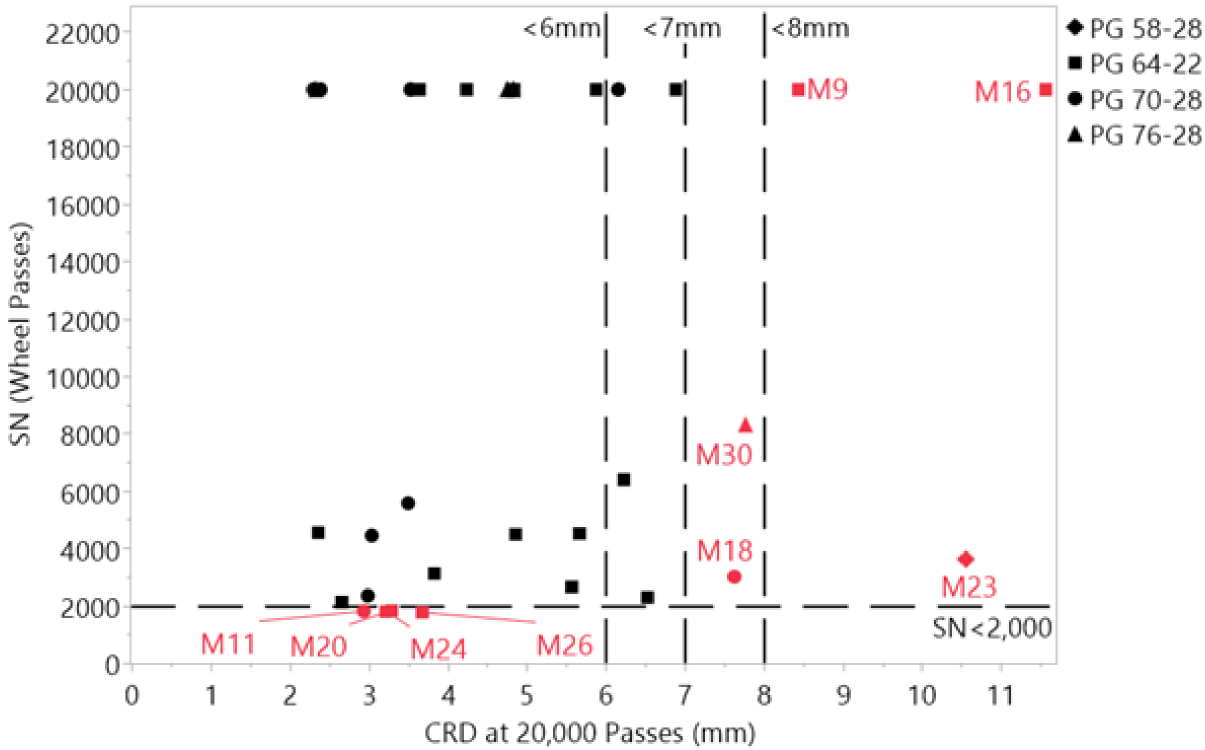


Figure 86. Combined CRD and SN Criteria.

3.2.5 IDT-HT Results

IDT-HT is a quick and simple test that can be used to assess the rutting resistance of asphalt mixes during production. Various studies have indicated a correlation between the IDT-HT strength and the rutting parameters obtained from the HWTT. In this section, we explore the correlation between IDT-HT strength and the CRD from the HWTT. Due to limited material availability, the IDT-HT was conducted for eighteen plant-produced mixes only and their results compared with the corresponding CRD results from the HWTT, to determine a rutting criterion based on IDT-HT tensile strength. The IDT-HT results of the tested mixes are presented in Table 27.

Table 27. IDT-HT Results for Selected PMLC Mixes

Mix ID	PG	Tensile Strength (kPa)	COV (%)
M1	PG76-28	33.9	8.1
M2	PG64-22	65.3	8.1
M3	PG64-22	54.1	0.8
M5	PG76-28	44.5	6.7
M7	PG64-22	35.0	10.9
M8	PG64-22	37.7	6.6
M9	PG64-22	46.5	5.6
M21	PG64-22	57.9	13.2
M23	PG58-28	39.1	7.0
M24	PG64-22	57.0	3.1
M25	PG64-22	78.7	2.9
M27	PG70-28	40.6	0.1
M28	PG64-22	36.6	2.8
M29	PG70-28	42.6	8.6
M30	PG76-28	31.7	9.8
M31	PG64-22	54.7	7.3
M32	PG64-22	56.2	7.3
M33	PG70-28	66.7	8.9

Figure 87 shows a box plot illustrating the distribution of IDT-HT strength for neat binders including PG64-22 and PG58-28 and modified binders such as PG70-28 and PG 76-28. The top whisker denotes the highest value, while the bottom whisker indicates the lowest value. The shaded area of the box represents the interquartile range, with the lower edge marking the 1st quartile and the upper edge marking the 3rd quartile. Within the shaded area, the median is shown as a straight line, and the mean value is represented by a small circle.

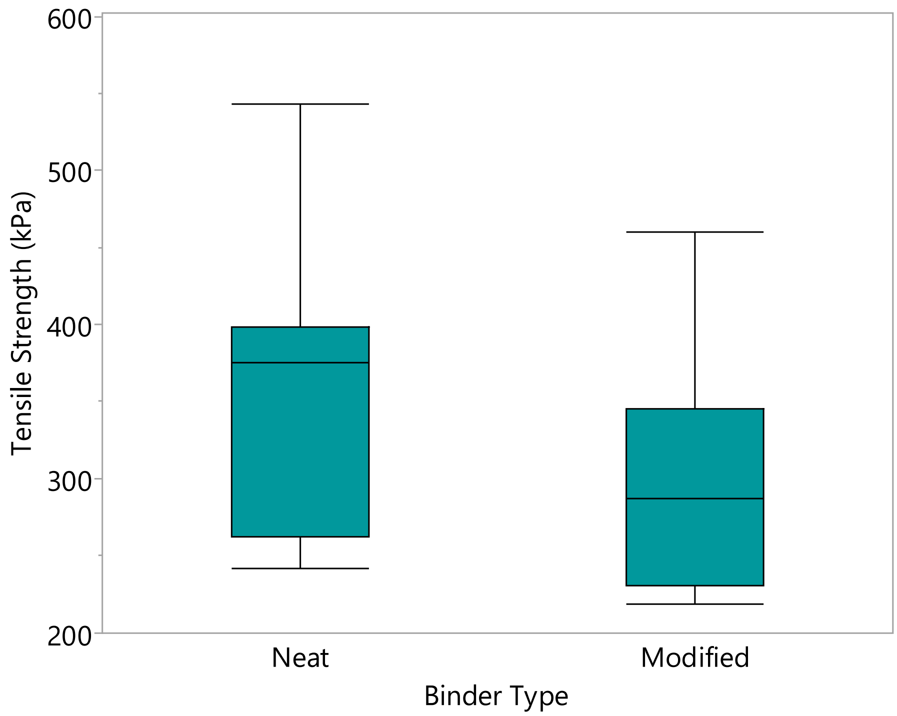


Figure 87. Boxplot showing the Variation of IDT-HT Tensile Strength for Neat and Modified Binders

As depicted in Figure 87 the mixes with neat binders exhibited a higher median IDT-HT strength compared to the mixes with modified binders. However, the boxplots reveal overlap between the two binder types.

A one-way ANOVA was carried out resulting in a p-value of 0.22 which indicates that the binders' PG is not statistically significant. In other words, there was no significant difference between the mean IDT-HT strength values for neat and modified binders.

From Figure 88, the relationship between the CRD_{20K} on the x-axis and the IDT-HT strength on the y-axis for all mixes was evaluated. There was a negative trend between the IDT-HT strength and the CRD_{20k} implying that there was an inversely proportional relationship between the parameters. Overall, there was a weak R^2 between the two parameters from the power relationship.

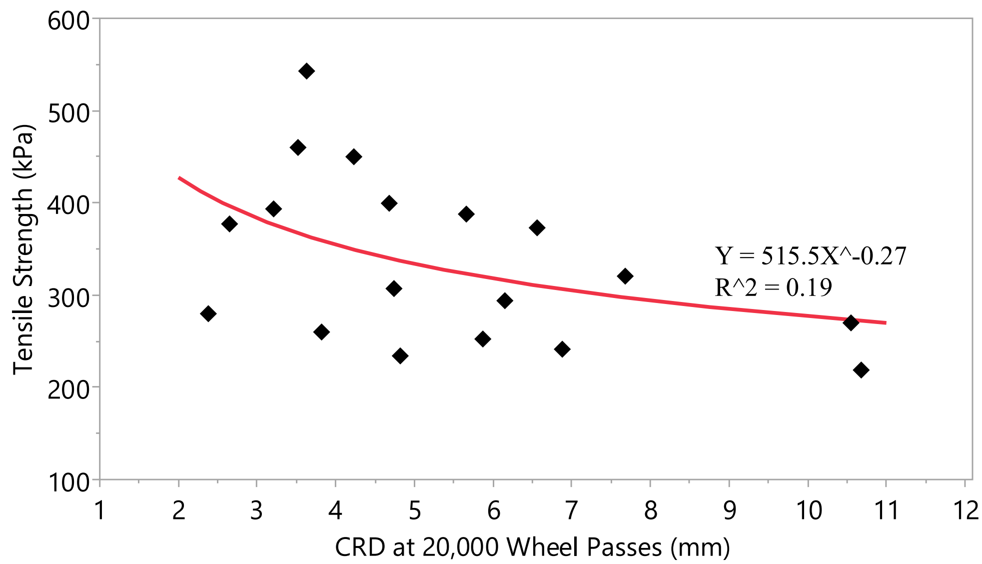


Figure 88. CRD_{20K} versus Tensile Strength (kPa) for all mixes.

Based on the correlation in the figure above, a minimum IDT-HT strength of 240kPa (35psi) is proposed as an acceptance threshold during production for all binder types. The proposed IDT-HT threshold satisfies a maximum CRD_{20K} of 8mm for all binder types. Other states have also proposed IDT-HT acceptance thresholds. For instance, (Yin & West, 2021) carried out a study for ALDOT asphalt mixes and proposed a minimum IDT-HT strength of 138Ka (20psi) as a performance test criterion for rutting tested at 50°C. (Bennert et al., 2018) also carried out a study on NJDOT asphalt mixes and recommended minimum IDT-HT strengths of 172kPa (25psi), 207kPa (30psi), and 324kPa (47psi), based on the asphalt mix type for quality control testing with testing performed at 44°C.

Figure 89 shows the relationship between CRD_{20K} on the x-axis and IDT-HT strength on the y-axis for both neat and modified binders. Neat binders include PG64-22 and PG58-28 mixes while modified binders include PG70-28 and 76-28 mixes. Overall, the modified binders resulted in lower strength. However, the R^2 values were low and there was a big overlap between the results of the modified and neat binders, which didn't justify developing different thresholds based on binder type.

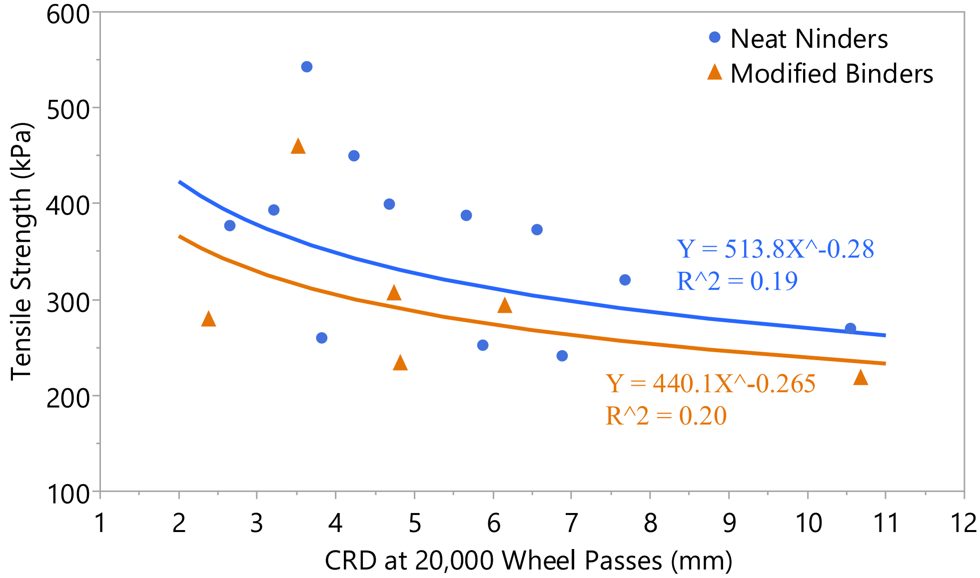


Figure 89. CRD_{20K} versus Tensile Strength.

3.2.6 Effect of Freeze-thaw Damage on Rutting and stripping resistance

3.2.6.1 HWTT before and after Freeze-thaw Damage

The objective of this task was to assess the effect of freeze-thaw damage using AASHTO T 283 on the HWTT and IDT-HT results. To this end, nine plant-produced mixes were selected, i.e. mixes M7, M9, M10, M15, M21, M23, M24, M29, and M33. This section presents and discusses the results from HWTT and IDT-HT conducted on the selected mixes before and after AASHTO T283 moisture conditioning. The findings are discussed with respect to key performance indicators such as TRD, Rutting Resistance Index (RRI) and IDT-HT tensile strength, for rutting, and SS, SN, and SIP for moisture.

Figure 90, Figure 91, and Figure 92 illustrates the HWTT rut depth curves of all nine mixes before and after AASHTO T-283 conditioning. The mixes were grouped based on mix type, i.e. S3, S4, and S5, in Figure 90, Figure 91, and Figure 92, respectively, to facilitate comparison between mixes.

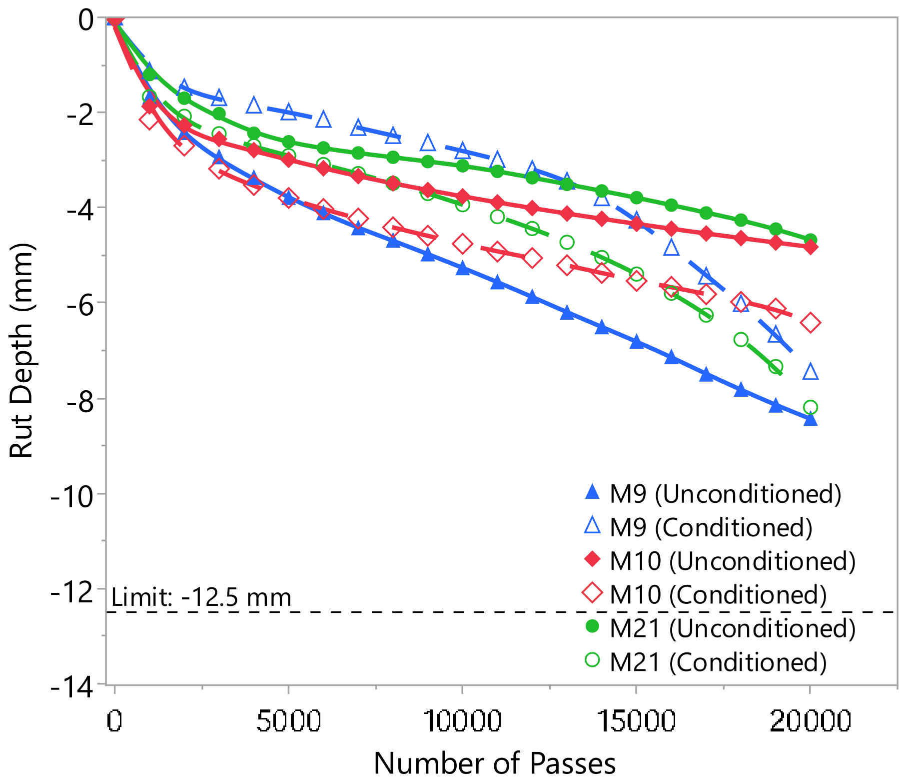


Figure 90. Rut Depth Curves of Dry and Wet Samples for S3 Mix Type.

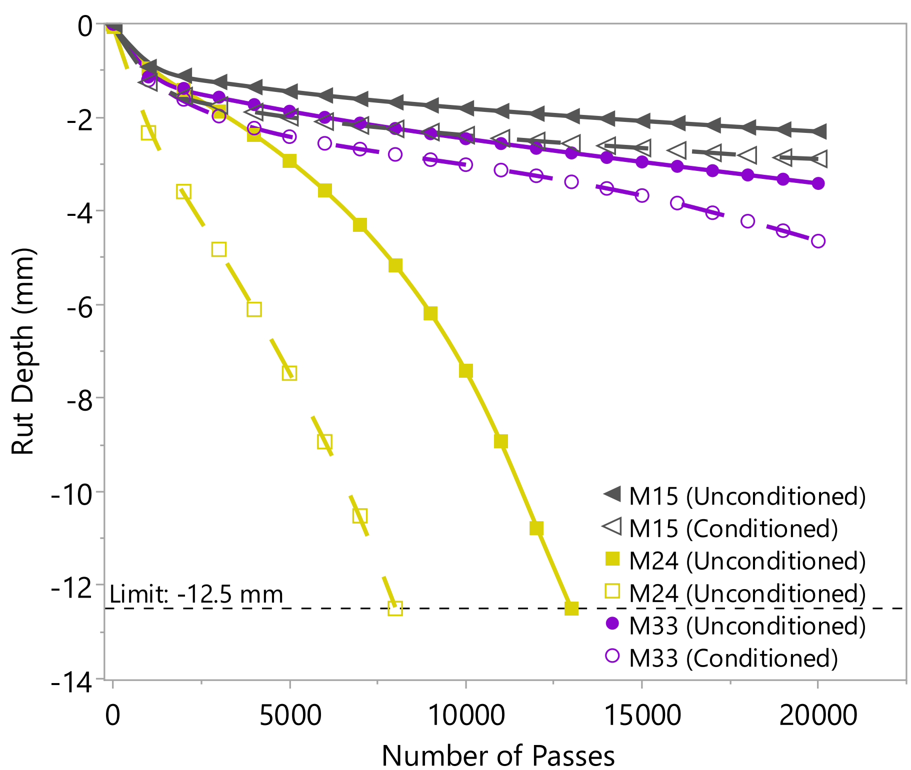


Figure 91. Rut Depth Curves of Dry and Wet Samples for S4 Mix Type.

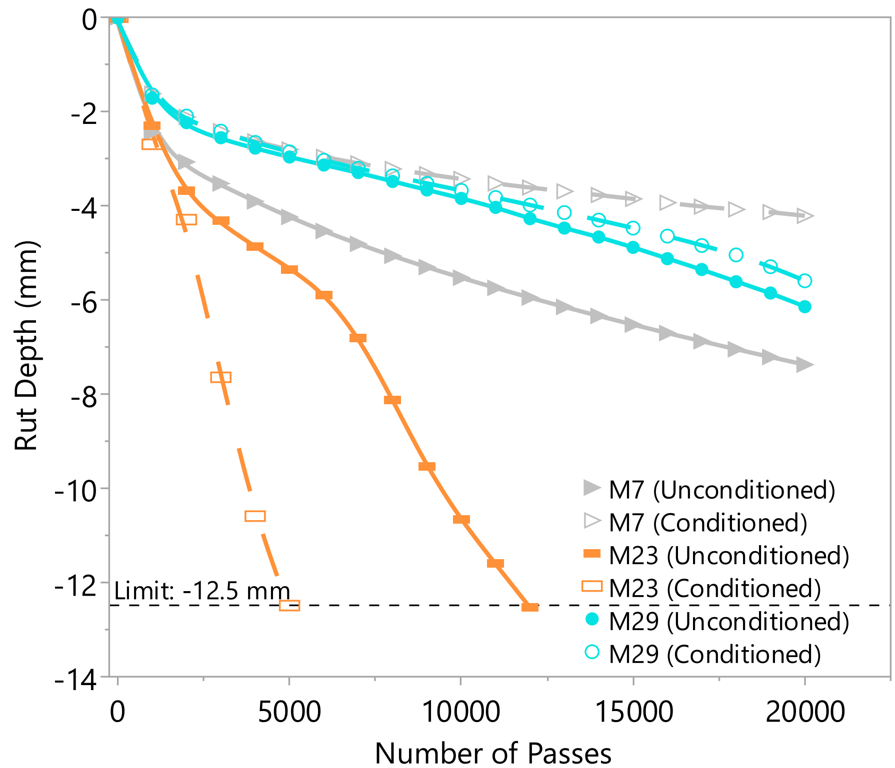


Figure 92. Rut Depth Curves of Dry and Wet Samples forS5 Mix Type.

Table 28 summarizes the HWTT results, for all mixes, including rutting parameters such as TRD at 10,000 and 20,000 passes, Rutting Resistance Index (RRI), and number of passes at 12.5-mm rut depth for all mixes in this study. The RRI was also calculated based on Equation 9.

$$RRI = N_{max} \times (1 - RD_{max}) \qquad (9)$$

Where *RRI* is rutting resistance index (inch), N_{max} = number of wheel passes at completion of test; and RD_{max} = rut depth at completion of test (inch).

Table 28. HWTT Rut Results.

Mix Type	Mix	Condition	CS	$TRD_{10,000}$	$TRD_{20,000}$	RRI	No. of passes @12.5 mm
S3	M9	Unconditioned	0.00030	5.5	7.4	14,200	-
S3	M9	Conditioned	0.00016	2.8	4.2	16,700	-
S3	M10	Unconditioned	0.00008	3.9	4.8	16,220	-
S3	M10	Conditioned	0.00016	4.8	6.4	15,000	-
S3	M21	Unconditioned	0.00016	3.1	4.4	16,600	-
S3	M21	Conditioned	0.00029	3.9	8.2	13,600	-
S4	M24	Unconditioned	0.00051	7.8	**	6,500	13,000
S4	M24	Conditioned	0.00087	16.2*	**	4,000	8,000
S4	M15	Unconditioned	0.00006	1.8	2.3	18,200	-
S4	M15	Conditioned	0.00008	2.4	2.9	17,800	-
S4	M33	Unconditioned	0.00009	2.4	3.5	17,200	-
S4	M33	Conditioned	0.00011	3.0	4.6	16,400	-
S5	M7	Unconditioned	0.00008	5.8	7.4	14,200	-
S5	M7	Conditioned	0.00007	3.4	4.2	16,800	-
S5	M23	Unconditioned	0.00052	10.7	**	6,000	12,000
S5	M23	Conditioned	0.00129	24.9*	**	2,500	5,000
S5	M29	Unconditioned	0.00020	3.8	6.1	15,200	-
S5	M29	Conditioned	0.00019	3.9	5.9	15,400	-

** Extrapolated*

*** Test stopped before 20,000 passes.*

As shown by the HWTT results, all unconditioned mixes had excellent rutting performance with a TRD of less than 12.5 mm at 20,000 passes, except for mixes M23 and M24, which reached 12.5 mm rut depth at 13,000 and 12,000 passes, respectively.

Following AASHTO T-283 conditioning, all mixes showed lower RRI indicating less rutting resistance except for mixes M7 and M9, which showed a slight improvement in rutting resistance. These results indicate that freeze-thaw damage, as simulated by AASHTO T-283, in most cases, had a negative effect on the rutting resistance. However, the effect of freeze-thaw damage on the RRI was not significant except for mixes M23 and M24.

The RRI and TRD parameters do not separate the rutting induced by permanent deformation and that caused by mixture stripping. The CRD parameter was introduced to define the rutting caused by permanent deformation only. Figure 93 shows the CRD values of each mix at 10,000 passes for all mixtures used in this study. From Figure 93, all mixes had higher CRD values, indicating lower rutting resistance, after freeze-thaw damage except for mixes M7 and M9.

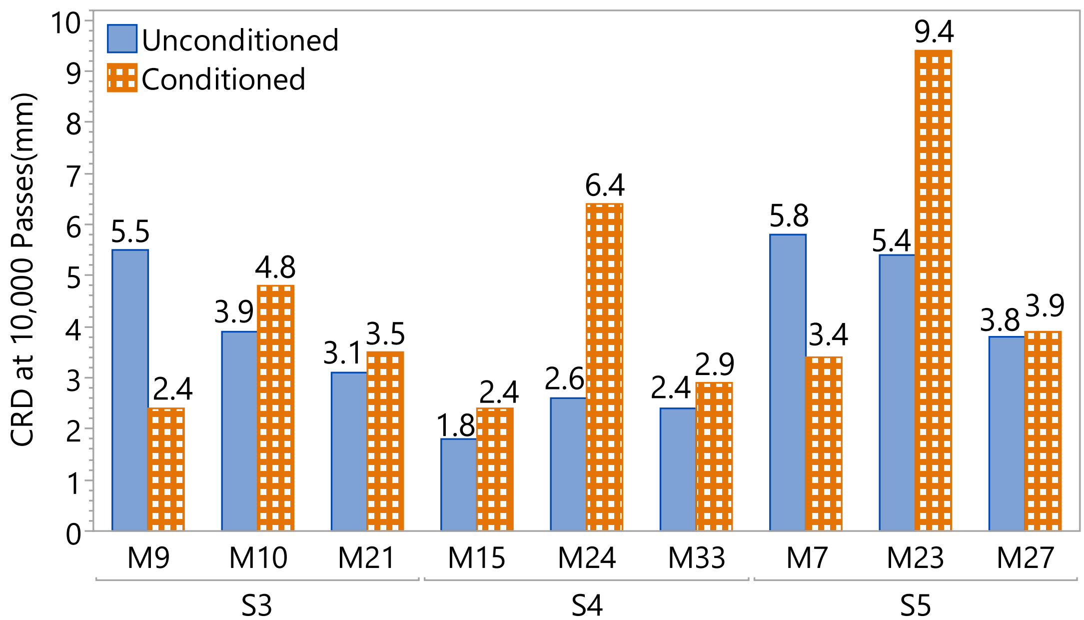


Figure 93. CRD Before and After AASHTO T-283 Conditioning

Table 29 summarizes the stripping-related parameters derived from HWTT, including SS, SIP, and SN under both unconditioned and conditioned conditions for all mixtures. The effect of AASHTO T-283 conditioning was evident on the stripping susceptibility of the mixes, where five out of the nine mixes tested showed higher susceptibility to stripping after freeze-thaw damage. Among these, three mixes (i.e. M9, M21, M33) showed no signs of stripping in the unconditioned state, however, stripping was noted in these mixes after being subjected to AASTHO T-283 conditioning prior to the HWTT. For some mixtures (i.e. M7, M10, M15, M29), there was no evidence of stripping in both the unconditioned and conditioned states.

By considering a moisture susceptibility performance criterion, requiring a minimum SN of 2,000 wheel passes for all mixtures in Oklahoma, as discussed in the section 3.2.2, all mixes in this study met the SN criterion before and after freeze-thaw damage.

Table 29. HWTT Moisture Results.

NMAS	Mix	Condition	SS	SIP	SN
S3	M9	Unconditioned	N/A	N/A	N/A
S3	M9	Conditioned	0.0006	13,597	4,544
S3	M10	Unconditioned	N/A	N/A	N/A
S3	M10	Conditioned	N/A	N/A	N/A
S3	M21	Unconditioned	N/A	N/A	N/A
S3	M21	Conditioned	0.00052	14,143	5,427
S4	M24	Unconditioned	0.00165	8,066	1,806
S4	M24	Conditioned	0.00175	2,942	1,104
S4	M15	Unconditioned	N/A	N/A	N/A
S4	M15	Conditioned	N/A	N/A	N/A
S4	M33	Unconditioned	N/A	N/A	N/A
S4	M33	Conditioned	0.0002	14,471	7,340
S5	M7	Unconditioned	N/A	N/A	N/A
S5	M7	Conditioned	N/A	N/A	N/A
S5	M23	Unconditioned	0.00141	6,610	3,635
S5	M23	Conditioned	0.0035	2,052	2,041
S5	M29	Unconditioned	N/A	N/A	N/A
S5	M29	Conditioned	N/A	N/A	N/A

3.2.6.2 IDT-HT before and after Freeze-thaw Damage

The IDT-HT Tensile Strength (TS) results, before and after freeze-thaw damage, are presented in Table 30 and Figure 94. Unconditioned strengths ranged from approximately 35 to 67 psi, with generally consistent variability (COV <10%). All mixes except for mix M23 after freeze-thaw damage, met the proposed minimum IDT-HT strength of 35psi (240kPa) for rutting as discussed previously.

Table 30. The Tensile Strength Values of IDT-HT Test for Both Unconditioned and Conditioned Samples.

NMAS	Mix	Unconditioned	Unconditioned	Conditioned	Conditioned	HTSR
NMAS	Mix	TS (psi)	COV (%)	TS (psi)	COV (%)	HTSR
S3	M9	46.5	5.6	72.3	11.8	1.5
S3	M10	41.8	3.0	39.6	0.2	0.9
S3	M21	57.9	13.2	45	2.2	0.8
S4	M24	57.1	3.1	40.8	6.2	0.7
S4	M15	38.3	4.6	41	0.3	1.1
S4	M33	66.7	8.8	48.7	8.4	0.7
S5	M7	35.0	10.8	52.9	5.9	1.5
S5	M23	39.1	7	29.9	4.9	0.8
S5	M9	40.6	2.1	36.4	1.9	0.9

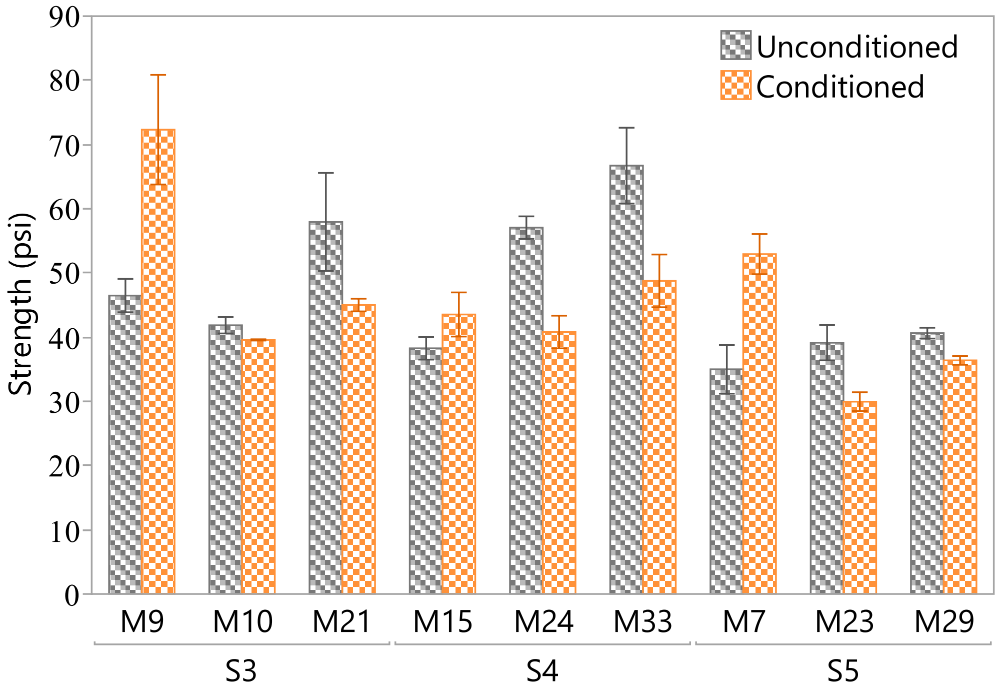


Figure 94. IDT-HT Strength Before and After AASHTO T-283 Conditioning

Table 30 and Figure 95 present the ratio of IDT-HT tensile strength (HTSR) between conditioned and unconditioned specimens. Figure 95 shows that moisture conditioning/freeze-thaw had different effects on the HTSR across the mixtures. Most mixes showed a HTSR of less than one indicating reduced tensile strength with moisture damage. Mix 15 showed no notable effect of moisture damage on tensile strength, whereas mixes M7 and M9 showed notable increase in tensile strength after freeze-thaw damage as given by their HTSR of 1.5.

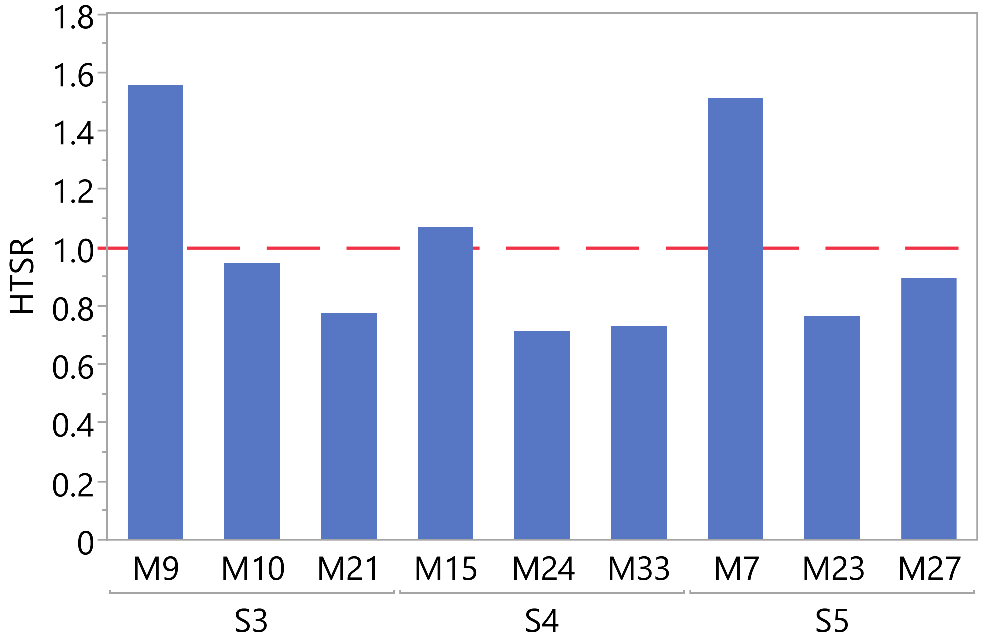


Figure 95. IDT-HT Strength Ratio due to Freeze-Thaw Damage.

Figure 96 illustrates the correlation between the HTSR from IDT-HT and the CRD Ratio from HWTT. The CRD Ratio is calculated as the ratio between unconditioned and conditioned specimens at 10,000 passes. A CRD ratio above one indicates lower CRD after conditioning, i.e. improved rutting resistance. A strong positive correlation was observed (R^2 = 0.84), indicating that both HWTT and IDT-HT captured similar trends in the effect of freeze-thaw damage on rutting resistance. Both mixes M7 and M9 showed TS ratios higher than one as well CRD ratio higher than one, confirming improved rutting resistance following moisture conditioning, as evidenced by both the IDT-HT and HWTT. Conversely, mixes with TS ratios below 1.0 generally had CRD ratios less than 1.0, confirming their reduced rutting resistance following moisture conditioning.

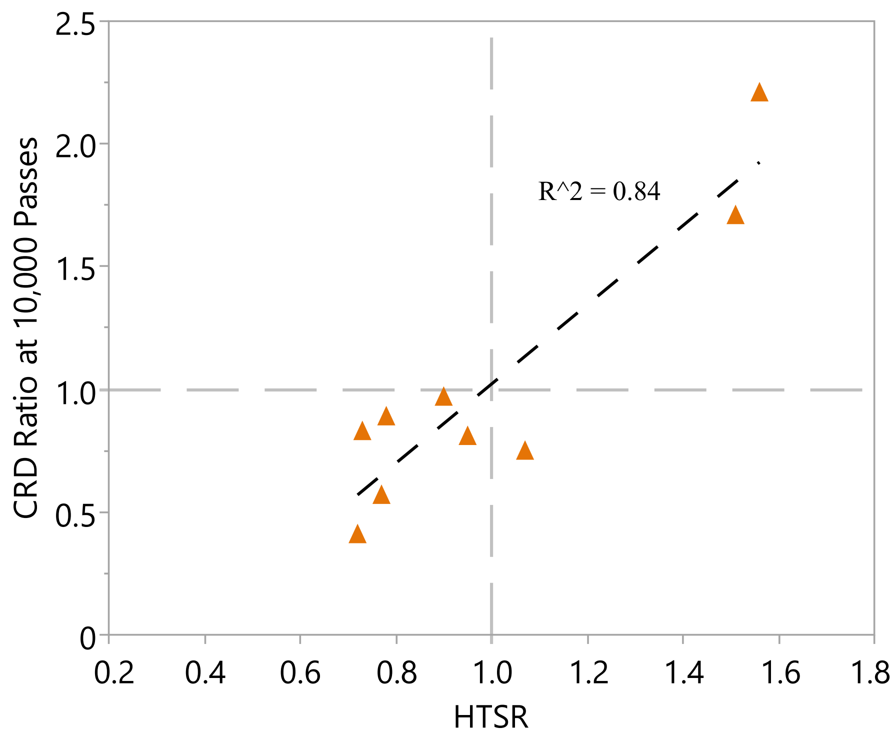


Figure 96. Correlation between TS Ratio and $CRD_{10,000}$ Ratio.

Figure 97 shows the correlation between HTSR from IDT-HT on the x-axis and Tensile Strength Ratio (TSR) from AASHTO T283 test on the y-axis. The TSR values are reported as part of the mix design. HTSR from IDT-HT provides a measure of the effect of moisture on the high-temperature rutting performance, whereas the TSR provides a measure of the effect of moisture on the strength at intermediate temperature. A moderate positive correlation (R^2 = 0.79) was found, indicating that both ratios captured roughly similar trends in moisture susceptibility. The stripping-susceptible mixes (red circles) grouped in the lower-left area, showing reduced strength in both tests, while the non-stripping mixes (blue circles) stayed higher on the plot. This shows both tests reflect moisture damage in a consistent way.

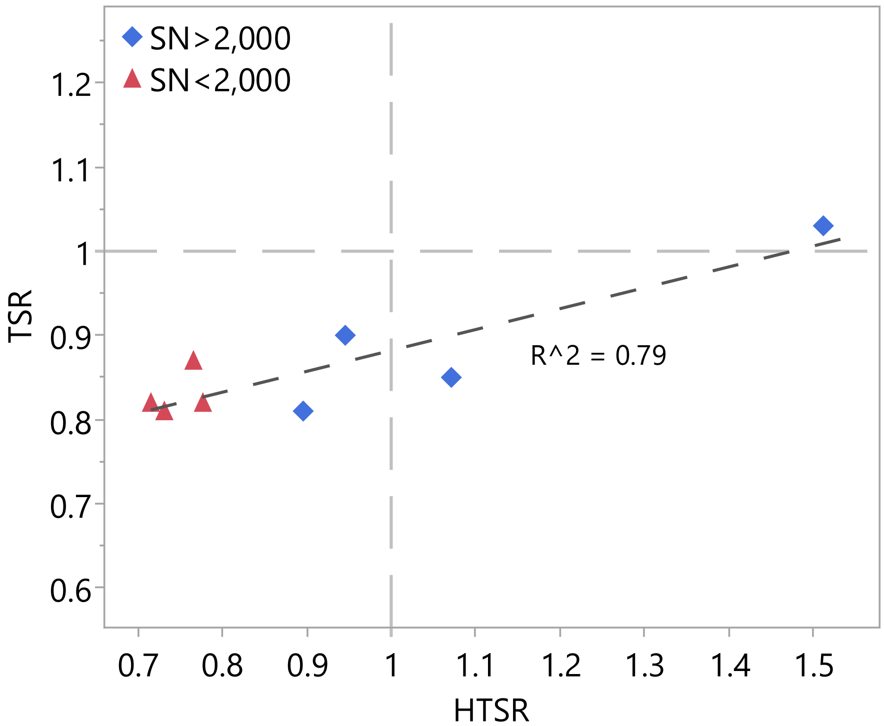


Figure 97. Correlation between HTSR and TSR

3.3 Combined Effect of Moisture and Aging

The objective of this task was to assess the effect of aging on the moisture sensitivity. To achieve this objective, seven plant mixes were selected to represent different mix designs. Plant-produced mixes representing STA condition were first subjected to long-term aging conditioning to produce LTA mixes. STA and LTA mixes were then tested to determine their moisture sensitivity according to AASHTO T-283. The unconditioned (dry) and moisture conditioned (wet) STA and LTA specimens were tested to determine their strength and cracking resistance. Figure 98 Provides an overview of the test methodology. The IDT was conducted to determine the mixtures' strength at different conditions, and the IDEAL-CT was conducted to evaluate the cracking resistance using the CT_{Index} parameter.

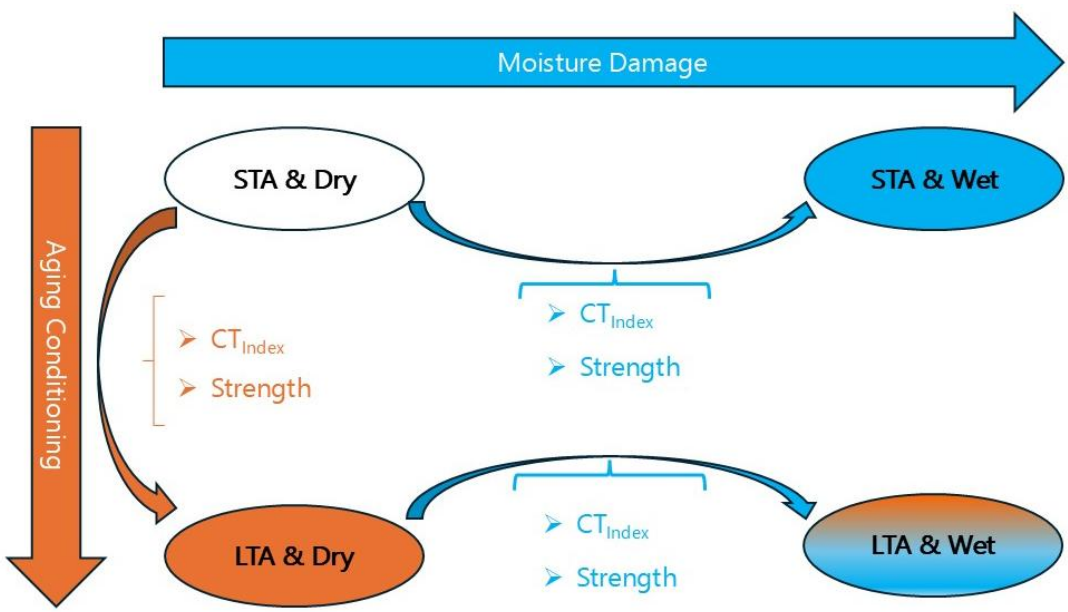
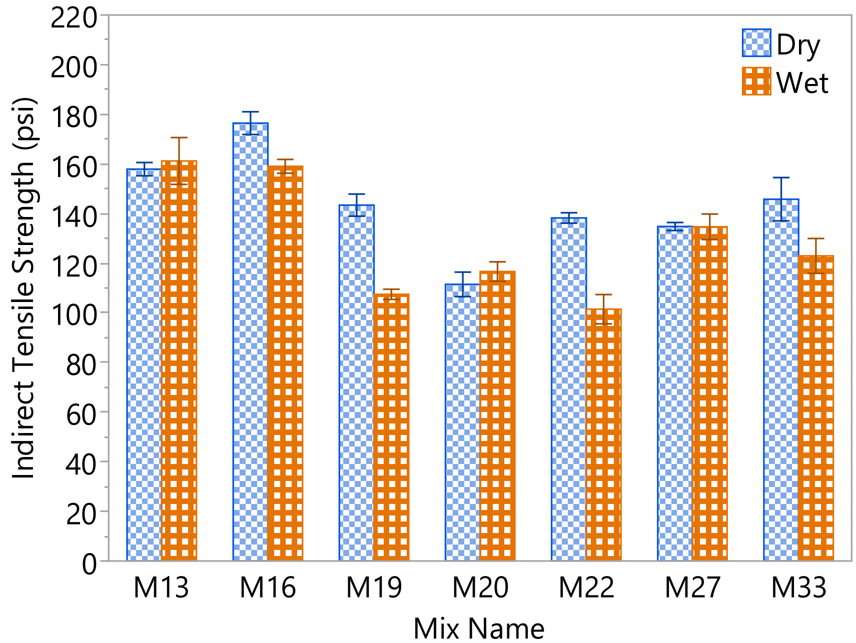


Figure 98. Research Study Framework of Moisture Susceptibility of Mixtures under Aging Conditions.

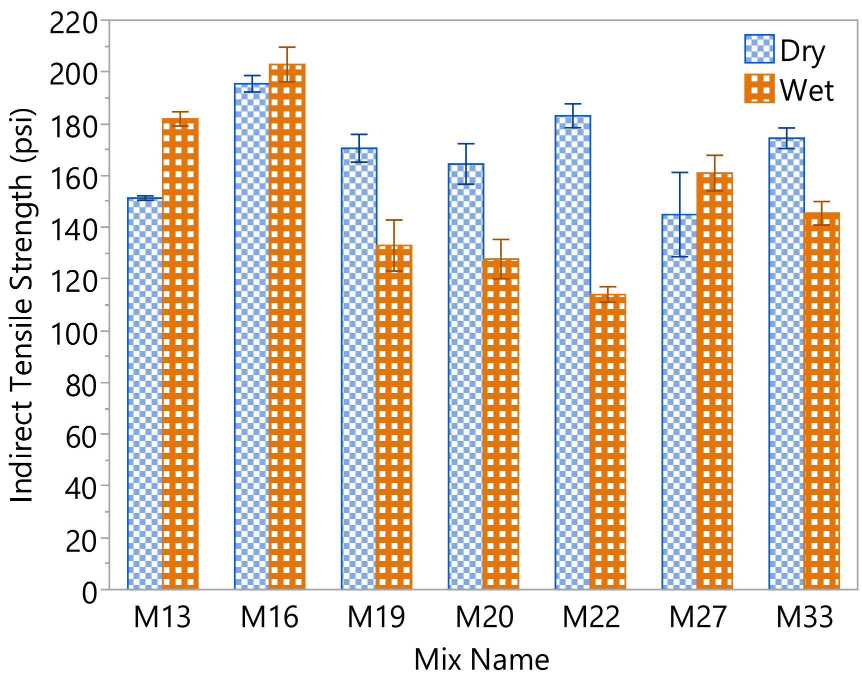
3.3.1 Indirect Tensile Strength (ITS)

Figure 99 shows the mean ITS for the STA and LTA mixes under both dry and wet conditions, with the error bars representing the standard error. A two-way ANOVA was conducted on the ITS results as shown in Table 31. Based on a threshold p-value of 0.05, all mixes showed significant differences in their mean ITS based on aging, except for mix M13. The effect of moisture conditioning on the ITS was statistically significant for all mixes except for mixes M16 and M27.

The ANOVA also showed the interaction between moisture and aging. It was noted that three out of seven mixes, i.e. mixes M19, M27, and M33, showed no interaction between moisture and aging, which indicates that for these three mixes, the response of the mix to moisture damage was the same in both the STA and LTA conditions. For the remaining four mixes, i.e. mixes M13, M16, M20, and M22, the response of the mix to moisture damage was dependent on the aging condition.



(a)



(b)

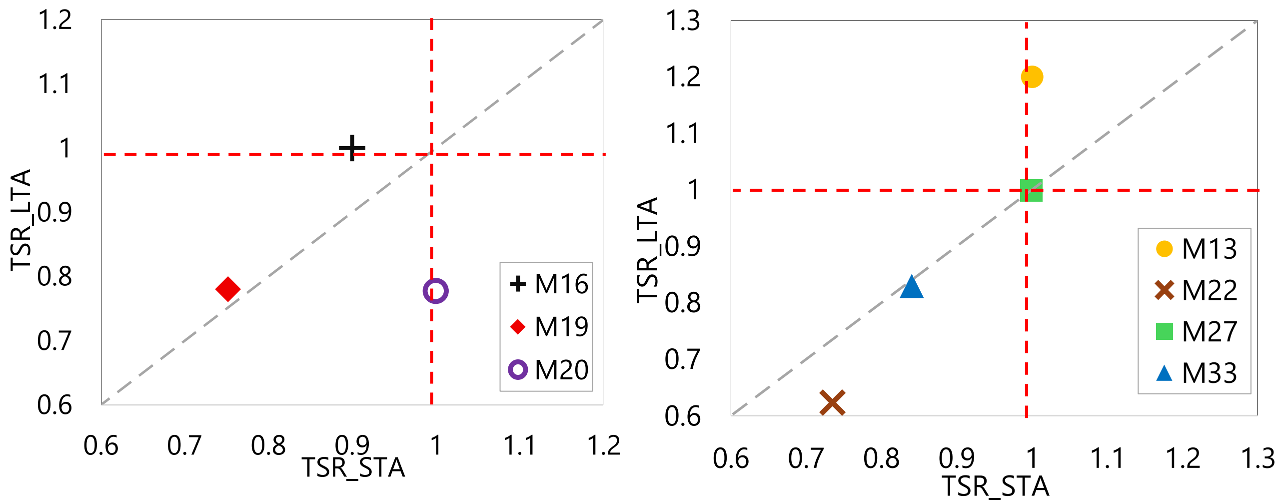
Figure 99. ITS Results for; (a) STA, and (b) LTA Mixes.

Table 31. Two-Way ANOVA on ITS results.

Mix ID	Moisture Effect	Aging Effect	Moisture*Aging Effect
M13	0.004*	0.110	0.011*
M16	0.165	<0.001*	0.007*
M19	<0.001*	0.001*	0.817
M20	0.008*	<0.001*	0.002*
M22	<0.001*	<0.001*	0.001*
M27	0.184	0.015*	0.183
M33	0.001*	0.001*	0.474

*Statistically Significant

To further analyze the ITS results, the TSR was calculated for both STA and LTA specimens as shown in Figure 100. The mixes were grouped based on their binder's PG where Figure 100(a) shows mixes with a PG 64-22 binder, and Figure 100(b) show mixes with a PG 70-28 binder. All mixes showed TSR values higher than 0.7, which is considered acceptable as per ODOT specifications. The TSR results verify the two-way ANOVA results presented earlier. Mixes, M13, M16, M20, and M22, showed difference in TSR, i.e. moisture susceptibility, based on their aging condition. The effect of aging on moisture susceptibility, however, did not follow a clear trend. For mixes M16 and M13, the TSR increased with aging, denoting better moisture resistance, whereas for mixes M20 and M22, the TSR decreased with aging, denoting reduced moisture resistance. It can be seen in Mixes M19, M27, and M33 that there is no difference between TSR values due to different aging conditions.



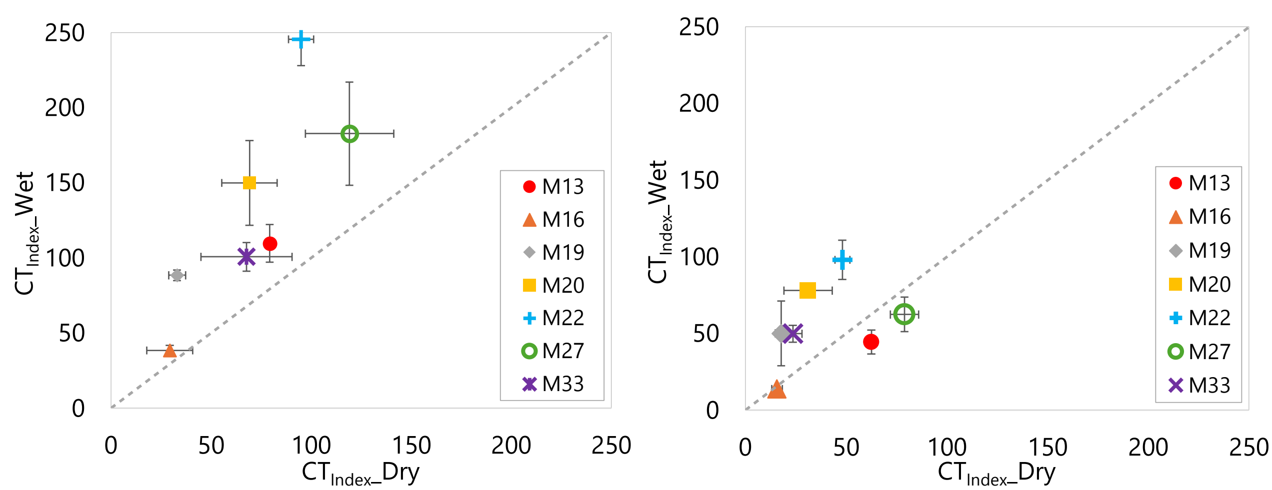
(a) (b)

Figure 100. TSR Values at STA versus LTA for mixes using; (a) PG64-22, (b) PG70-28.

3.3.2 IDEAL-CT

3.3.2.1 CT_{Index}

The CT_{Index} results of the dry and wet specimens were compared in Figure 101 under both STA and LTA conditions. The error bars shown in Figure 101 indicates one standard error. It was noted that moisture conditioning resulted in an increase in the CT_{Index}, for all mixes in the STA condition and for most mixes in the LTA condition.



(a) (b)

Figure 101. Average CT_{Index} Values of Dry and Wet Samples under; (a) STA, and (b) LTA.

A two-way ANOVA test was conducted to further evaluate the change in CT_{Index} with moisture and aging conditioning as well as interaction between moisture and aging as shown in Table 32. All mixtures showed statistically significant difference in CT_{Index} with aging. With regards to effect of moisture damage, most mixes showed statistically significant difference in their CT_{Index} in the dry and wet conditions. It was also noted that four out of seven mixes, i.e. mixes M13, M19, M22, and M27, showed an interaction between moisture and aging, which indicates that for these mixes, the change of CT_{Index} with moisture damage was dependent on the aging condition.

Table 32. Two-way ANOVA of Effect of Moisture and Aging on CT_{Index}

Mix ID	Moisture Effect	Aging Effect	Moisture*Aging Effect
M13	0.311	<0.001*	0.005*
M16	0.457	0.005*	0.250
M19	<0.001*	<0.001*	0.032*
M20	0.002*	0.003*	0.197
M22	<0.001*	<0.001*	0.001*
M27	0.164	0.002*	0.035*
M33	0.007*	0.001*	0.652

*Statistically significant

3.3.2.2 Effect of Moisture Damage on IDEAL-CT Parameters

To further assess the effect of moisture conditioning on the CT_{Index}, the CT_{Index} parameters such as fracture energy and l_{75}/m_{75} were plotted at both STA and LTA conditions as shown in Figure 102. For the STA specimens, it can be noted that both fracture energy and l_{75}/m_{75} increased with moisture conditioning. The STA results indicate that moisture conditioning caused an average increase of 11 percent and 60 percent in the fracture energy and l_{75}/m_{75} parameters, respectively. However, for the LTA specimens, the change in the fracture energy and l_{75}/m_{75} parameters with moisture conditioning were less pronounced, compared to the change noted in the STA specimens. For the LTA specimens, moisture conditioning mostly resulted in an increase in fracture energy and

I_{75}/m_{75} however there were a few exceptions where the fracture energy and I_{75}/m_{75} showed a decrease with moisture conditioning.

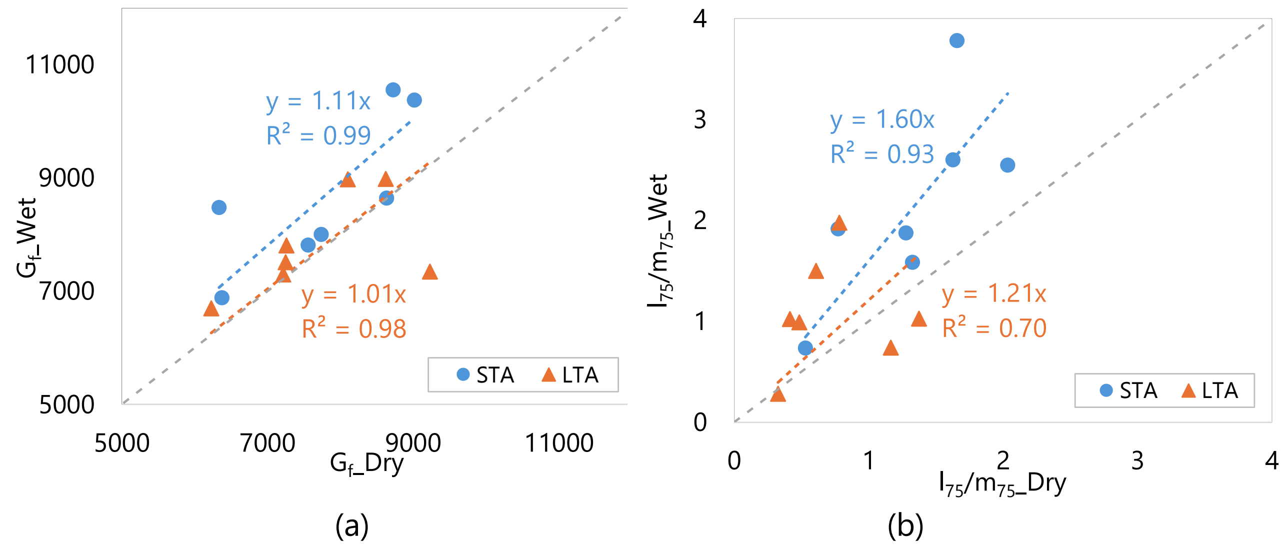
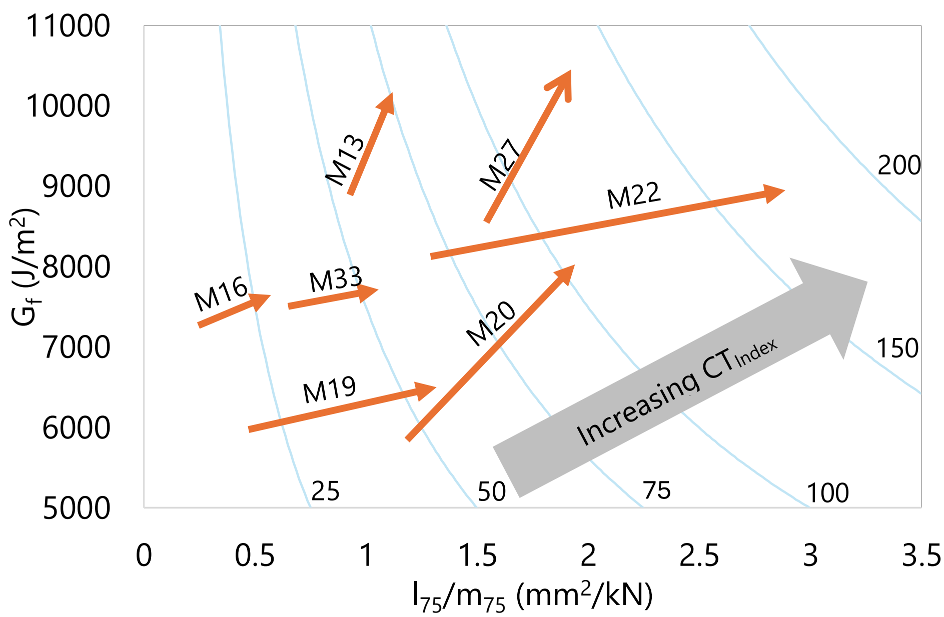


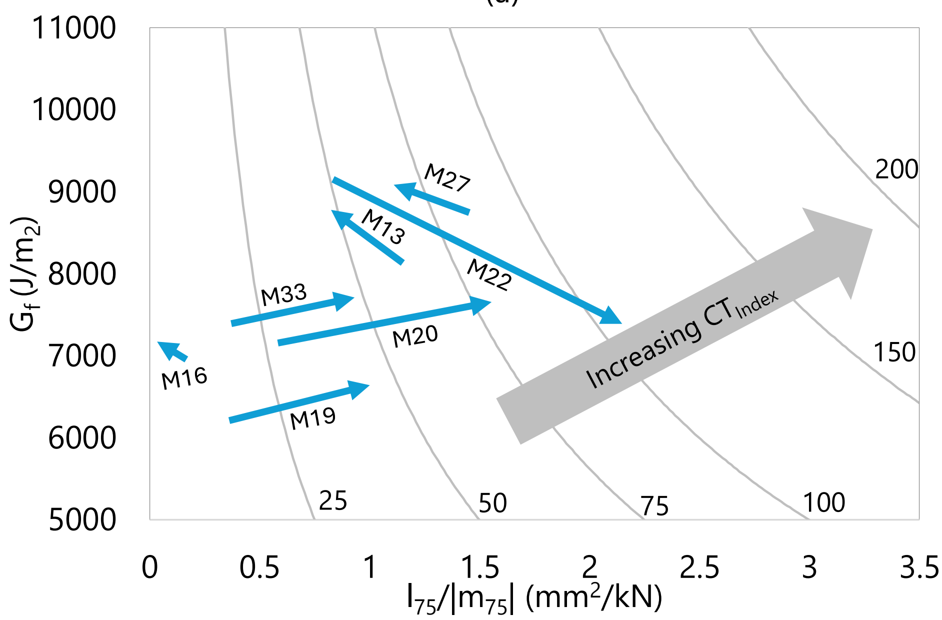
Figure 102. CT_{Index} Parameters at both dry and wet conditions; (a) Fracture Energy, (b) I_{75}/m_{75}.

The effect of moisture conditioning on the CT_{Index} of STA and LTA specimens is further illustrated using IDEAL-CT interaction diagrams as shown in Figure 103. The direction of the arrows represents the change in the mixtures' CT_{Index} with moisture conditioning. It is shown that all arrows for the STA specimens are pointing to the upper right corner of the diagram indicating an increase in the CT_{Index}. The slope of the arrows signifies the extent of increase in fracture energy and I_{75}/m_{75}. As noted previously, the average increase in I_{75}/m_{75} across all mixes was much higher compared to the increase in fracture energy. For LTA specimens, the results were not consistent among all mixes, where some mixes showed an increase in CT_{Index} while others showed a decrease in CT_{Index} as noted previously.

The increase in CT_{Index} due to moisture conditioning could be attributed to the mixture getting softer hence undergoing further deformation before failure, which resulted in an increase in the l_{75}/m_{75} parameter.



(a)



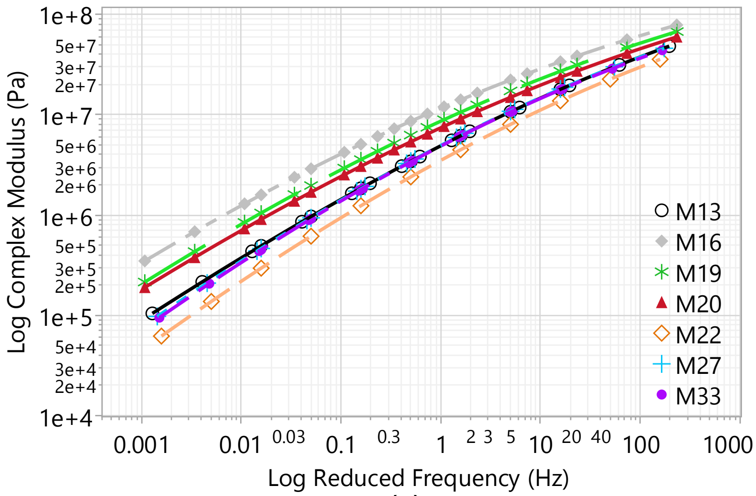
(b)

Figure 103. CT_{Index} Interaction Diagram for Effect of Moisture Conditioning on; (a) STA, (b) LTA samples.

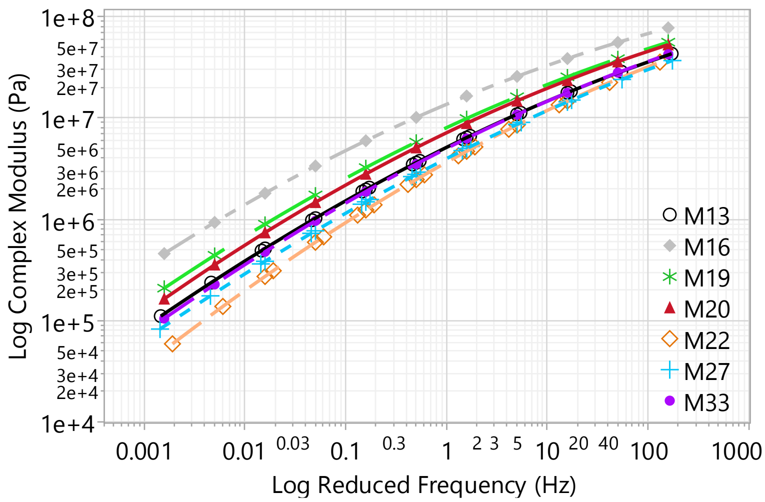
3.3.3 Frequency-Temperature Sweep Test

The results of the complex modulus master curves versus reduced frequency of all seven binders at different conditions such as STA-Dry, STA-Wet, and LTA-Dry are shown in Figure 104.

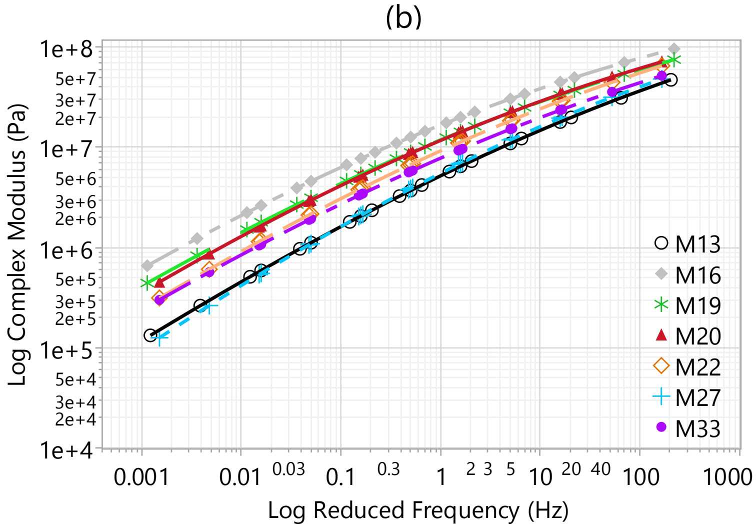
Figure 104 shows the complex modulus master curves, with neat binders (M16, M19, and M20) having higher complex moduli than the modified binders (M13, M22, M27, and M33) at all three STA-dry, STA-wet, and LTA-dry conditions.



(a)



(b)



(b)

Figure 104. Complex Modulus Master Curves under Different Conditions; (a) STA-Dry, (b) STA-Wet, and (c) LTA-Dry.

The phase angle master curves shown in Figure 105 reveal clear differences between the binder types. It can be noted that the neat binders (M16, M19, M20) show a smooth and consistent decrease in phase angle as frequency increases, reflecting a single relaxation mechanism typical of neat binders. The term relaxation is used to describe how a viscoelastic material responds to being deformed or placed under strain (D. Christensen et al., 2019). Single relaxation mechanism of neat binders means that the stress in these binders relaxes through one dominant molecular process, where the viscoelastic response is mainly governed by the base bitumen structure (Petersen et al., 1994; Stastna et al., 1994). In contrast, the polymer-modified binders (M13, M22, M27, M33) display a curved or shoulder-like trend at low frequencies before the phase angle starts to decrease at intermediate and high frequencies (Yusoff et al., 2012). This shape suggests the presence of multiple relaxation mechanisms. That is because the added polymer network introduces additional molecular motions that delay stress relaxation and increase elasticity, especially at low frequencies. As a result, polymer-modified binders maintain lower phase angles at low frequencies, indicating greater viscous–elastic balance, while still converging with neat binders at higher frequencies. This behavior suggests that polymer modification enhances elastic response and deformation resistance but also introduces thermos-rheological complexity (Polacco et al., 2022; Teltayev et al., 2022).

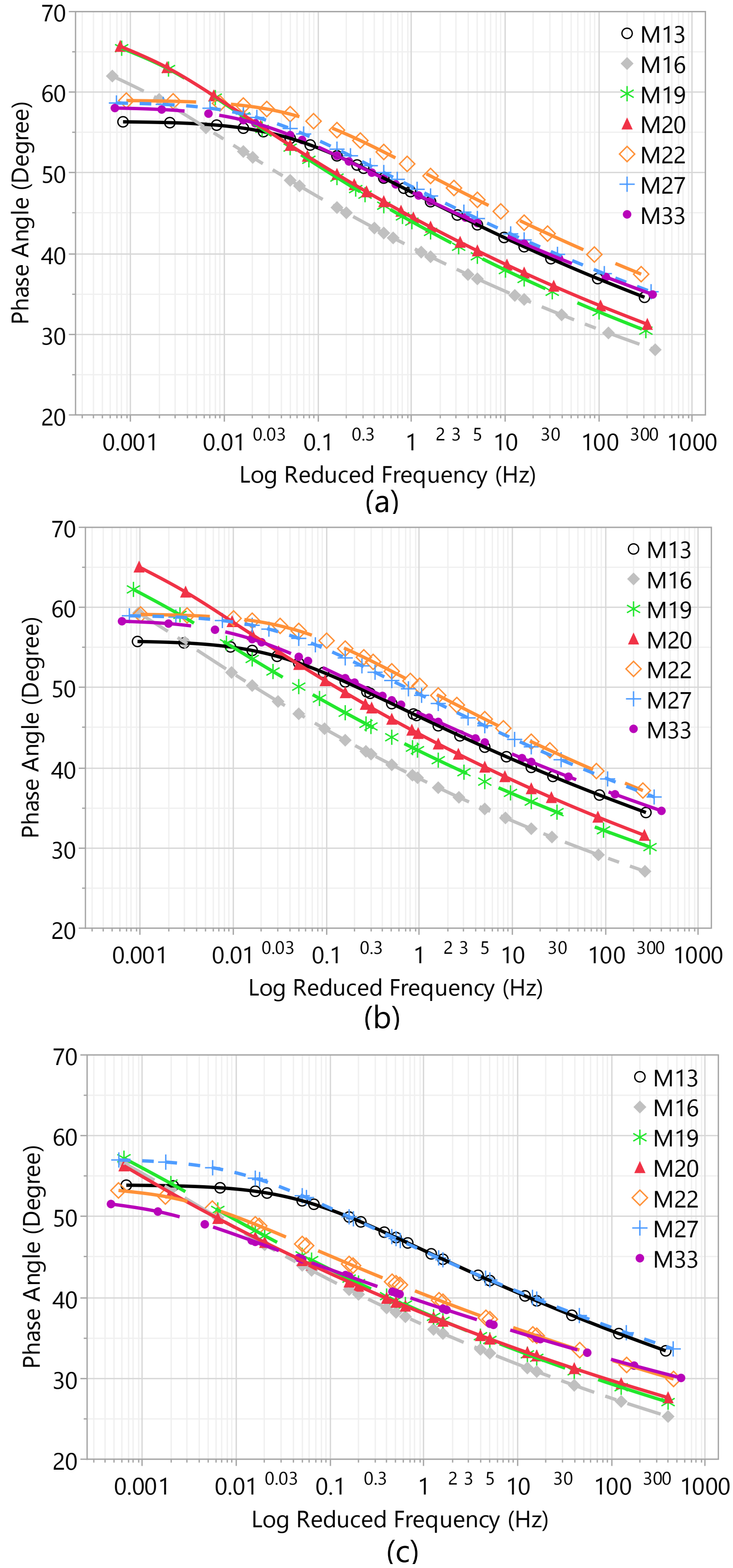


Figure 105. Phase angle master curves under different conditions: (a) STA-Dry, (b) STA-Wet, and (c) LTA-Dry.

To assess the effect of aging and moisture on the complex modulus and phase angles, the master curves for mix M20 using neat binder, and mix M22 using modified binder, are shown in Figure 106, as an example.

The complex modulus and phase angle master curves for the recovered binder from the other mixes, i.e. M13, M16, M19, M27, and M33, can be found in Appendix D. As can be seen in Figure 106, the moisture conditioning in the STA condition did not cause a notable effect on the complex modulus or phase angle, as evidenced by the master curves of the STA-dry and STA-wet conditions. This illustrates that the impact of moisture conditioning on the asphalt mixture performance is not related to a change in the binder properties and is primarily driven by stripping effect between the aggregate and the binder.

LTA results in an increase in the modulus and a decrease in phase angle, as noted by the master curves of the STA-dry and LTA-dry conditions. The effect of LTA on the complex modulus is more evident at lower frequencies. At higher frequencies, the binder approaches the glassy modulus so the curves start to flatten (Mensching et al., 2016).

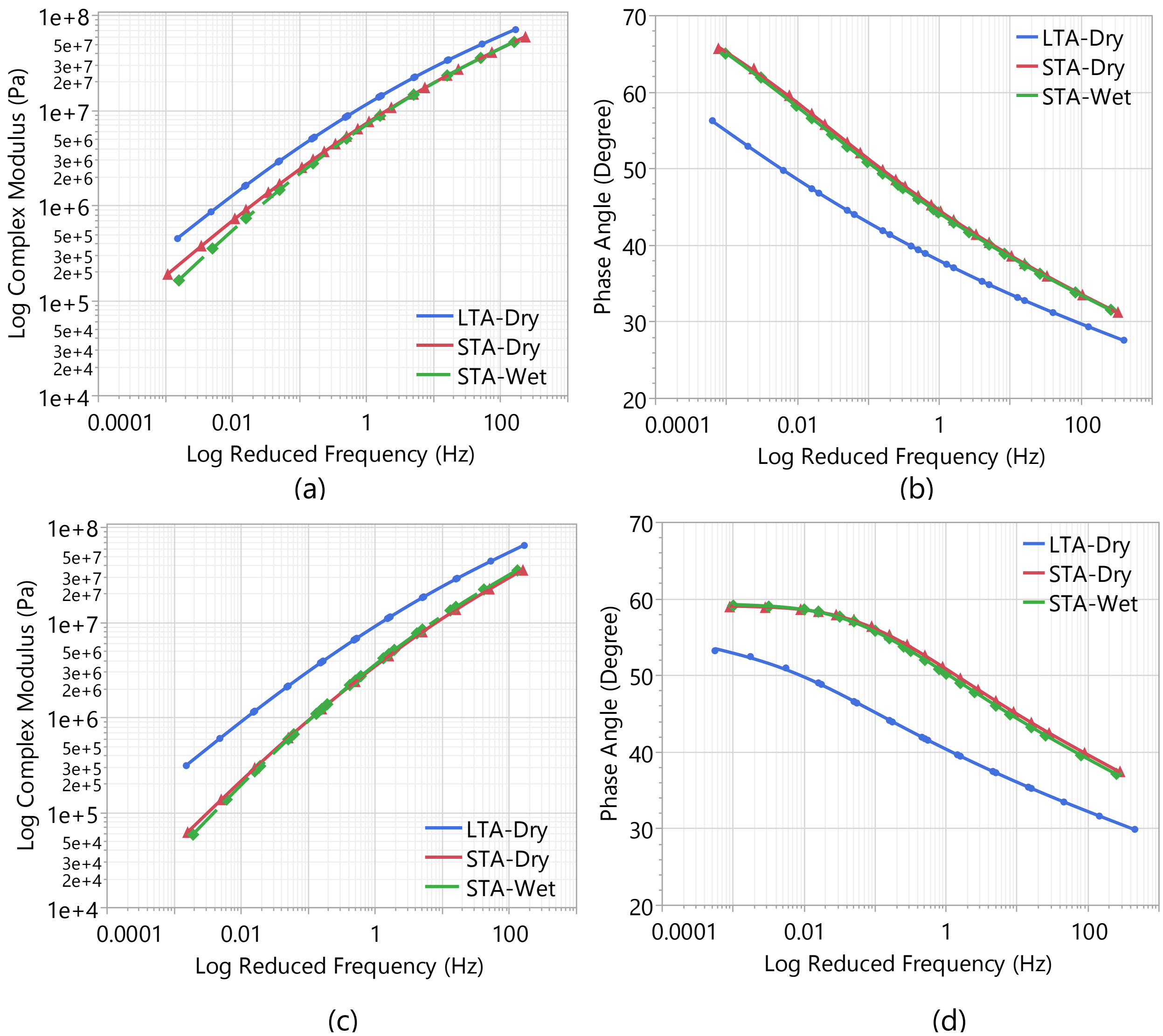


Figure 106. Complex Modulus and Phase Angle Master Curves under Different Conditions; (a) M20, and (b) M22.

The CAM model parameters for both complex modulus and phase angle master curves are presented in Table 33 for all binders at different conditions. K and m are fitting constants. As can be seen, the R^2 values for all types of binders are higher than 99.30%, indicating that the CAM method can efficiently model both complex modulus and phase angle master curves of binders. From the fitting constants, the values for STA-Dry and STA-Wet samples are very close, indicating that there was no notable difference between these two conditions.

Table 33. The CAM Model Parameters for Complex Modulus and Phase Angle Master Curves.

Mix	Condition	Complex Modulus Master Curves	Complex Modulus Master Curves	Complex Modulus Master Curves	Phase Angle Master Curves	Phase Angle Master Curves	Phase Angle Master Curves
Mix	**Condition**	**m**	**k**	**R^2**	**m**	**k**	**R^2**
M13	**LTA-Dry**	0.60	0.054	0.996	0.70	0.156	0.999
M16	**LTA-Dry**	0.69	0.062	0.998	0.77	0.135	0.999
M19	**LTA-Dry**	0.69	0.057	0.997	0.74	0.134	0.999
M20	**LTA-Dry**	0.90	0.053	0.996	0.77	0.137	0.999
M22	**LTA-Dry**	0.60	0.049	0.996	0.70	0.153	0.999
M27	**LTA-Dry**	0.63	0.050	0.995	0.80	0.150	0.999
M33	**LTA-Dry**	0.58	0.043	0.994	0.72	0.137	0.999
M13	**STA-Dry**	0.63	0.056	0.997	0.73	0.162	0.998
M16	**STA-Dry**	0.72	0.062	0.996	0.81	0.135	0.998
M19	**STA-Dry**	0.75	0.064	0.998	0.80	0.141	0.997
M20	**STA-Dry**	0.76	0.061	0.998	0.76	0.141	0.997
M22	**STA-Dry**	0.66	0.055	0.996	0.87	0.149	0.998
M27	**STA-Dry**	0.65	0.054	0.996	0.84	0.149	0.998
M33	**STA-Dry**	0.65	0.052	0.996	0.90	0.142	0.998
M13	**STA-Wet**	0.62	0.053	0.997	0.92	0.133	0.997
M16	**STA-Wet**	0.73	0.064	0.997	0.82	0.122	0.998
M19	**STA-Wet**	0.74	0.059	0.997	0.90	0.128	0.999
M20	**STA-Wet**	0.76	0.060	0.998	1.05	0.128	0.998
M22	**STA-Wet**	0.66	0.055	0.996	0.90	0.144	0.999
M27	**STA-Wet**	0.66	0.052	0.995	0.84	0.140	0.996
M33	**STA-Wet**	0.65	0.050	0.993	0.94	0.136	0.998

Figure 107 presents the complex modulus and phase angle master curves for the M33 binder in the STA-wet condition, as an example, to illustrate the CAM model rheological parameters such as crossover frequency (f_c), crossover modulus (G^*_c), and rheological index (R-value). f_c is the frequency at which phase angle equals to 45°. The complex shear modulus corresponding to f_c is defined as G_c^*. R-value is the difference between the log of the glassy modulus (assumed as 10^9 Pa in this study) and the log of $|G_c^*|$ (D. A. Anderson et al., 1994b). Binders which are more susceptible to cracking have a lower crossover frequency and higher R-value, larger difference between glassy modulus (10^9) and G_c*.

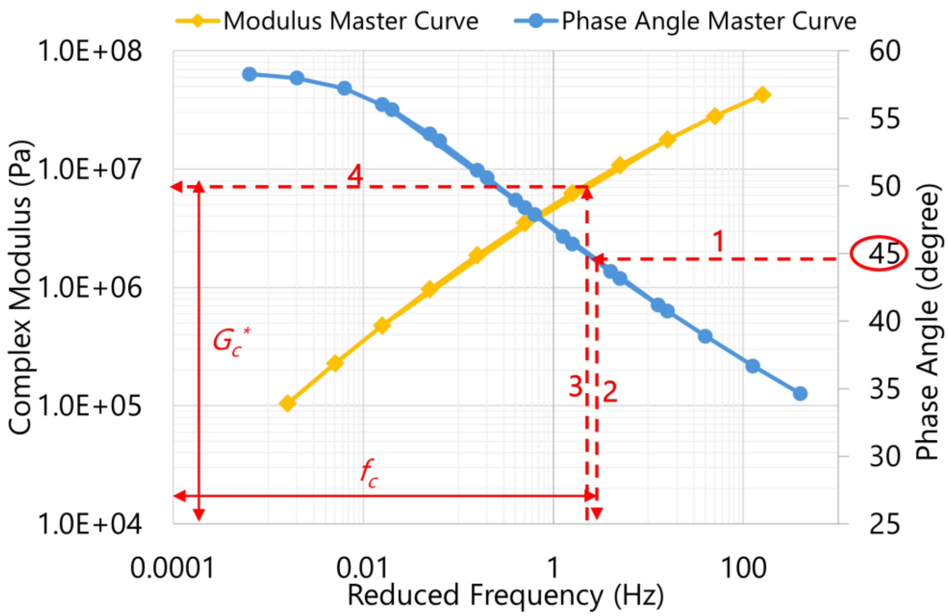


Figure 107. Illustration of Rheological Parameters on Master Curve.

Figure 108 shows the R-value plotted against f_c for all recovered binders at the STA-Dry and LTA-Dry conditions to illustrate the effect of aging. It is shown that aging results in an increase in the R-value, which indicate a flatter master curve, and a decrease in crossover frequency (D. Christensen et al., 2019), This means a larger portion of the master curve falls within the elastic behavior region (delay of viscous behavior) (Mensching et al., 2016).

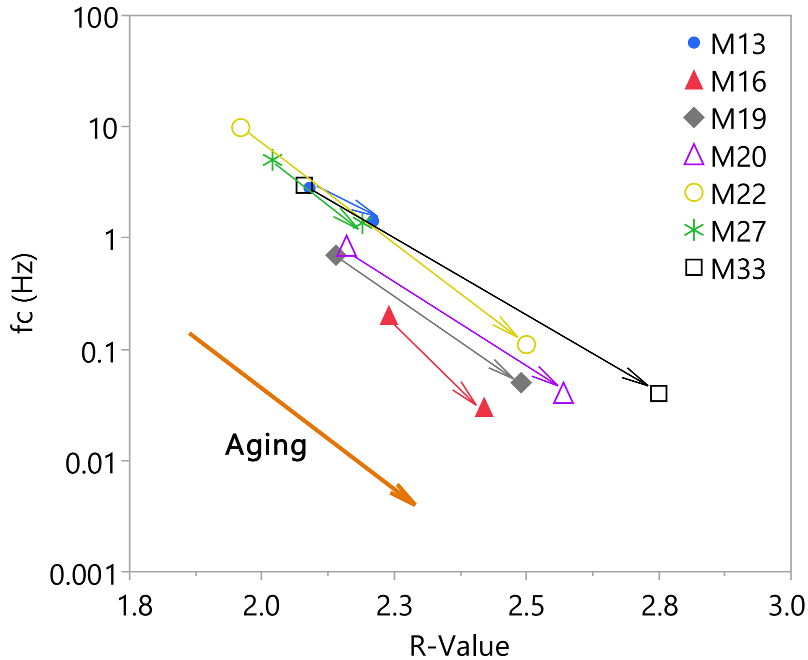


Figure 108. Crossover Frequency vs. Rheological Index for STA-Dry and LTA-Dry Conditions

Figure 109 shows the R-value plotted against fc for all recovered binders at the STA-Dry and STA-Wet conditions to illustrate the effect of moisture conditioning. It is noted that the moisture conditioning tends to shift the binders toward higher R-value and lower cross over frequency, however the effect was not significant. The results of a one-way ANOVA is shown in Table 34 highlighting that the effect of aging was statistically significant o the R-value and crossover frequency, unlike the effect of moisture which was not statically significant.

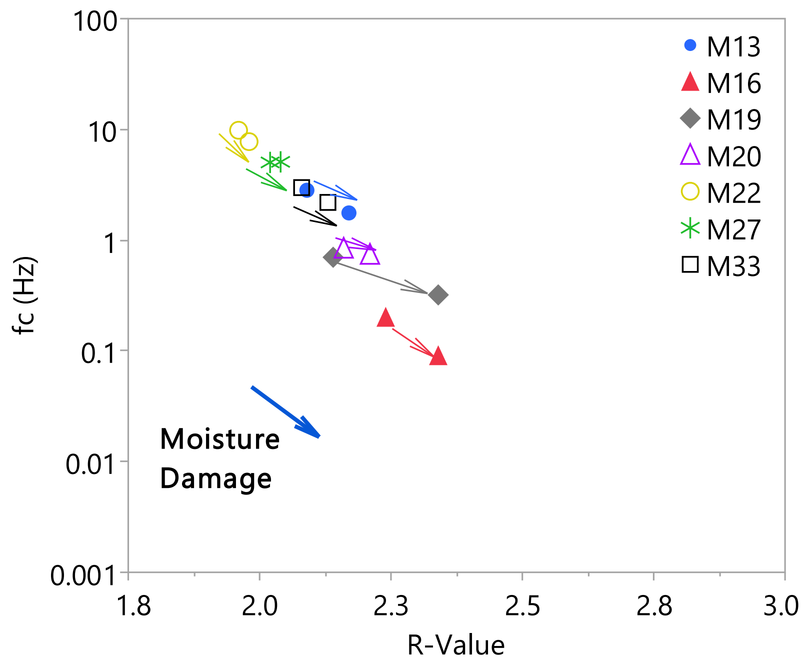
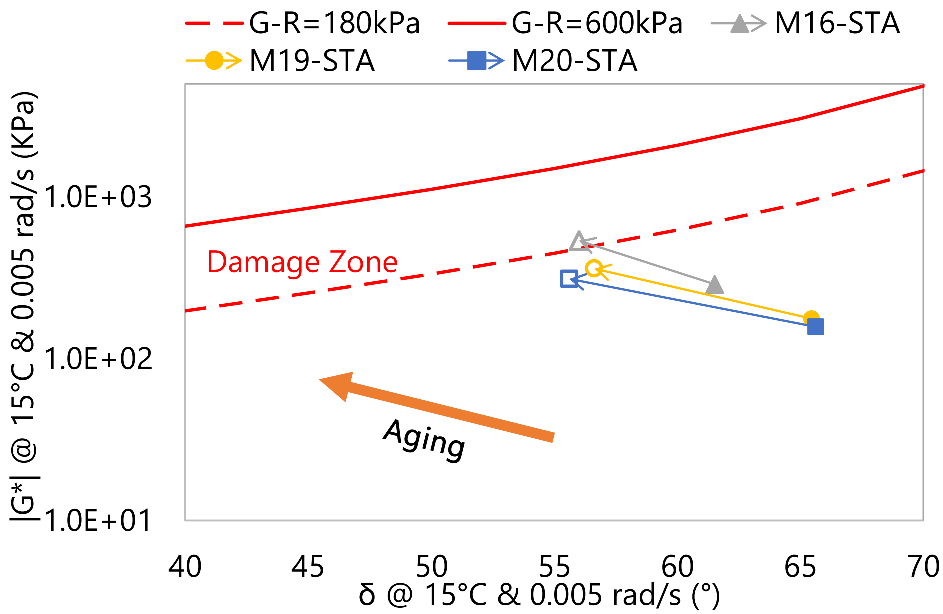


Figure 109. Crossover Frequency vs. Rheological Index for STA-Dry and STA-Wet conditions.

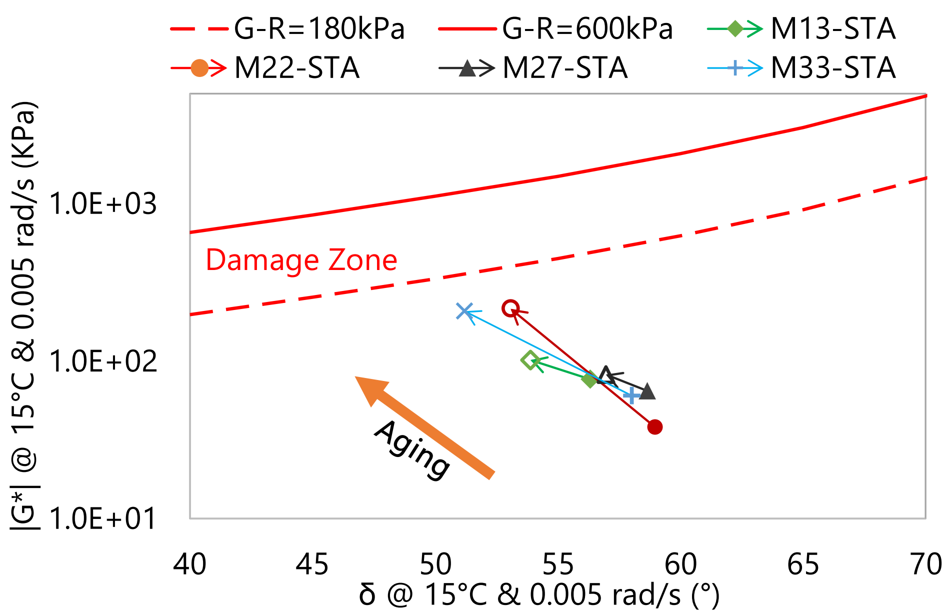
Table 34. P-value Results from One-Way ANOVA of Effect of Aging and Moisture Conditioning on the Rheological Parameters.

Condition	f_c	R-value
Aging	0.045*	0.001*
Moisture	0.709	0.259

Figure 110 and Figure 111 show the G-R parameter results on a Black Space Diagram (BSD), where the binders' |G*| values at 15°C and 0.005 rad/s are plotted on the y-axis, and δ values at the same condition is plotted on the x-axis. The damage zone is represented by two curves indicating G-R parameters of 180 KPa and 600 KPa, which signify the onset of block cracking, and ultimate failure, respectively. G-R parameters below 180 kPa generally indicate ductile performance, whereas values above 600 kPa correspond to severe embrittlement and block-cracking risk (Deef-Allah & Abdelrahman, 2025). The BSD tracks and identifies the effects of aging on the neat and modified binder as shown in Figure 110(a) and Figure 110(b), respectively. The arrows in the figures shows the effect of aging. The BSDs show that aging negatively impacts the G-R values, with a significant increase (p-value=0.008) in all binders. As binders age, the G-R parameters shift to the upper left corner (high-stiffness/low-δ), which indicates more pronounced elastic behavior (Airey, 2002).



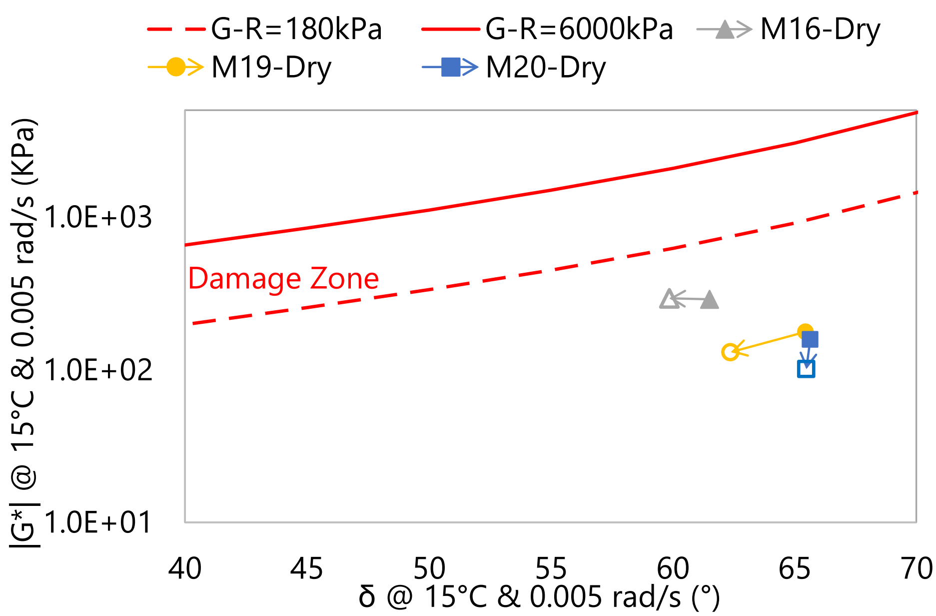
(a)



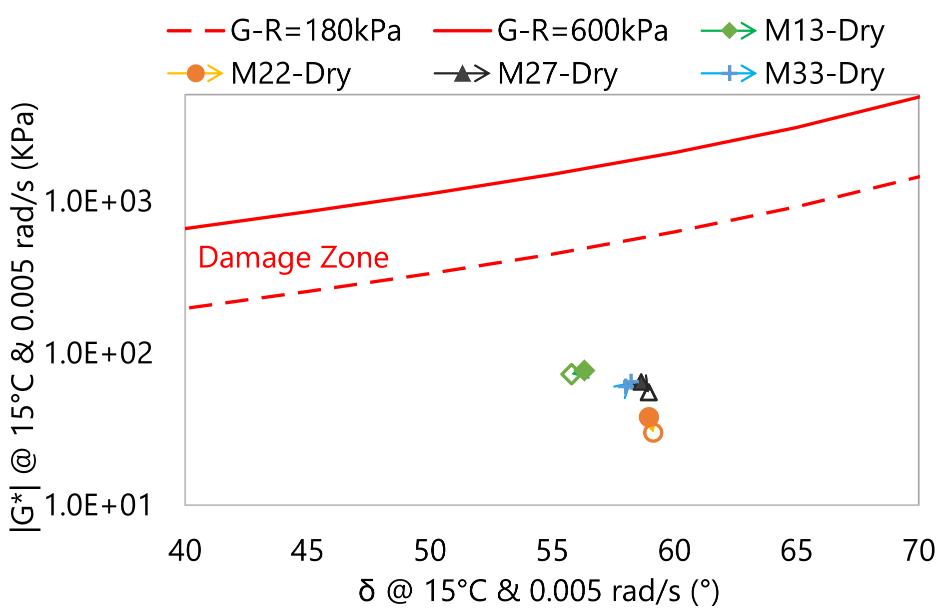
(b)

Figure 110. Binder Glover-Rowe Parameter Results to Show Effect of Aging for; (a) Neat Binders (PG64-22), and (b) Polymer-Modified Binders (PG70-28).

Figure 111 shows the effect of moisture conditioning on G-R parameter using a BSD for both neat and modified binders. The arrows in Figure 111 show the effect of moisture conditioning. It is shown that no significant impact or clear trend is observed due to moisture conditioning. This lack of effect on the G-R parameter due to moisture damage is more evident in modified binders.



(a)



(b)

Figure 111. Binder Glover-Rowe Parameter Results to Show Effect of Moisture for; (a) Neat Binders (PG64-22), and (b) Polymer-Modified Binders (PG70-28).

The G-R results of the neat and polymer-modified binders are summarized in Figure 112. The rheological analysis using the G-R parameter clearly differentiates the performance of the binder types and highlights the damaging effect of aging. The polymer-modified binders, such as M13, M22, M27, and M27, consistently exhibit better resistance to cracking, maintaining G-R values lower than the neat binders across all conditioning states. Specifically, the influence of LTA resulted in increase in the G-R parameter for all mixes about 233%, on average, indicating that oxidative aging stiffens the binder and makes it highly susceptible to fracture. Conversely, the STA-Dry and STA-Wet samples show similar G-R values, not significantly difference, suggesting that the binder properties were not affected by moisture damage, as previously noted.

All binders showed a G-R parameter less than the threshold of 180 kPa, except for the binder from mix M16 at LTA condition, which showed a G-R value of about 200 KPa; slightly above the current threshold.

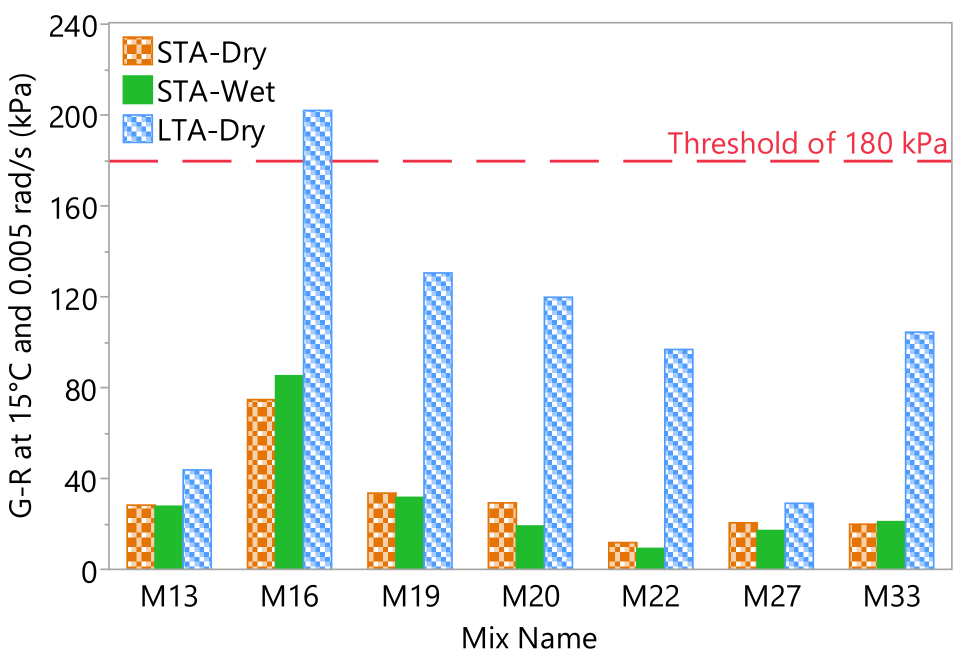


Figure 112. G-R Parameter Results of Binders under Different Conditions.

3.4 Effect of Lab Storage on Mixtures

As a side task to this study, the effect of lab storage on the mixture aging was investigated using IDEAL-CT testing. STA and LTA mixes were allowed to sit for one year in the lab and retested using IDEAL-CT. The tensile strength for the STA and LTA specimens at year 0 and year 1 are shown in Table 35. The results clearly indicate that the ITS of all mixes, except for mix M13 at LTA condition, increased with lab storage by 13%, on average, indicating the notable effect of lab storage on mix strength.

Table 35. Effect of lab storage on ITS.

Mix ID	STA	STA	STA	LTA	LTA	LTA
Mix ID	ITS-Dry (year 0)	ITS-Dry (year 1)	Change in ITS	ITS Dry (year 0)	ITS Dry (year 1)	Increase in ITS
M13	129	158	+22%	163	151	-7%
M16	163	176	+8%	176	195	+11%
M20	103	112	+9%	128	164	+28%
M22	128	138	+8%	155	183	+18%
M27	122	135	+11%	124	145	+17%
M33	126	146	+16%	152	174	+15%
Average	-	-	+12%	--		+14%

The results of the CT_{Index}, of all mixes at the STA and LTA conditions before and after lab-storage are presented in Figure 113. The results clearly show that the lab storage resulted in a notable reduction in the CT_{Index} for all mixtures at both the STA and LTA conditions. Based on the ANOVA results presented in Table 36, it is evident that the effect of lab storage on the CT_{Index} was significant.

The results of the stiffness of all mixes at the STA and LTA conditions before and after lab storage are presented in Figure 114. The results show that stiffness increased notably with lab storage for both STA and LTA specimens. Based on the statistical analysis presented in Table 36, it is clear that the effect of lab storage on the stiffness values was significant.

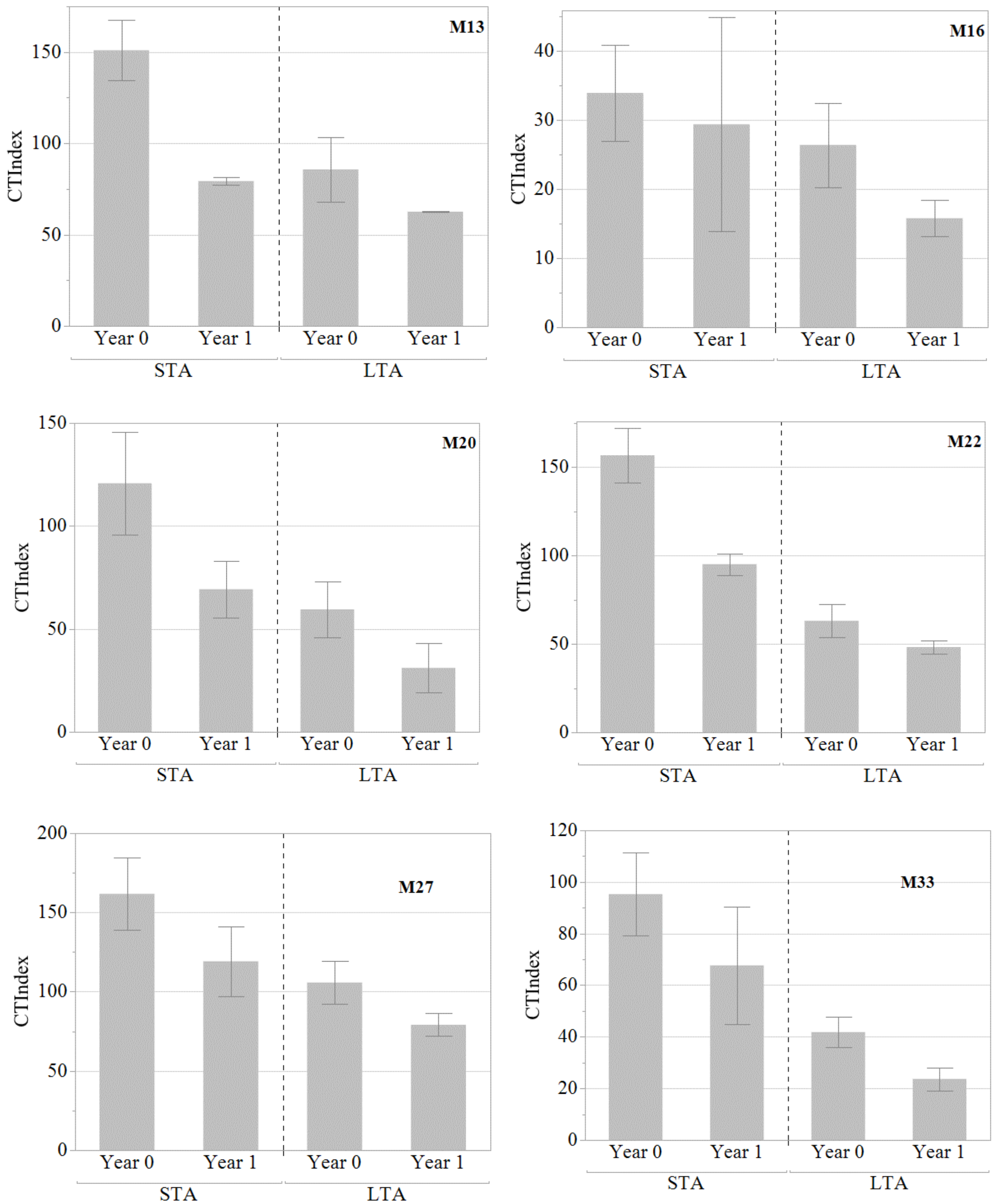


Figure 113. The Results of CT_{Index} Values for Both STA and LTA Tested Samples Before and After Lab Storage.

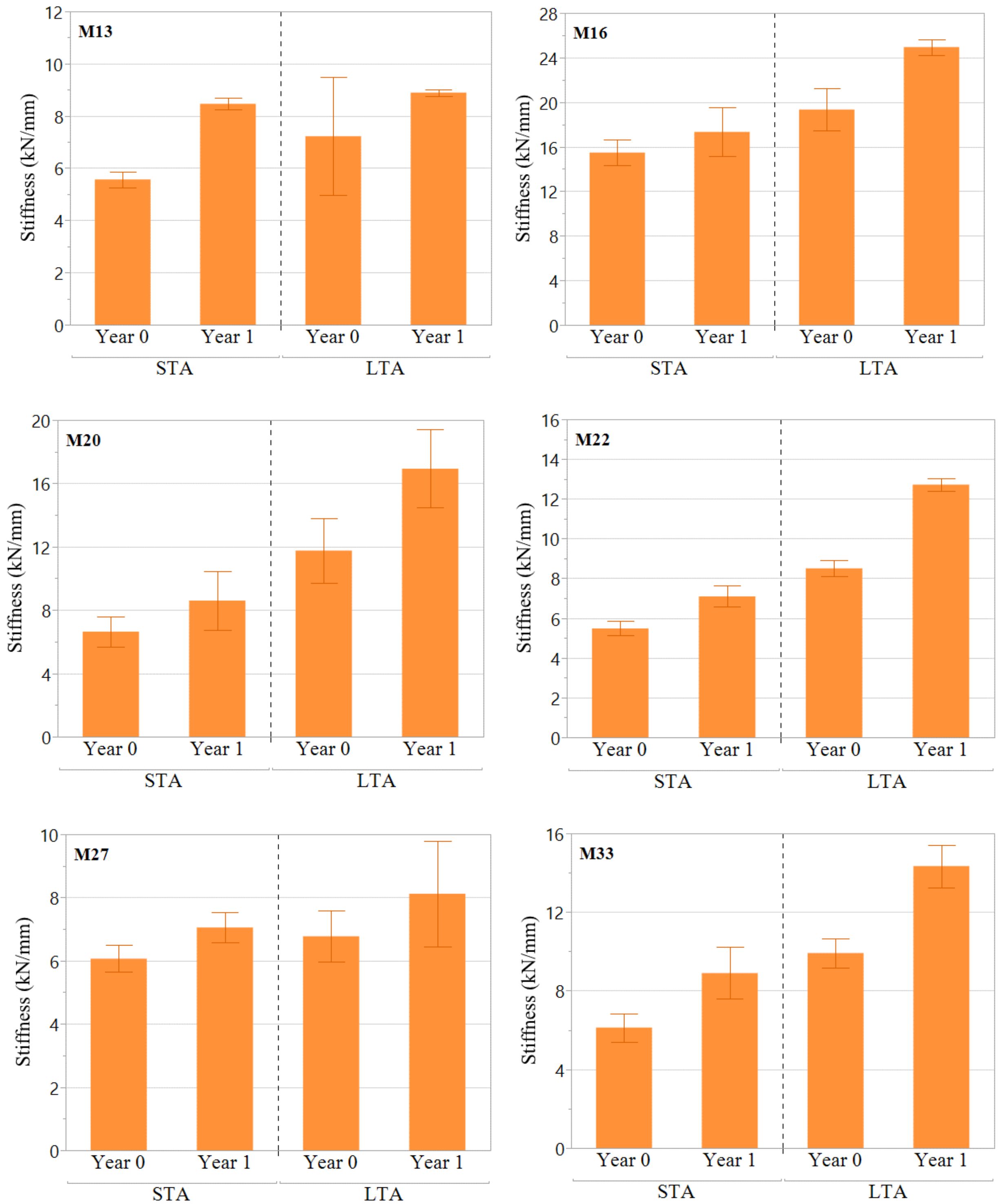


Figure 114. The Results of Stiffness values for both STA and LTA Samples Tested before and after lab storage.

Table 36. One-Way ANOVA Results of the Effect of Lab Storage on the Mixtures' Strength and Stiffness.

Parameters	**p-value**	**p-value**
Parameters	**STA**	**LTA**
CT_{Index}	0.0332*	0.0214*
Strength	0.0236*	0.0062*
Stiffness	0.0422*	0.0311*

*Statistically significant

3.5 Superpave Gyratory Compaction Parameters

3.5.1 Effect of Aging on Compaction Parameters

The Superpave Gyratory compaction data for all mixes were analyzed to determine if different compaction parameters such as CS, CEI, LP, and number of gyrations at 92% of G_{mm} can be used to provide a measure of mixture aging. Figure 115 presents violin plots illustrating the distribution of Superpave compaction parameters under two different aging conditions. Each violin plot combines a box-and-whisker representation with a Kernel Density Estimation (KDE), offering a comprehensive view of the central tendency, variability, and data distribution. Figure 115 (a) presents the CS parameter which represents the rate of densification during the linear compaction phase, reflecting how quickly the mix compacts between the initial and design gyrations (Leiva & West, 2008). It is shown that the median CS increases slightly with aging, with the LTA condition showing less variability, as indicated by a narrower interquartile range and spread. The KDE shape suggests a more peaked distribution for the LTA condition and a flatter, more dispersed pattern for the STA. A few outliers are present but do not strongly affect the overall trend. These results suggest that aging influences the spread of CS values as well as the central tendency. Figure 115 (b) represents the CEI distributions, which reflects the variation in a mixture's shear resistance during compaction (Bahia et al., 1998). It is shown that the median CEI increases with aging, however not significantly. The interquartile range under LTA is wider, and the KDE shows greater distributional spread. As shown in Figure 115(c), the LP parameter, which defined as the first instance of two consecutive gyrations resulting in the same sample height, demonstrates minimal changes between aging conditions. Both the medians and overall spread remain comparable, with only a slight increase in variability under LTA. These results suggest that LP is relatively insensitive to aging effects, and the compaction curve shape remains consistent across conditions.

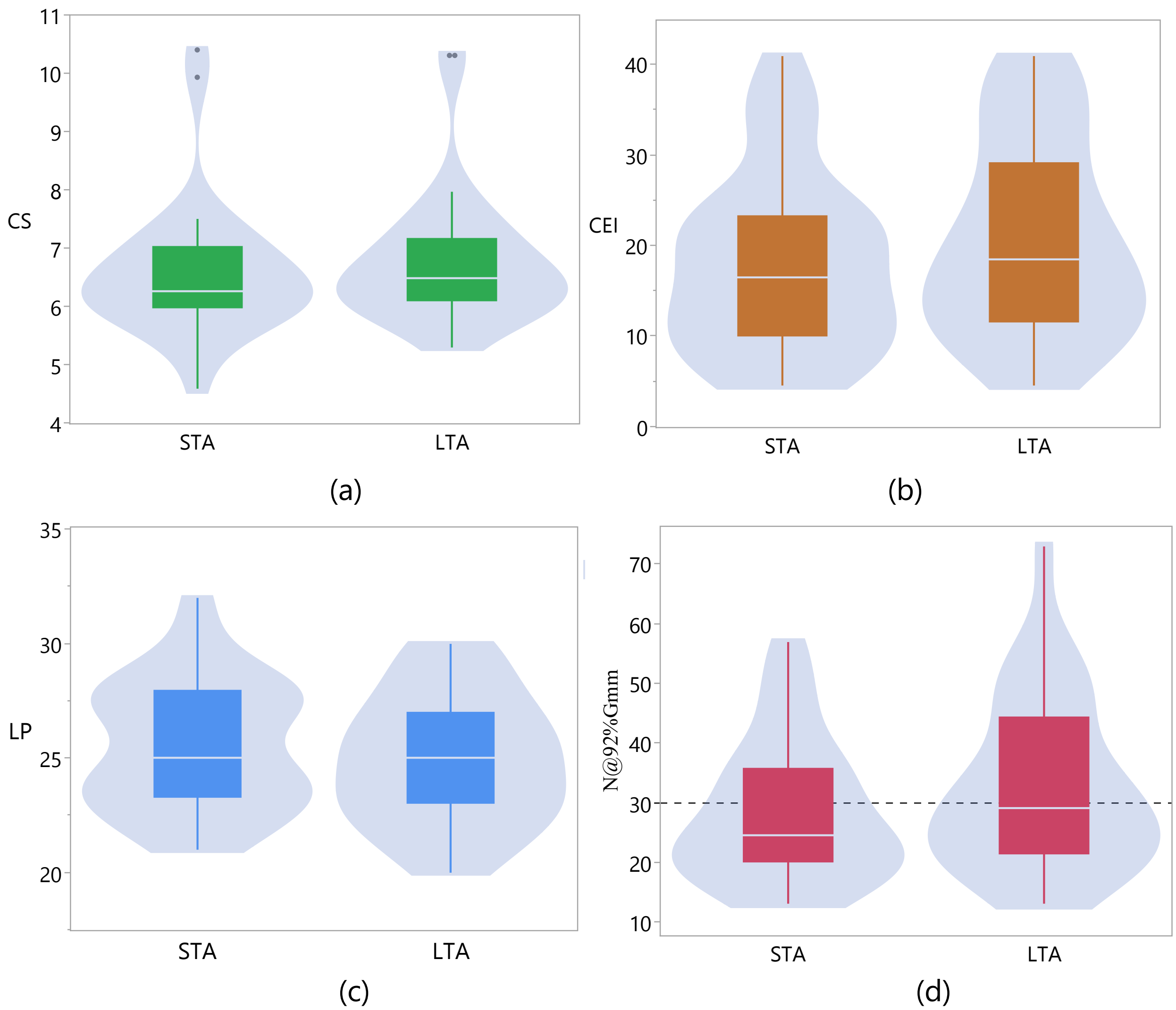


Figure 115. Boxplots of Results of; (a) CS, (b) CEI, (c) LP, and (d) N@92%G_{mm} for all Mixtures at both STA and LTA Conditions.

Figure 115 (d) highlights a notable difference in the N@92%Gmm parameter, between the STA and LTA conditions. Mixtures subjected to LTA required more gyrations to reach 92% G_{mm}, with both the median and spread increasing under LTA. This trend reflects the stiffening of the mixture and reduced compactability after extended aging. However, due to the variability in compaction parameters between aging conditions (represented by the height of the boxes), determining the statistical significance of these differences requires further rigorous statistical testing, which is discussed in the next sections.

3.5.2 Combined Effect of Aging and Mix Composition

Figure 116 and Figure 117 show the boxplots effects of mixture parameters such as binder type, aggregate gradation, RAP content and mix types on the compaction parameters under different aging conditions. Different binder grades used in this study were grouped based on modification: a) unmodified binder (PG 64-22); and b) modified binders (PG 70-28 and PG 76-28). For RAP content, mixtures were categorized as: a) without RAP (0%); and b) RAP-containing mixes (20% and 25%). Aggregate gradation was grouped by NMAS into a) coarse mixes (19.0 and 12.5 mm); and b) fine mixes (9.5 and 4.75 mm). Based on production method, mixture type was classified as: a) HMA; and b) WMA.

From Figure 116(a), For CS, modified binders show high variability under STA that decreases after LTA. Also, it can be seen that the median CS is slightly increasing from STA to LTA for modified binders. The variability in the CS for unmodified binders remains consistent between STA and LTA conditions. For CEI, aging has little effect on unmodified binders, while modified binders show increased CEI and variability after LTA, indicating higher compaction resistance due to binder stiffening. LP remains unaffected by aging for both unmodified and modified binders, indicating this parameter is less sensitive to aging. For the N@92% G_{mm} parameter, unmodified binders show no difference between STA and LTA conditions, while modified binders show slightly higher N@92% G_{mm} after LTA, reflecting reduced compactability. From Figure 116(b), course-graded mixtures consistently exhibit higher values for all compaction parameters, indicating greater resistance to compaction. These mixes require more energy and effort to reach target density. Fine-graded mixes, in contrast, compact more efficiently, with lower energy demand and fewer gyrations, making them more favorable in terms of workability.

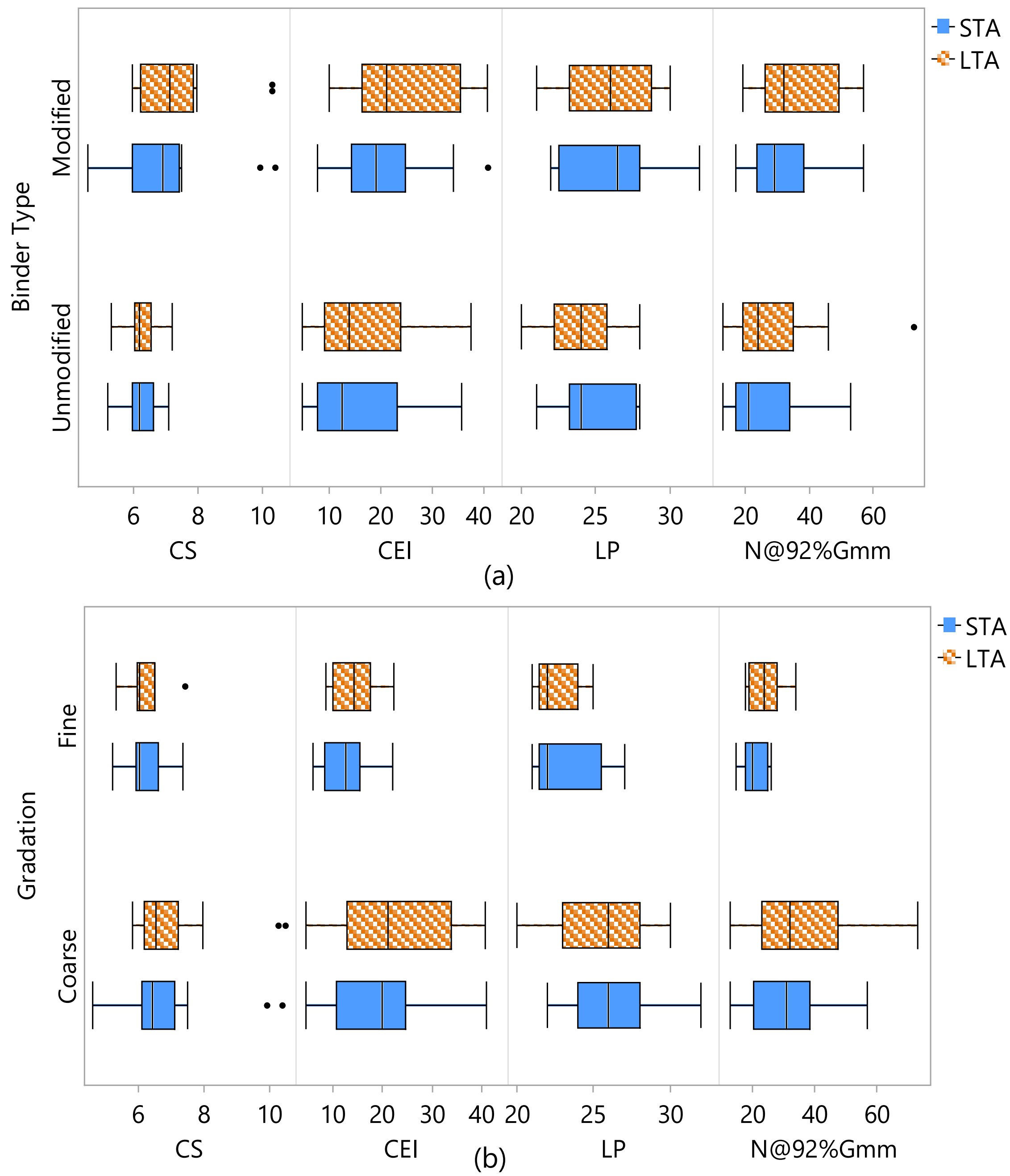


Figure 116. Compaction Parameters at STA and LTA Conditions versus; (a) Binder Type, and (b) Gradation.

It can be seen in Figure 117(a) that incorporating RAP seems to increase CEI, LP, and Gyration values, indicating reduced compactability. The presence of aged binder in the RAP stiffens the mix, making densification more difficult. CS remains mostly unchanged.

WMA mixtures in Figure 117(b) tend to show slightly lower CS, and LP compared to HMA. The CEI is similar across HMA and WMA mixes, though HMA mixes show slightly more variability.

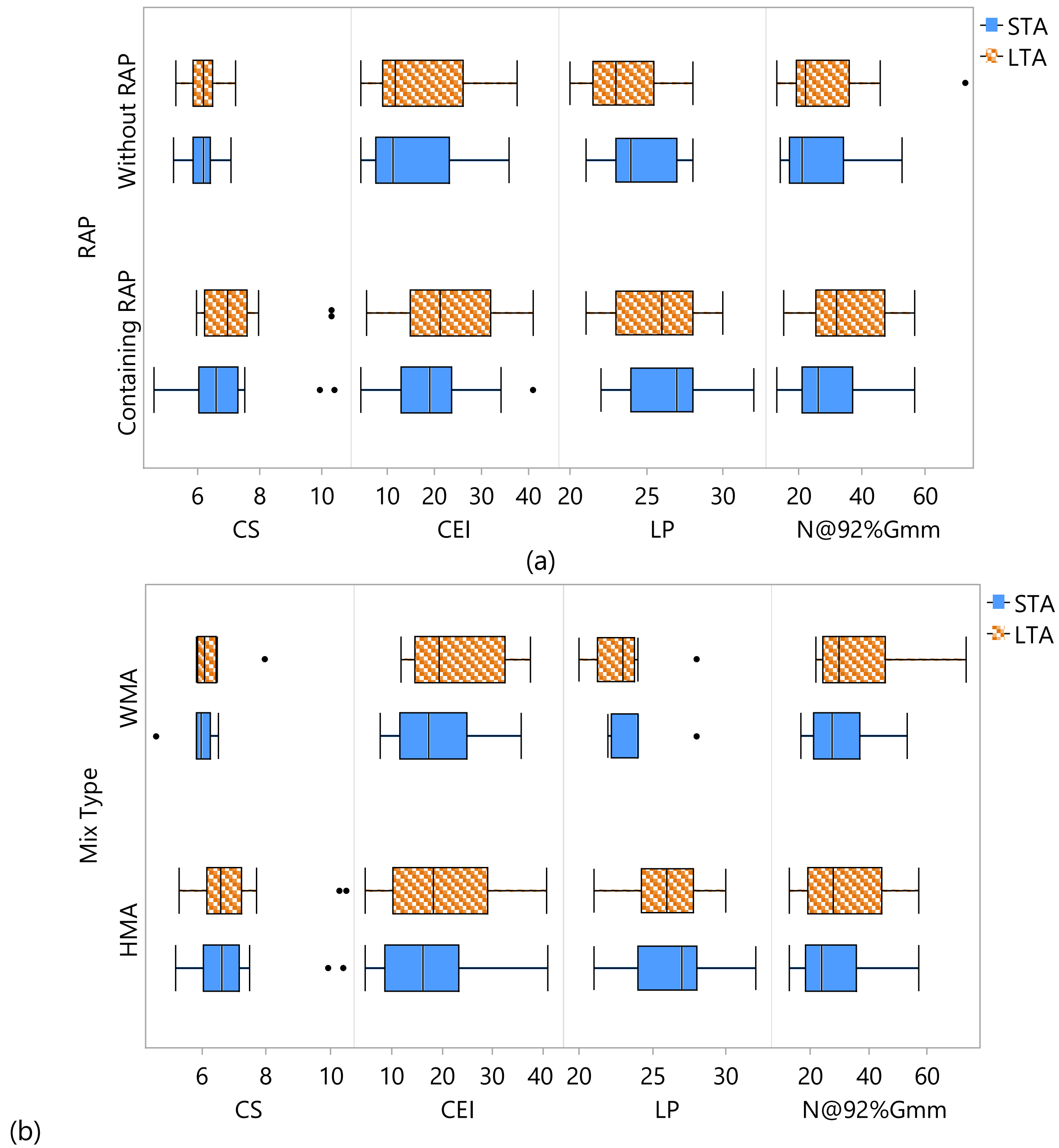


Figure 117. Compaction Parameters at STA and LTA Conditions versus; (a) RAP, and (b) Mix Type.

A two-way ANOVA was conducted to evaluate the impact of aging and various mix design variables as well as interaction of aging and mixture variables on compaction parameters, and the results are presented in Table 37. From the results, aging does not have a statistically significant effect on any of the compaction parameters. On the other hand, gradation, mix type, and RAP content appear to have significant effects on the different compaction parameters. The results also indicate that, among all mix design variables, only aggregate gradation had a statistically significant effect on all compaction parameters. The LP parameter appears to be the most sensitive, among all compaction parameters, since it was able to capture differences between mixes due to gradation, RAP content, and mix type.

Table 37. Two-Way ANOVA Results for Effects of Aging and Mix Compositions on Compaction Parameters

Variables	**Compaction Parameters**	**Compaction Parameters**	**Compaction Parameters**	**Compaction Parameters**
Variables	**CS**	**CEI**	**LP**	**N@92%G_{mm}**
Aging Effect	0.629	0.492	0.584	0.382
Binder Type	0.178	0.302	0.143	0.221
Gradation	0.024*	0.015*	<0.001*	0.013*
RAP	0.232	0.611	0.002*	0.997
Mix Type	0.007*	0.615	0.001*	0.384
Aging* Binder Type	0.939	0.686	0.551	0.971
Aging*Gradation	0.710	0.764	0.827	0.824
Aging*RAP	0.807	0.951	0.770	0.818
Aging*Mix Type	0.600	0.877	0.761	0.721

**Statistically significant*

4 SUMMARY, CONCLUSIONS, AND FUTURE WORK

4.1 Summary

This report provides a comprehensive account of the tasks conducted under this research project, which focuses on assessing the effect of aging on plant-produced asphalt mixtures. The primary aim of this research is to gain a thorough understanding of how the properties of asphalt mixes change over time, encompassing the various phases from production through the service life within the pavement structure.

The research team investigated the impact of LTA on the cracking resistance of asphalt mixtures using thirty-six asphalt mixtures, sampled from different asphalt plants in Oklahoma. The mixes contained a wide range of different raw materials, including different mix types, binders' PG, mix production technologies, and RAP content.

To evaluate the effect of aging on the cracking resistance, the CT_{Index} parameter from the IDEAL-CT test was used. To simulate field aging, a LTA conditioning protocol which involves heating the plant-produced mixes in the lab for 8 hours at 135^{o}C was used according AASHTO R121, to simulate 3-5 years of service. The change in CT_{Index} with LTA was determined and a rigorous statistical analysis using ANOVA and Welch's ANOVA was conducted to evaluate the effect of different mix variables on the aging rate. A pair-wise comparison was also performed using Games-Howell and Tukey's test to provide further insight into the comparison between the mixes. For further evaluation, PCA analysis was also employed to explain the complex interplay between these IDEAL-CT parameters and their combined influence on cracking resistance.

In addition to this, the report outlines a detailed methodology for predictive modeling, offering a detailed and systematic approach to using the mix design parameters of asphalt mixes to predict the performance of asphalt binders, with a specific focus on their cracking resistance.

This study also provides findings on the rutting and moisture susceptibility performance of asphalt mixes in Oklahoma using HWT test. It proposes limiting criteria for both rutting and moisture susceptibility for the different mixes used in Oklahoma. Additionally, the report examines the use of the IDT-HT as a surrogate QC/QA test during asphalt mix production and provides proposed limiting thresholds for the IDT-HT test.

Moreover, the study evaluates the influence of freeze-thaw moisture conditioning on the rutting and stripping of asphalt mixes. The specimens were subjected to a freeze-thaw cycle following AASHTO T283 procedures. The conditioned samples were tested using the HWTT and IDT-HT to evaluate the effect of freeze-thaw moisture conditioning on the rutting and stripping parameters.

Moreover, the combined effect of aging and moisture on the asphalt mixtures and binder was evaluated. STA and LTA mixtures were subjected to AASHTO T-283 moisture conditioning, and the impact of moisture damage was determined at both aging conditions. The mixtures were tested using IDEAL-CT and IDT to determine their cracking resistance, and strength, respectively. The mixes were subjected to extraction and recovery, and the recovered binders were tested using a DSR at different temperatures and frequencies.

To evaluate the effect of aging conditioning and mix parameters on the compactability of the plant mixes, the SGC data for both STA and LTA samples were collected. From the obtained data and compaction curves, different compaction parameters such as CS, CEI, LP, and number of gyrations required to reach 92% G_{mm} were calculated, and analyzed in relation to different mixture variables and aging.

4.2 Findings and Conclusions

4.2.1 Cracking Resistance

- Overall, the CT_{Index} of all mixes was sensitive to aging. LTA caused an average reduction in the CT_{Index} of 43%. The reduction in CT_{index} values was primarily driven by the reduction in m_{75} and l_{75}, whereas no significant change was noted in the fracture energy with aging.
- NMAS, RAP content, and binder's PG have significant effects on the CT_{Index} values of asphalt mixes under STA and LTA conditions. In contrast, the effect of mix type (HMA and WMA) on the CT_{Index} was not significant for both STA and LTA conditions.
- There is no statistical correlation between the rate of aging and mixture variables such as total AC content, binder's PG, and mix type.
- The tensile strength and CT_{Index} of mixes varied significantly within one year of lab storage, highlighting the effect of lab storage on aging.
- The ANN and PLS models provided the best prediction of CT_{Index} values compared to the other predictive models used in this study.
- PCA identified two principal components (PC1 and PC2) that captured a significant portion (94%) of the variance in the data. PC1 was primarily influenced by l_{75} and m_{75}, with less impact from fracture energy. An increase in PC1 indicated an increase in both l_{75} and fracture energy, and a decrease in m_{75}. On the other hand, PC2 showed a strong association with fracture energy with less impact from m_{75} and no impact from l_{75}. An increase in PC2 indicated a decrease in both fracture energy and m_{75}.
- The principal components were used to describe the shape of the IDEAL-CT load-displacement curve where an increase in PC1 indicated a widening of the curve's base, whereas an increase in PC2 indicated a decrease in the area under the curve

and less steep post-peak slope with minimum effect on the displacement. The CT_{Index} alone could not be used to describe the shape of the IDEAL-CT load-displacement curve.

- K-means clustering identified six distinct clusters within the mixture data. These clusters exhibited different CT_{Index} values and load-displacement curve characteristics. The clusters also revealed relationships between the CT_{Index} and the mix composition, i.e. AC, binder modification, RAP content, and mix type.

- Clusters exhibiting favorable cracking performance in long-term aged samples generally show peak loads in the range of 10–15 kN. Curves within this peak-load range tend to provide a balanced combination of fracture energy, l_{75}, and m_{75}, leading to optimal IDEAL-CT parameters and CT_{Index} values.

4.2.2 Rutting Resistance

- No clear trend was noted between the TRD during production and that during mix design.

- The CRD provides a better measure of rutting compared to TRD as it allows isolation of rutting due to plastic deformation from that due to stripping. A maximum CRD of 6 mm was proposed at 10,000 passes, 15,000 passes, and 20,000 passes for PG64-22, PG70-28 and PG76-28, respectively.

- The IDT-HT at 50^{o}C can be used as a surrogate QC/QA rutting performance test during production. A minimum IDT-HT strength of 240 kPa (35 psi) is proposed as an acceptance threshold during production for all binder types.

- The effect of AASHTO T-283 conditioning on the rutting performance was captured using both HWTT and IDT-HT, where both tests showed similar trends with freeze-thaw damage. Most mixes showed reduced or comparable rutting resistance with freeze-thaw.

4.2.3 Moisture Resistance

- A moisture susceptibility performance criterion requiring a minimum SN of 2,000-wheel passes, which coincides with a SIP of 10,000 passes, is proposed for Oklahoma. When applied to the 33 PMLC mixes, 87% of the mixes pass this criterion.
- There was no correlation between the TSR from the AASHTO T 283 and the SIP from the HWTT.
- A strong exponential correlation was found between the SN and SIP parameters, whereas a reasonable power relationship was found between the SN and SS parameters.
- The effect of freeze-thaw damage on stripping was evident for some mixes which showed a lower SN and SIP after AASHTO T-283 conditioning.
- Aging affected the moisture sensitivity of mixes; however no clear trend was noted. LTA significantly increased the G-R parameter for all binders by about 233%, highlighting its damaging effect on cracking resistance.
- Moisture conditioning led to a significant increase in CT_{Index}. This increase was attributed to mixes getting softer resulting in more deformation, and higher air voids due to swelling. Hence, CT_{Index} parameter is not considered a suitable measure of moisture resistance.
- Moisture conditioning at the mixture level caused significant reduction in tensile strength. However, moisture conditioning at the binder level did not show a significant detrimental effect on either the complex modulus, phase angle, or the G-R parameter, suggesting no appreciable change in the binders' stiffness or cracking resistance due to moisture exposure.

- Polymer-modified binders showed extra relaxation mechanisms, seen as a shoulder-like trend in the phase angle master curves at low frequencies. This behavior indicates higher elasticity and a better viscous-elastic balance compared to neat binders. G-R results also showed that polymer-modified binders consistently maintain better resistance to cracking than neat binders across all conditions.

4.2.4 Compaction Effort

- Aging affected compactability, with LTA generally leading to higher CEI values and an increased number of gyrations at 92% G_{mm}. However, statistical analysis showed that these changes were not significant across all compaction parameters, suggesting a moderate but inconclusive influence of aging under the conditions evaluated.
- Mix compositions had a more pronounced and statistically significant impact on compaction, compared to the effect of aging. Coarse-graded aggregates and RAP content increased resistance to compaction.
- Among all mix parameters studied in this research, aggregate gradation had the strongest influence on asphalt mixture compaction.

ACKNOWLEDGMENTS

The author would like to acknowledge the support of the Office of Research and Implementation and the Materials Division at the Oklahoma Department of Transportation. The author would also like to thank Lincoln Beard, Dennis Muncy, and Richard Steger from Ingevity for their help with the HWTT.

REFERENCES

AASHTO. (2011a). AASHTO T 324 Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Hot Mix Asphalt (HMA).

AASHTO. (2011b). Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Hot Mix Asphalt (HMA) (AASHTO T 324). Standard Specifications for Transportation Materials and Methods of Sampling and Testing.

AASHTO. (2012). Standard Method of Test for Determining the Rheological Properties of Asphalt Binder Using a Dynamic Shear Rheometer (DSR) (AASHTO T 315). American Association of State Highway and Transportation Officials (AASHTO).

AASHTO. (2014). Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage (AASHTO T 283). American Association of State Highway and Transportation Officials (AASHTO).

AASHTO. (2019). Standard Practice for Reducing Samples of Asphalt Mixtures to Testing Size (AASHTO R 47). American Association of State Highway and Transportation Officials (AASHTO).

AASHTO. (2022). Standard Method of Test for Percent Air Voids in Compacted Dense and Open Asphalt Mixtures (AASHTO T 269-14). American Association of State Highway and Transportation Officials (AASHTO).

AASHTO. (2024). Standard Practice for Long-Term Laboratory Conditioning of Asphalt Mixtures (AASHTO R 121). American Association of State Highway and Transportation Officials (AASHTO).

Airey, G. D. (2002). Use of Black Diagrams to Identify Inconsistencies in Rheological Data. Road Materials and Pavement Design, 3(4), 403–424.

Albayati, A. H., & Abduljabbar, M. H. (2019). The Simulation of Short-Term Aging Based on the Moisture Susceptibility of Asphalt Concrete Mixtures. Results in Engineering, 2, 100012.

Ali, S. A., Zaman, M., Ghabchi, R., Rahman, M. A., Ghos, S., & Rani, S. (2022). Effect of Additives and Aging on Moisture-Induced Damage Potential of Asphalt Mixes Using Surface Free Energy and Laboratory-Based Performance Tests. International Journal of Pavement Engineering, 23(2), 285–296. https://doi.org/10.1080/10298436.2020.1746813

Amelian, S., Abtahi, S. M., & Hejazi, S. M. (2014). Moisture Susceptibility Evaluation of Asphalt Mixes Based on Image Analysis. Construction and Building Materials, 63, 294–302.

Anderson, D. A., Christensen, D. W., Bahia, H. U., Dongre, R., Sharma, M. G., Antle, C. E., & Button, J. (1994a). Binder characterization and evaluation. Volume 3: Physical characterization. Strategic Highway Research Program, National Research Council.

Anderson, D. A., Christensen, D. W., Bahia, H. U., Dongre, R., Sharma, M. G., Antle, C. E., & Button, J. (1994b). Binder Characterization and Evaluation. Volume 3: Physical Characterization. Strategic Highway Research Program, National Research Council.

Anderson, R. M., King, G. N., Hanson, D. I., & Blankenship, P. B. (2011a). Evaluation of the Relationship Between Asphalt Binder Properties and Non-Load Related Cracking. Journal of the Association of Asphalt Paving Technologists, 80, 615–664.

Anderson, R. M., King, G. N., Hanson, D. I., & Blankenship, P. B. (2011b). Evaluation of the Relationship between Asphalt Binder Properties and Non-Load Related Cracking. Journal of the Association of Asphalt Paving Technologists, 80. https://trid.trb.org/View/1135500

Arabani, M., & Majd Rahimabadi, M. (2023). Assessing the Effect of Short and Long-Term Aging on Moisture Damage of Hot Mix Asphalt Using Two Different Methods. International Journal of Pavement Engineering, 24(1), 2168661.

Asgharzadeh, S. M., Tabatabaee, N., Naderi, K., & Partl, M. (2013). An Empirical Model for Modified Bituminous Binder Master Curves. Materials and Structures, 46(9), 1459–1471.

ASTM. (2019a). Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature (ASTM D8225-19). ASTM International.

ASTM. (2019b). Standard Test Methods for Automated Extraction of Asphalt Binder from Asphalt Mixtures (ASTM D8159). American Society for Testing and Materials.

ASTM. (2021). Standard Practice for Recovery of Asphalt Binder from Solution Using the Rotary Evaporator (ASTM D5404/D5404M). American Society for Testing and Materials.

Bahia, H. U., Friemel, T. P., Peterson, P. A., Russell, J. S., & Poehnelt, B. (1998). Optimization of constructibility and resistance to traffic: A new design approach for HMA using the Superpave compactor. Journal of the Association of Asphalt Paving Technologists, 67. https://trid.trb.org/View/542933

Bennert, T., Haas, E., & Wass, E. (2018). Indirect Tensile Test (IDT) to Determine Asphalt Mixture Performance Indicators during Quality Control Testing in New Jersey. Transportation Research Record, 2672(28), 394–403. https://doi.org/10.1177/0361198118793276

Bennert, T., Haas, E., Wass, E., & Berger, B. (2021). Indirect Tensile Testing for Balanced Mixture Design and Quality Control Performance Testing. Asphalt Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions, 89, 363–389. https://www.researchwithrutgers.com/en/publications/indirect-tensile-testing-for-balanced-mixture-design-and-quality-

Bennert, T., Haas, E., Wass, E., Tulanowski, D., & Zavery, Z. (2022). Statewide Assessment of Balanced Mixture Design for New York State's Asphalt Mixtures. Transportation Research Record: Journal of the Transportation Research Board, 2676(11), 340–351.

Bennert, T., Sheehy, E., Hanson, D., Gresavage, S., & Fee, F. (2014). Implementation of Performance-Based Specifications for Asphalt Mix Design and Production Quality Control for New Jersey. Application of Asphalt Mix Performance-Based Specifications, 13, 13–31.

Bhalla, D. (2017). Regression: Transform Negative Values. https://www.listendata.com/2015/09/regression-transform-negative-values.html

Box, G. E., & Cox, D. R. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society Series B: Statistical Methodology, 26(2), 211–243.

Boz, I., Habbouche, J., Diefenderfer, S. D., Coffey, G., Seitllari, A., & Ozbulut, O. E. (2023). Evaluating the Rutting Potential of Asphalt Mixtures with Simple and Practical Tests. Transportation Research Record, 03611981231207089.

Brownlee, J. (2016). Parametric and Non-parametric Machine Learning Algorithms. Machine Learning Mastery. https://machinelearningmastery.com/parametric-and-nonparametric-machine-learning-algorithms/

Caro, S., Masad, E., Bhasin, A., & Little, D. N. (2008). Moisture Susceptibility of Asphalt Mixtures, Part 1: Mechanisms. International Journal of Pavement Engineering, 9(2), 81–98. https://doi.org/10.1080/10298430701754341

Carroll, R. J., & Schneider, H. (1985). A Note on Levene's Tests for Equality of Variances. Statistics and Probability Letters, 3(4), 191–194.
Chao, J., & Jinxi, Z. (2018). Prediction Model for Asphalt Pavement Temperature in High-Temperature Season in Beijing. Advances in Civil Engineering, 2018(1), 1837952. https://doi.org/10.1155/2018/1837952
Chen, C. (2020). Validation of Laboratory Cracking Tests for Field Top-Down Cracking Performance [PhD Thesis]. Auburn University.
Chen, C., Yin, F., Turner, P., West, R. C., & Tran, N. (2018a). Selecting a Laboratory Loose Mix Aging Protocol for the NCAT Top-Down Cracking Experiment. Transportation Research Record, 28(2672), 359–371.
Chen, C., Yin, F., Turner, P., West, R. C., & Tran, N. (2018b). Selecting a Laboratory Loose Mix Aging Protocol for the NCAT Top-Down Cracking Experiment. Transportation Research Record, 2672(28), 359–371.
Chen, H., Adwani, D., Zorigtbaatar, N., Zhou, F., & Karki, P. (2025). Assessment of Binder Test Methods for Detecting Cracking Susceptibility in Asphalt Mixtures. International Journal of Fatigue, 194, 108827.
Chen, X., & Huang, B. (2008). Evaluation of Moisture Damage in Hot Mix Asphalt Using Simple Performance and Superpave Indirect Tensile Tests. Construction and Building Materials, 22(9), 1950–1962.
Chen, Z., Zhang, H., Duan, H., & Shi, C. (2021). Improvement of Thermal and Optical Responses of Short-Term Aged Thermochromic Asphalt Binder by Warm-Mix Asphalt Technology. Journal of Cleaner Production, 279, 123675.
Christensen, D., & Bonaquist, R. (2007). Using the Indirect Tension Test to Evaluate Rut Resistance in Developing Hot-Mix Asphalt Mix Designs. In Practical approaches to hot-mix asphalt mix design and production quality control testing (Vol. 62).
Christensen, D., Mensching, D., Rowe, G., Anderson, R. M., Hanz, A., Reinke, G., & Anderson, D. (2019). Past, Present, and Future of Asphalt Binder Rheological Parameters: Synopsis of 2017 Technical Session 307 at the 96th Annual Meeting of the Transportation Research Board. Transportation Research Circular, E-C241. https://trid.trb.org/View/1601364
Christensen, D. W. (2023). Interlayer Mixture Design (No. WHRP 0092-21-04). Wisconsin Department of Transportation, Research and Library Unit.
Christensen, D. W., & Anderson, D. A. (1992). Interpretation of Dynamic Mechanical Test Data for Paving Grade Asphalt Cements (With Discussion). Journal of the Association of Asphalt Paving Technologists, 61, 67–116.
Christensen, D. W., Anderson, D. A., & Rowe, G. M. (2017). Relaxation Spectra of Asphalt Binders and the Christensen–Anderson Rheological Model. Road Materials and Pavement Design, 18(sup1), 382–403.
Christensen, D. W., Bonaquist, R., & Jack, D. P. (2000). Evaluation of Triaxial Strength as a Simple Test for Asphalt Concrete Rut Resistance (FHWAPA-2000-010+97-04(19)). Federal Highway Administration.
Chugh, A. (2020). MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared – Which Metric Is Better? Analytics Vidhya. https://medium.com/analytics-vidhya/parametric-and-nonparametric-models-in-machine-learning-a9f63999e233
Das, K. R., & Imon, A. (2016). A Brief Review of Tests For Normality. American Journal of Theoretical and Applied Statistics, 5(1), 5–12.

Deef-Allah, E., & Abdelrahman, M. (2025). Quality Assurance Framework for Recovered Binders and Aggregates From Asphalt Mixtures Incorporating Recycled Materials. Recycling, 10(2), 71.

Diab, A., Enieb, M., & Singh, D. (2019). Influence of Aging on Properties of Polymer-Modified Asphalt. Construction and Building Materials, 196, 54–65. https://doi.org/10.1016/j.conbuildmat.2018.11.105

Diefenderfer, S. D., & Bowers, B. F. (2019). Initial Approach to Performance (Balanced) Mix Design: The Virginia Experience. Transportation Research Record: Journal of the Transportation Research Board, 2673(2), 335–345. https://doi.org/10.1177/0361198118823732

Elkashef, M., Harvey, J., Jones, D., & Jiao, L. (2022). The Impact of Silo Storage on the Fatigue and Cracking Resistance of Asphalt Mixes. Journal of Construction and Building Materials, 326, 126880.

Esposito Vinzi, V., & Russolillo, G. (2013). Partial Least Squares Algorithms and Methods. WIREs Computational Statistics, 5(1), 1–19. https://doi.org/10.1002/wics.1239

Faheem, A. F., Bahia, H. U., & Ajideh, H. (2005). Estimating Results of a Proposed Simple Performance Test for Hot-Mix Asphalt From Superpave Gyratory Compactor Results. Transportation Research Record, 1929(1), 104–113.

Farrokhzade, F., Sabouri, M., & Tabatabaee, N. (2022). Aging Characteristics of Neat and Modified Asphalt Binders Based on Rheological Evaluations at Intermediate Temperatures. Construction and Building Materials, 322, 126387.

Fathi, A., Mazari, M., Saghafi, M., Hosseini, A., & Kumar, S. (2019). Parametric Study of Pavement Deterioration Using Machine Learning Algorithms. Airfield and Highway Pavements 2019, 31–41. https://doi.org/10.1061/9780784482476.004

Fisher, R. (1925). Statistical Methods for Research Workers. Oliver and Boyd.

Frost, J. (2018). Benefits of Welch's ANOVA Compared to the Classic One-Way ANOVA. Statistics By Jim: Making Statistics Intuitive. https://statisticsbyjim. com/anova/welchs-anova-compared-to-classic-one-wayanova

Games, P. A., & Howell, J. F. (1976). Pairwise Multiple Comparison Procedures with Unequal N's and/or Variances: A Monte Carlo Study. Journal of Educational Statistics, 1(2), 113–125. https://doi.org/10.3102/10769986001002113

Gawali, S. (2021). Skewness and Kurtosis: Quick Guide. https://www.analyticsvidhya.com/blog/2021/05/shape-of-data-skewness-and-kurtosis/

Ghanizadeh, A. R., Safi Jahanshahi, F., Khalifeh, V., & Jalali, F. (2020). Predicting Flow Number of Asphalt Mixtures Based on the Marshall Mix Design Parameters Using Multivariate Adaptive Regression Spline (MARS). International Journal of Transportation Engineering, 7(4), 433–448.

Glover, C., Davison, R. R., Domke, C. H., Ruan, Y., Juristyarini, P., Knorr, D. B., & Jung, S. H. (2005). Development of a New Method for Assessing Asphalt Binder Durability With Field Validation [Publication FHWA/TX-05/1872-2, Report 0-1872-2]. U.S. Department of Transportation.

Gopalakrishnan, K., & Kim, S. (2011). Support Vector Machines Approach to HMA Stiffness Prediction. Journal of Engineering Mechanics, 137(2), 138–146. https://doi.org/10.1061/(ASCE)EM.1943-7889.0000214

Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2018). Use of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) for

Multivariate Association Between Bioactive Compounds and Functional Properties in Foods: A Critical Perspective. Trends in Food Science & Technology, 72, 83–90.
Grilli, A., Gnisci, M. I., & Bocci, M. (2017). Effect of Ageing Process on Bitumen and Rejuvenated Bitumen. Construction and Building Materials, 136, 474–481.
Gul, M. A., Islam, M. K., Awan, H. H., Sohail, M., Al Fuhaid, A. F., Arifuzzaman, M., & Qureshi, H. J. (2022). Prediction of Marshall stability and Marshall flow of asphalt pavements using supervised machine learning algorithms. Symmetry, 14(11), 2324.
Guo, N., You, Z., Zhao, Y., Tan, Y., & Diab, A. (2014). Laboratory Performance of Warm Mix Asphalt Containing Recycled Asphalt Mixtures. Construction and Building Materials, 64, 141–149. https://doi.org/10.1016/j.conbuildmat.2014.04.002
Gurudath, S. (2020). Parametric and Non-Parametric Models in Machine Learning. https://medium.com/analytics-vidhya/parametric-and-nonparametric-models-in-machine-learning-a9f63999e233
Hamzah, M. O., Kakar, M. R., & Hainin, M. R. (2015). An Overview of Moisture Damage in Asphalt Mixtures. Journal Teknologi (Sciences & Engineering), 73(4), 125–131. https://doi.org/10.11113/jt.v73.4479
Hawraa, K., & Baaj, H. (2018). Evaluation of the Impact of Silo Storage on Thermal Cracking of the Hot Mix Asphalt With RAP. TAC: Transportation Association of Canada.
Heidel, E. (2023). Transformations for ANOVA. https://www.scalestatistics.com/transformations-for-anova.html
Hu, W., Jia, X., Huang, B., & Park, H. (2017). Evaluation of Compactability of Asphalt Mixture Utilizing Asphalt Vibratory Compactor. Construction and Building Materials, 139, 419–429.
Huang, B., Shu, X., Dong, Q., & Shen, J. (2010). Laboratory Evaluation of Moisture Susceptibility of Hot-Mix Asphalt Containing Cementitious Fillers. Journal of Materials in Civil Engineering, 22(7), 667–673. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000047
Ikeagwuani, C. C. (2022). Determination of Unbound Granular Material Resilient Modulus with MARS, PLSR, KNN and SVM. International Journal of Pavement Research and Technology, 15(4), 803–820. https://doi.org/10.1007/s42947-021-00054-w
Ishaq, M. A., & Giustozzi, F. (2022). Correlation Between Rheological Tests on Bitumen and Asphalt Low Temperature Cracking Tests. Construction and Building Materials, 320, 126109.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning with Applications in R (2nd ed.). Springer.
Kamboozia, N., Saed, S. A., & Rad, S. M. (2021). Rheological Behavior of Asphalt Binders and Fatigue Resistance of SMA Mixtures Modified With Nano-Silica Containing RAP Materials Under the Effect of Mixture Conditioning. Construction and Building Materials, 303, 124433.
Kaseer, F., Cucalon, L. G., Arámbula-Mercado, E., Martin, A. E., & Epps, J. (2018). Practical Tools for Optimizing Recycled Materials Content and Recycling Agent Dosage for Improved Short- and Long-Term Performance of Rejuvenated Binder Blends and Mixtures. Journal of the Association of Asphalt Paving Technologists, 87, 513–550.
Khosla, N. P., & Harikrishnan, K. I. (2007). Tensile strength – A design and evaluation tool for Superpave mixtures (FHWA/NC/2006-24). The North Carolina Department of Transportation.

Kim, J.-H. (2009). Estimating Classification Error Rate: Repeated Cross-Validation, Repeated Hold-Out and Bootstrap. Computational Statistics & Data Analysis, 53(11), 3735–3745.

Korstanje, J. (2021). Partial Least Squares: A Deep-Dive Into Partial Least Squares Regression and Partial Least Squares Discriminant Analysis With Full Worked-Out Examples in R and Python. https://towardsdatascience.com/partial-least-squares-deep-dive-123456

Kriz, P., Grant, D., Gale, M., Ontario, S., Lavorato, S., & Pahalan, A. (2014). Reclaimed Asphalt Pavement-Virgin Binder Diffusion in Asphalt Mixes. 59th Annual Conference of the Canadian Technical Asphalt Association (CTAA).

Kuhn, M. (2015). Caret: Classification and regression training. Astrophysics Source Code Library, Ascl-1505.

Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Leiva, F., & West, R. C. (2008). Analysis of Hot-Mix Asphalt Lab Compactability Using Lab Compaction Parameters and Mix Characteristics. Transportation Research Record, 2057(1), 89–98. https://doi.org/10.3141/2057-11

Levene, H. (1960). Robust Tests for Equality of Variances. In I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, & H. B. Mann (Eds.), Contributions to Probability and Statistics, Essays in Honor of Harold Hotelling (pp. 278–292).

Li, B., Yang, J., Li, X., Liu, X., Han, F., & Li, L. (2015). Effect of Short-Term Aging Process on the Moisture Susceptibility of Asphalt Mixtures and Binders Containing Sasobit Warm Mix Additive. Advances in Materials Science and Engineering. https://doi.org/10.1155/2015/645135

Li, Q., Song, S., Wang, J., Xu, A., Li, X., & Li, B. (2022). Characterization of Residual Moisture Distribution in Cold-Recycled Mixtures with Foamed Asphalt Using Hyperspectral Technology. Journal of Materials in Civil Engineering, 34(11), 04022288.

Lin, P., Yan, C., Huang, W., Li, Y., Zhou, L., Tang, N., Lv, Q., & others. (2019). Rheological, Chemical and Aging Characteristics of High Content Polymer Modified Asphalt. Construction and Building Materials, 207, 616–629.

Ling, C., Hanz, A., & Bahia, H. (2016). Measuring Moisture Susceptibility of Cold Mix Asphalt with a Modified Boiling Test Based on Digital Imaging. Construction and Building Materials, 105, 391–399. https://doi.org/10.1016/j.conbuildmat.2015.12.093

Ling, M., Luo, X., Hu, S., Gu, F., & Lytton, R. L. (2017). Numerical Modeling and Artificial Neural Network for Predicting J-Integral of Top-Down Cracking in Asphalt Pavement. Transportation Research Record, 2631(1), 83–95.

Liu, J., Liu, F., Zheng, C., Zhou, D., & Wang, L. (2022). Optimizing Asphalt Mix Design Through Predicting the Rut Depth of Asphalt Pavement Using Machine Learning. Construction and Building Materials, 356, 129211.

López-Montero, T., Crucho, J., Picado-Santos, L., & Miró, R. (2018). Effect of Nanomaterials on Ageing and Moisture Damage Using the Indirect Tensile Strength Test. Construction and Building Materials, 168, 31–40.

Lottman, R. P. (1978). Predicting Moisture-Induced Damage to Asphaltic Concrete (NCHRP Report No. 192). Transportation Research Board.

Majidifard, H., Jahangiri, B., Buttlar, W. G., & Alavi, A. H. (2019). New Machine Learning-based Prediction Models for Fracture Energy of Asphalt Mixtures. Measurement, 135, 438–451.

Margaritis, A., Soenen, H., Fransen, E., Pipintakos, G., Jacobs, G., Blom, J., & Van Den Bergh, W. (2020). Identification of Ageing State Clusters of Reclaimed Asphalt Binders Using PCA and HCA Based on Chemo-Rheological Parameters. Construction and Building Materials, 244, 118276.

Martin, A. E., Mercado, E. A., Gil, J. M., Kim, Y.-R., Kommidi, S. R., Tran, N., Yin, F., Moraes, R., Rodezno, C., Verma, M., Bairgi, B. K., & King, G. (with National Cooperative Highway Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine). (2024). Capturing the Durability of High Recycled Binder Ratio Asphalt Mixtures. National Academies Press. https://doi.org/10.17226/27991

Maschauer, D., Mirwald, J., & Hofko, B. (2022). Viennese Ageing Procedure (VAPro): Adaptions and Further Development to Address Low-Temperature Performance of Aged Asphalt Mixtures. Road Materials and Pavement Design, 23(sup1), 147–161. https://doi.org/10.1080/14680629.2022.2028272

Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. Journal of Applied Science and Technology Trends, 1(2), 140–147.

Menon, K. (2023). The Complete Guide to Skewness and Kurtosis. https://www.simplilearn.com/tutorials/statistics-tutorial/skewness-and-kurtosis

Mensching, D. J., Jacques, C. D., & Daniel, J. S. (2016). Applying the Glover-Rowe Parameter to Evaluate Low-Temperature Performance of Hot Mix Asphalt LTPP Sections. Journal of Materials in Civil Engineering, 28(10), 04016096. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001695

Mensching, D. J., Rowe, G. M., Daniel, J. S., & Bennert, T. (2015). Exploring Low-Temperature Performance in Black Space. Road Materials and Pavement Design, 16(sup2), 230–253. https://doi.org/10.1080/14680629.2015.1029693

Meroni, F., Flintsch, G. W., Habbouche, J., Diefenderfer, B. K., & Giustozzi, F. (2021). Three-Level Performance Evaluation of High RAP Asphalt Surface Mixes. Construction and Building Materials, 309, 125164.

Mirzaiyanrajeh, D., Dave, E. V., Sias, J. E., & Ramsey, P. (2022). Developing a Prediction Model for Low Temperature Fracture Energy of Asphalt Mixtures using Machine Learning Approach. International Journal of Pavement Engineering, 24(2), 2024185.

Mohammad, L., & Kim, M. (2011). Development of Performance-Based Specifications for Louisiana Asphalt Mixtures. https://rosap.ntl.bts.gov/view/dot/23840/dot_23840_DS1.pdf

Molenaar, A., Scarpas, A., Liu, X., & Erkens, S. (2002). Semi-Circular Bending Test; Simple But Useful? Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions, 71, 794–815.

Molinaro, A. (2005). Prediction Error Estimation: A Comparison of Resampling Methods. Bioinformatics, 21(15), 3301–3307.

Moraes, R., Velasquez, R., & Bahia, H. (2017). Using Bond Strength and Surface Energy to Estimate Moisture Resistance of Asphalt-Aggregate Systems. Construction and Building Materials, 130, 156–170. https://doi.org/10.1016/j.conbuildmat.2016.10.043

Moraes, R., Yin, F., Chen, C., Andriescu, A., Mensching, D. J., & Tran, N. (2023). Evaluation of Long-Term Oven Aging Protocols on Field Cracking Performance of Asphalt Binders Containing Reclaimed Asphaltic Materials (RAP/RAS). Road Materials and Pavement Design, 24(sup1), 437–450. https://doi.org/10.1080/14680629.2023.2181004

NCAT. (2020). Using the Hamburg Wheel Track Test for Balanced Mix Design. https://www.eng.auburn.edu/research/centers/ncat/newsroom/2020-fall/hamburg.html

Nguyen, L. N., Le, T.-H., Nguyen, L. Q., & Tran, V. Q. (2023). Machine learning approaches for predicting Cracking Tolerance Index (CTIndex) of asphalt concrete containing reclaimed asphalt pavement. Plos One, 18(10), e0287255.

Nikolaides, A. (2014). Highway Engineering: Pavements, Materials and Control of Quality. CRC Press.

ODOT. (2019). Standard Specifications for Highway Construction. Oklahoma Department of Transportation. https://www.odot.org/c_manuals/specbook/2019-FULL-SPEC-Web-Version.pdf

OHDL. (2014). Hamburg Rut Testing of Compacted Hot-Mix Asphalt (HMA) (OHDL 55). Oklahoma Department of Transportation. https://oklahoma.gov/odot/business-center/materials/ohdl.html

Olagunju, M., Mousavi Rad, S., Sadeghi, M., Vivanco, D., & Elkashef, M. (2025). Evaluating the Rutting and Moisture Resistance of Asphalt Mixes in Oklahoma Using Indirect Tensile Test at High Temperature Strength and Hamburg Wheel Tracking Test Corrected Rut Depth and Stripping Number. Transportation Research Record, 03611981251324206. https://doi.org/10.1177/03611981251324206

Onofri, A. (2019). Stabilising Transformations: How Do I Present my Results? https://www.statforbiology.com/2019/stat_general_reportingresults/

Ozer, H., Al-Qadi, I. L., Lambros, J., El-Khatib, A., Singhvi, P., & Doll, B. (2016). Development of the Fracture-based Flexibility Index for Asphalt Concrete Cracking Potential using Modified Semi-circle Bending Test Parameters. Construction and Building Materials, 115, 390–401. https://doi.org/10.1016/j.conbuildmat.2016.03.144

Ozer, H., Al-Qadi, I. L., Singhvi, P., Bausano, J., Carvalho, R., Li, X., & Gibson, N. (2018). Prediction of Pavement Fatigue Cracking at an Accelerated Testing Section Using Asphalt Mixture Performance Tests. International Journal of Pavement Engineering, 19, 264–278.

Pasandín, A. R., Pérez, I., Oliveira, J. R., Silva, H. M., & Pereira, P. A. (2015). Influence of Ageing on the Properties of Bitumen from Asphalt Mixtures with Recycled Concrete Aggregates. Journal of Cleaner Production, 101, 165–173.

Pearson, K. L. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11), 559–572.

Pellinen, T., & Xiao, S. (2006). Stiffness of Hot-Mix Asphalt (FHWA/IN/JTRP-2005/20, 2644; Issue FHWA/IN/JTRP-2005/20, 2644, p. FHWA/IN/JTRP-2005/20, 2644). Purdue University. https://doi.org/10.5703/1288284313390

Petersen, J. C., Robertson, R. E., Branthaver, J. F., Harnsberger, P. M., Duvall, J. J., Kim, S. S., & Bahia, H. U. (1994). Binder Characterization and Evaluation, Volume 1. Strategic Highway Research Program, National Research Council.

Peterson, R. L., Mahboub, K. C., Anderson, R. M., Masad, E., & Tashman, L. (2004). Comparing Superpave Gyratory Compactor Data to Field Cores. Journal of Materials in Civil Engineering, 16(1), 78–83. https://doi.org/10.1061/(ASCE)0899-1561(2004)16:1(78)

Pirouz, D. M. (2006). An overview of partial least squares. Available at SSRN 1631359. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1631359

Polacco, G., Cappello, M., Cuciniello, G., & Filippi, S. (2022). Apparent Molecular Weight Distributions in Bituminous Binders. Materials, 15(13), 4700.
Pouranian, M. R., & Haddock, J. E. (2020). Effect of Aggregate Gradation on Asphalt Mixture Compaction Parameters. Journal of Materials in Civil Engineering, 32(9), 04020244. https://doi.org/10.1061/(ASCE)MT.1943-5533.0003315
Punith, V., Xiao, F., Putman, B., & Amirkhanian, S. N. (2012). Effects of Long-Term Aging on Moisture Sensitivity of Foamed WMA Mixtures Containing Moist Aggregates. Materials and Structures, 45, 251–264.
Radeef, H., Hassan, N. A., Abidin, A. R. Z., Mahmud, M. Z. H., Yaacob, H., Mashros, N., & Mohamed, A. (2021). Effect of Aging and Moisture Damage on the Cracking Resistance of Rubberized Asphalt Mixture. Materials Today: Proceedings, 42, 2853–2858.
Radeef, H., Hassan, N. A., Abidin, A. Z., Mahmud, M., Ismail, C., Abbas, H., Al-Saffar, Z., & Redha, S. (2022). Impact of Aging and Moisture Damage on the Fracture Properties of Plastic Waste Modified Asphalt. IOP Conference Series: Earth and Environmental Science, 971, 012009.
Rahman, S., Hossain, M., & Nelson, P. (2011). A Study of Effectiveness of Thin Surface Treatments Using Hamburg Wheel Tracking Device. Transportation and Development Institute Congress, Integrated Transportation and Development for a Better Tomorrow, 418–427.
Rodezno, C., Moraes, R., Yin, F., & Fortunatus, M. (2021). Recycled Asphalt Binder Study. Wisconsin Highway Research Program, Report No. WHRP 0092-19-04.
Romero-Zambrana, P. (2023). Variability of the IDEAL-CT Test for Pavement Cracking to Achieve a Balanced Asphalt Mix Design ((No. MPC-589)). Mountain-Plains Consortium. https://rosap.ntl.bts.gov/view/dot/66633
Rousseeuw, P. J. (1985). Multivariate Estimation with High Breakdown Point. Mathematical Statistics and Applications B, 3, 217–230.
Rowe, G. M. (2011). Evaluation of the Relationship Between Asphalt Binder Properties and Non-Load-Related Cracking (Prepared Discussion). Journal of the Association of Asphalt Paving Technologists, 80, 615–664.
Safaei, F., Castorena, C., & Kim, Y. R. (2016). Linking Asphalt Binder Fatigue to Asphalt Mixture Fatigue Performance Using Viscoelastic Continuum Damage Modeling. Mechanics of Time-Dependent Materials, 20(3), 299–323.
Safjan, K. (2023). Comprehensive Guide to Interpreting R2, MSE, and RMSE for Regression Models. https://safjan.com/interpreting-r2-mse-rmse-for-regression-models/
Sandamal, K., Shashiprabha, S., Muttil, N., & Rathnayake, U. (2023). Pavement Roughness Prediction Using Explainable and Supervised Machine Learning Technique for Long-Term Performance. Sustainability, 15(12), 9617.
Seitllari, A., Boz, I., Habbouche, J., & Diefenderfer, S. D. (2023). Using mechanistic–empirical based analysis to evaluate rutting performance thresholds for balanced mix design tests. Construction and Building Materials, 400, 132762. https://doi.org/10.1016/j.conbuildmat.2023.132762
Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). Biometrika, 52(3/4), 591–611.
Sharanappa, S. (2022). Types of Transformations in Machine Learning. Medium. https://medium.com/@sunils0506/types-of-transformations-in-machine-learning-5e397a6c9ab4

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-Means Clustering Algorithm. IEEE Access, 8, 80716–80727.
Stastna, J., Zanzotto, L., & Ho, K. (1994). Fractional complex modulus manifested in asphalts. Rheologica Acta, 33(4), 344–354.
Statistical Methods. (2023). Levene's Test for Homogeneity of Variance in R. https://stat-methods.com/home/levenes-test-for-homogeneity-of-variance/
Statistics Solutions. (2013). Statistics Solutions. ANOVA (Analysis of Variance). https://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/anova/
Stephanie, G. (2023). Welch's ANOVA: Definition, Assumptions. https://www.statisticshowto.com/welchs-anova/
Teltayev, B., Amirbayev, E., & Radovskiy, B. (2022). Evaluating the Effect of Polymer Modification on the Low-Temperature Rheological Properties of Asphalt Binder. Polymers, 14(13), 2548. https://doi.org/10.3390/polym14132548
Tian, X., Li, G., Lu, X., Guo, C., & Hu, H. (2022). Study on the Short-Term Aging Behavior of Asphalt Based on PCA and LSM Analysis. Journal of Materials in Civil Engineering, 34(8), Article 8. https://doi.org/10.1061/(ASCE)MT.1943-5533.0004333
Tran, N. T., & Takahashi, O. (2018). Evaluating the Rutting Resistance of Wearing Course Mixtures with Different Fine Aggregate Sources Using the Indirect Tensile Strength Test. Journal of Testing and Evaluation, 48(4), Article 4. https://doi.org/10.1520/JTE20180152
Upadhya, A., Thakur, M. S., Al Ansari, M. S., Malik, M. A., Alahmadi, A. A., Alwetaishi, M., & Alzaed, A. N. (2022). Marshall Stability Prediction with Glass and Carbon Fiber Modified Asphalt Mix Using Machine Learning Techniques. Materials, 15(24), 8944.
Vamsikrishna, G., & Singh, D. (2023). Comparison of Rutting Resistance of Plant Produced Asphalt Mixes Using Hamburg Wheel Tracker and Surrogate Simple Performance Tests: IDEAL-RT and HT-IDT. Journal of Materials in Civil Engineering, 35(12), Article 12. https://doi.org/10.1061/JMCEE7.MTENG-16205
Vavrik, W. R. (2000). Asphalt Mixture Design Concepts to Develop Aggregate Interlock [PhD Thesis]. University of Illinois at Urbana-Champaign.
Wang, K., Yuan, Y., Han, S., & Yang, H. (2018). Application of Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) and Principal Component Analysis (PCA) for Quick Identifying of the Bitumen Produced by Different Manufacturers. Road Materials and Pavement Design, 19(8), 1940–1949.
Welch, B. (1938). The Significance of the Difference Between Two Means When the Population Variances are Unequal. Biometrika, 29, 350–362.
Weldegiorgis, M. T., & Tarefder, R. A. (2015). Towards a Mechanistic Understanding of Moisture Damage in Asphalt Concrete. Journal of Materials in Civil Engineering, 27(3), 04014128. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001087
West, R., Yin, F., Rodezno, C., & Taylor, A. (2021). Balanced Mixture Design Implementation Support (Final Report WHRP 0092-20-04). Auburn University, National Center for Asphalt Technology. https://rosap.ntl.bts.gov/view/dot/62612
Wu, C., Li, L., Wang, W., & Gu, Z. (2021). Experimental characterization of viscoelastic behaviors of nano-TiO\textsubscript2/CaCO\textsubscript3 modified asphalt and asphalt mixture. Nanomaterials, 11(1), 106.
Xiao, F., Amirkhanian, S., & Juang, C. H. (2009). Prediction of Fatigue Life of Rubberized Asphalt Concrete Mixtures Containing Reclaimed Asphalt Pavement Using Artificial

Neural Networks. Journal of Materials in Civil Engineering, 21(6), 253–261. https://doi.org/10.1061/(ASCE)0899-1561(2009)21:6(253)
Xiao, R., & Huang, B. (2022). Moisture Damage Mechanism and Thermodynamic Properties of Hot-Mix Asphalt under Aging Conditions. ACS Sustainable Chemistry & Engineering, 10(45), 14865–14887. https://doi.org/10.1021/acssuschemeng.2c04281
Xiao, R., Polaczyk, P., & Huang, B. (2022). Measuring Moisture Damage of Asphalt Mixtures: The Development of A New Modified Boiling Test Based on Color Image Processing. Measurement, 190, 110699. https://doi.org/10.1016/j.measurement.2022.110699
Xu, M., Zhang, Y., Zhao, P., & Liu, C. (2020). Study on Aging Behavior and Prediction of SBS Modified Asphalt with Various Contents Based on PCA and PLS Analysis. Construction and Building Materials, 265, 120732.
Yeo, I. K., & Johnson, R. A. (2000). A New Family of Power Transformations to Improve Normality or Symmetry. Biometrika, 87(4), 954–959.
Yin, F., Arambula, E., Lytton, R., Martin, A. E., & Cucalon, L. G. (2014). Novel Method for Moisture Susceptibility and Rutting Evaluation Using Hamburg Wheel Tracking Test. Transportation Research Record, 2446(1), 1–7. https://doi.org/10.3141/2446-01
Yin, F., Chen, C., Moraes, R., Hanz, A., Hehir, J., Knudtson, D., Auburn University. National Center for Asphalt Technology, & Mathy Technology and Engineering (MTE) Services, Inc. (2023). Impact of Polymer Modification on IDEAL-CT and I-FIT for Cracking Resistance Evaluation of Asphalt Mixtures (NRRA202303). https://rosap.ntl.bts.gov/view/dot/67165
Yin, F., Chen, C., West, R., Martin, A. E., & Arambula-Mercado, E. (2020). Determining the Relationship Among Hamburg Wheel-Tracking Test Parameters and Correlation to Field Performance of Asphalt Pavements. Transportation Research Record: Journal of the Transportation Research Board, 2674(4), 281–291. https://doi.org/10.1177/0361198120912430
Yin, F., Kaseer, F., Arámbula-Mercado, E., & Epps Martin, A. (2017). Characterising the Long-Term Rejuvenating Effectiveness of Recycling Agents on Asphalt Blends and Mixtures with High RAP and RAS Contents. Road Materials and Pavement Design, 18(sup4), 273–292.
Yin, F., Taylor, A. J., & Tran, N. (2020). Performance Testing for Quality Control and Acceptance of Balanced Mix Design.
Yin, F., & West, R. (2021). Balanced Mix Design Resource Guide. https://member.asphaltpavement.org/Shop/Product-Catalog/Product-Details?productid=695C89AA-A56B-EB11-A812-000D3A984636
Yusoff, N. I. M., Hainin, M. R., & Airey, G. D. (2012). A Comparative Study of Phase Angle Predictive Equations Using Bituminous Binder Data. Arabian Journal for Science and Engineering, 37(6), 1571–1583.
Yusoff, N. I. M., Jakarni, F. M., Nguyen, V. H., Hainin, M. R., & Airey, G. D. (2013). Modelling the Rheological Properties of Bituminous Binders Using Mathematical Equations. Construction and Building Materials, 40, 174–188.
Zack. (2019). How to Check ANOVA Assumptions. https://www.statology.org/anova-assumptions/
Zack. (2021). RMSE vs. R-squared: Which Metric Should You Use? https://www.statology.org/rmse-vs-r-squared/

Zhang, H., Xu, G., Chen, X., Wang, R., & Shen, K. (2019). Effect of Long-Term Laboratory Aging on Rheological Properties and Cracking Resistance of Polymer-Modified Asphalt Binders at Intermediate and Low Temperature Range. Construction and Building Materials, 226, 767–777.

Zhang, R., Sias, J. E., & Dave, E. V. (2022). Evaluation of the Cracking and Aging Susceptibility of Asphalt Mixtures Using Viscoelastic Properties and Master Curve Parameters. Journal of Traffic and Transportation Engineering (English Edition), 9(1), 106–119.

Zhao, P., Gao, D., Ren, R., Han, K., Yang, Z., Meng, W., & Fan, W. (2021). Short-Term Aging Performance Evaluation of Asphalt Based on Principal Component and Cluster Analysis. Journal of Testing and Evaluation, 49(1), Article 1. https://doi.org/10.1520/JTE20180781

Zhou, F., Im, S., Sun, L., & Scullion, T. (2017). Development of an IDEAL Cracking Test for Asphalt Mix Design and QC/QA. Road Materials and Pavement Design, 18(sup4), 405–427. https://doi.org/10.1080/14680629.2017.1389082

Zhou, F., Newcomb, D., Gurganus, C., Banihashemrad, S., Park, E. S., Sakhaeifar, M., & Lytton, R. L. (2016). Experimental Design for Field Validation of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures. NCHRP Project, 9.

Zhou, Z., Gu, X., Dong, Q., Ni, F., & Jiang, Y. (2020). Low- and Intermediate-Temperature Behaviour of Polymer-Modified Asphalt Binders, Mastics, Fine Aggregate Matrices, and Mixtures with Reclaimed Asphalt Pavement Material. Road Materials and Pavement Design, 21(7), 1872–1901. https://doi.org/10.1080/14680629.2019.1642137

Ziari, H., Behbahani, H., Izadi, A., & Nasr, D. (2013). Long Term Performance of Warm Mix Asphalt Versus Hot Mix Asphalt. Journal of Central South University, 20, 256–266.

APPENDIX A

Overview of Predictive Models used in this Report.

Linear Models

Simple Linear Regression (SLR)

The simple linear regression (SLR) model is the simplest regression model in which we have only one predictor X (Smadi and Abu-Afouna, 2012). This model, which is common in practice, is written as Equation (A1);

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t , \quad t = ,\dots,n, \qquad (A1)$$

where Y_t and X_t are the values of the response and predictor variables in the t^{th} trial, respectively, β_0 and β_1 are unknown parameters and ϵt are usually assumed to be i^{id} from $N(0, \sigma^2_\epsilon)$, especially for inference purposes. The terms $\{\epsilon_t\}$ satisfying such conditions are named in the time series context as the white noise (WN) process (Wei, 2006). The variable X is usually assumed fixed and non-random. For several predictors, the SLR model generalizes to what is known as the multiple linear regression model. For the estimation of the SLR model, there are two common methods, the first is the ordinary least squares (OLS) method, which relies on minimizing the sum of squares of errors $\sum\epsilon^2{}_t$. For the SLR model, the OLS estimators of β_0 and β_1 are;

$$\beta_0 = Y - \beta_1 X \qquad (A2)$$

and

$$\beta_1 = [\textstyle\sum_t(X_t - X)(Y_t - Y)] / [\sum t\,(X_t - X)^2] = S_{XY} / S_{XX} \qquad (A3)$$

It is known that these estimators are unbiased and best linear unbiased estimators (BLUE). The second method is the maximum likelihood method, which under the assumptions of independence and normality of ϵt produces again β_0 and β_1 above (Kutner et al., 2005). In turn, the fitted SLR model is written as

$$Y_t = \beta_0 - \beta_1 X \qquad (A4)$$

so that the estimated errors or residuals, denoted by et, are defined as $e_t = Y_t - Y_t,\ t = 1$, …, n. Once the regression model is fitted, an important step in model building and diagnosis is to check for the assumptions of the model, namely; independence, normality, and constant variance of errors. The residuals of the fitted model play a primary role in this purpose. Several graphs and tests for residuals may be used to examine those assumptions. These techniques are usually known as residual analysis. In particular, to examine the independence of error terms we use the plot of residuals against time or order (usually named as the residual plot). Besides, an important test specifically designed for testing the lack of randomness in residuals is the Durbin-Watson test. For a detailed account of various methods for the assessment of the assumptions of regression models, see (Kutner et al., 2005).

Partial Least Squares (PLS)

PLS is a projection method that models the relationship between the response Y and the independent variable X. The PLS regression is particularly suitable when X has many dimensions and there is an obvious multi-collinearity among these dimensions. Therefore, the PLS regression is quite suitable and popular for infrared spectrum analysis, since an infrared spectrum includes absorption bands at different wavenumbers (variables), and their intensities at most wavenumbers are collinear with each other (Ferrão et al., 2011).

Generally, the PLS regression constructs a linear model that predicts the Y based on measurements of the X. To do so, the PLS regression searches for a set of components (namely latent vectors) that performs a simultaneous decomposition of the X and Y with a target in which these components explain as much as possible of the covariance between the X and Y. This step is similar to the principle Component Analysis (PCA) (Abdi, 2010). After the decomposition, a regression step is followed where the latent vectors obtained from the X are used to predict the Y. In the PLS regression, the X and Y matrix are decomposed following Equations (A5) and (A6):

$$X=TP'+E \tag{A5}$$

$$Y=UQ'+F \tag{A6}$$

For this study, the variable X is the input matrix of infrared spectra and the response Y is the REOB contents of tested materials (0%, 3%, 6%, 10%, and 15%). By analogy with PCA, T and U are the score matrix, and P and Q are the loading matrix for X and Y, respectively. E and F are the unexplained residues. The decomposition is implemented by maximizing the covariance between the T and U, which can provide the best correlation. After the decomposition, the PLS regression model can generate a weight vector. The REOB content will be able to be predicted by linearly co-adding the peak intensities of infrared spectra of the REOB blends in accordance with the weight vector (De Peinder et al., 2009).

Elastic NET (ENET)

The elastic net (ENET) is a modified version of the lasso that is resistant to extremely high correlations between the predictors is the elastic net (ENET) (Friedman et al., 2010). The ENET was introduced for the analysis of high dimensional data in order to avoid the unpredictability of the lasso solution routes when predictors have significant correlations (Zou and Hastie, 2005). The ENET may be expressed as follows and employs a combination of the $\ell 1$ (lasso) and $\ell 2$ (ridge regression) penalties:

$$\beta(enet) = (1 + \frac{\lambda_2}{n})\{\arg\min\|y - X\beta\|_2^2 + \lambda_2 \|\beta\|_2^2 + \lambda_1 \|\beta\|_1\} \quad (A7)$$

On setting $\alpha = \lambda_2/(\lambda_1 + \lambda_2)$, the ENET estimator (A7) is seen to be equivalent to the minimizer of:

$$\beta(enet2) = \arg\min\|y - X\beta\|_2^2, subject\ to\ P_\alpha(\beta) = (1 - \alpha)\|\beta\|_1 + \alpha \|\beta\|_2^2 \leq s \quad (A8)$$

that is, the ENET penalty is P_α (β) (Zou and Hastie, 2005). When α =1, the ENET reduces to simple ridge regression; for α =0, it reduces to the lasso. Automatic variable selection is carried out by the ENET's $\ell 1$ portion, while improved prediction is achieved by the $\ell 2$ portion, which promotes grouped selection and stabilizes the solution paths in relation to random sampling. The ENET can choose groups of correlated features even when the groups are unknown beforehand by creating a grouping effect during variable selection, which causes a group of highly correlated variables to typically have coefficients of similar magnitude. In contrast to the lasso, the elastic net chooses more than n variables when p >> n. The elastic net does not, however, possess the oracle quality.

Non-Linear Models

Multivariate Adaptive Regression Spline (MARS)

MARS was first developed by Friedman (Friedman, 1991) to help fit the relationship between a group of predictors and a dependent variable. The divide-and-conquer approach of MARS, which divides the training data sets into several areas and assigns each region its own regression line, is quick and efficient (Sekulic and Kowalski, 1992; Steinberg, 2001). Data are the driving force behind MARS.

How to most effectively ascertain the underlying link between the dependent variable and the vector of predictors, such as pavement surface distress, traffic volume, pavement age, and environmental variables, is the fundamental challenge in pavement performance and degradation modeling. The MARS method looks for interactions between all variables as well as all possible univariate hinge sites. It accomplishes this by using sets of variables together referred to as basis functions (LeBlanc and Tibshirani, 1994). The method is comparable to the application of splines.

The general form of a MARS predictor is as follows:

$$Y = \beta_0 + \sum\sum [\beta_{jb} (+)Max(0,x_j - H_{bj}) + \beta_{jb}(-)Max(0, H_{bj} - x_j)\} \qquad (A9)$$

For P predictor variables and B basis function. The basis functions $Max(0, x - H)$ and Max(0, H − x) are univariate and do not have to be present if their β coefficients are 0. The H values are called “hinges” or “knots”. An example MARS model is:

$$y = 28 - 0.5\ Max(0,x_1 - 6) + 2.5\ Max(0.6 - x_1) + Max(0,x_2 - 7) + Max(0,2 - x_3) - Max(0,x_3 - 13) \qquad (A10)$$

Since $Max(a, b) = -Min(-a, -b)$ and $a + Max(b, c) = Max(a+b, a+c)$, a MARS predictive model is always expressed as a sum of Max and Min operators on piecewise linear univariate expressions. The MARS algorithm proceeds as follows: a forward stepwise search for the basis function takes place with the constant basis function, the only one present initially. At each step, the split that minimized some “lack of fit” criterion from all the possible splits on each basis function is chosen. This continues until the model reaches some predetermined maximum number of basic functions, which should be about twice the number expected in the model to aid the subsequent backward stepwise deletion of the basis function.

The backward stepwise function involves removing basis functions one at a time until the “lack of fit” criterion is a minimum. In the backward stepwise deletion, the last important basis functions are eliminated one at a time. The lack of fit measure used is based on the generalized cross-validation (GCV) [Craven and Wahba, 1979]:

$$GCV = A * \sum_i (y_i - f(x)) / N \qquad (A11)$$

with $A = [1-C(M)/N]-2$ and being the complexity function (Craven and Wahba, 1979). The GCV criterion is the average residual error multiplied by a penalty to adjust for the variability associated with the estimation of more parameters in the model.

Artificial Neural Network (ANN)

A popular AI methodology for pattern recognition is ANN. This strategy is based on biological neural networks seen in nature (Bishop, 2006). An ANN model may make conclusions using a sizable collection of artificial neurons through the supervised training process. A biological brain uses a huge number of interconnected biological neurons to solve pattern recognition issues, which is quite similar to how ANN works (Pham, 2017).

Multiple nodes make up an ANN model, which mimics the actual neurons seen in the human brain.

Its adaptability and capacity for learning enable the ANN to achieve its benefits. There is no constraint on the quantity of input and output variables in practice, and it has been demonstrated that an ANN with just one hidden layer is capable of approximating extremely complicated nonlinear functions. However, the ANN has a major flaw: it takes a lot of trial and error to correctly configure a network with the right number of neurons and activation functions.

Through axons, a neuron may process information and communicate with other neurons. A weight value is featured by each connection or axon. As a result, ANN may learn a discriminating function by changing the weights' values. An input layer, a hidden layer, and an output layer are frequently seen in ANN models (Figure A1).

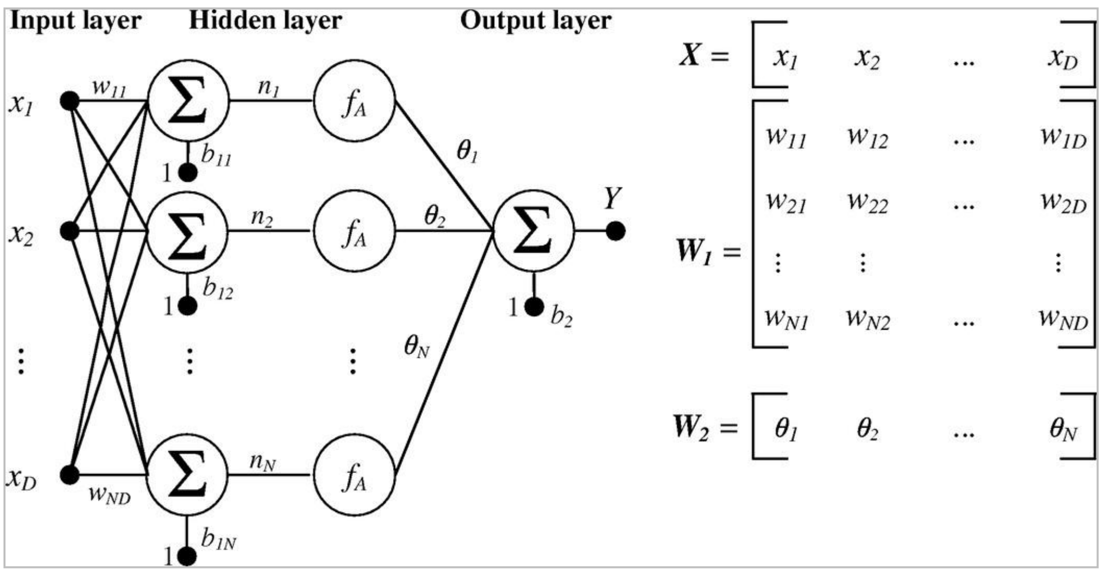


Figure A1. The ANN Model Structure (Tran and Hoang, 2016).

Providing that the learning task is to approximate a classification function $f : X \epsilon R^D \rightarrow Y \epsilon R^C$, where D denotes the number of input attributes and *C = 2* represents the number of class labels, the ANN structure is shown in the equation A12 (Tran and Hoang, 2016):

$$f(X) = b_2 + W_2 \times (f_A(b_1 + W_1 \times X)) \qquad (A12)$$

$f\ (X) = b_2 + W_2 \times (f_A(b_1 + W_1 \times X))$ where W1 and W2 denote weight matrices of the hidden layer and the output layer, respectively. *b = [b11, b12, ..., b1N]* represents a bias vector of the hidden layer; b2 is a bias vector of the output layer; f_A denotes an activation function (e.g., log-sigmoid). Generally, the weight matrices and the bias vectors of the ANN can be effectively trained using the error backpropagation framework (Bishop, 2006; Hoang et al., 2017).

Generally, the weight matrices and the bias vectors of an ANN are trained via a process that employs the framework of error backpropagation (Tran and Hoang, 2016). Moreover, the Mean Square Error (MSE) is employed as the objective function for training an ANN structure for function approximation tasks (Freeman and Skapura, 1991):

$$MSE= W_1,W_2,b_1,b_2.\ 1/M \sum e^2_i \qquad (A13)$$

where M represents the number of data samples and e_i is an output error. $e_i=Y_{(i,P)}\ -Y_{(i,A)}$ $e_i = Y_{i,P} - Y_{i,A}$ ($Y_{i,P}$ and $Y_{i,A}$ denote the predicted and actual outputs, resp.).

Support Vector Machine (SVM)

SVM is regarded as a well-liked CI technique. This approach is used in line with statistical learning theory, which is widely used in many branches of science and engineering, including challenges involving classification and regression (Faizollahzadeh Ardabili et al., 2018). Instead of reducing the local training error, SVM tries to lower the generalized upper bound error. One of the key benefits of the SVM model over more established machine learning techniques is this. Additionally, to address the drawbacks of the traditional ANN method, which uses empirical risk minimization when modeling a given variable, the SVM model employs the Structural Risk Minimization Principle (SRMP) and has a high generalization capacity.

Problems involving classification and regression may be solved using an unsupervised learning technique like the SVM. To address the shortcomings of the conventional ANN method, the SVM model incorporates SRMP and exhibits complete generalization ability. To simulate a specific variable, empirical risk reduction is used (Faizollahzadeh Ardabili et al., 2018). The SVM is regarded as a linear classification method that seeks to identify the dataset's most trustworthy line. Support vector regression, which is generally referred to as binary, may be used with this approach for actual outputs (non-binary). We attempted to address the challenge of SVR parameter setting in this work.

The basic function of SVR is minimizing Equation (A14) (Smola and Schölkopf, 2004):

$$\min 1/2\ w^T w + C \sum^{N}_{(i=1)}(\delta^{+} + \delta^{-}) \qquad (A14)$$

whose δ and C parameters will be explained in the SVM parameters section; the value w is the weight vector.

The particle filter is a random-based state estimator operating through noises. It affects x_k and y_k, and the values of the noise and equations are shown in Equation A15. Furthermore, the measurement noise is defined as the dimensions and weights (Carpenter et al., 1999):

$$X_k = f_k(x_{k-1}, u_k, w_k)$$
$$Yk = h_k(x_k, u_k, v_k) \qquad (A15)$$

where x_k represents the sluice state, y_k is the output, f_k is the process function, h_k is the measurement function, u_k is the input and w_k and v_k are noises that affect the equations.

APPENDIX B

Example Calculations of Shape Factor "n"

To calculate the “shape factor” (n) the gradation of the aggregate is required. The maximum dimension of the aggregate particle (D_{max}) is determined, then the ratio of the specific dimension (d_i) of an aggregate particle to its maximum dimension (di/ D_{max}) is calculated. The logarithm of each of the aggregate percentage passing is then taken, it is denoted by $\log(P_i)$, and the logarithm of the (d_i/ D_{max}) for each sieve size is also found, it’s denoted by $\log(d_i / D_{max})$. A graph showing $\log (d_i / D_{max})$ against $\log (P_i)$ is plotted, and a regression line is fitted between the two variables which gives the slope of the line. The slope of the line is the “shape factor”. Table B1, Table B2, and Figure B1 are examples of the calculation for the shape factor.

Table B1. Example calculation of the Shape factor for mix M1

Sieve #	Sieve Size (mm)	Mix 1 JMF
3/4 in	19	100.0
1/2 in	12.5	96.0
3/8 in	9.5	90.0
#4	4.75	66.0
#8	2.36	41.0
#16	1.18	27.0
#30	0.6	19.0
#50	0.3	12.0
#100	0.15	5.0
#200	0.075	2.5
D_{max}		19

Table B2. Results for the d_i/D_{max}, $Log(d_i/D_{max})$ and Log (P_i)

di/D_{max}	$Log(d_i/D_{max})$	Log (P_i)
1	0	2
0.65789474	-0.18184359	1.98227123
0.5	-0.30103	1.95424251
0.25	-0.60205999	1.81954394
0.12421053	-0.9058416	1.61278386
0.06210526	-1.20687159	1.43136376
0.03157895	-1.50060235	1.2787536
0.01578947	-1.80163235	1.07918125
0.00789474	-2.10266234	0.69897
0.00394737	-2.40369234	0.39794001

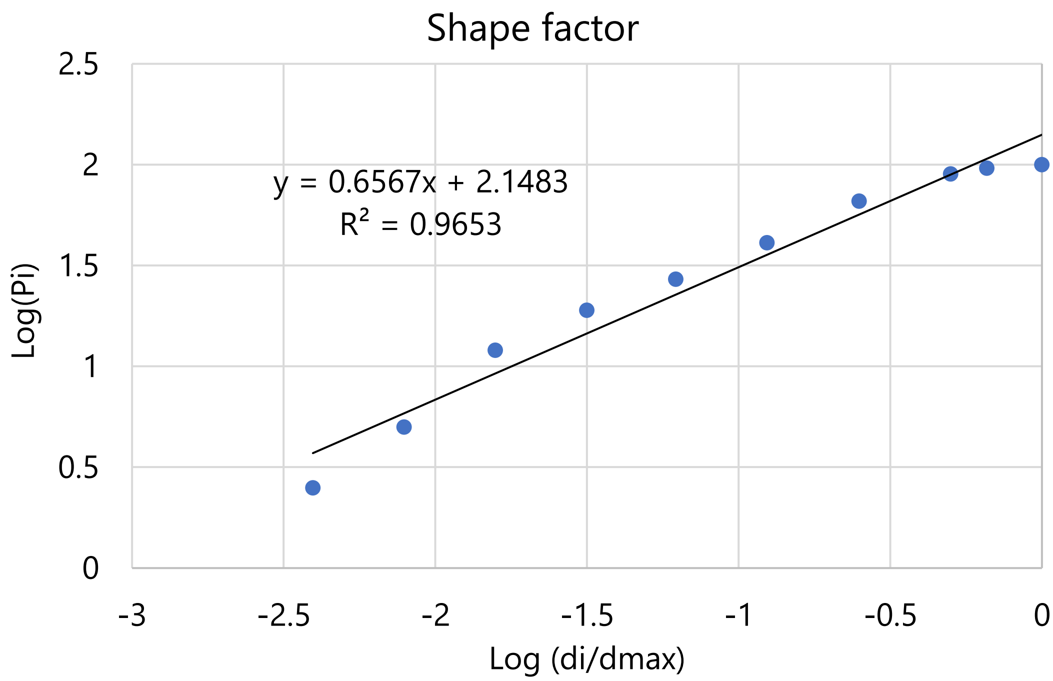


Figure B1. Plot showing Determination of Shape Factor (n = 0.6567) for Mix M1.

APPENDIX C

HWTT Results and Figures

Table C1. Rut depth for Mix M1.

Pass Count	Rut Depth (mm)
0	0
1000	1.16
2000	1.74
3000	2.16
4000	2.45
5000	2.72
6000	2.96
7000	3.17
8000	3.38
9000	3.59
10000	3.80
11000	3.97
12000	4.14
13000	4.28
14000	4.43
15000	4.55
16000	4.67
17000	4.80
18000	4.87
19000	4.92
20000	4.82

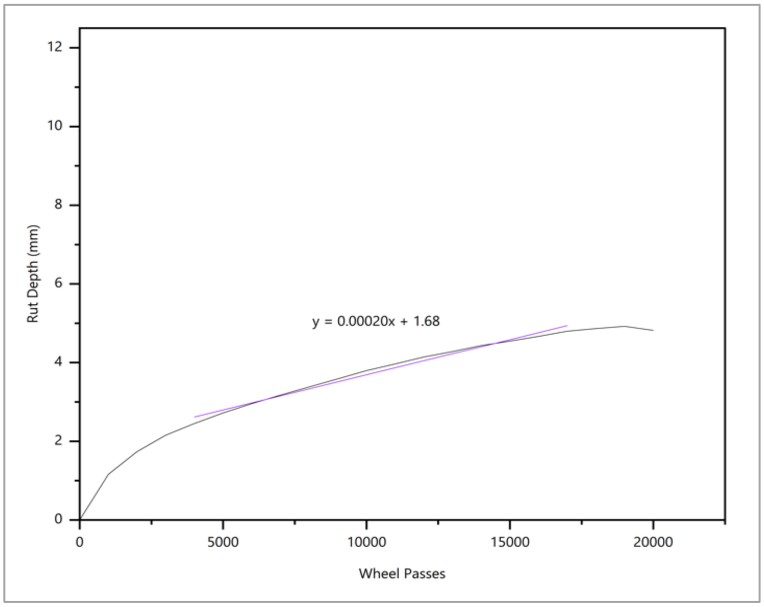


Figure C1. Rut Depth vs Pass Count for Mix 1.

Table C2. Rut depth for Mix M2.

Pass Count	Rut Depth (mm)
0	0
1000	1.48
2000	1.88
3000	2.16
4000	2.39
5000	2.58
6000	2.74
7000	2.87
8000	2.99
9000	3.10
10000	3.20
11000	3.31
12000	3.40
13000	3.51
14000	3.60
15000	3.70
16000	3.81
17000	3.89
18000	4.01
19000	4.12
20000	4.23

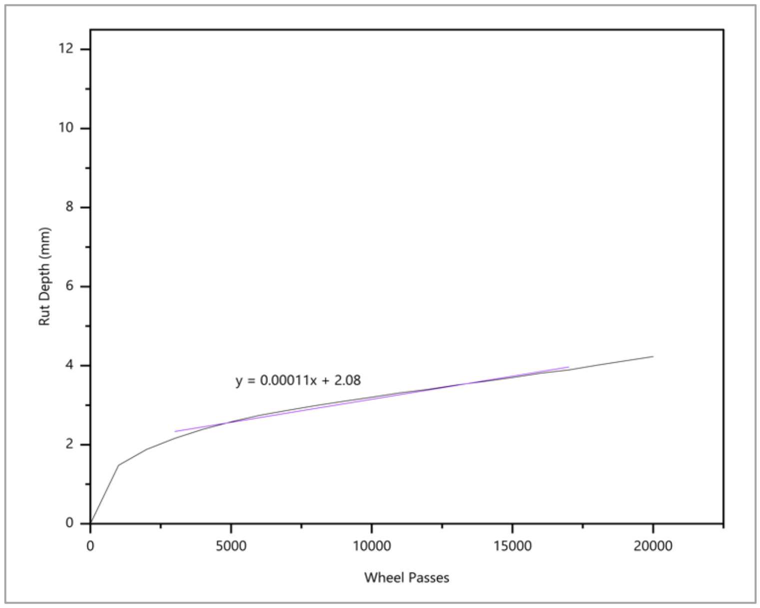


Figure C2. Rut Depth vs Pass Count for Mix 2.

Table C3. Rut depth for Mix M3.

Pass Count	Rut Depth (mm)
0	0
1000	1.57
2000	2.36
3000	2.87
4000	3.35
5000	3.66
6000	3.92
7000	4.13
8000	4.32
9000	4.53
10000	4.69
11000	4.87
12000	5.04
13000	5.25
14000	5.49
15000	5.80
16000	6.20
17000	6.70
18000	7.23
19000	7.89
20000	8.60

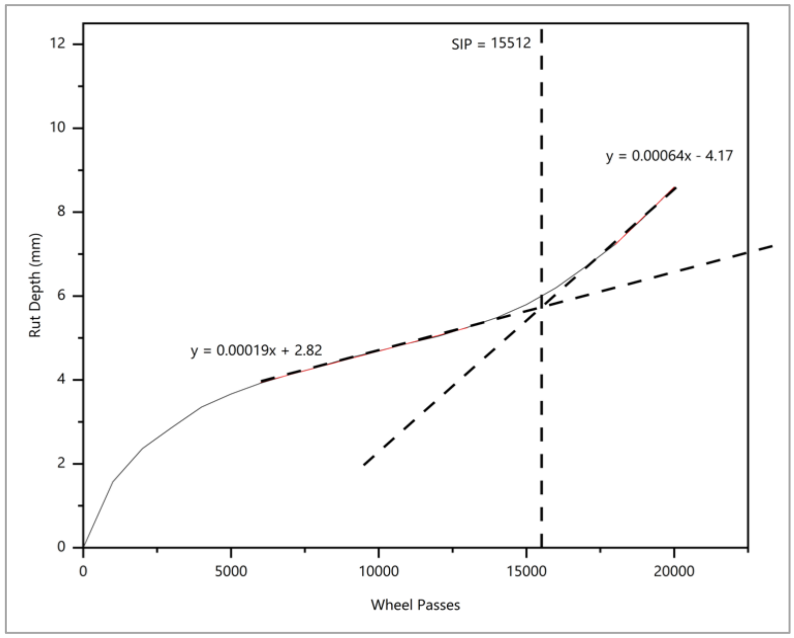


Figure C3. Rut Depth vs Pass Count for Mix 3.

Table C4. Rut depth for Mix M4.

Pass Count	Rut Depth (mm)
0	0
1000	2.04
2000	2.54
3000	2.89
4000	3.16
5000	3.37
6000	3.57
7000	3.73
8000	3.88
9000	4.03
10000	4.16
11000	4.32
12000	4.59
13000	4.88
14000	5.23
15000	5.86
16000	7.37
17000	9.31
18000	11.57
19000	16.08

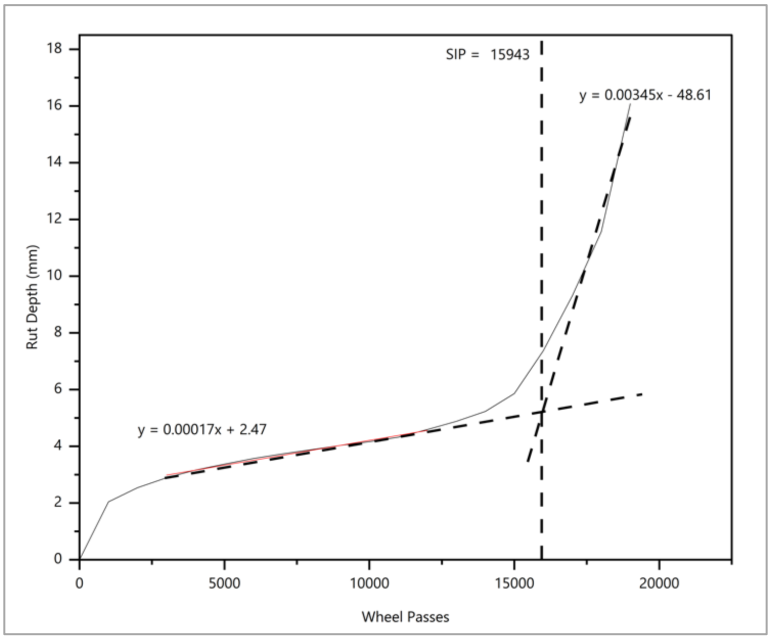


Figure C4. Rut Depth vs Pass Count for Mix 4.

Table C5. Rut depth for Mix M5.

Pass Count	Rut Depth (mm)
0	0
1000	1.26
2000	1.98
3000	2.34
4000	2.69
5000	2.94
6000	3.18
7000	3.41
8000	3.58
9000	3.73
10000	3.84
11000	3.95
12000	4.05
13000	4.14
14000	4.23
15000	4.33
16000	4.42
17000	4.51
18000	4.59
19000	4.66
20000	4.74

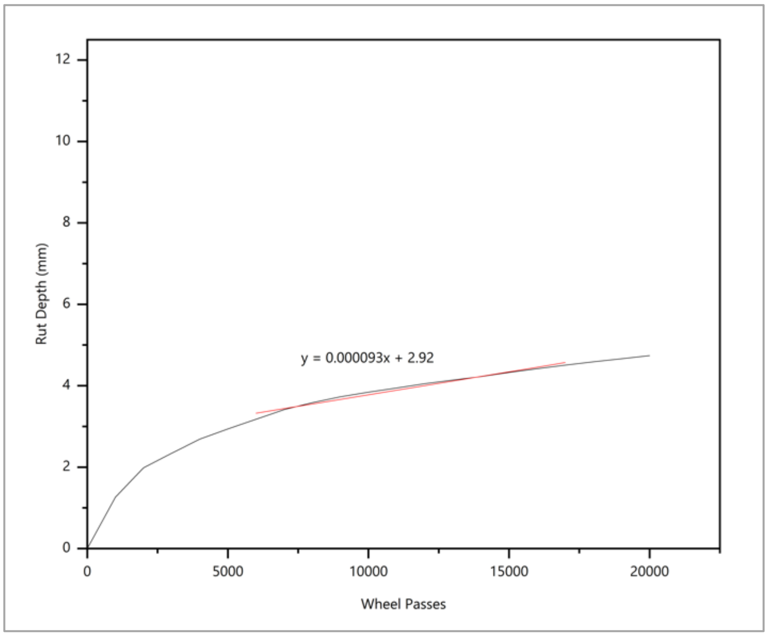


Figure C5. Rut Depth vs Pass Count for Mix 5.

Table C6. Rut depth for Mix M6.

Pass Count	Rut Depth (mm)
0	0
1000	2.03
2000	2.76
3000	3.25
4000	3.65
5000	4.01
6000	4.37
7000	4.77
8000	5.28
9000	5.92
10000	6.98
11000	9.65
12000	12.27
13000	14.68
14000	17.45
15000	18.86

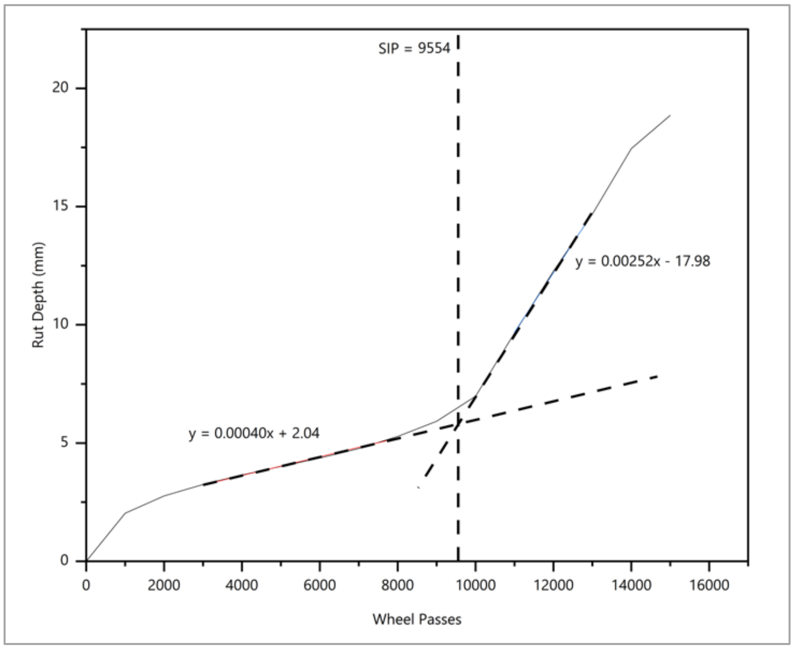


Figure C6. Rut Depth vs Pass Count for Mix 6.

Table C7. Rut depth for Mix M7.

Pass Count	Rut Depth (mm)
0	0
1000	2.94
2000	3.30
3000	2.74
4000	4.17
5000	4.27
6000	5.42
7000	4.13
8000	5.53
9000	6.12
10000	6.28
11000	6.27
12000	5.88
13000	6.51
14000	6.33
15000	6.47
16000	6.55
17000	6.64
18000	6.75
19000	6.80
20000	6.91

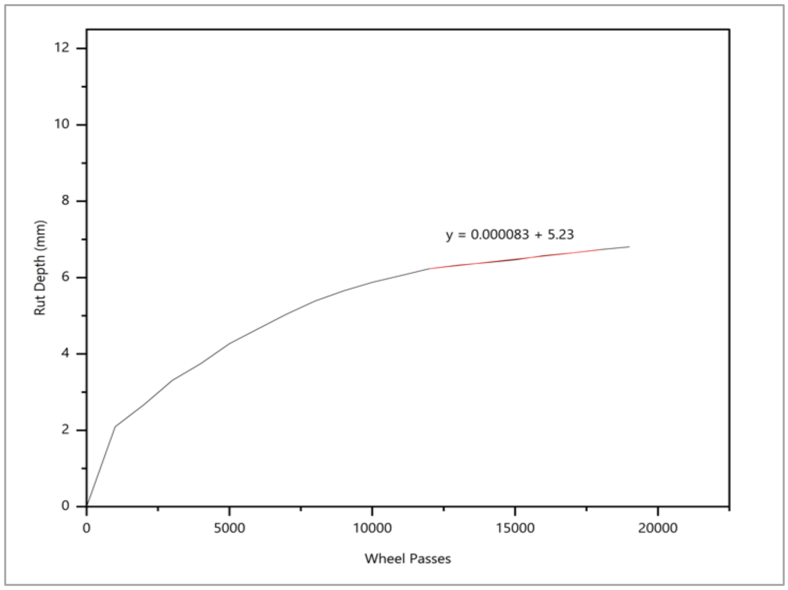


Figure C7. Rut Depth vs Pass Count for Mix 7.

Table C8. Rut depth for Mix M8.

Pass Count	Rut Depth (mm)
0	0
1000	1.42
2000	1.83
3000	2.12
4000	2.34
5000	2.53
6000	2.76
7000	2.94
8000	3.12
9000	3.17
10000	3.65
11000	4.48
12000	6.53
13000	9.97
14000	14.16
15000	16.29
16000	17.76
17000	18.91

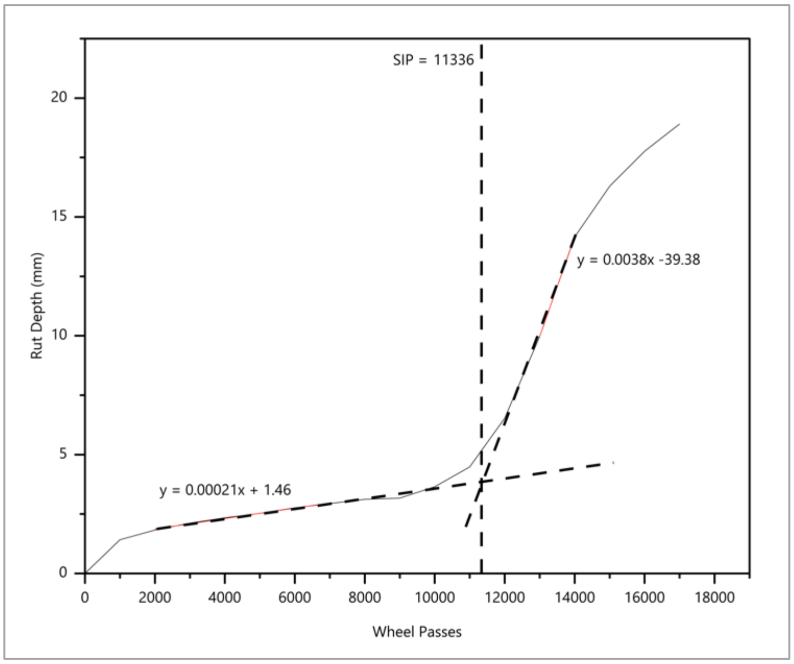


Figure C8. Rut Depth vs Pass Count for Mix 8.

Table C9. Rut depth for Mix M9.

Pass Count	Rut Depth (mm)
0	0
1000	1.98
2000	3.05
3000	3.48
4000	3.42
5000	3.82
6000	4.22
7000	4.62
8000	4.11
9000	4.92
10000	5.62
11000	6.20
12000	4.93
13000	6.31
14000	6.57
15000	6.85
16000	7.15
17000	7.48
18000	7.86
19000	8.17
20000	8.43

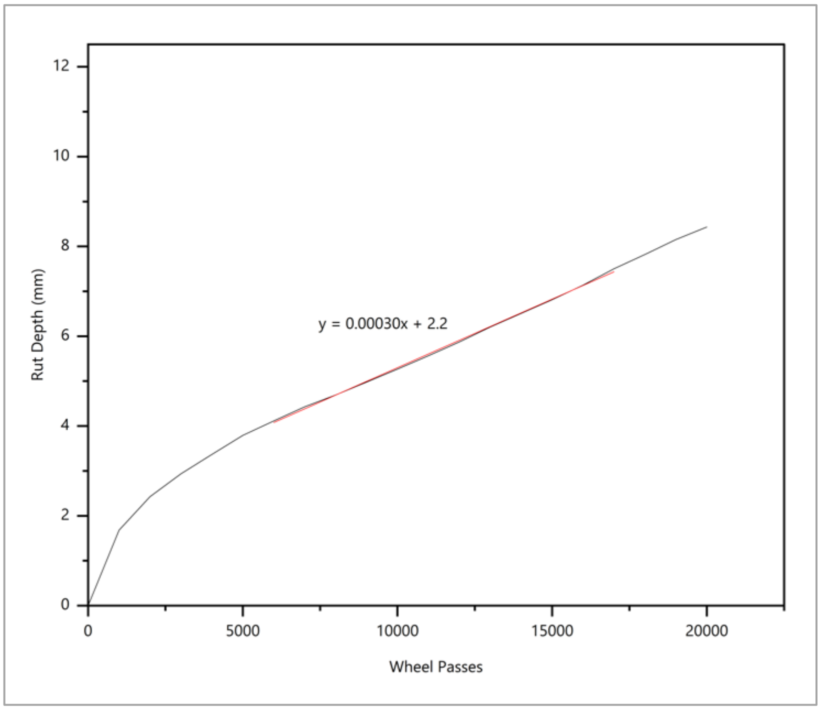


Figure C9. Rut Depth vs Pass Count for Mix 9.

Table C10. Rut depth for Mix M10.

Pass Count	Rut Depth (mm)
0	0
1000	1.64
2000	2.12
3000	2.49
4000	2.82
5000	3.08
6000	3.30
7000	3.47
8000	3.60
9000	3.75
10000	3.87
11000	3.95
12000	4.06
13000	4.14
14000	4.22
15000	4.32
16000	4.40
17000	4.48
18000	4.56
19000	4.63
20000	4.73

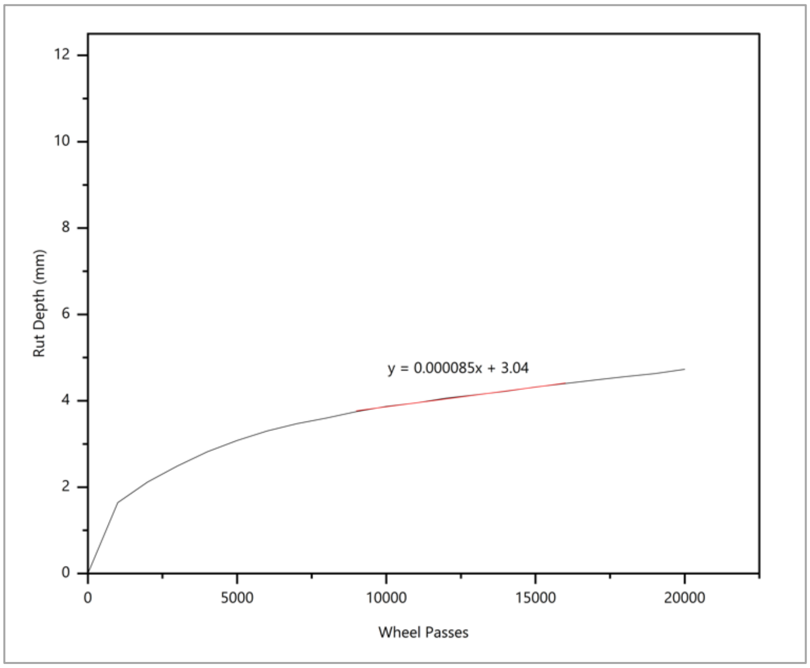


Figure C10. Rut Depth vs Pass Count for Mix 10.

Table C11. Rut depth for Mix M11.

Pass Count	Rut Depth (mm)
0	0
1000	1.23
2000	1.53
3000	1.75
4000	1.96
5000	2.18
6000	2.43
7000	2.73
8000	3.38
9000	4.96
10000	9.28
11000	16.85

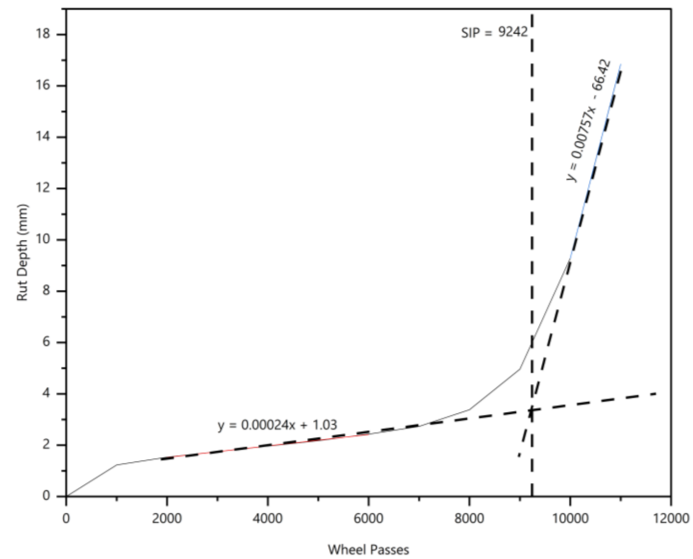


Figure C11. Rut Depth vs Pass Count for Mix 11.

Table C12. Rut depth for Mix M12.

Pass Count	Rut Depth (mm)
0	0
1000	0.98
2000	1.62
3000	1.91
4000	2.11
5000	2.28
6000	2.42
7000	2.55
8000	2.67
9000	2.78
10000	2.89
11000	3.00
12000	3.10
13000	3.18
14000	3.21
15000	3.25
16000	3.39
17000	3.68
18000	4.26
19000	5.16
20000	6.37

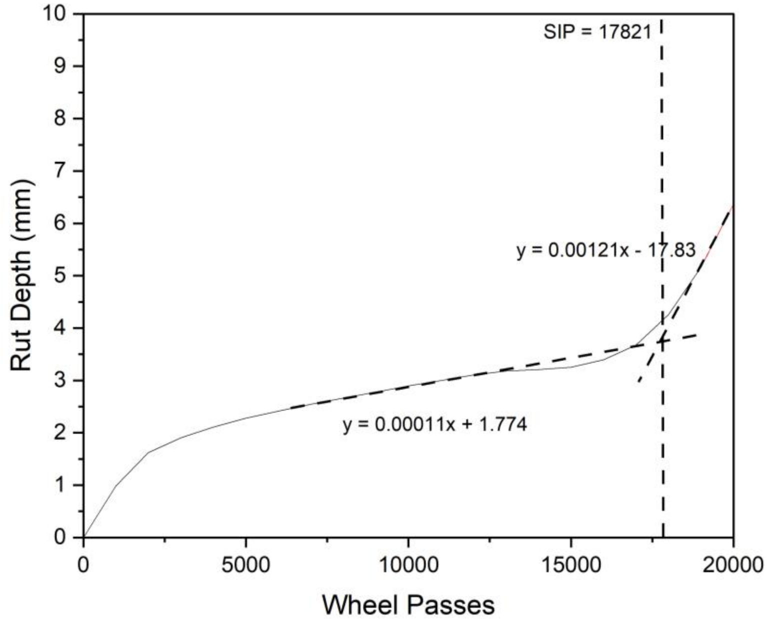


Figure C12. Rut Depth vs Pass Count for Mix 12.

Table C13. Rut depth for Mix M13.

Pass Count	Rut Depth (mm)
0	0
1000	1.08
2000	1.21
3000	1.34
4000	1.44
5000	1.52
6000	1.60
7000	1.68
8000	1.73
9000	1.78
10000	1.85
11000	1.91
12000	1.92
13000	1.99
14000	2.05
15000	2.09
16000	2.14
17000	2.18
18000	2.21
19000	2.24
20000	2.27

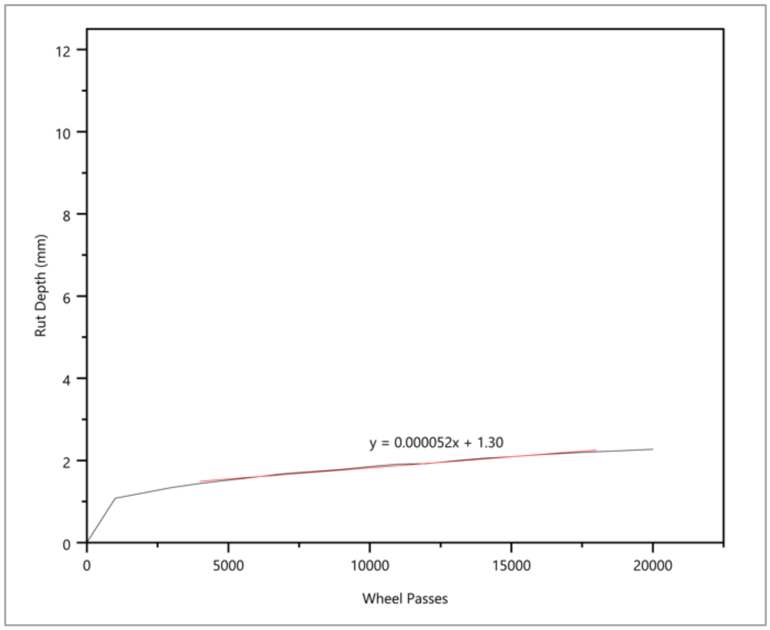


Figure C13. Rut Depth vs Pass Count for Mix 13.

Table C14. Rut depth for Mix M14.

Pass Count	Rut Depth (mm)
0	0
1000	1.25
2000	1.53
3000	1.73
4000	1.89
5000	2.03
6000	2.19
7000	2.38
8000	2.53
9000	2.58
10000	2.64
11000	2.57
12000	2.75
13000	2.90
14000	3.21
15000	3.50
16000	4.13
17000	4.97
18000	6.29
19000	7.44
20000	8.46

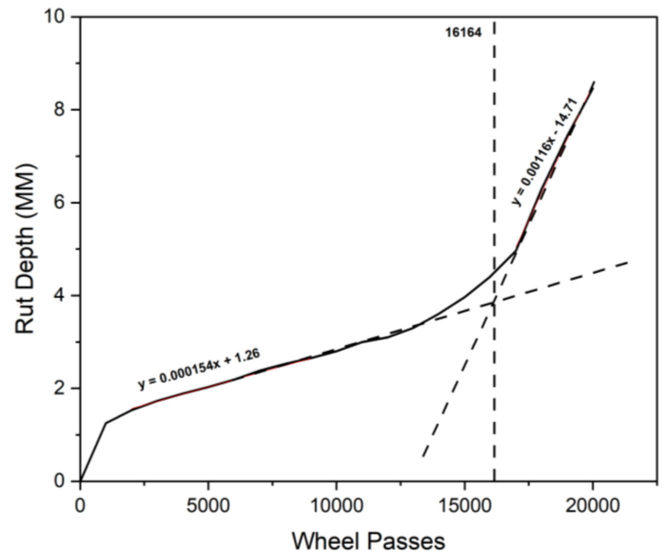


Figure C14. Rut Depth vs Pass Count for Mix 14.

Table C15. Rut depth for Mix M15.

Pass Count	Rut Depth (mm)
0	0
1000	0.96
2000	1.14
3000	1.27
4000	1.37
5000	1.44
6000	1.52
7000	1.60
8000	1.67
9000	1.74
10000	1.80
11000	1.86
12000	1.93
13000	1.98
14000	2.03
15000	2.08
16000	2.14
17000	2.18
18000	2.21
19000	2.26
20000	2.32

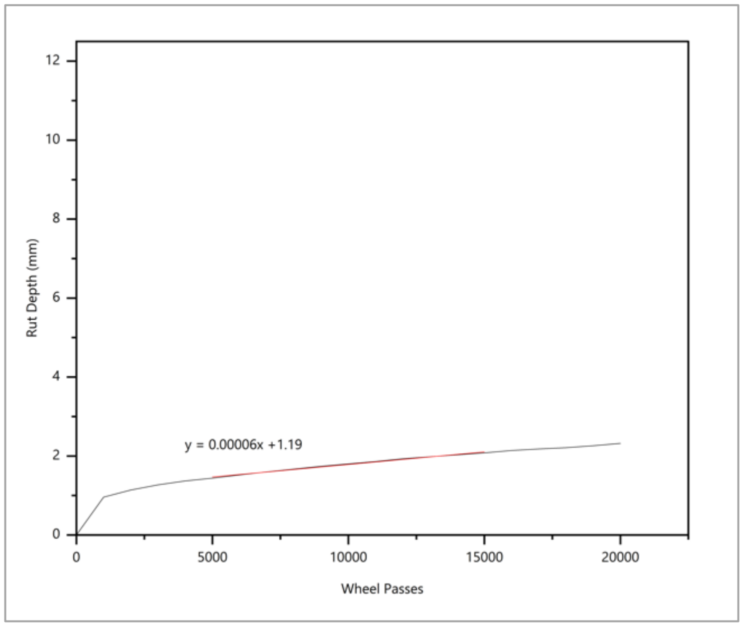


Figure C15. Rut Depth vs Pass Count for Mix 15.

Table C16. Rut depth for Mix M16.

Pass Count	Rut Depth (mm)
0	0
1000	1.44
2000	2.13
3000	2.73
4000	3.29
5000	3.82
6000	4.34
7000	4.85
8000	5.35
9000	5.84
10000	6.33
11000	6.82
12000	7.31
13000	7.79
14000	8.29
15000	8.77
16000	9.26
17000	9.76
18000	10.25
19000	10.74
20000	11.24

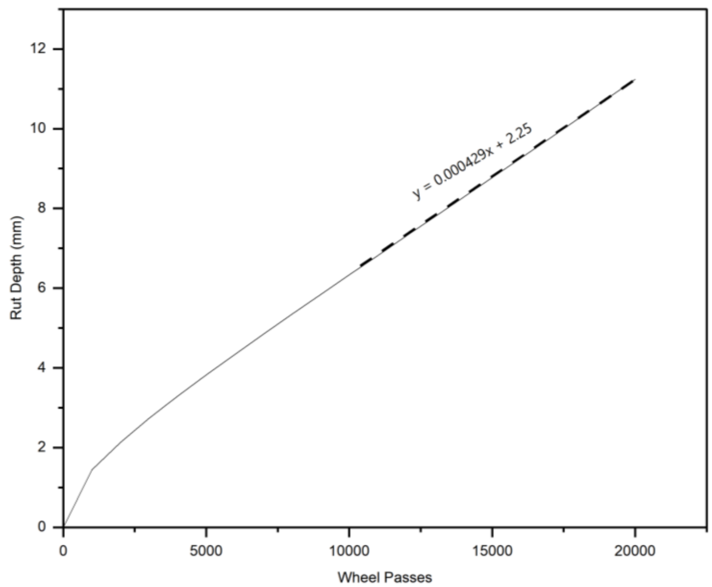


Figure C16. Rut Depth vs Pass Count for Mix 16.

Table C17. Rut depth for Mix M17.

Pass Count	Rut Depth (mm)
0	0
1000	1.75
2000	2.25
3000	2.57
4000	2.82
5000	3.03
6000	3.21
7000	3.39
8000	3.58
9000	3.81
10000	3.95
11000	3.88
12000	4.16
13000	4.47
14000	4.84
15000	5.55
16000	6.33
17000	7.52
18000	9.71
19000	11.20
20000	12.95

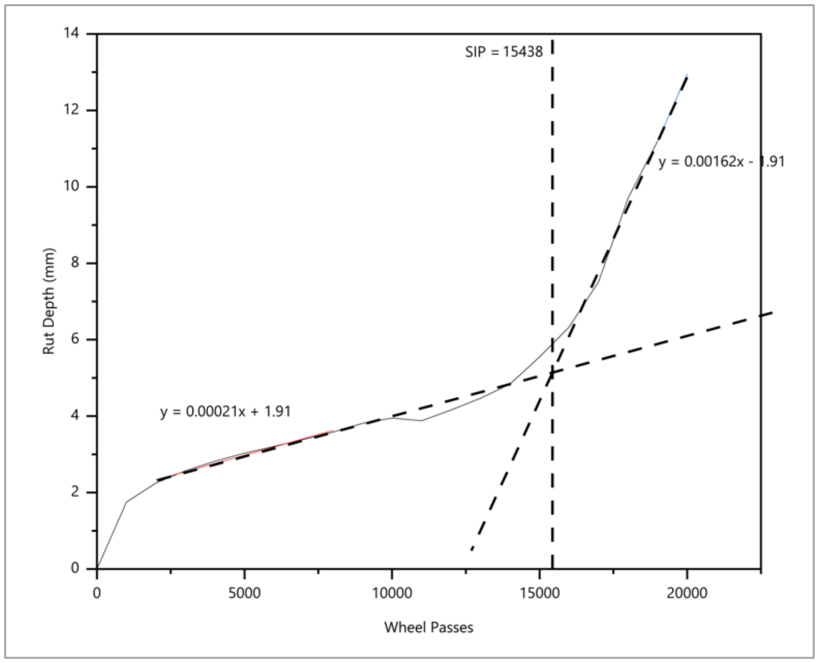


Figure C17. Rut Depth vs Pass Count for Mix 17.

Table C18. Rut depth for Mix M18.

Pass Count	Rut Depth (mm)
0	0
1000	2.08
2000	2.94
3000	3.54
4000	4.00
5000	4.48
6000	4.98
7000	5.49
8000	5.83
9000	6.48
10000	7.18
11000	8.73
12000	11.36
13000	15.74

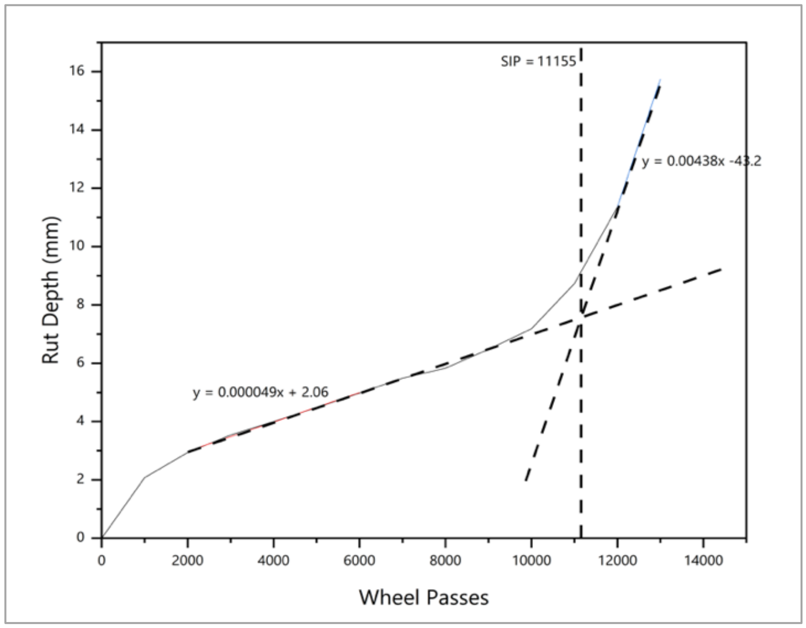


Figure C18. Rut Depth vs Pass Count for Mix 18.

Table C19. Rut depth for Mix M19.

Pass Count	Rut Depth (mm)
0	0
1000	1.31
2000	2.00
3000	2.56
4000	3.08
5000	3.47
6000	3.87
7000	4.33
8000	4.92
9000	5.67
10000	6.88
11000	8.21
12000	9.97
13000	13.46

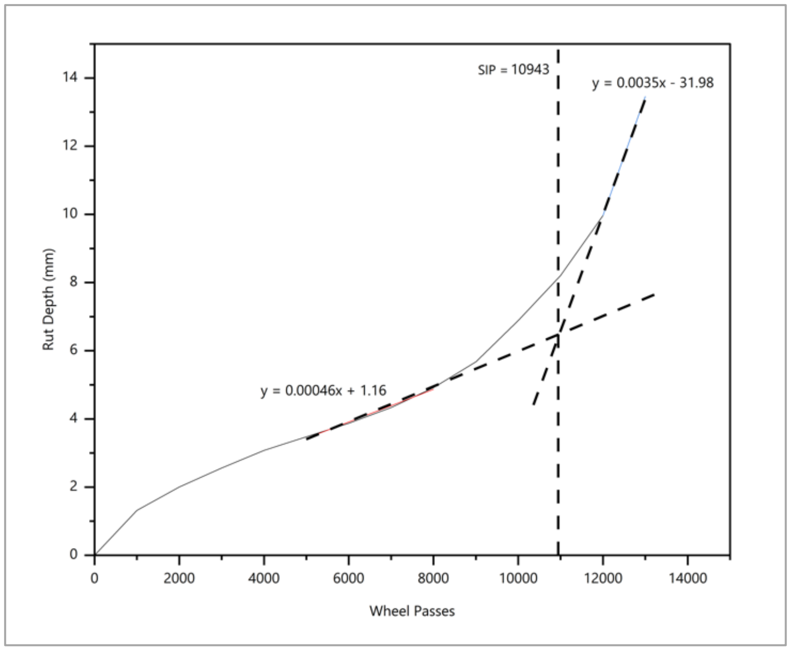


Figure C19. Rut Depth vs Pass Count for Mix 19.

Table C20. Rut depth for Mix M20.

Pass Count	Rut Depth (mm)
0	0
1000	1.23
2000	1.58
3000	1.84
4000	2.04
5000	2.24
6000	2.73
7000	3.30
8000	3.71
9000	4.23
10000	4.83
11000	7.06
12000	8.10
13000	9.18
14000	10.80
15000	12.61
16000	14.53
17000	17.22

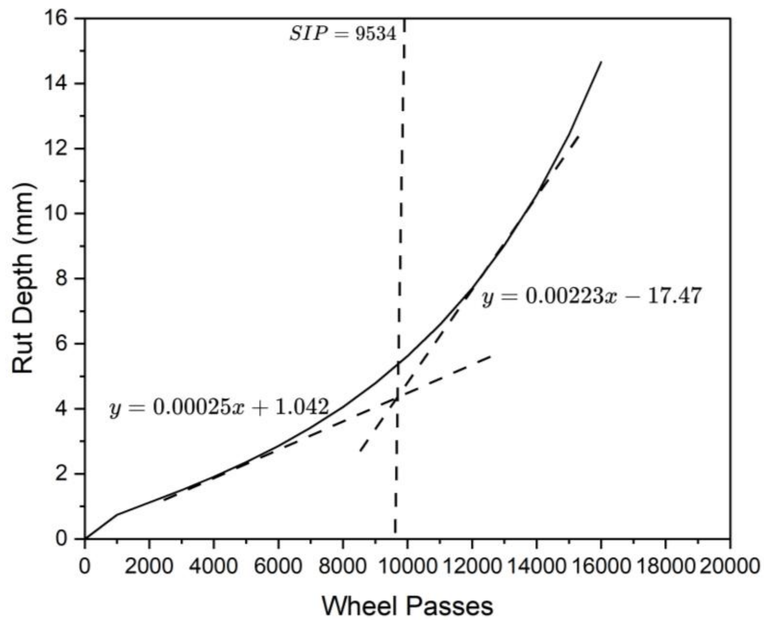


Figure C20. Rut Depth vs Pass Count for Mix 20.

Table C21. Rut depth for Mix M21.

Pass Count	Rut Depth (mm)
0	0
1000	1.20
2000	1.70
3000	2.02
4000	2.44
5000	2.61
6000	2.74
7000	2.85
8000	2.94
9000	3.03
10000	3.12
11000	3.24
12000	3.37
13000	3.51
14000	3.65
15000	3.79
16000	3.94
17000	4.11
18000	4.26
19000	4.45
20000	4.68

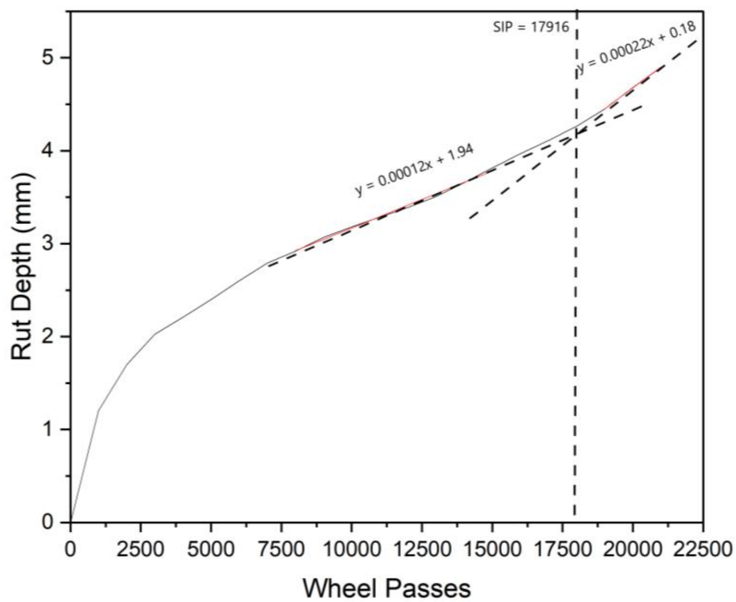


Figure C21. Rut Depth vs Pass Count for Mix 21.

Table C22. Rut depth for Mix M22.

Pass Count	Rut Depth (mm)
0	0
1000	1.30
2000	1.62
3000	1.85
4000	2.02
5000	2.19
6000	2.32
7000	2.49
8000	2.72
9000	3.08
10000	4.06
11000	5.04
12000	6.00
13000	7.27
14000	9.69
15000	12.24
16000	14.97

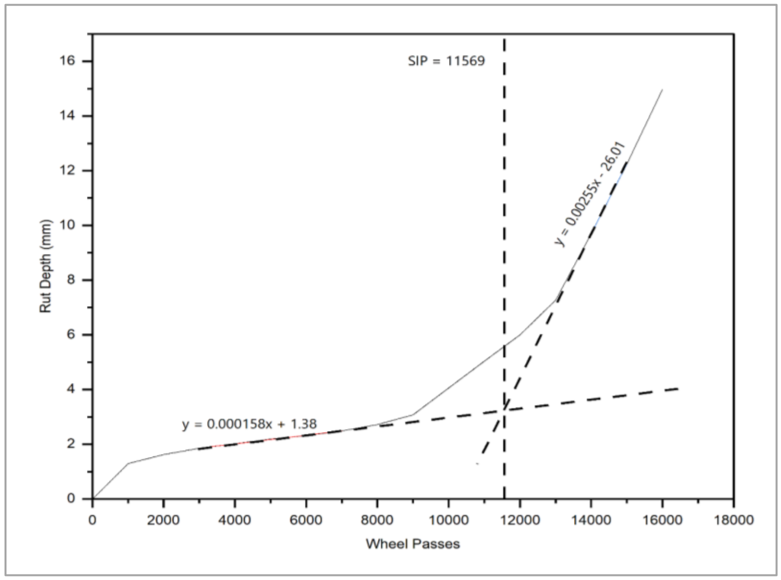


Figure C22. Rut Depth vs Pass Count for Mix 22.

Table C23. Rut depth for Mix M23.

Pass Count	Rut Depth (mm)
0	0
1000	2.30
2000	3.68
3000	4.32
4000	4.87
5000	5.36
6000	5.90
7000	6.82
8000	8.14
9000	9.55
10000	10.66
11000	11.60
12000	12.54
13000	13.57
14000	14.65
15000	16.21
16000	17.81
17000	19.11

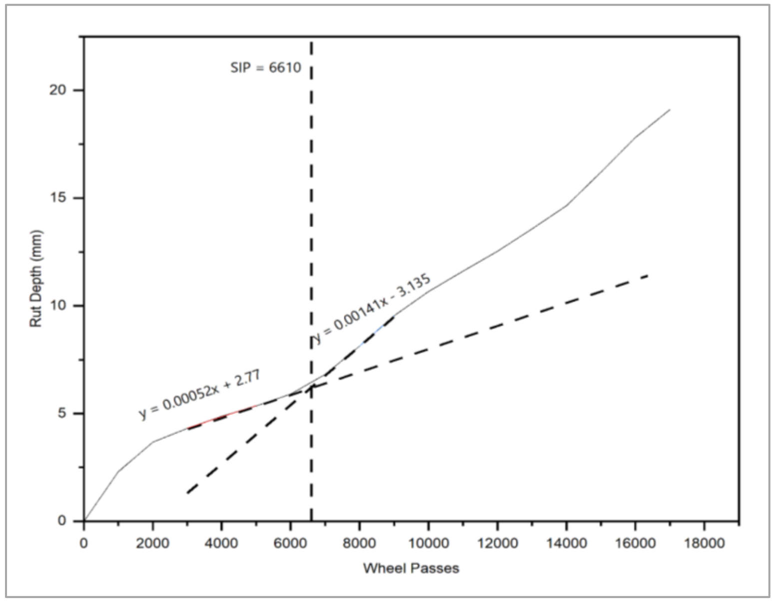


Figure C23. Rut Depth vs Pass Count for Mix 23.

Table C24. Rut depth for Mix M24.

Pass Count	Rut Depth (mm)
0	0
1000	1.23
2000	1.69
3000	2.08
4000	2.44
5000	2.82
6000	3.34
7000	3.98
8000	4.83
9000	5.91
10000	7.79
11000	9.44
12000	10.85
13000	12.88
14000	14.14
15000	15.74
16000	17.08
17000	18.22

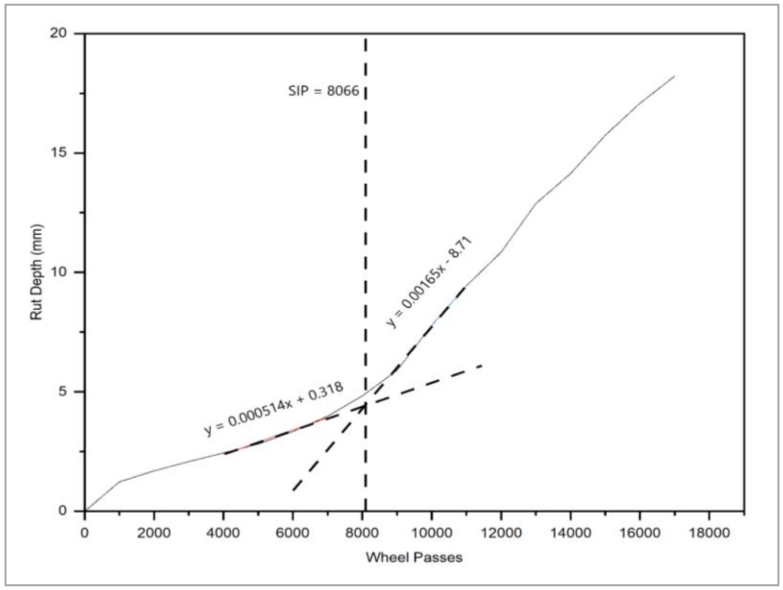


Figure C24. Rut Depth vs Pass Count for Mix 24.

Table C25. Rut depth for Mix M25.

Pass Count	Rut Depth (mm)
0	0
1000	1.55
2000	2.12
3000	2.18
4000	2.37
5000	2.52
6000	2.68
7000	2.82
8000	2.94
9000	3.04
10000	3.10
11000	3.19
12000	3.25
13000	3.32
14000	3.39
15000	3.46
16000	3.51
17000	3.53
18000	3.54
19000	3.60
20000	3.63

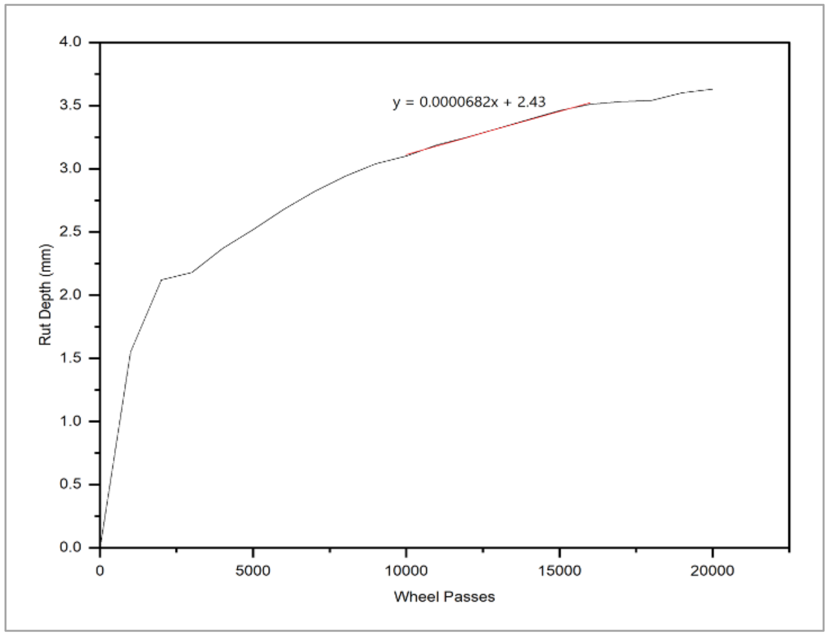


Figure C25. Rut Depth vs Pass Count for Mix 25.

Table C26. Rut depth for Mix M26.

Pass Count	Rut Depth (mm)
0	0
1000	0.95
2000	1.36
3000	1.83
4000	2.11
5000	2.44
6000	2.88
7000	3.42
8000	4.18
9000	5.12
10000	6.17
11000	7.55
12000	9.29
13000	11.01
14000	12.84
15000	14.64
16000	15.91
17000	16.83
18000	17.65
19000	18.10
20000	18.91

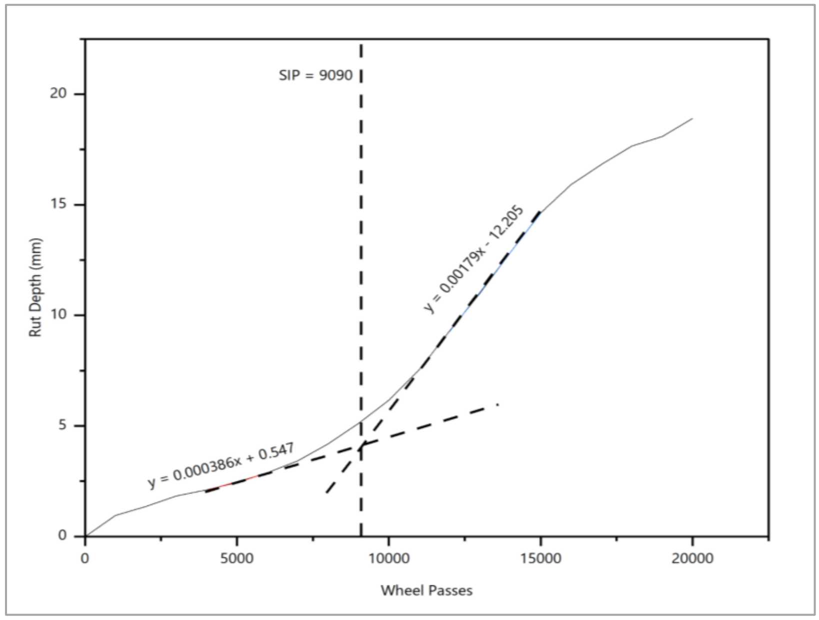


Figure C26. Rut Depth vs Pass Count for Mix 26.

Table C27. Rut depth for Mix M27.

Pass Count	Rut Depth (mm)
0	0
1000	1.08
2000	1.30
3000	1.42
4000	1.53
5000	1.62
6000	1.71
7000	1.78
8000	1.86
9000	1.90
10000	1.96
11000	2.02
12000	2.06
13000	2.10
14000	2.15
15000	2.19
16000	2.25
17000	2.28
18000	2.32
19000	2.37
20000	2.38

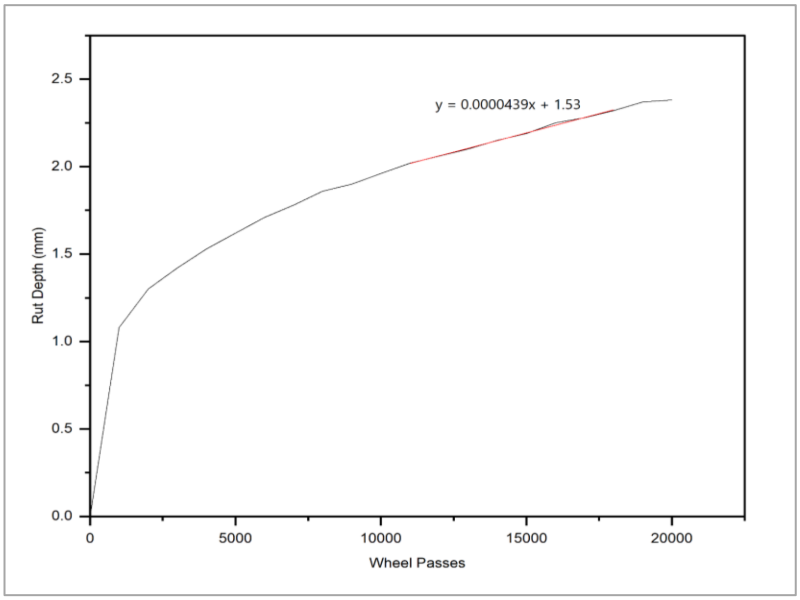


Figure C27. Rut Depth vs Pass Count for Mix 27.

Table C28. Rut depth for Mix M28.

Pass Count	Rut Depth (mm)
0	0
1000	1.78
2000	2.53
3000	3.03
4000	3.69
5000	4.04
6000	4.27
7000	4.44
8000	4.59
9000	4.70
10000	4.80
11000	4.88
12000	4.96
13000	5.08
14000	5.19
15000	5.31
16000	5.42
17000	5.53
18000	5.64
19000	5.76
20000	5.87

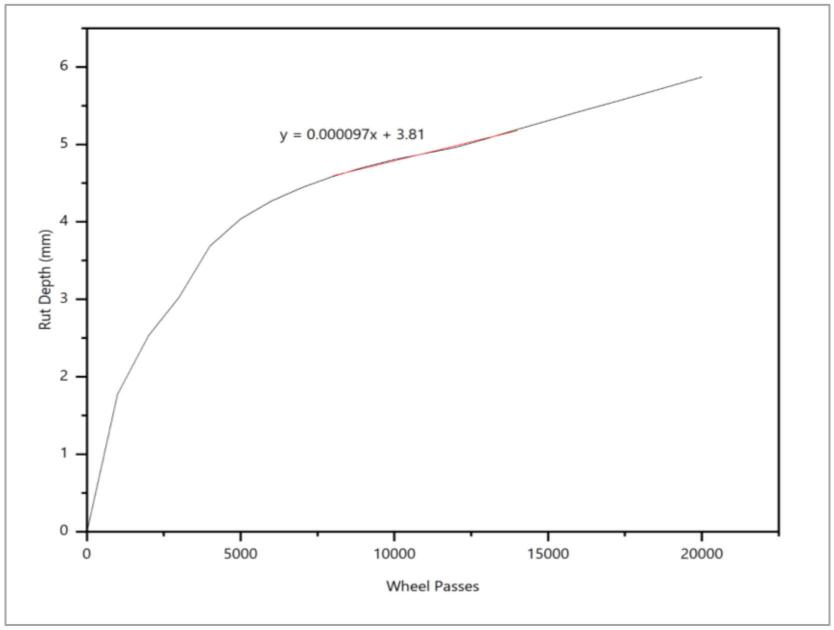


Figure C28. Rut Depth vs Pass Count for Mix 28.

Table C29. Rut depth for Mix M29.

Pass Count	Rut Depth (mm)
0	0
1000	1.72
2000	2.24
3000	2.56
4000	2.78
5000	2.97
6000	3.14
7000	3.30
8000	3.49
9000	3.67
10000	3.85
11000	4.04
12000	4.28
13000	4.48
14000	4.67
15000	4.89
16000	5.13
17000	5.36
18000	5.62
19000	5.86
20000	6.15

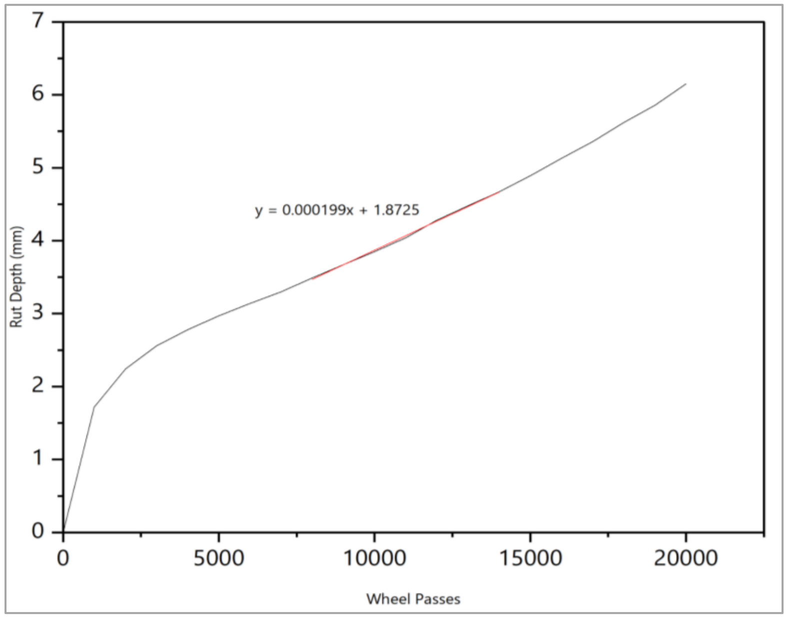


Figure C29. Rut Depth vs Pass Count for Mix 29.

Table C30. Rut depth for Mix M30.

Pass Count	Rut Depth (mm)
0	0
1000	1.77
2000	2.49
3000	3.31
4000	3.75
5000	4.12
6000	4.48
7000	4.87
8000	5.30
9000	5.71
10000	6.12
11000	6.54
12000	6.94
13000	7.29
14000	7.63
15000	7.93
16000	8.44
17000	8.96
18000	9.50
19000	10.26
20000	10.68

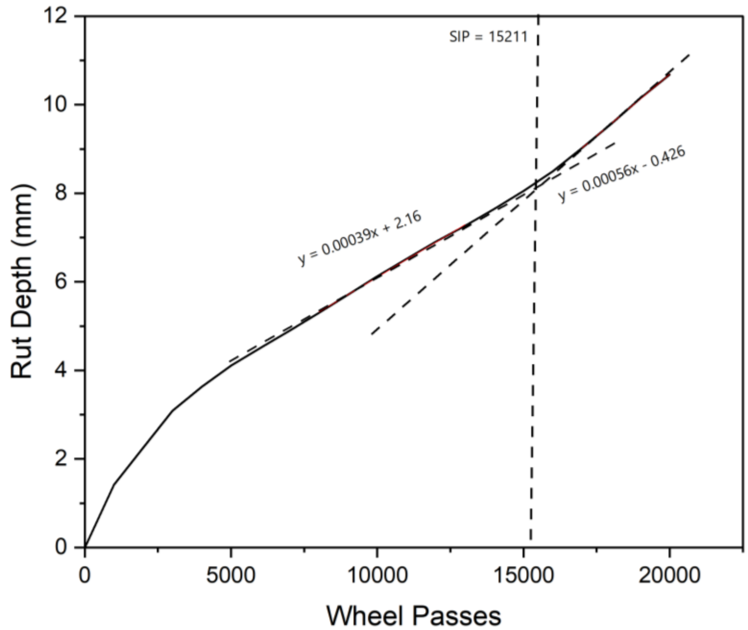


Figure C30. Rut Depth vs Pass Count for Mix 30.

Table C31. Rut depth for Mix M31.

Pass Count	Rut Depth (mm)
0	0
1000	1.04
2000	1.67
3000	1.86
4000	2.08
5000	2.33
6000	2.68
7000	3.28
8000	3.96
9000	4.92
10000	6.08
11000	7.25
12000	8.72
13000	10.95
14000	13.24
15000	15.02
16000	16.44
17000	17.28
18000	17.56
19000	18.15
20000	18.82

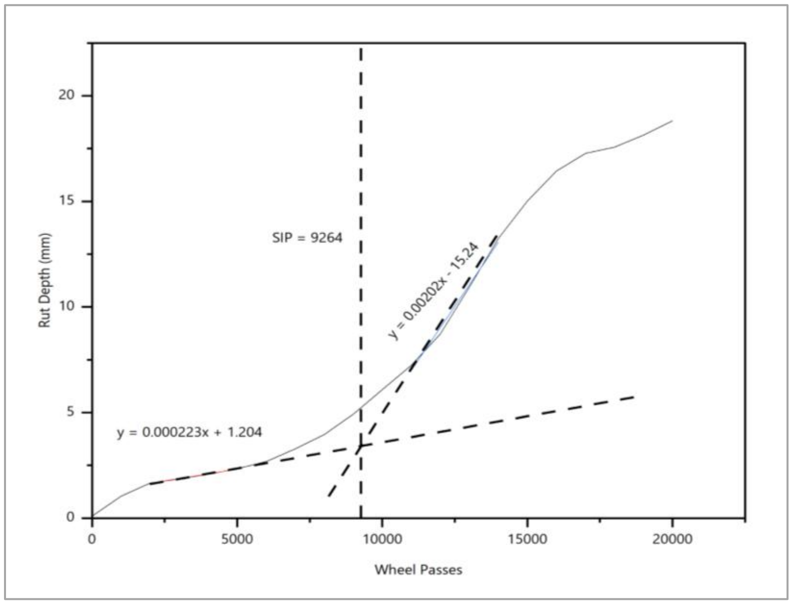


Figure C31. Rut Depth vs Pass Count for Mix 31.

Table C32. Rut depth for Mix M32.

Pass Count	Rut Depth (mm)
0	0
1000	2.50
2000	3.23
3000	3.66
4000	3.98
5000	4.25
6000	4.48
7000	4.70
8000	4.92
9000	5.19
10000	5.72
11000	6.14
12000	6.56
13000	7.33
14000	8.08
15000	9.00
16000	10.26
17000	11.62
18000	11.96
19000	14.15
20000	15.59

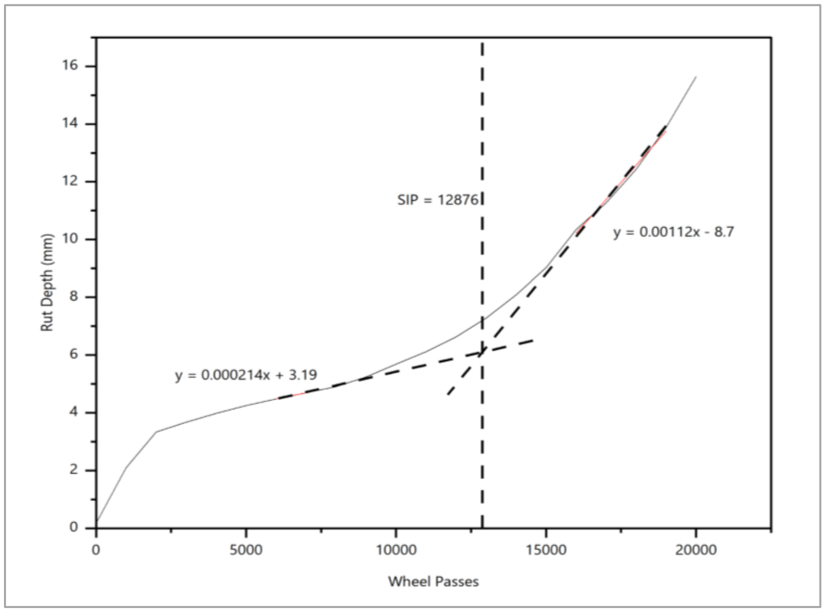


Figure C32. Rut Depth vs Pass Count for Mix 32.

Table C33. Rut depth for Mix M33.

Pass Count	Rut Depth (mm)
0	0
1000	1.17
2000	1.47
3000	1.62
4000	1.77
5000	1.91
6000	2.01
7000	2.13
8000	2.24
9000	2.31
10000	2.42
11000	2.52
12000	2.61
13000	2.69
14000	2.80
15000	2.93
16000	3.05
17000	3.19
18000	3.22
19000	3.36
20000	3.52

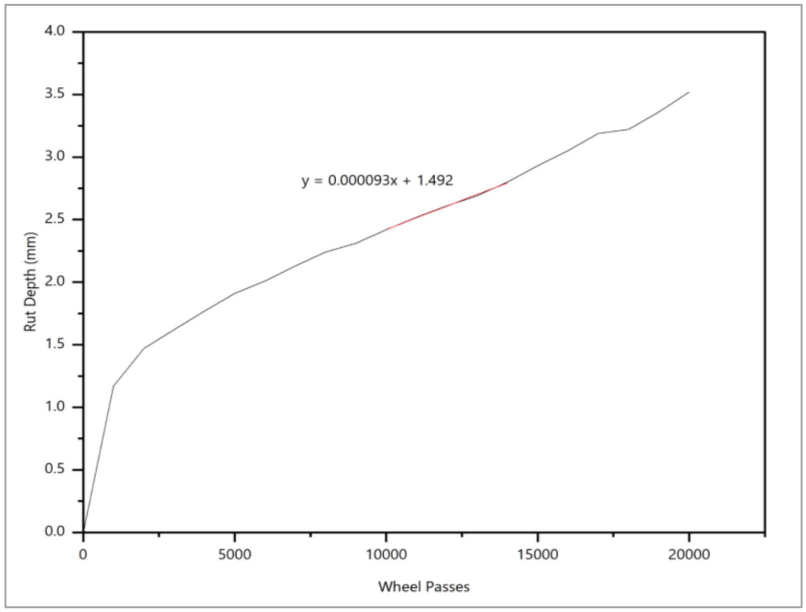


Figure C33. Rut Depth vs Pass Count for Mix 33.

APPENDIX D

Frequency-Temperature Sweep Test Results and Figures

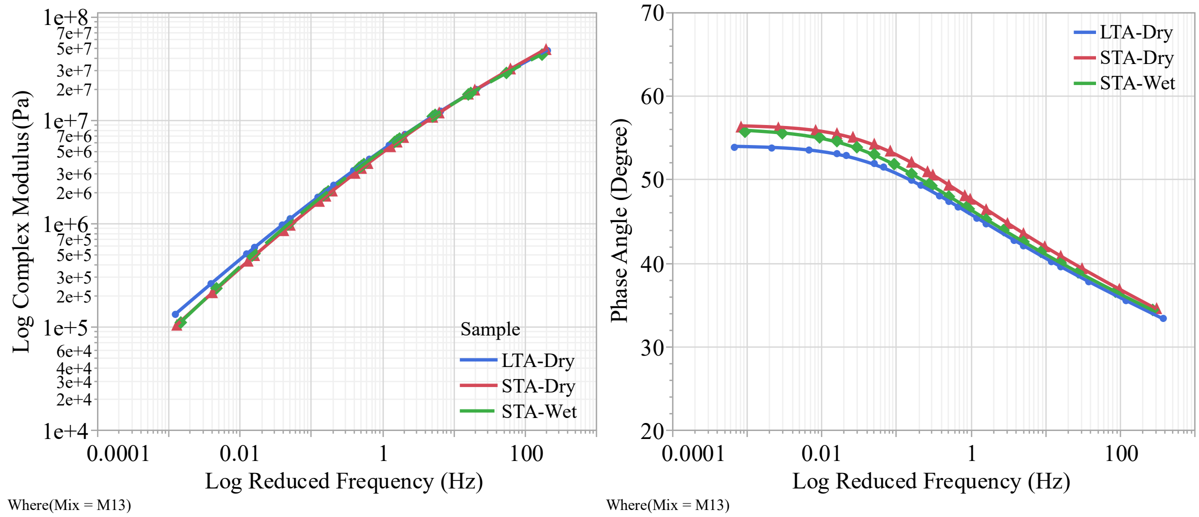


Figure D1. Complex Modulus and Phase Angle Master Curves under Different Conditions for M13.

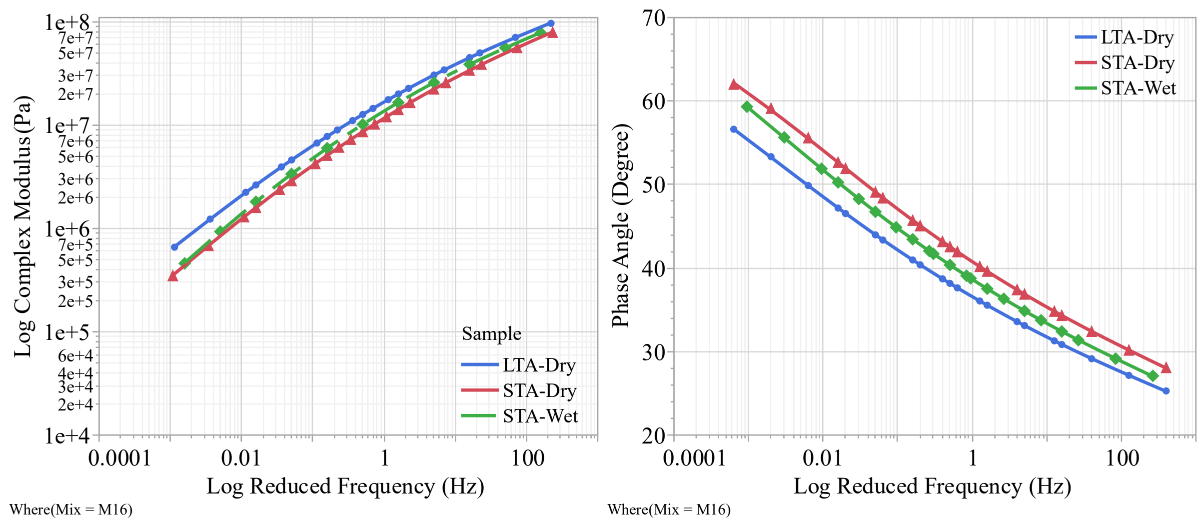


Figure D2. Complex Modulus and Phase Angle Master Curves under Different Conditions for M16.

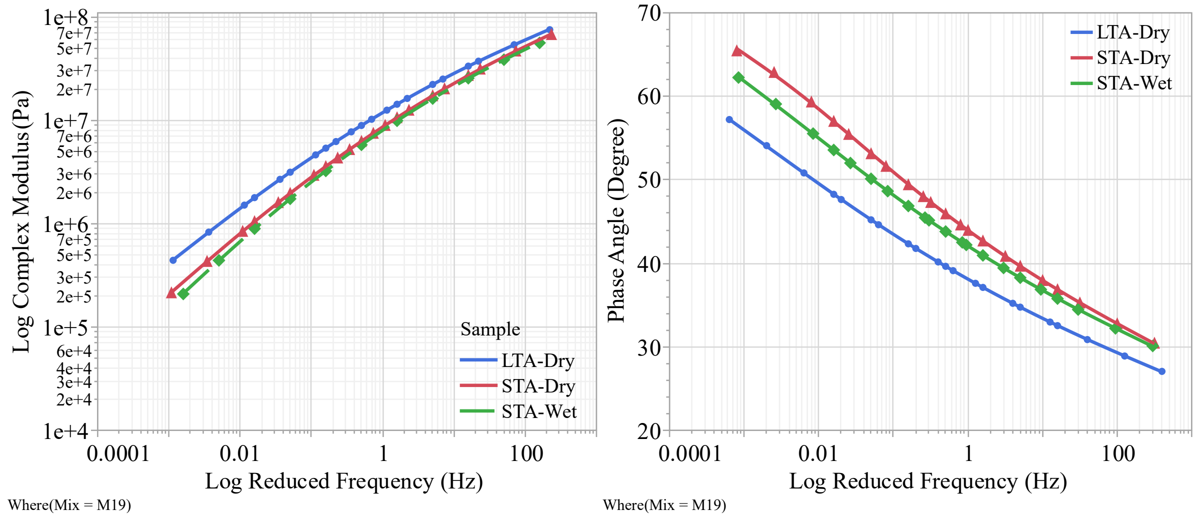


Figure D3. Complex Modulus and Phase Angle Master Curves under Different Conditions for M19.

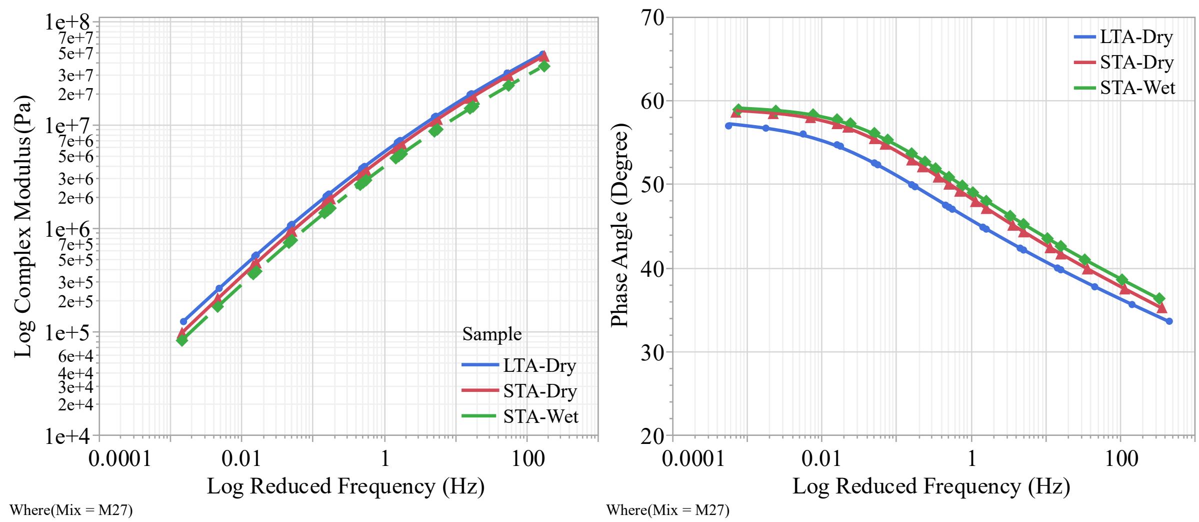


Figure D4. Complex Modulus and Phase Angle Master Curves under Different Conditions for M27.

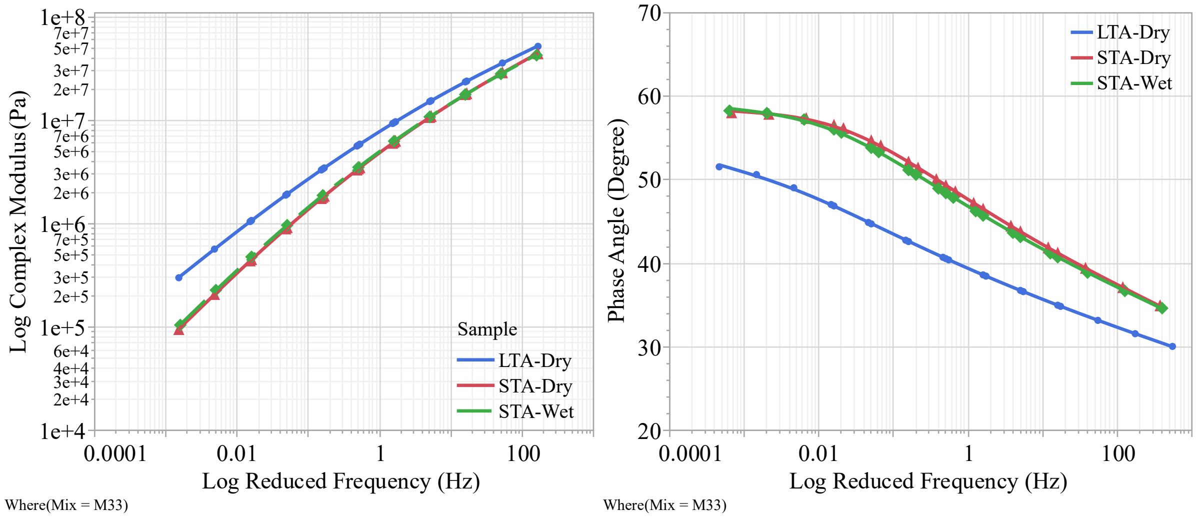


Figure D5. Complex Modulus and Phase Angle Master Curves under Different Conditions for M33.

APPENDIX E

Simplified Guideline for HWTT Analysis

The following steps outline a practical method to calculate CRD, SN, and LCST following the NCAT methodology.

1. **Run the HWTT and export raw data**

Collect the Rut Depth (RD) versus load cycles data for each mix.

2. **Smooth the rut-depth curve**

Apply a light smoothing method (moving average or low-order polynomial) to reduce measurement noise. The goal is to keep the original curve shape unchanged.

3. **Identify the creep portion of the curve**

In ORIGIN or similar software, a linear model is fitted to the early steady-state region of the curve (creep phase). This helps confirm that the curve shape is consistent with expected post-compaction and creep behavior. The slope is CS Parmeter. Then, fit another linear model to the stripping phase region of the curve. The slope is SS parameter. The SIP is graphically represented at the intersection of the fitted lines that characterize the creep phase and the stripping phase as shown in Figure E1.

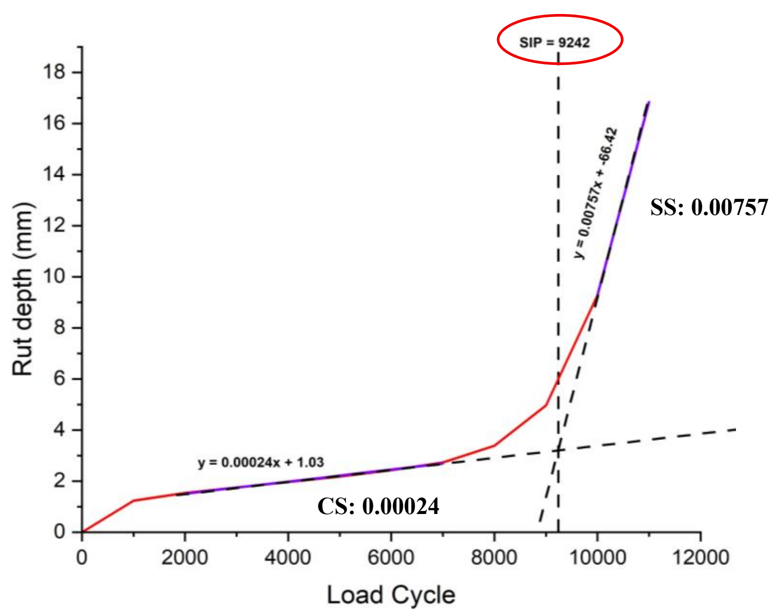


Figure E1. Typical HWTT output of rut depth versus load cycle

4. **Fit the complete HWTT curve using NCAT's visco-plastic model**

Fit the three-parameter deformation model (Equation E1) described in the NCAT study (Yin et al, 2014).

$$RD_{LC} = \rho * [\ln(LC_{ult}/LC)]^{-(1/\beta)} \qquad E1$$

RD_{LC} = rut depth at certain number of load cycles (mm); LC = Number of load cycles; LC_{ult}, ρ, and β = model coefficients.

The fitted curve separates two behaviors illustrated in Figure E2:
• Negative curvature (rutting dominated by visco-plastic deformation)
• Positive curvature (onset of moisture-induced stripping)

5. **Determine the Stripping Number (SN)**

To identify SN, the critical number of wheel passes where the curvature of the fitted curve changes from negative to positive, take the second derivative of Equation (E2), and set it to zero. Then, solve SN formula (Equation E2):

$$SN = LC_{ult} * e^{-(1+1/\beta)} \qquad E2$$

SN represents the maximum number of load cycles that the asphalt mixture can resist in the HWTT before the adhesive fracture between the asphalt binder and the aggregate occurs.

This marks the onset of stripping and replaces the traditional SIP, which can vary depending on the chosen fitting lines.
LC_{ST} is defined as the difference between the number of wheel passes at a rut depth of 12.5 mm ($LC_{12.5}$) and the stripping number (SN), representing the remaining load cycles the mixture can sustain after stripping initiates.

1. **Calculate Corrected Rut Depth (CRD)**

Then, the rut depth data up to SN is fitted with the modified Tseng-Lytton model (Equation E3) and utilized to extrapolate the rut depth at wheel passes beyond SN.

$$CRD = RD_{ult} * e^{(-(\propto/N)^\wedge\lambda)} \qquad E3$$

RD_{ult}, ∝, and λ = model coefficients

This allows to isolate deformation caused by permanent mixture rutting from deformation caused by stripping.
CRD at 10,000 passes (CRD_{10K}) or 20 000 passes (CRD_{20K}) reflects true rutting performance, independent of moisture-induced failure.

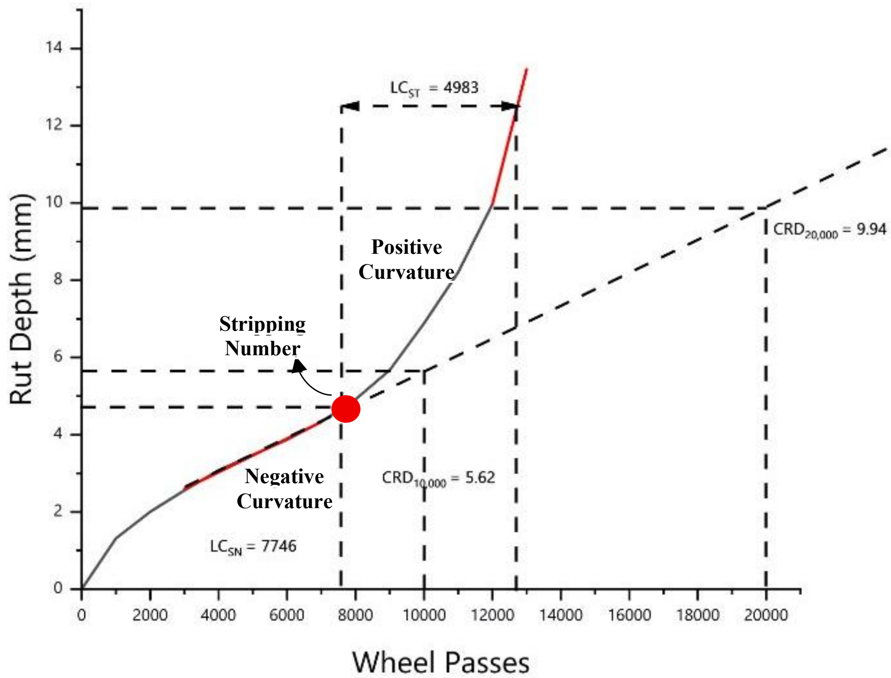


Figure E2. HWTT NCAT parameters determination

Table E1: SN, LC_{ST}, and $LC_{12.5}$ values of all mixes.

Mix	SN	$LC_{12.5}$	LC_{ST}
M1	No SN	No SN	No SN
M2	No SN	No SN	No SN
M3	6,403	26,237	19,834
M4	4,499	18,442	13,923
M5	No SN	No SN	No SN
M6	2,716	12,363	9,648
M7	No SN	No SN	No SN
M8	3,133	13,216	10,083
M9	No SN	No SN	No SN
M10	No SN	No SN	No SN
M11	1,817	10,649	8,832
M12	5,581	20,976	15,395
M13	No SN	No SN	No SN
M14	4,558	20,913	16,355
M15	No SN	No SN	No SN
M16	No SN	No SN	No SN
M17	4,457	19,751	15,294
M18	3,018	12,372	9,355
M19	2,664	12,766	10,102
M20	1,839	15,035	13,195
M21	7,667	34,363	26,696
M22	2,361	15,276	12,915
M23	3,635	12,071	8,435
M24	1,806	12,755	10,949
M25	No SN	No SN	No SN
M26	1,777	13,791	12,014
M27	No SN	No SN	No SN
M28	No SN	No SN	No SN
M29	No SN	No SN	No SN
M30	8,318	23,877	15,559
M31	2,036	13,698	11,662
M32	4,527	18,184	13,656
M33	No SN	No SN	No SN