Performance Evaluation of HWY-2017-29 Project Asphalt Mixtures and Pavement

NCDOT Project 2022-06
FHWA/NC/2022-06
March 2025

Y. Richard Kim, Ph.D., P.E.
Dept. of Civil, Construction, &
Environmental Engineering
North Carolina State University

RESEARCH & DEVELOPMENT

Performance Evaluation of HWY-2017-29 Project Asphalt Mixtures and Pavement

FINAL REPORT

Submitted to:

North Carolina Department of Transportation
(Project No. RP2022-06)

Submitted by:

Y. Richard Kim, Ph.D., P.E., F.ASCE
Jimmy D. Clark Distinguished University Professor
Campus Box 7908
Department of Civil, Construction & Environmental Engineering
North Carolina State University
Raleigh, NC 27695-7908
Tel: 919-515-7758, Fax: 919-515-7908
kim@ncsu.edu

Ala' Hudaib
Graduate Research Assistant
Department of Civil, Construction & Environmental Engineering
North Carolina State University
Raleigh, NC 27695-7908
aahudaib@ncsu.edu

Dong Ho Yoo
Graduate Research Assistant
Department of Civil, Construction & Environmental Engineering
North Carolina State University
Raleigh, NC 27695-7908
Dyoo4@ncsu.edu

Department of Civil, Construction, and Environmental Engineering
North Carolina State University
Raleigh, NC

February 2025

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA/NC/2022-06	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Performance Evaluation of HWY-2017-29 Project Asphalt Mixtures and Pavement		5. Report Date March 2025 6. Performing Organization Code:
7. Author(s) Y. Richard Kim, Ph.D., P.E., Ala' Hudaib, Dong Ho Yoo		8. Performing Organization Report No.
9. Performing Organization Name and Address Campus Box 7908, Dept. of Civil, Construction, & Environmental Engineering, NC State University, Raleigh, NC 27695-7908		10. Work Unit No. 11. Contract or Grant No.
12. Sponsoring Agency Name and Address North Carolina Department of Transportation Research and Development 1020 Birch Ridge Dr Raleigh, NC 27610		13. Type of Report and Period Covered Final Report, August 1, 2021 – Dec. 31, 2024 14. Sponsoring Agency Code RP2022-06
15. Supplementary Notes		

16. Abstract

This report presents the findings of the NCDOT Project RP2022-06, which evaluated the performance of asphalt mixtures that were used in two construction projects. Laboratory testing was conducted on RS9.5C, RI19.0C, and RB25.0C asphalt mixtures to assess fundamental performance metrics, including resistance to cracking and rutting using Asphalt Mixture Performance Tester as well as balanced mix design tests. The study evaluated the impact of variations in binder content and air voids on asphalt mixture performance, revealing apparent performance differences across samples. The study investigated the impact of short-term aging protocols and found that using two hours instead of four hours for short-term aging yields better agreement between lab-mixed lab-compacted and plant-mixed lab-compacted samples. This finding is essential for improving the accuracy of laboratory predictions and aligning them more closely with real-world performance. Index-volumetrics relationships and performance-volumetrics relationships were calibrated and verified, providing essential tools for predicting the performance of quality control/quality assurance samples and evaluating material behavior under different conditions. The project also determined aging model coefficients for the included mixtures, enabling the incorporation of aging effects in pavement performance simulations. Pavement performance simulations were performed using FlexPAVE™ version 2.2 software and the results highlighted the importance of intermediate and base mixtures on the full-depth asphalt performance. S_{app} and RSI values were predicted by inputting QC/QA data to the developed IVR functions. The predicted index values showed that RS9.5C remained within the Standard traffic tier, RI19.0C met the Standard tier for S_{app} but variable tiers for RSI (10% Standard, 64% Heavy, 18% Very Heavy, 8% Extremely Heavy), and 23% of RB25.0C samples fell below Standard tier according to S_{app} values. Additionally, the report provides a framework for assessing the impact of current tolerance limits for binder content and in-place density on fatigue and rutting performance. Based on the QMS limits, the 95% confidence intervals for S_{app} across all three mixtures remained within the Standard traffic tier, while RSI confidence intervals varied, with RS9.5C spanning Standard to Heavy and RI19.0C ranging from Heavy to Extremely Heavy. This framework can be applied to a more representative sample size of North Carolina asphalt mixtures to develop recommendations for amending the current Quality Management System limits. These findings contribute to optimizing quality control measures and improving the long-term performance of North Carolina's highway infrastructure.

17. Key Words Asphalt mixture, asphalt pavement, AMPT, BMD, short-term aging, IVR, PVR, aging model, QMS limits, quality control	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 95	22. Price N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

DISCLAIMER

The contents of this report reflect the views of the authors and are not necessarily the views of North Carolina State University. The authors are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the North Carolina Department of Transportation at the time of publication. This report does not constitute a standard, specification, or regulation.

ACKNOWLEDGEMENTS

This research was sponsored by the North Carolina Department of Transportation. The Steering and Implementation Committee was comprised of Shihai Zhang, P.E., Charles Colgate, Wiley Jones, P.E., Tony Collins, James B. Phillips, P.E., Brian Hunter, Boyd Tharrington, Andrew D. Wargo, Ph.D., P.E., Neil Mastin, P.E., Todd W. Whittington, P.E., and Mustan Kadibhai, P.E. (PM). These advisors have given invaluable direction and support to the research team throughout the project.

EXECUTIVE SUMMARY

This report presents the findings of the NCDOT Project RP2022-06, which evaluated the performance of asphalt mixtures that were used in two construction projects, namely, east end connector and Rockingham bypass projects. The included mixtures consisted of RS9.5C for the surface layer, RI19.0C for the intermediate layer, and RB25.0C for the base layer, all commonly used in North Carolina highway projects. The research employed comprehensive testing methods, including the suite of Asphalt Mixture Performance Tester (AMPT) tests as well as balanced mix design (BMD) tests, to understand how construction variability influences key performance outcomes. Throughout the study, plant-mixed lab-compacted (PMLC) and lab-mixed lab-compacted (LMLC) samples were tested to determine their sensitivity to changes in binder content and air void content. It was demonstrated how variations in binder content and in-place air voids impacted cracking resistance and rutting potential. At the same time, the complexities involved in predicting the effects of binder content and air void content on cracking performance from the quality assurance (QA) data when these two factors change simultaneously. The research compared the performance of LMLC samples that were short-term aged for two and four hours to the performance of PMLC samples. The results showed better agreement between the performance of LMLC samples aged for two hours and the PMLC samples' performance.

Performance testing was then conducted on LMLC samples prepared at four different volumetric conditions called "four corners". The results were used to calibrate the index-volumetrics relationships (IVRs), and the accuracy of the function's predictions was verified for LMLC and PMLC samples. S_{app} and RSI values were predicted by inputting QC/QA data to the developed IVR functions. The predicted index values showed that RS9.5C remained within the Standard traffic tier, RI19.0C met the Standard tier for S_{app} but variable tiers for RSI (10% Standard, 64% Heavy, 18% Very Heavy, 8% Extremely Heavy), and 23% of RB25.0C samples fell below Standard tier according to S_{app} values. These findings highlight the impact of production variability on mixture performance and potential fatigue cracking risks. FlexPAVE™ simulations were performed on the four corners of the surface mixtures showing insignificant differences between the different conditions. The results highlighted the significance and role of the intermediate and base mixtures on the full-depth asphalt pavement performance. The researchers then utilized the calibrated IVR functions to quantify the impact of the limits in the NCDOT Quality Management System (QMS) for binder content and in-place density on fatigue and rutting performance. The suggested framework showed that the QMS limits yielded +/- 21%, +/- 32%, and +/- 28% change in S_{app} values based on 95% confidence interval ranges for the surface, intermediate, and base mixtures, respectively. The results also showed +/- 46%, and +/- 51% change in RSI values based on 95% confidence interval ranges for the surface, and intermediate mixtures, respectively. Based on the QMS limits, the 95% confidence intervals for S_{app} across all three mixtures remained within the Standard traffic tier, while RSI confidence intervals varied, with RS9.5C spanning Standard to Heavy and RI19.0C ranging from Heavy to Extremely Heavy. These results highlight how QMS limits for binder and air void content influence traffic designation for asphalt mixtures. The developed framework provides a systematic approach for evaluating the impact of QMS limits on fatigue and rutting performance that can be used with a large enough representative sample of North Carolina standard mixtures to develop recommendations for the performance-based QMS limits in a way that ensures improvement to asphalt mixtures performance and asphalt pavement durability.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Item	Definition
AM-PRS	Asphalt Mixture Performance-Related Specifications
AMPT	Asphalt Mixture Performance Tester
AQCs	Acceptance quality characteristics
JMF	Job mix formula
LVDT	Linear variable differential transformer
NCSU	North Carolina State University
NCDOT	North Carolina Department of Transportation
NMAS	Nominal maximum aggregate size
IVR	Index-volumetric relationship
PVR	Performance-volumetric relationship
QA	Quality assurance
RAP	Reclaimed asphalt pavement
RAS	Reclaimed asphalt shingles
RSI	Rutting Strain Index
SSR	Stress sweep rutting
S-VECD	Simplified viscoelastic continuum damage
VFA	Voids filled with asphalt
VMA	Voids in mineral aggregate
QMS	Quality management system

CHAPTER 1. INTRODUCTION

The acceptance quality characteristics (AQCs) commonly used in most quality-related specifications focus on the volumetric properties of asphalt mixtures. This preference arises from the fact that these properties are easy to measure, relatively consistent, and have a recognized influence on pavement durability. However, a key challenge is that the relationship between volumetric properties and actual pavement performance varies depending on the quality of materials used. Consequently, relying solely on volumetric control introduces performance uncertainties. To address this limitation, performance-related specifications (PRS) are recommended. PRS integrates material and construction characteristics that are more directly linked to performance. The main advantage of PRS lies in its emphasis on AQCs that correlate with essential engineering properties, which can be material-specific but identifiable through supplementary testing beyond standard quality assurance (QA) procedures.

Given the potential of PRS, the Federal Highway Administration (FHWA) has invested in developing asphalt mixture PRS (AM-PRS). These are implemented through:

1. Fundamental material testing using the Asphalt Mixture Performance Tester (AMPT)
2. Analysis based on mechanistic performance prediction models
3. Cracking and rutting index parameters
4. Index-volumetrics relationships (IVRs) and performance-volumetrics relationships (PVRs)
5. Index-based and predictive AMPT-based balanced mix design (so called BMD+) methods
6. Software tools such as FlexMAT™, FlexPAVE™, and FlexMIX™ for material analysis, pavement performance prediction, and mix design
7. Pay tables accounting for differences in pavement life between as-designed and as-constructed pavements

These PRS tools provide a unified framework linking asphalt mixture design, pavement design, and construction QA through consistent testing methods and models. In particular, IVRs and PVRs allow the estimation of mixture performance for a wide range of volumetric conditions using mixture performance and the volumetric properties of the so-called ‘four corners’ conditions. This ability of IVRs and PVRs will be useful in developing pay factors for AQCs and determining acceptance limits based on pavement performance.

In the HWY-2017-29 research project, PRS concepts were tested using samples from a newly constructed full-depth asphalt pavement on Carr Road in Durham, NC. The pavement consisted

of a 76.2 mm (3 inches) thick asphalt concrete (AC) surface layer using RS9.5C mix, a 101.6 mm (4 inches) thick intermediate layer with RI19.0C mix, and a 111.76 mm (4.4 inches) thick base layer with RB25.0C mix. The NCSU research team collected plant-mixed, lab-compacted (PMLC) samples and conducted AMPT tests. The results were then analyzed using FlexPAVE version 1.1 to predict fatigue cracking and rutting performance. The team developed PVRs for RS9.5C and RI19.0C using FlexPAVE predictions and volumetric properties of the four corners samples. However, aging effects were not factored into FlexPAVE at that time since the aging model was not incorporated in FlexPAVE at that time.

Simple index tests have gained popularity among state highway agencies (SHAs) in the context of balanced mix design (BMD) for asphalt mixtures. Among these, the IDEAL-CT and Hamburg wheel-tracking (HWT) tests are widely accepted. This research project leveraged materials and construction samples from an actual project to give the North Carolina Department of Transportation (NCDOT) an opportunity to assess the feasibility of using IDEAL-CT and HWT (collectively referred to as BMD tests) as part of their QA processes.

The objectives of this research are as follows:

1. Assess how construction variability affects asphalt mixture performance
2. Determine how well AMPT performance tests, IDEAL-CT, and HWT predict asphalt mixture performance under various volumetric conditions
3. Evaluate the impact of Quality Management System (QMS) limits on asphalt mixture performance

CHAPTER 2. TEST METHODS, MODELS, SOFTWARE PROGRAMS, AND MATERIALS

Test methods, models, and software programs

Asphalt Mixture Performance Tester (AMPT) test methods

The AM-PRS framework is built upon AMPT performance tests, comprehensive mechanistic models, and advanced software tools designed for mixture-level analysis and pavement performance predictions. Within the framework of the AM-PRS, asphalt mixtures are subjected to three AMPT performance tests: the dynamic modulus ($|E^*|$) test (AASHTO TP 132) to characterize the mixture's stiffness as a function of loading frequency and temperature, the cyclic fatigue test (AASHTO T 400) to characterize the mixture's cracking resistance, and the stress sweep rutting (SSR) test (AASHTO TP 134) to characterize the mixture's rutting resistance. This suite of AMPT tests takes a total of five days (two days for specimen fabrication and three days for AMPT testing) and is typically referred to as "BMD+ tests".

Dynamic modulus test

The dynamic modulus is a fundamental performance-related property used to evaluate asphalt mixture stiffness. The dynamic modulus ($|E^*|$) is determined by calculating the ratio of the stress amplitude to the strain amplitude when the material is subjected to sinusoidal loading within the linear viscoelastic range of 50-75 microstrain. Asphalt materials are classified as thermorheologically simple materials. This classification means that the combined effects of loading frequency and temperature can be described using a single parameter called reduced frequency. The reduced frequency is used to generate a dynamic modulus mastercurve through time-temperature shift factors by utilizing the two springs, two parabolic elements, and one dashpot (2S2P1D) model. In the 2S2P1D model, the two springs represent the elastic response, capturing the immediate deformation upon loading. The two parabolic elements account for time-dependent viscoelastic behavior, modeling the delayed response under applied loading. The single dashpot represents the viscous component, describing the material's permanent deformation over time (Olard and Di Benedetto, 2003). The 2S2P1D model is used in the AM-PRS framework instead of a sigmoidal function due to its ability to represent the low-temperature behavior of asphalt concrete. A summary of the equations governing the 2S2P1D model, as described in AASHTO T 411, is provided in Equations (1) to (5). A typical dynamic modulus mastercurve is shown in Figure 1.

$$E'_{2S2P1D} = E_{00} + \frac{E'_1}{\left(\frac{E'_1}{E_0 - E_{00}}\right)^2 + \left(\frac{E'_2}{E_0 - E_{00}}\right)^2} \quad (1)$$

$$E''_{2S2P1D} = \frac{E'_2}{\left(\frac{E'_1}{E_0 - E_{00}}\right)^2 + \left(\frac{E'_2}{E_0 - E_{00}}\right)^2} \quad (2)$$

$$E'_1 = (E_0 - E_{00})\left[1 + \delta \times (\omega_R \tau_E)^{-k} \times cos\left(\frac{k\pi}{2}\right) + (\omega_R \tau_E)^{-h} \times cos\left(\frac{h\pi}{2}\right)\right] \quad (3)$$

$$E'_2 = (E_0 - E_{00})\left[\delta \times (\omega_R \tau_E)^{-k} \times sin\left(\frac{k\pi}{2}\right) + (\omega_R \tau_E)^{-h} \times sin\left(\frac{h\pi}{2}\right) + (\omega_R \tau_E \beta)^{-1}\right] \quad (4)$$

$$|E^*| = \sqrt{(E'_{2S2P1D})^2 + (E''_{2S2P1D})^2} \quad (5)$$

where

E' = storage modulus,
E'' = loss modulus,
E_0 = maximum storage modulus value,
E_{00} = minimum storage modulus value,
ω_R = reduced angular frequency, and
$\tau_E, \beta, \delta, k, h$ = fitting coefficients.

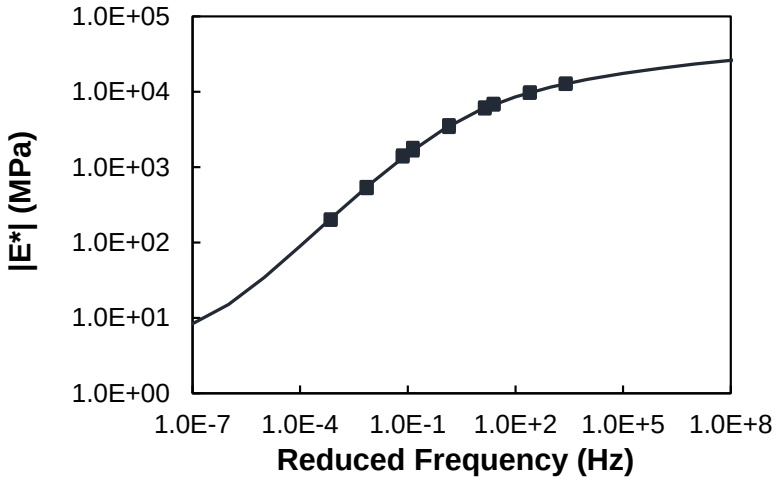

Figure 1: Dynamic Modulus Mastercurve.

Dynamic modulus tests are conducted at three different loading rates (10 Hz, 1 Hz, and 0.1 Hz) and three temperatures (4°C, 20°C, and 35°C or 40°C). On-specimen strain is measured through three linear variable differential transducers (LVDTs) mounted on the specimen. Dynamic modulus, phase angle, and data quality indicators are then determined. The test is typically conducted using three specimens.

Direct tension cyclic fatigue test

The direct tension cyclic fatigue test is an actuator displacement-controlled test that applies repeated cyclic loadings until failure. During the test, key data are collected, including the

applied stress, axial strain on the specimen, and the number of loading cycles until failure. The primary outputs of this test include the damage characteristic curve, often referred to as the material integrity (*C*) versus damage (*S*) curve, and the failure criterion, D^R. The damage characteristic curve, shown in Figure 2, represents the fundamental relationship between material integrity (*C*) and the accumulation of damage (*S*). The damage characteristic curve is represented by a power model that is unique for each asphalt mixture and is shown in Equation (6). The failure criterion, D^R, quantifies the average reduction in pseudo stiffness as the mixture accumulates damage and approaches failure. D^R is calculated as the slope of the fitted linear relationship between cumulative (*1-C*) to failure and the number of cycles to failure (Wang and Kim 2017) as shown in Figure 3.

$$C = 1 - C_{11}S^{C_{12}} \tag{6}$$

where
C = material integrity,
S = damage, and
C_{11}, C_{12} = fitting coefficients of the power model.

By integrating the damage characteristic curve with the mixture's linear viscoelastic properties, fatigue damage accumulation under cyclic loading can be characterized and predicted. Additionally, the D^R failure criterion identifies the onset of failure. These properties are further utilized in pavement structural analysis models to evaluate the fatigue performance and lifespan of asphalt pavements subjected to repetitive loading. The apparent damage capacity fatigue index, S_{app}, can then be determined. This parameter reflects the combined effects of the material's stiffness and toughness on its fatigue resistance and is shown in Equation

(*7*). A higher S_{app} value indicates a higher capacity to withstand fatigue damage. The test is performed on three replicates and specimen-specific as well mixture-specific properties are determined.

$$S_{app} = 1000^{\frac{\alpha}{2}-1} \frac{a_{T(S_{app})}{}^{\frac{1}{\alpha+1}} \left(\frac{D^R}{C_{11}}\right)^{\frac{1}{C_{12}}}}{|E*|_{LVE,S_{app}}{}^{\frac{\alpha}{4}}} \tag{7}$$

where
$a_{T(S_{app})}$ = shift factor from AMPT test temperature to S_{app} temperature,
α = slope of the relaxation modulus mastercurve,
D^R = failure criterion,
C_{11}, C_{12} = damage characteristic curve fitting coefficients, and
$|E*|_{LVE,S_{app}}$ = dynamic modulus at 10 Hz and S_{app} temperature.

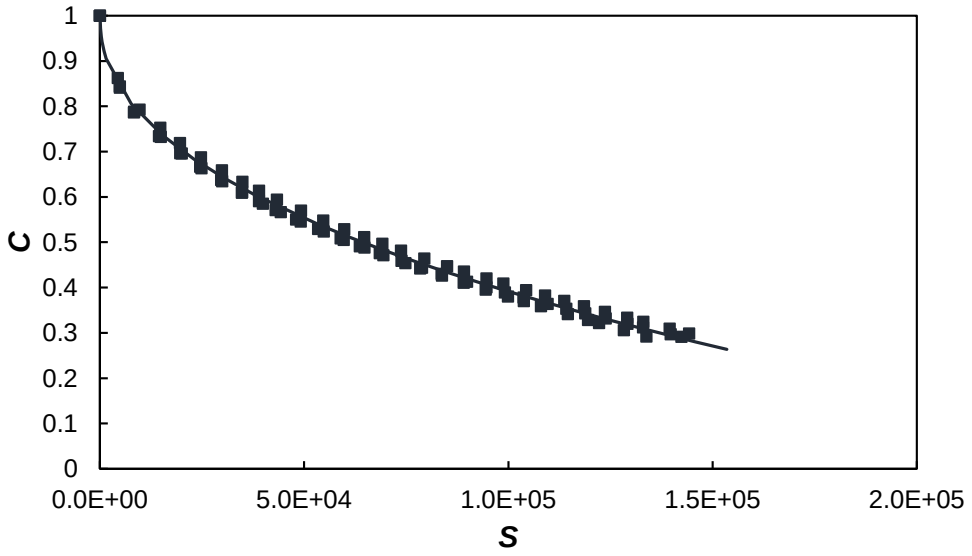

Figure 2: Damage Characteristic Curve.

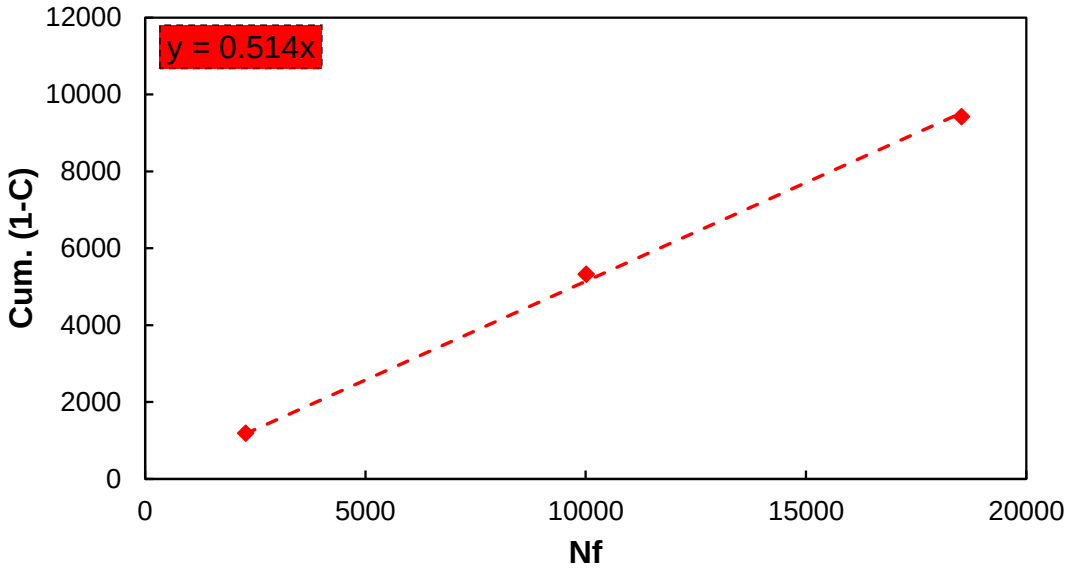

Figure 3: Failure Criterion, D^R.

Stress Sweep Rutting (SSR) test

The SSR test evaluates the permanent deformation characteristics of asphalt mixtures by measuring their response to deviatoric stress, loading duration, and temperature variations. The results from four SSR tests, two conducted at high temperatures and two at low temperatures, are used to develop the shift permanent deformation model. This model describes how an asphalt mixture deforms permanently under stress and can be integrated into pavement structural analysis models for predicting the long-term rutting performance of asphalt pavements.

During SSR testing, the asphalt specimens are evaluated at two different temperatures under a constant confining pressure of 69 kPa (10 psi). Each test involves three loading blocks, each

consisting of 200 cycles, performed at three different deviatoric stress levels. For each cycle, the load pulse duration is 0.4 seconds, followed by a rest period of either 1.6 seconds for the low-temperature test or 3.6 seconds for the high-temperature test. The test temperatures are selected based on AASHTO TP 134 guidelines and depend on the geographic location where the asphalt mixture will be applied.

The determination of the rutting strain index (RSI) involves dividing the asphalt layer into multiple sublayers to accurately capture the distribution of permanent strain throughout the pavement depth. This method ensures that variations in stress and strain, caused by traffic loading, and temperature along the pavement depth and time are properly accounted for. The vertical stress at different depths is first computed using pavement response models, which serve as input for calculating the permanent strain at each sublayer. This strain is determined using results from the SSR test and the permanent deformation shift model. To quantify the overall rutting potential, the permanent strains across all sublayers are integrated to compute the average permanent strain throughout the asphalt layer. The RSI reflects the average level of permanent strain experienced by the asphalt mixture and is defined as the ratio of the permanent deformation in the asphalt layer to the thickness of the layer at the end of a 20-year pavement service life, with 30 million 18-kip standard axle load repetitions applied under a standard pavement structure (Ghanbari et al., 2020). Lower RSI values indicate greater resistance to rutting.

Balanced mix design (BMD) test methods

Indirect Tensile Asphalt Cracking Test (IDEAL-CT)

The IDEAL-CT is a simple and cost-effective test conducted in accordance with ASTM D8225. The test is used to evaluate the cracking resistance of asphalt mixtures. The test is performed at room temperature with minimal specimen preparation, making it suitable for both mix design evaluation and quality control. During the test, a vertical compressive load is applied to the specimen at a constant displacement rate of 50 mm/min until failure. The test produces a load-displacement curve, and the Cracking Tolerance Index (CT-Index) is derived from parameters such as peak load, post-peak slope, and total fracture energy. The CT-Index equation is shown in Equation (8). A higher CT-Index value indicates greater resistance to cracking.

$$CT - Index = \frac{t}{62} \times \frac{G_f}{|m_{75}|} \times \frac{l_{75}}{D} \times 10^6 \quad (8)$$

where
G_f = fracture energy calculated by dividing the work of fracture by the specimen's cross-sectional area,
l_{75} = displacement at 75% of the peak load in the post-peak stage,
m_{75} = slope in the post-peak stage,
t = specimen thickness, and
D = specimen diameter.

Hamburg Wheel Tracker Test (HWTT)

The Hamburg Wheel Tracker Test is standardized under AASHTO T 324. This test evaluates the rutting resistance and moisture susceptibility of asphalt mixtures by simulating wheel loading in a controlled water bath. The test involves subjecting an asphalt specimen to repeated passes of a steel wheel at a specified loading rate under submerged conditions, simulating the combined effects of traffic loading and environmental moisture.

During the test, the steel wheel moves back and forth across the specimen, and the depth of permanent deformation (rutting) is continuously recorded. The stripping inflection point (SIP) can also be determined to evaluate moisture damage, indicating when water-induced stripping accelerates the deformation rate. The lower rut depths and higher SIP values reflect mixtures with superior rutting resistance and moisture durability.

Performance-volumetrics relationships (PVRs) and index-volumetrics relationships (IVRs)

Although the as-constructed pavement performance that is required for determining pay factors is best predicted using AMPT performance test results, the use of AMPT performance test results for QA purposes poses practical problems that are difficult to overcome. For example, the time required to fabricate performance test specimens and conduct all the performance tests is too long for the test results to be used efficiently in typical SHA QA processes. Also, any adjustments to the mixture during construction should be based on AQCs that can easily be changed by a contractor. Due to these limitations, the AM-PRS employs AQCs that are commonly used in most agencies' QA processes for asphalt pavements (e.g., volumetric properties, in-place density, etc.). Then, the as-constructed pavement performance is predicted based on these AQCs.

The PVR function is the relationship between volumetrics and performance and is based on testing mixtures under four different volumetric conditions, i.e., *four corners*. The PVR concept has been applied and verified in several PRS shadow projects, including NCDOT Project HWY-2017-29. The results presented in the HWY-2017-29 illustrated the accuracy and reasonableness of performance predictions using PVRs.

The IVR function is the relationship between performance indexes and the volumetric properties of asphalt mixtures. The underlying concept of the IVR is similar to that of the PVR. IVRs can be developed using the BMD+ index values (i.e., S_{app} and RSI) and the volumetric properties determined from the same four corners that are used in the PVR development. However, the effect of air voids on mixture performance cannot be addressed by CT-Index due to the well-established finding that the change in CT-Index as a function of air voids is opposite to the well-known effect of air voids on the cracking performance of asphalt mixtures. Therefore, IDEAL-CT testing was done at a constant air void content of 4%, resulting in *two points*, not four corners, for the development of the CT-Index IVR.

To establish the four corners, the process begins by identifying the properties of the materials used in the mixture, including gradations and specific gravities. Next, two aggregate gradations are created: one finer and the other coarser than the original mixture's gradation, while still remaining within the same gradation category. The properties of the mixture corresponding to each gradation are then determined. Once these gradations are finalized, two different volumetric conditions are chosen for each gradation, forming the four corner conditions. It is important that

these selected conditions effectively capture the impact of variations in mixture properties on performance. Finally, the mixtures at selected four corners undergo fatigue and rutting performance testing. It is important to note that when evaluating the four corners used in fatigue testing from a rutting perspective, a conflicting effect on rutting performance can be observed. Specifically, corners with high binder content and low air void content or low binder content and high air void content can produce insufficient difference in rut depths for the prediction of rut depths in a wide range of mixture conditions. To mitigate this issue, alternative corners for rutting evaluation were selected that eliminate this conflict. These selected corners maintain either high binder content with high air void content or low binder content with low air void content, providing a wide range of rutting behavior. The main advantage of the four corners concept is that it allows for performance prediction without the need for any additional performance testing after the IVRs and PVRs have been calibrated, using volumetric parameters.

Software programs

In order to aid state agencies' implementation of the AM-PRS, the NCSU research team has developed a set of software programs to support the tasks. These programs include FlexMAT for Cracking and FlexMAT for Rutting for material-level cracking and rutting analyses, respectively, FlexPAVE for pavement performance analysis, and FlexMIX for BMD+. FlexMAT for Cracking and FlexMAT for Rutting are Excel-based programs that are used to process the data files generated by the AMPT for fatigue and rutting analyses, respectively. After importing the appropriate data files, FlexMAT for Cracking performs complex analysis algorithms with minimal inputs from the user to generate a dynamic modulus mastercurve, time-temperature shift factors, S_{app} values, and the input file for FlexPAVE. Similarly, SSR test data are imported into FlexMAT for Rutting to calibrate the shift permanent deformation model, determine the RSI values, and generate the input file for FlexPAVE.

FlexPAVE is a pavement performance analysis program that employs VECD theory to account for the effects of loading rate and temperature on asphalt pavement responses and distress mechanisms. This software program utilizes three-dimensional finite element analysis with moving loads to compute the mechanical responses under various traffic loads. Also, the Enhanced Integrated Climate Model is used in FlexPAVE to introduce realistic climatic conditions to the pavement response and performance calculations. The computed strain and stress values then are used to calculate the fatigue damage and rut depth values.

Materials

Surface, intermediate, and base asphalt mixtures, identified as RS9.5C, RI19.0C, and RB25.0C, respectively, were utilized in this research project. The surface mixture was sampled from the east end connector project, while the intermediate and base mixtures were sampled from the Rockingham bypass project. The research team sampled 5 different samples for each of the surface and intermediate mixtures, and 4 different samples for the base mixtures. The samples were obtained on different days and times throughout the production of these mixtures. The team also sampled component materials that were used in the production of these mixtures to prepare lab-mixed, lab-compacted (LMLC) specimens. Some of the properties of the sampled mixtures are presented in Table 1. Also, some of the properties of the different LMLC conditions can be seen in Table 2. It should be noted that the G_{mm} measured by the research team was used for

calculating volumetric parameters, rather than the value reported in the JMF. In this report, mixture IDs have two letters followed by two numbers. The first letter indicates the type of mix design gradation, i.e., "C" for coarse gradation or "F" for fine gradation. Note that all three mixtures tested in this project are fine-graded mixtures. The second letter "F" indicates a gradation that is close to the upper gradation limits and "C" indicates a gradation that is close to the maximum density line. As for the two numbers, the first one represents the air void content at N_{des}, while the second represents the air void content of the test specimen. For example, FC33 indicates that the mix design gradation is fine while the gradation for this corner is coarser than the mix design's, with an air void content at N_{des} of 3% and a target air void content of the test specimen is 3%.

Table 1: Mixture Information.

	RS9.5	RI19.0C	RB25.0C
Project	East End Connector	Rockingham Bypass	
NMAS (mm)	9.5	19.0	25.0
Gradation Type	Fine	Fine	Fine
Virgin Binder Grade	PG 64-22	PG 58-28	PG 58-28
Pay Binder Grade	PG 64-22	PG 64-22	PG 64-22
Binder Content (%)	5.8	4.8	4.4
RAP/RAS Content (%)	30/0	25/3	25/3
VMA	17.1	15.2	14.3
VFA	77.0	73.8	72.0
G_{mm} (JMF)	2.478	2.499	2.516
G_{mm} (NCSU)	2.417-2.443	2.492-2.511	2.516-2.552

Table 2: RS9.5C LMLC Conditions Properties.

		Binder Content (%)	RAP Content (%)	VMA_{IP}	VFA_{IP}	G_{mm}
Fatigue Testing	FC33	5.8	35	15.8	77.9	2.455
	FC44	5.1	35	15.7	73.3	2.466
	FC57	4.8	35	17.6	59.2	2.480
	FF34	6.9	35	20.1	78.6	2.406
	FF44	6.5	35	18.8	80.3	2.418
	FF58	6.2	35	17.6	59.2	2.430
Rutting Testing	FC37	5.8	35	19.7	59.8	2.455
	FC44	5.1	35	15.4	75.0	2.466
	FC53	4.8	35	14.0	77.5	2.480
	FF37	6.9	35	22.6	67.7	2.406
	FF44	6.5	35	18.9	80.0	2.418
	FF53	6.2	35	17.4	83.7	2.430

Table 3: RI19.0C LMLC Conditions Properties.

		Binder Content (%)	RAP/RAS Content (%)	VMA_{IP}	VFA_{IP}	G_{mm}
Fatigue Testing	FC33	5.2	25/3	14.8	80.4	2.478
	FC44	4.6	25/3	14.9	69.8	2.500
	FC57	4.0	25/3	15.9	56.0	2.522
	FF33	6.7	25/3	17.7	85.3	2.414
	FF44	5.8	25/3	17.4	75.3	2.443
	FF57	5.0	25/3	18.0	61.1	2.473
Rutting Testing	FC37	5.2	25/3	18.8	60.1	2.478
	FC44	4.6	25/3	14.7	71.3	2.500
	FC55	4.0	25/3	14.1	64.5	2.522
	FF37	6.7	25/3	21.6	66.6	2.414
	FF44	5.8	25/3	17.4	75.3	2.443
	FF54	5.0	25/3	15.4	74.0	2.473

Table 4: RB25.0C LMLC Conditions Properties.

	Binder Content (%)	RAP/RAS Content (%)	VMA_{IP}	VFA_{IP}	G_{mm}
FC33	4.6	25/3	13.7	76.0	2.511
FC57	4.0	25/3	17.5	48.5	2.536
FF33	5.7	25/3	16.2	78.0	2.461
FF44	5.0	25/3	19.4	53.7	2.487

CHAPTER 3. PERFORMANCE TESTING OF CONSTRUCTION SAMPLES

RS9.5C mixture

The research team determined the binder content of the plant-mixed, lab-compacted (PMLC) samples using the ignition oven and performed a washed sieve analysis on the recovered aggregates. The binder content and target air void content for the PMLS samples are provided in Table 5 and the gradation of the five PMLC samples as well as the gradation of the JMF are provided in Figure 4. As can be seen from the table, the binder content ranged from 6.0% to 6.4% with the reported binder content in the JMF being 5.8%. For the PMLC samples gradation, the maximum difference in percent passing for a given sieve size was 4.1 between the PMLC samples and JMF. The target air void content for the PMLC sample testing was fixed at 7% as the measured in-place air void content for the PMLC samples was close to 7%.

Table 5: Binder Content of RS9.5C PMLC Samples.

	Binder Content (%)	Target Air Voids (%)
PMLC Sample 1	6.1	7
PMLC Sample 2	6.2	7
PMLC Sample 3	6.4	7
PMLC Sample 4	6.0	7
PMLC Sample 5	6.0	7

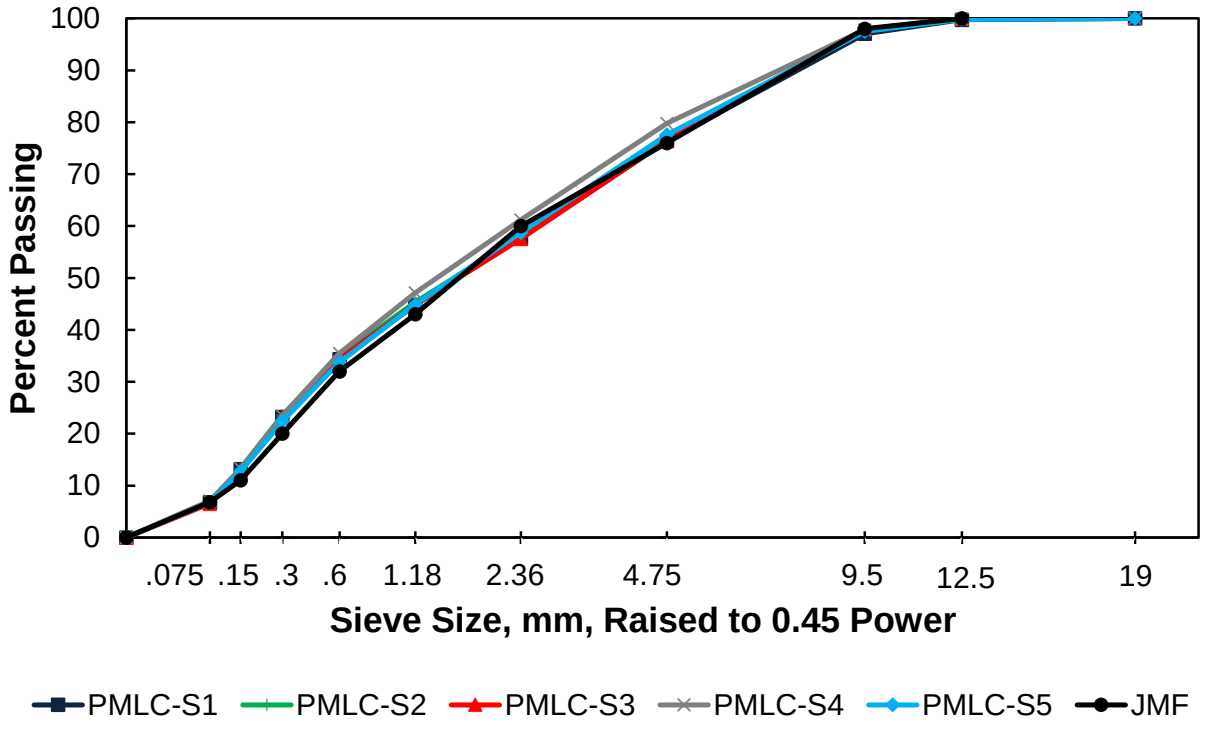

Figure 4: RS9.5C PMLC Samples and JMF Gradation.

BMD+ and BMD testing were performed for all the PMLC samples with a target air void content of 7%. The dynamic modulus, cyclic fatigue, and SSR results are shown in Figure 5 to Figure 8.

In Figure 5, the dynamic modulus results of the PMLC samples were comparable and met the reproducibility limits. Figure 6 and Figure 7 show that the trends in the D^R and S_{app} values are similar. PMLC-S2, with the second-highest binder content, showed the highest D^R and S_{app} values. While PMLC-S3 had the highest binder content, the D^R and S_{app} values were comparable to PMLC-S1 and PMLC-S4. PMLC-S5 exhibited the lowest D^R and S_{app} values. Overall, the D^R values ranged from 0.50 to 0.57, and the S_{app} values ranged from 15.8 to 21.8. Regarding the SSR results, PMLC-S2 and PMLC-S3, which had the highest binder contents, exhibited the highest RSI values. These samples were followed by PMLC-S5, PMLC-S1, and PMLC-S4, respectively. The lowest RSI value observed in PMLC-S4 may be attributed to its slightly coarser gradation compared to the other PMLC samples. The RSI values ranged from 9.0% to 14.6%.

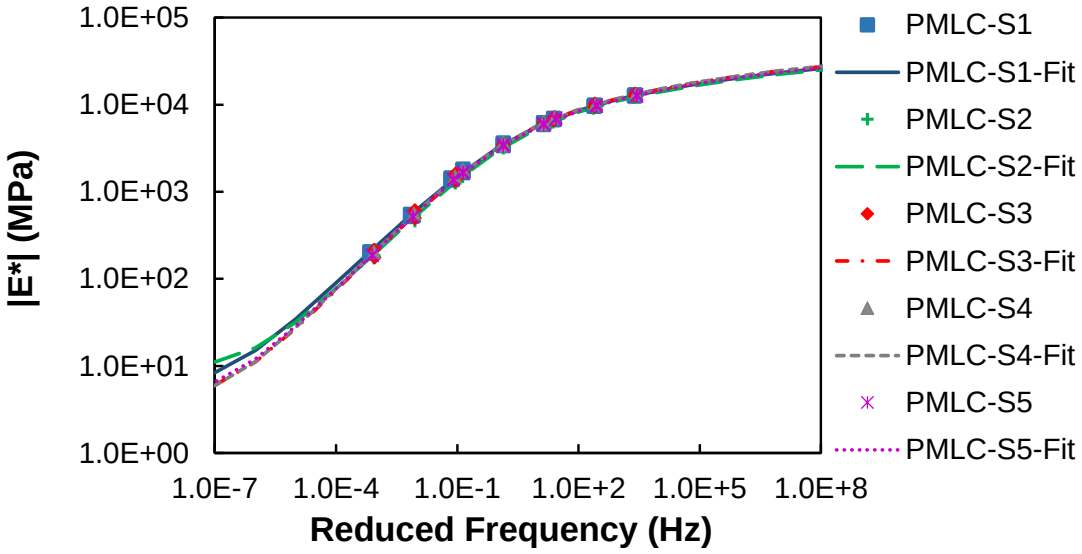

Figure 5: RS9.5C PMLC Samples Dynamic Modulus Mastercurve.

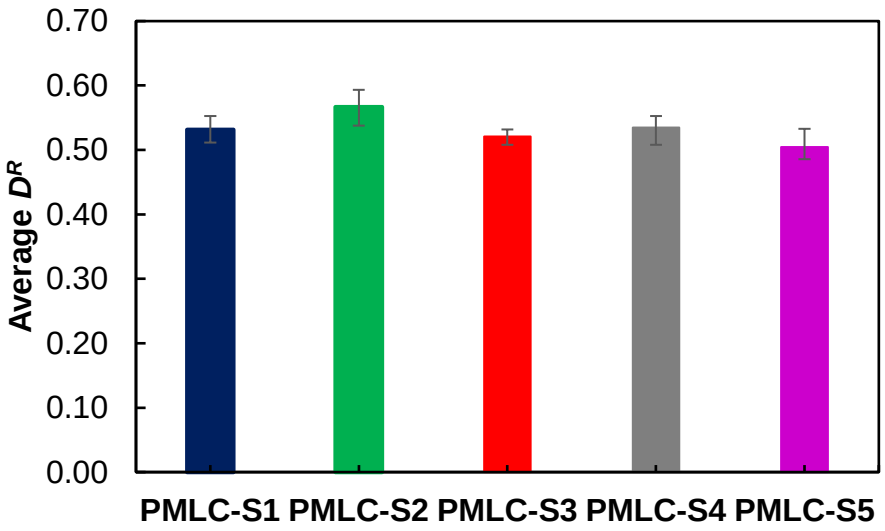

Figure 6: RS9.5C PMLC Samples Average D^R.

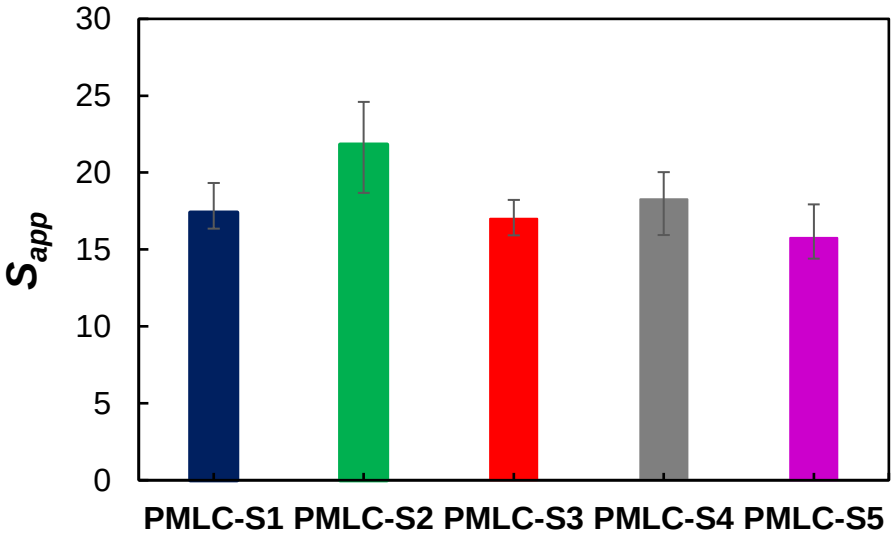

Figure 7: RS9.5C PMLC Samples S_{app} Values.

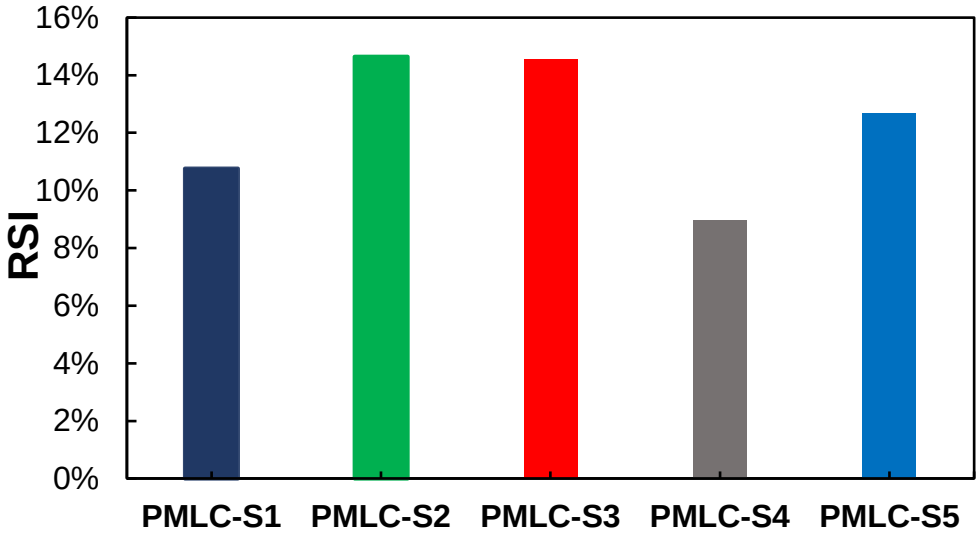

Figure 8: RS9.5C PMLC Samples RSI Values.

The CT-Index results presented in Figure 9 ranged from 62.3 to 40.2. The results show a better agreement with the binder content compared to S_{app} results. HWT rut depth varied from 6.64 mm to 4.98 mm, as can be seen in Figure 10. PMLC-S1 and PMLC-S3 were the only samples that did not exhibit stripping. HWT results did not show a direct correlation with the binder content; however, it should be noted that HWT test is a rutting as well as a stripping test at the same time.

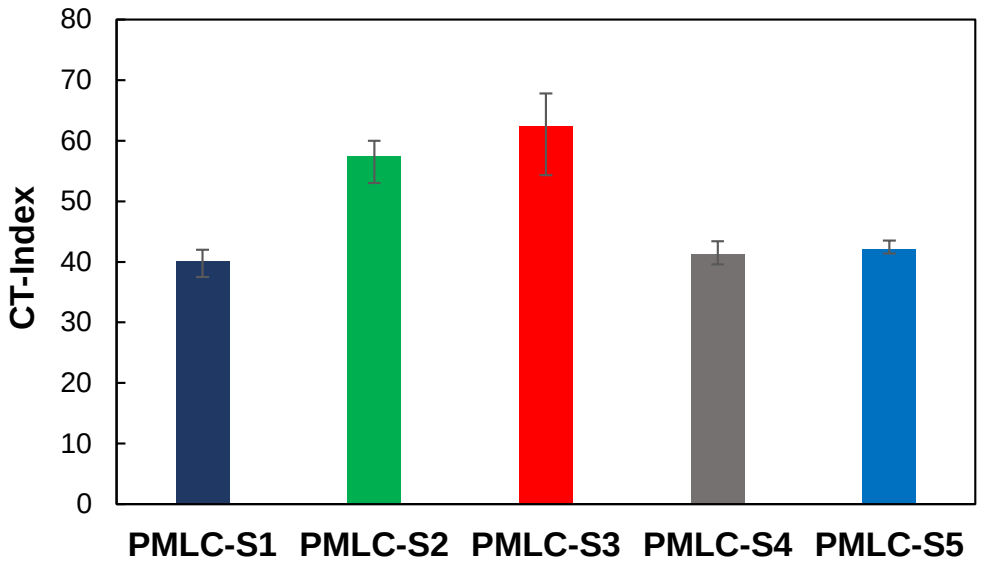

Figure 9: RS9.5C PMLC Samples IDEAL-CT Results.

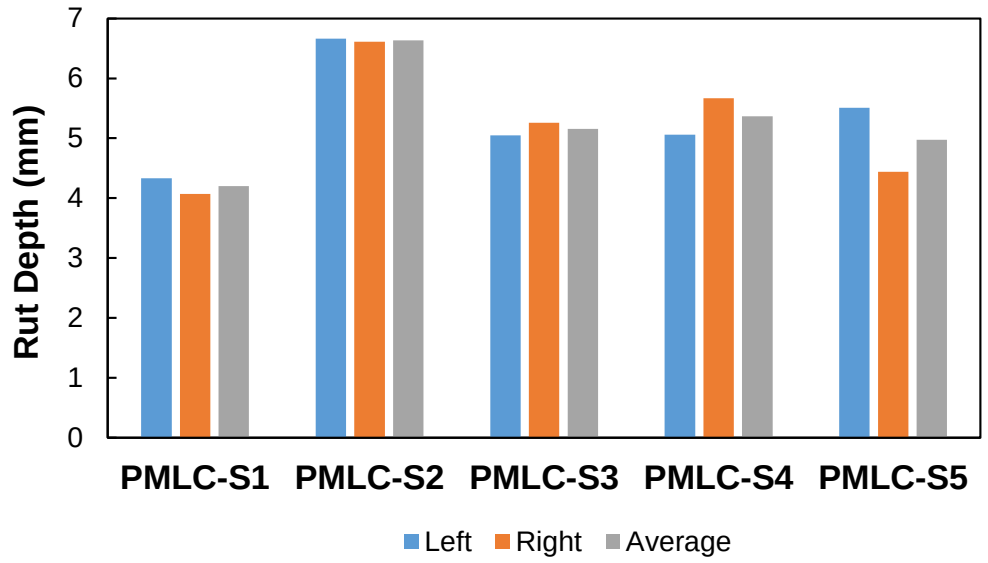

Figure 10: RS9.5C PMLC Samples HWT Results.

RI19.0C mixture

The binder content of the PMLC samples ranged from 3.7% to 5.3% with the JMF's binder content being 4.8%. The maximum difference in percent passing for a given sieve size was 8.0 between the PMLC samples and JMF. The binder content and target air void content are summarized in Table 6, and the gradation of the PMLC samples is presented in Figure 11.

Table 6: Binder Content of RI19.0C PMLC Samples.

	Binder Content (%)	Target Air Voids (%)
PMLC Sample 1	5.3	5.8
PMLC Sample 2	4.0	6.5
PMLC Sample 3	4.0	6.1
PMLC Sample 4	4.0	6.1
PMLC Sample 5	3.7	6.7

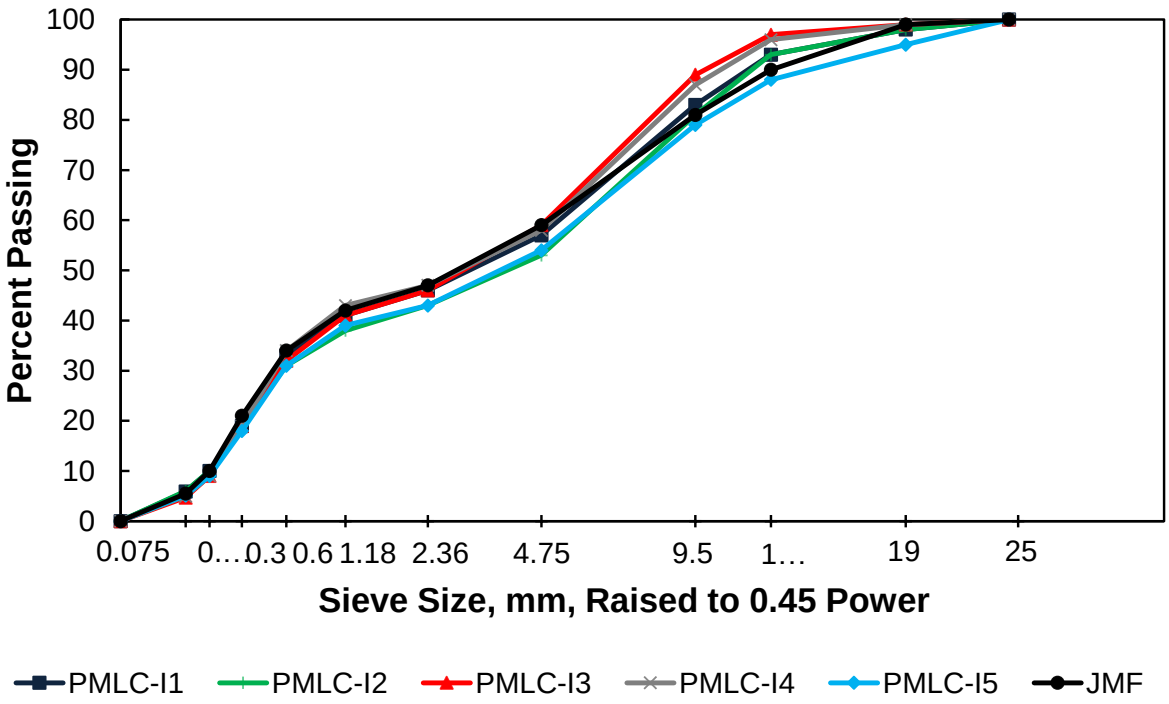

Figure 11: RI19.0C PMLC Samples and JMF Gradation.

BMD+ and BMD testing were performed for all the PMLC samples with a target air void content equal to in-place air void content shown in Table 6, with the exception of IDEAL-CT where the target air void content was 7%. The dynamic modulus, cyclic fatigue, and SSR results are presented in Figure 12 to Figure 15. The dynamic modulus results did not vary significantly between the different PMLC samples and met the reproducibility limits. The D^R values ranged from 0.55 to 0.59, and the S_{app} results ranged from 13.3 to 18.5. PMLC-I3, which had a medium binder and air void contents compared to the other PMLC samples, exhibiting the lowest S_{app} value is believed to be due to material and test variabilities. PMLC-I1, with the highest binder content and lowest target air void content, and PMLC-I5, with the lowest binder content and highest target air void content, showed the highest RSI values, followed by PMLC-I4, PMLC-I2, and PMLC-I3, respectively. Overall, the RSI values ranged from 1.6% to 2.2%.

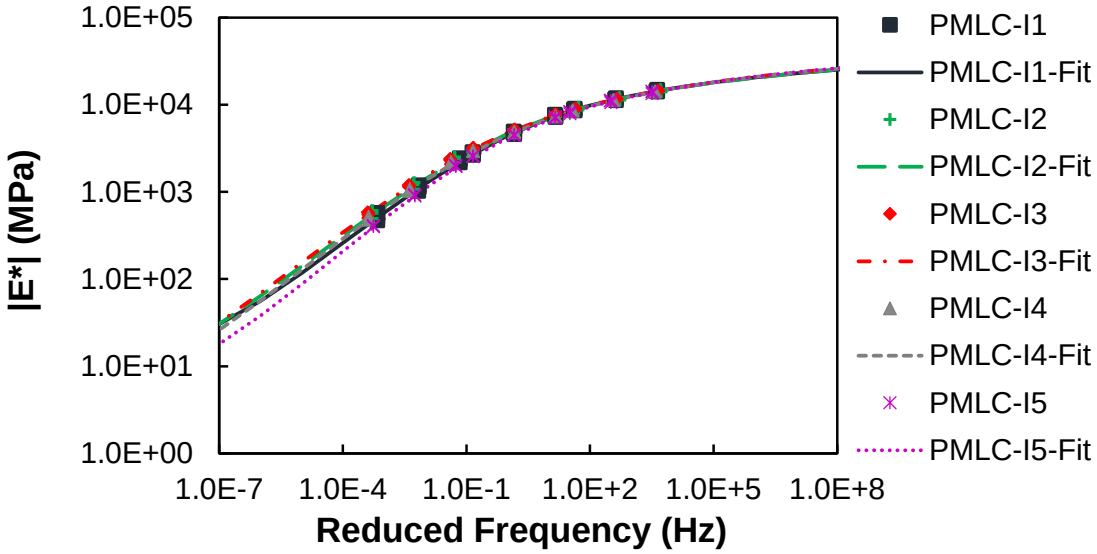

Figure 12: RI19.0C PMLC Samples Dynamic Modulus Mastercurve.

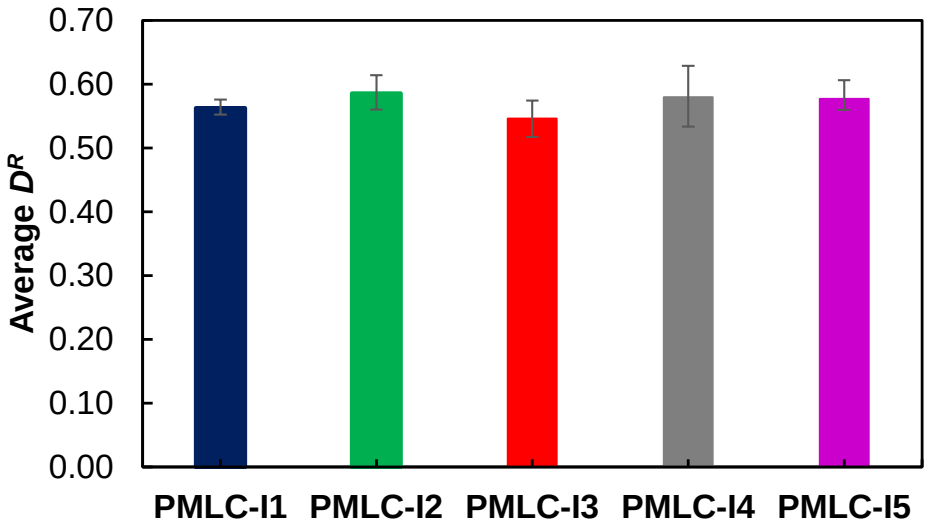

Figure 13: RI19.0C PMLC Samples Average D^R.

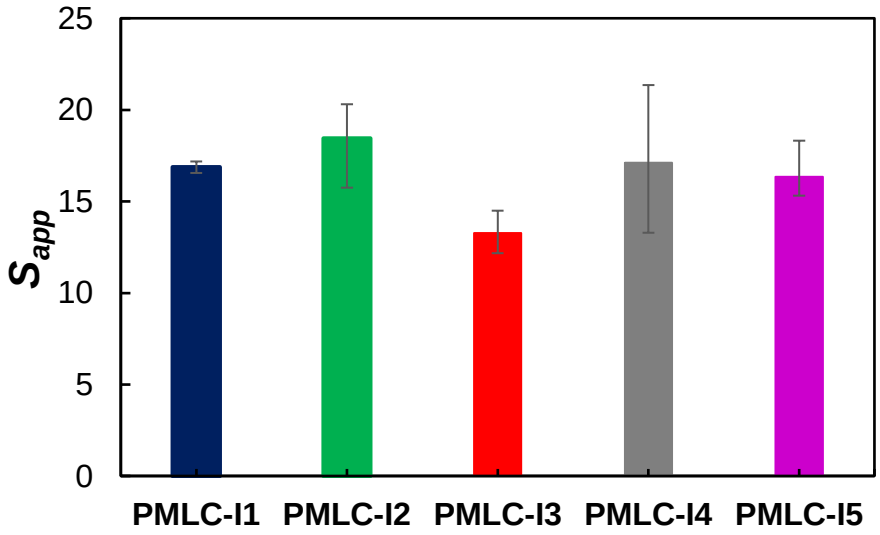

Figure 14: RI19.0C PMLC Samples S_{app} Values.

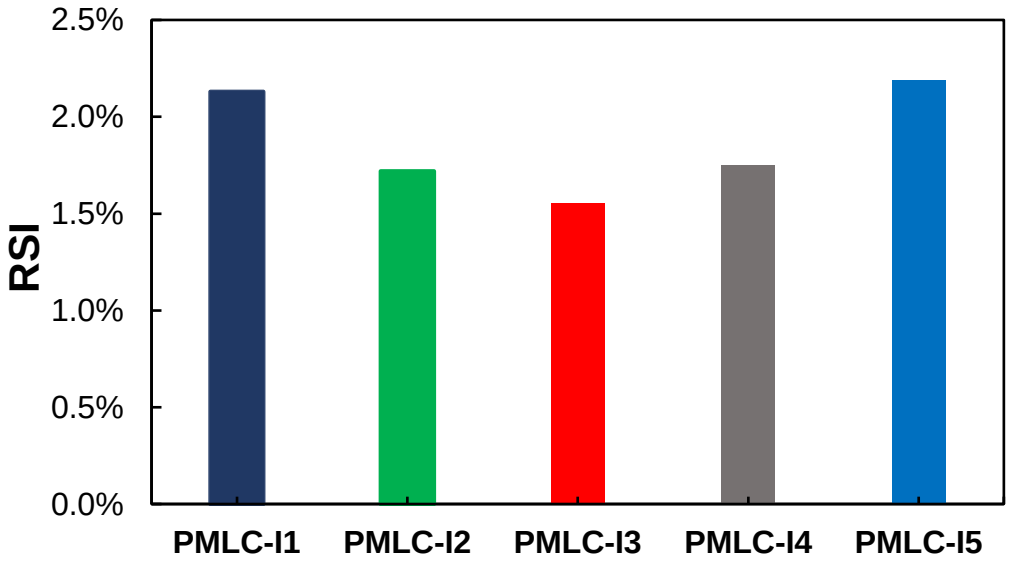

Figure 15: RI19.0C PMLC Samples RSI Values.

The CT-Index results presented in Figure 16 ranged from 15.1 to 24.3. PMLC-I1, with the highest binder content, had the highest CT-Index, while PMLC-I3 showed the lowest CT-Index. HWT rut depth ranged from 1.31 mm to 2.59 mm, as can be seen from Figure 17. PMLC-I5, with the highest target air void content and lowest binder content, had the highest rut depth, while PMLC-I3 showed the lowest rut depth.

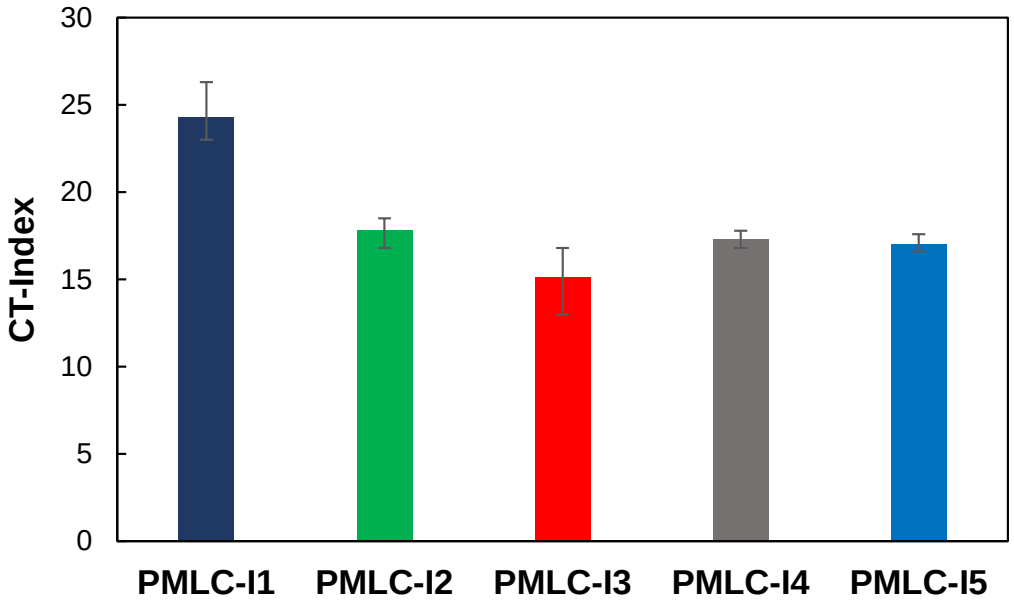

Figure 16: RI19.0C PMLC Samples IDEAL-CT Results.

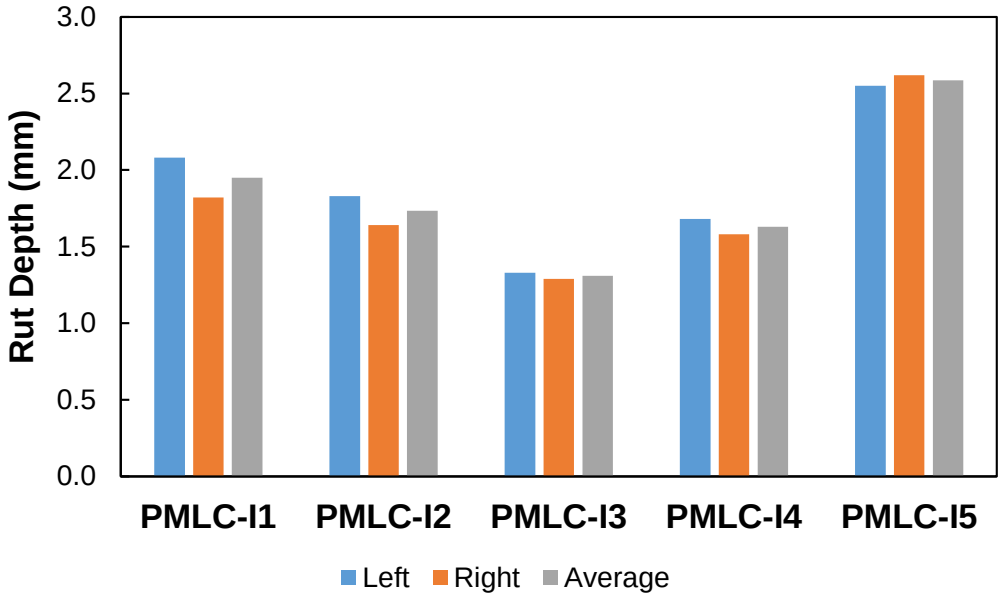

Figure 17: RI19.0C PMLC Samples HWT Results.

RB25.0C mixture

When dealing with mixtures that have a nominal maximum aggregate size (NMAS) of 25 mm, the dynamic modulus test is typically performed on large specimens (100 mm in diameter and 150 mm in height) according to the procedure described in AASHTO T 378. Large specimens with 100 mm in diameter and 130 mm in height are used to perform the cyclic fatigue test for the aforementioned type of mixes according to AASHTO T 400. The main advantage of using large

specimens to perform these tests is that it is easier to obtain the target air void content when compared to using small specimens, which have a diameter of 38 mm and a height of 110 mm. According to the NCHRP IDEA project findings by the NCSU research team, the success rate of 38 mm cores to be within the target air void content is about 50%, which means that two gyratory samples are needed to obtain three 38 mm specimens for testing. Another finding from the NCHRP IDEA project is that the rate of end failure increases as the NMAS increases. These findings were the reasons why the NCSU team did not recommend the 38 mm specimens for the dynamic modulus and cyclic fatigue tests for 25 mm NMAS mixtures. However, using large specimens is accompanied by using a larger amount of mixture and glue. More importantly, the required time for gluing increases from less than an hour to more than 16 hours as specimen size increases from 38 mm diameter to 100 mm diameter. Temperature conditioning time also increases.

Considering all the reasons described above, the research team made the decision to use 38 mm specimens to run the dynamic modulus and cyclic fatigue tests. This decision is based on the fact that switching 38 mm specimen testing to 100 mm specimen testing only for 25 mm NMAS mixtures would cause more difficulties in SHA and contractor laboratories than making more gyratory samples in order to obtain a sufficient number of 38 mm cores within the target air void content and to obtain middle failures. In addition, having two sets of hardware and software programs for two different specimen geometries would increase the potential of errors. The measured air void content for small specimens (with 38 mm diameter) obtained from two gyratories for one of the obtained PMLC samples with a target air void content of 7.5% can be seen in Table 7. Four 38 mm cores out of eight cores had the air void content within ±0.5% of the target air void content, thus confirming the finding from the NCHRP IDEA project. When considering all the tested PMLC samples as well as the samples from four corners, the success rate ranged from 40% to 75%. Thus, the research team recommends preparing four gyratories for the dynamic modulus and cyclic fatigue tests to ensure having enough specimens within the target air void content and in case some end failure occurs during the cyclic fatigue test. Sixteen 38 mm cores can be obtained from four gyratories, and eight cores can meet the air voids requirement assuming a 50% success rate. Three cores can be used for each of the dynamic modulus and cyclic fatigue tests and two remaining cores can be used for additional cyclic fatigue tests in case end failure has resulted from the first three tests.

Table 7: Air Void Content of Small Specimens Obtained from Two Gyratories.

Specimen	Air Void Content (%)
Rep.1-1	8.0
Rep.1-2	7.6
Rep.1-3	7.6
Rep.1-4	6.9
Rep.2-1	6.4
Rep.2-2	8.5
Rep.2-3	6.9
Rep.2-4	7.1

As can be seen from Table 8, the four PMLC samples binder content varied from 3.3% for PMLC-B4 to 4.9% for PMLC-B3, with PMLC-B1 having the closest binder content to the JMF's of 4.4%. PMLC-B1 had the closest gradation to the one provided in the JMF, while PMLC-B4 had the largest deviation in gradation from the JMF's, as can be seen in Figure 18. It can be noticed from the measured binder content and the determined gradation that the four PMLC samples varied significantly, with a difference of 1.5% in binder content and as large as 24% difference in percent passing sieve #4.

It should be noted that the PMLC-B4 sample had poor film thickness and inadequate coating compared to the other PMLC samples, which had been noticed visually even before preparing the test specimens. This problem has also been reflected by the measured binder content. Figure 19 shows some aggregates with poor coating from PMLC-B4. Also, it should be mentioned that PMLC-B4 exhibited greater difficulty in obtaining valid results for the cyclic fatigue test.

The research team decided on the target air void content based on the acquired in-place density data. However, it was not possible to obtain the in-place density data for some sampling dates. The research team selected the target air void content for these dates based on the overall variation in the in-place density data. The target air void content for the four PMLC samples is shown in Table 8. The air void content for PMLC-B1 and PMLC-B3 are from the measured in-place density of those samples, whereas those for PMLC-B2 and PMLC-B4 are from the overall variation of the in-place density data.

Table 8: Binder Content and Target Air Void Content of RB25.0C PMLC Samples.

Sample ID	Binder Content (%)	Target Air Voids (%)
PMLC-B1	4.3	7.7
PMLC-B2	3.9	7.5
PMLC-B3	4.9	7.7
PMLC-B4	3.3	6.5

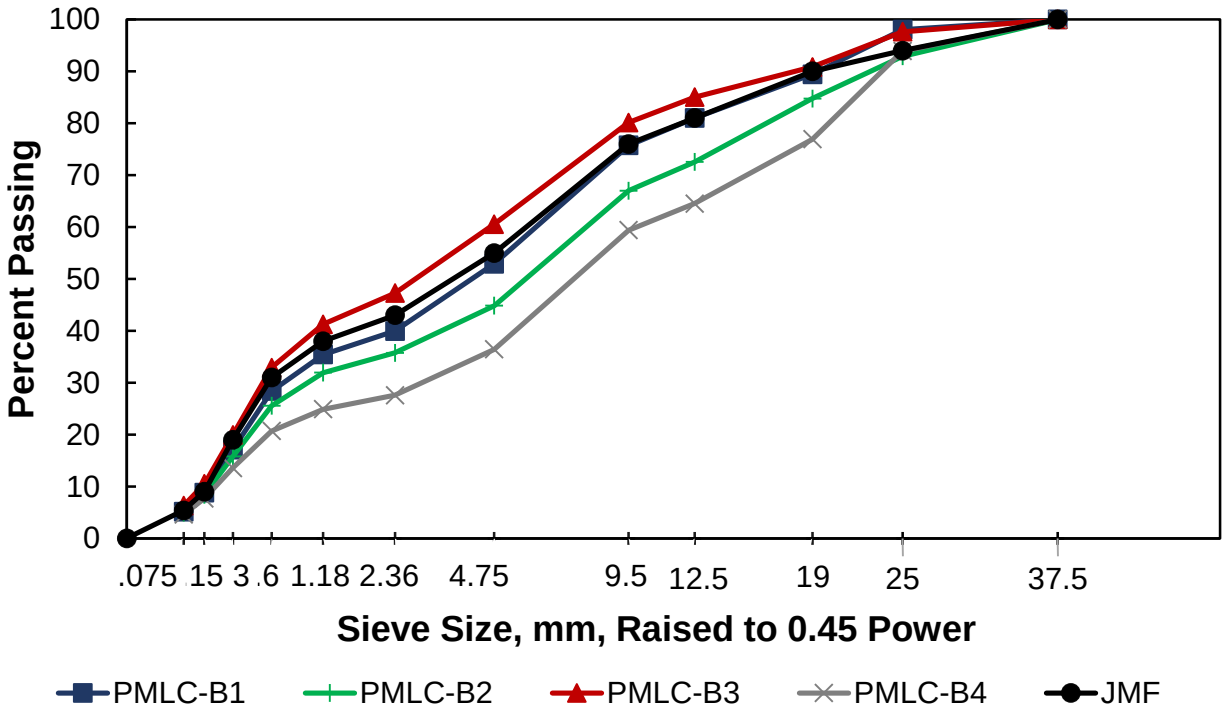

Figure 18: RB25.0C PMLC Samples and JMF Gradation.

Figure 19: RB25.0C PMLC-B4 Poor Coating and Film Thickness.

AMPT tests were conducted on the four PMLC samples for the base mixture and the results are presented in Figure 20 to Figure 23. As shown in Figure 20, the dynamic modulus results were comparable for the four PMLC samples; however, PMLC-B4 showed a higher dynamic modulus than the rest of the PMLC samples, which is supported by the lower binder content and the lower air void content. The highest D^R value from PMLC-B4 can be attributed to the lowest air void content, although it had the lowest binder content. Note that the S_{app} value for PMLC-B4 is the second lowest among the four PMLC samples; however, it exhibited greater variability compared to PMLC-B3, resulting in a statistically insignificant difference in the S_{app} results between the two samples. As discussed in CHAPTER 2, S_{app} balances between the modulus and toughness of mixtures, which may explain why the D^R value from PMLC-B4 is high whereas the S_{app} value is low.

PMLC-S1 had a higher D^R value than PMLC-B2 because PMLC-B1 had a higher binder content than PMLC-B2 with comparable air void content. However, PMLC-B3, which had the highest binder content, did not show a higher D^R value than PMLC-B1. The S_{app} for PMLC-B1 was the highest, although it had the second-highest binder content because it had the second-highest D^R value and the lowest dynamic modulus value. PMLC-B2 exhibiting a higher S_{app} value than PMLC-B3, which had the lowest S_{app} value despite having the highest binder content with comparable air void content to PMLC-B1 and B2, can be explained by the fact that PMLC-B3 had the highest C_{12} value. It should be noted that the C_{12} parameter is present in the S_{app} equation in the denominator of the power term, meaning that small differences in C_{12} values can yield large changes in the whole fatigue term $((\frac{D^R}{C_{11}})^{\frac{1}{C_{12}}})$ value. The sensitivity of S_{app} due to the change in C_{12} may need to be investigated further based on the PMLC-B2 and PMLC-B3 data. As for PMLC-B4, the competing effect between the highest D^R value in the numerator and the highest dynamic modulus value in the denominator yielded the second-lowest S_{app} value. Overall, the D^R values ranged from 0.52 to 0.65, and S_{app} values ranged from 8.5 to 17.9.

The above discussion presents the complexities involved in predicting the effects of binder content and air void content on cracking performance from the QA data when these two factors change simultaneously. Measuring performance directly from the QA samples would eliminate this problem.

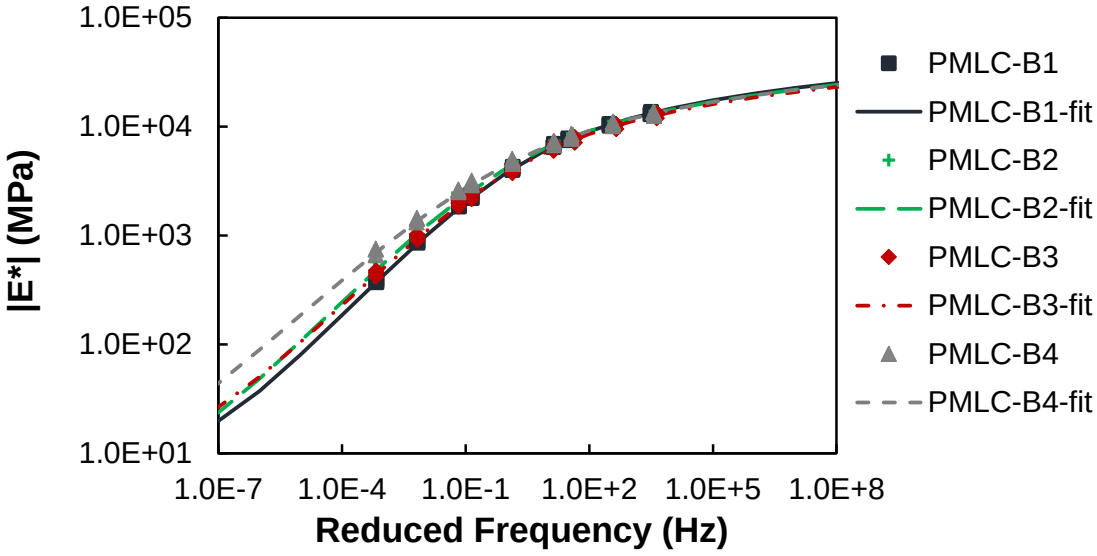

Figure 20: RB25.0C PMLC Samples Dynamic Modulus Mastercurve.

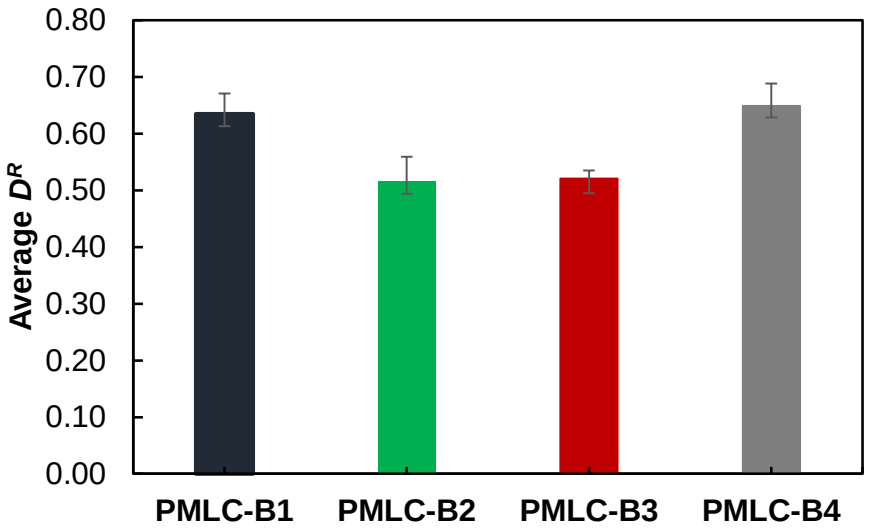

Figure 21: RB25.0C PMLC Samples Average D^R.

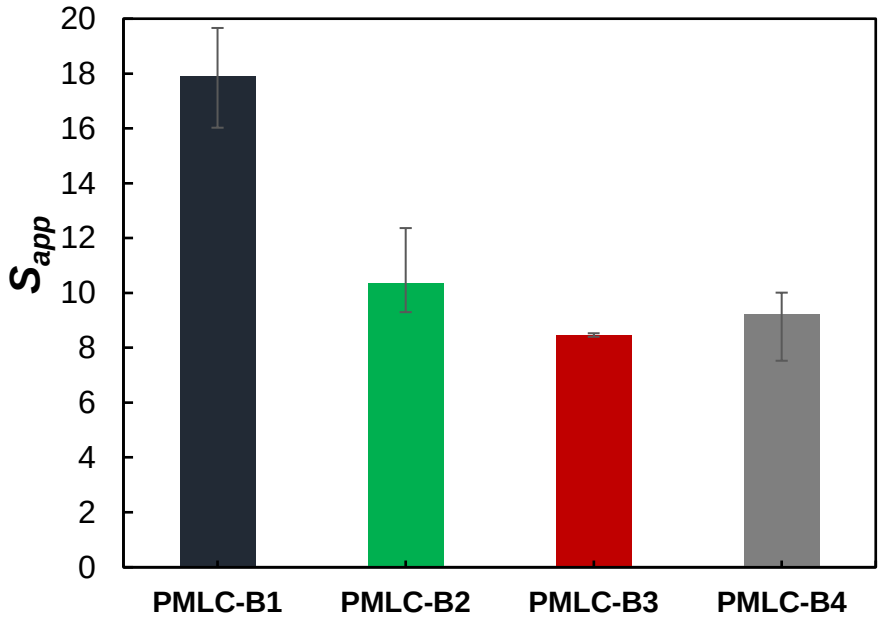

Figure 22: RB25.0C PMLC Samples S_{app} Values.

A stress sweep rutting (SSR) test was performed on the four PMLC samples and the results are presented in Figure 23. PMLC-B4, with the lowest binder content and air void content, exhibited the lowest RSI value. The remaining three PMLC samples had a comparable air void content, but PMLC-B3 had the highest binder content followed by PMLC-B1 and PMLC-B2, respectively. The ranking of the RSI values followed the same order as the ranking of the binder content, where PMLC-B3 had the highest RSI value, followed by PMLC-B1 and PMLC-B2, respectively.

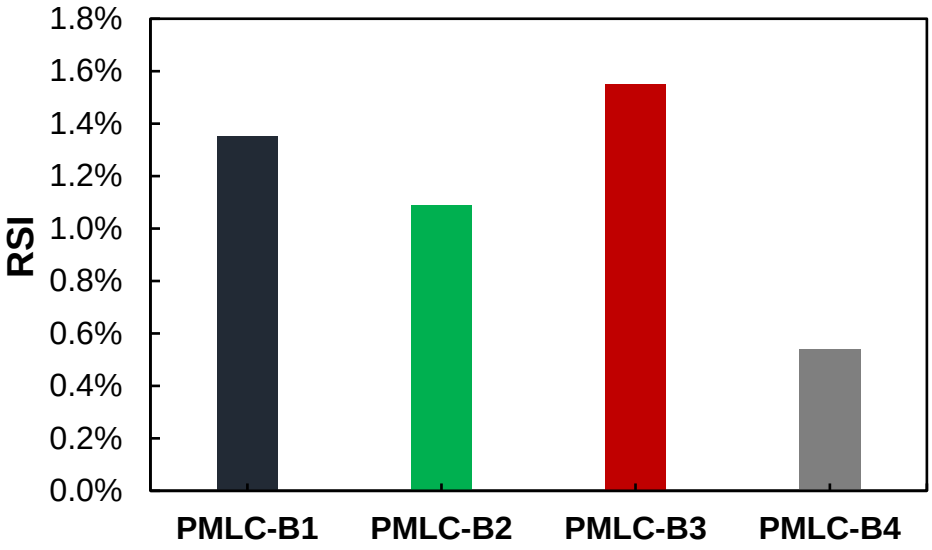

Figure 23: RB25.0C PMLC Samples RSI Values.

Short-term aging duration investigation

Before starting to prepare LMLC samples, the research team wanted to ensure that the prepared specimens would reflect the condition of the PMLC samples accurately. Consequently, the research team investigated the appropriate duration of short-term aging (STA). To perform this investigation, first, the researchers prepared a replicate of RS9.5C's JMF and performed AMPT testing on it. The JMF replicate was denoted as "LMLC-4STA" and was compared to PMLC-S1 and PMLC-S5 since they had comparable gradations and binder contents to the JMF mixture.

The dynamic modulus, cyclic fatigue, and SSR results for PMLC-S1, PMLC-S5, and LMLC-4STA are presented in Figure 24 to Figure 27. The presented dynamic modulus mastercurves show that LMLC-4STA exhibited a slightly higher modulus compared to the two PMLC samples. While the D^R and S_{app} values shown were comparable, the RSI value of the LMLC-4STA was significantly lower than the two PMLC samples. The LMLC-4STA RSI value was 32.5% and 42.8% lower than PMLC-S1 and PMLC-S5, respectively. The results suggested that the LMLC samples were undergoing more aging than the PMLC samples.

The research team prepared another JMF replicate with a short-term aging duration of two hours, referred to as LMLC-2STA. The test results for the dynamic modulus, cyclic fatigue, and SSR tests are presented in Figure 28 to Figure 31. The dynamic modulus mastercurve for LMLC-2STA showed better alignment with the results from the other PMLC samples, compared to LMLC-4STA, as it completely overlapped the PMLC-S1 mastercurve.

The D^R values were found to be similar across both short-term aging durations and the PMLC samples. However, the S_{app} value for LMLC-2STA unexpectedly showed a slight decrease compared to LMLC-4STA. In contrast, the RSI value for LMLC-2STA was higher than that of LMLC-4STA and demonstrated better agreement with the other PMLC samples, being only 4% and 18.6% lower than PMLC-S1 and PMLC-S5, respectively.

Given the drop in the dynamic modulus mastercurve and the increase in the RSI value for LMLC-2STA compared to LMLC-4STA, the research team concluded that the unexpected drop in the S_{app} value could be attributed to material and testing variabilities. Based on the improved agreement in the dynamic modulus and SSR test results between LMLC-2STA and the PMLC samples, the team decided to proceed with a two-hour short-term aging duration for LMLC testing in this project. The major reason for this decision is that one of the primary objectives of this study is to predict the performance of field QA samples using the LMLC mixture data.

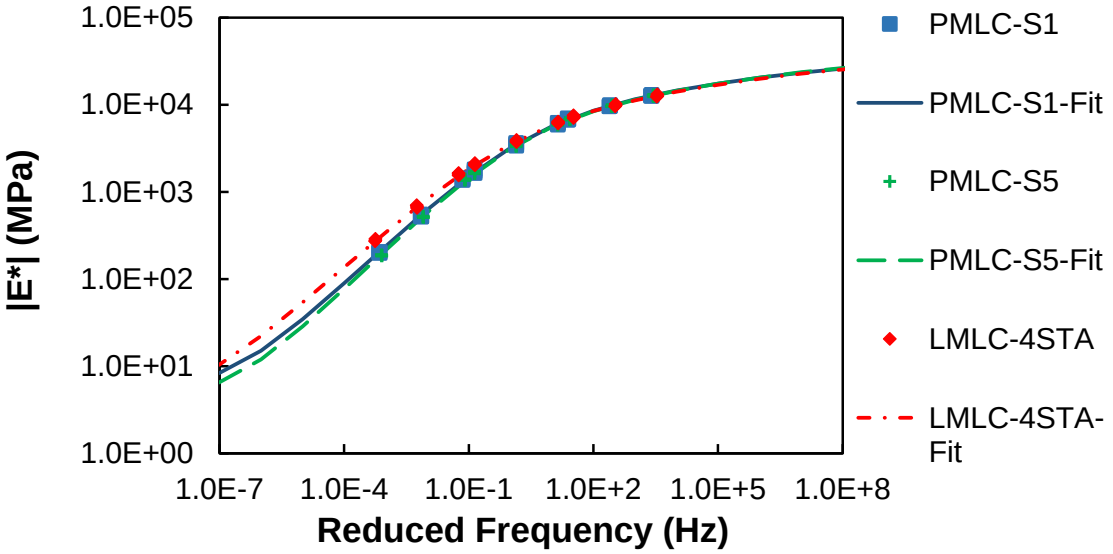

Figure 24: PMLC-S1, PMLC-S5, and LMLC-4STA Dynamic Modulus Mastercurve.

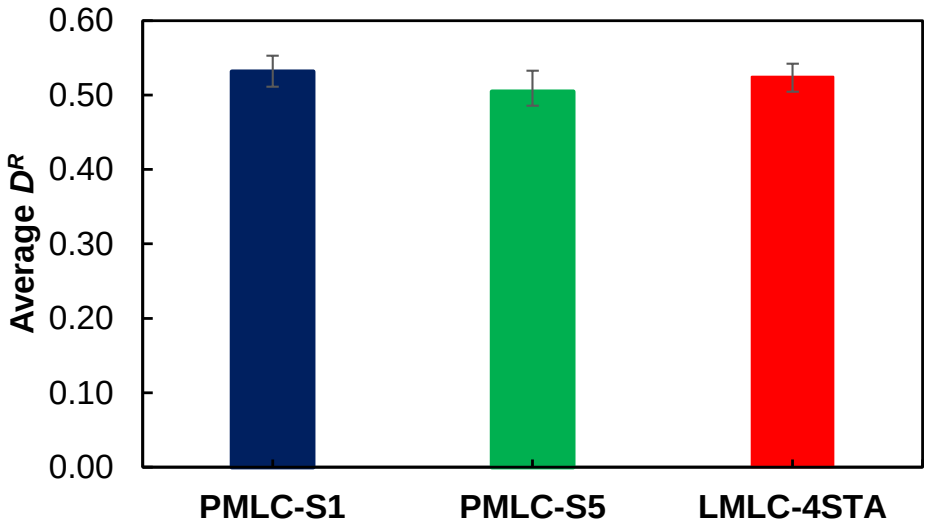

Figure 25: PMLC-S1, PMLC-S5, and LMLC-4STA D^R Values.

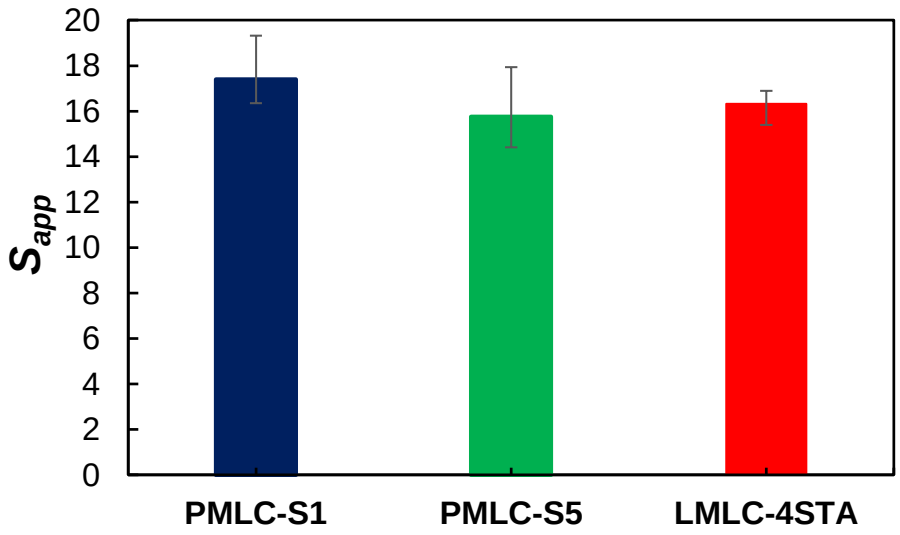

Figure 26: PMLC-S1, PMLC-S5, and LMLC-4STA S_{app} Values.

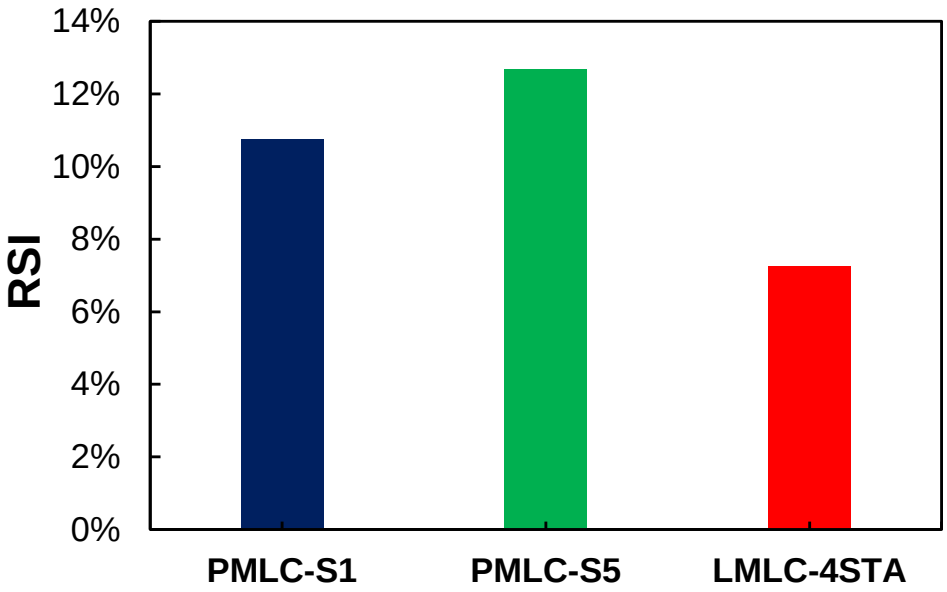

Figure 27: PMLC-S1, PMLC-S5, and LMLC-4STA RSI Values.

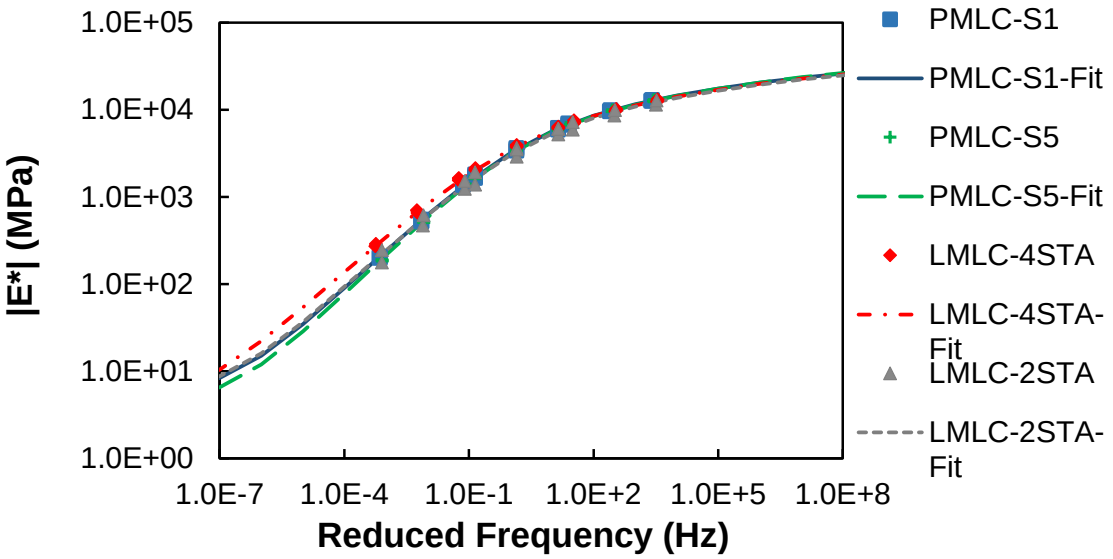

Figure 28: PMLC-S1, PMLC-S5, LMLC-4STA, and LMLC-2STA Dynamic Modulus Mastercurve.

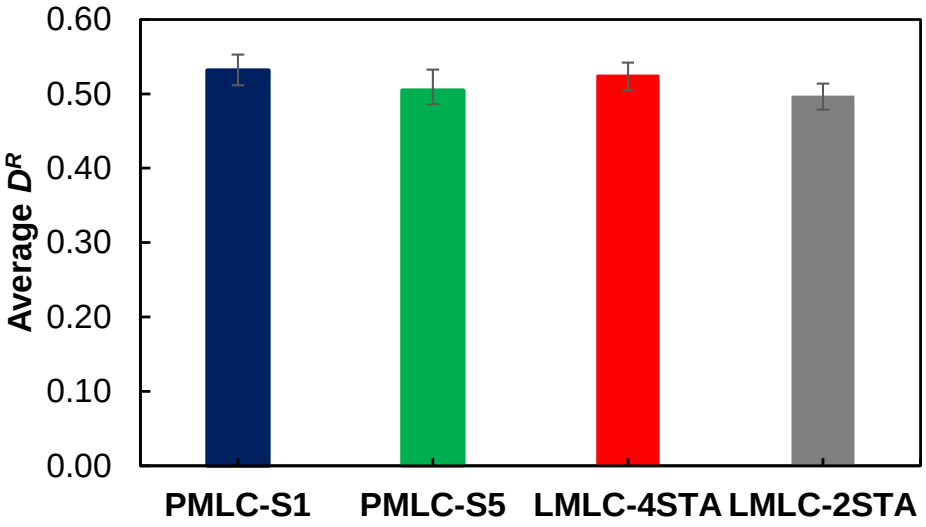

Figure 29: PMLC-S1, PMLC-S5, LMLC-4STA, and LMLC-2STA D^R Values.

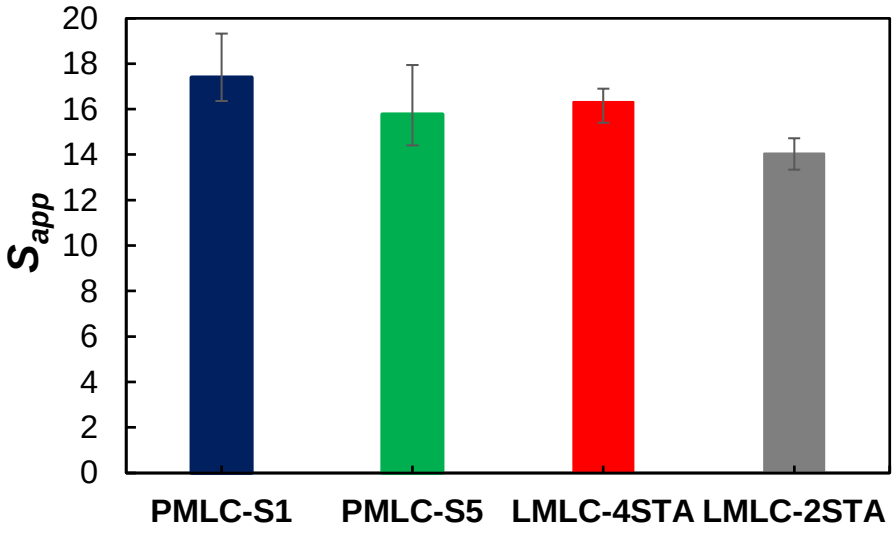

Figure 30: PMLC-S1, PMLC-S5, LMLC-4STA, and LMLC-2STA S_{app} Values.

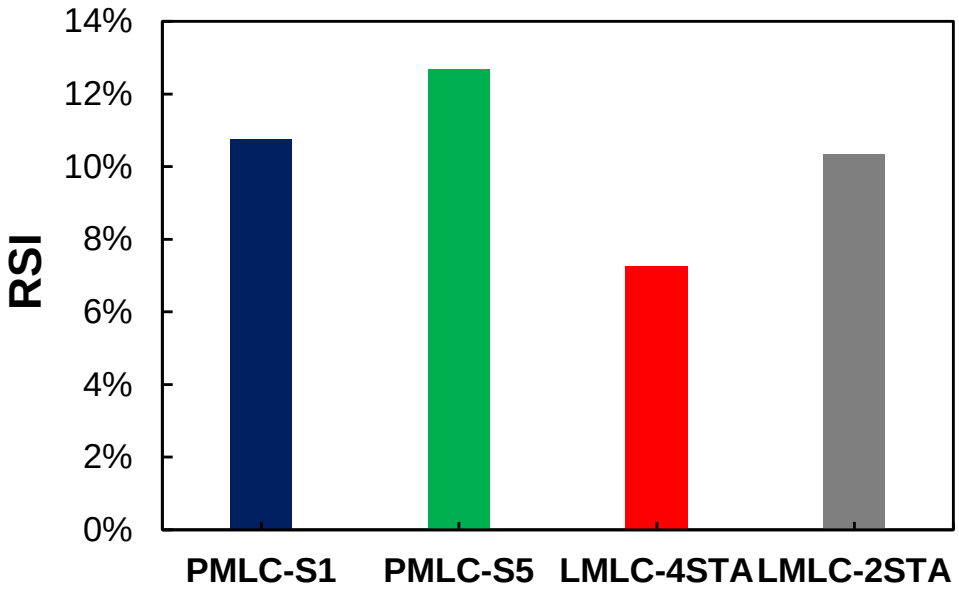

Figure 31: PMLC-S1, PMLC-S5, LMLC-4STA, and LMLC-2STA RSI Values.

CHAPTER 4. DEVELOPMENT OF INDEX-VOLUMETRICS RELATIONSHIPS

Development of the two gradations and volumetric domains

The first step of preparing the four corners and developing the IVR function is to prepare two trial aggregate blend gradations: one that closely follows the maximum density line and another that aligns with the upper limit for fine-graded mixtures or the lower limit for coarse-graded mixtures as specified by the SHA. The two gradations should be wide enough so that the VMA difference between the mix design specimens prepared at P_{bi} is equal to or higher than 1.0%. The required material properties to perform this step are summarized in Table 9. Once the two gradations have been finalized, mix design specimens with asphalt contents of P_{bi} ± X% for each of the chosen aggregate blend gradations should be prepared. The X% of binder content should be selected so that the asphalt content and air void content relationship cover the range of target air void content at N_{des} for the test specimens without a need for extrapolation. An example of an obtained AC-AV relationship is presented in Figure 32. The AC-AV relationship is then used to determine the optimum binder content at a given air void content at N_{des} to the nearest 0.1%.

Table 9: Material Properties Required to Develop the Two Gradations.

Virgin Aggregates	RAP/RAS	Asphalt Binder	Mixture
Gradation from wet sieve analysis (T 27), Bulk specific gravity (G_{sb}), %Absorption or apparent specific gravity of fine and coarse aggregate (G_{sa}) (T 84, T 85, and T 100).	Aggregate gradation (T 27), Binder content (T 308 or T 319), Binder specific gravity (T 228) Maximum specific gravity (T 209)	Specific gravity (T 228)	N_{des}

Note: All standards referenced in this table are derived from the American Association of State Highway and Transportation Officials (AASHTO) standards.

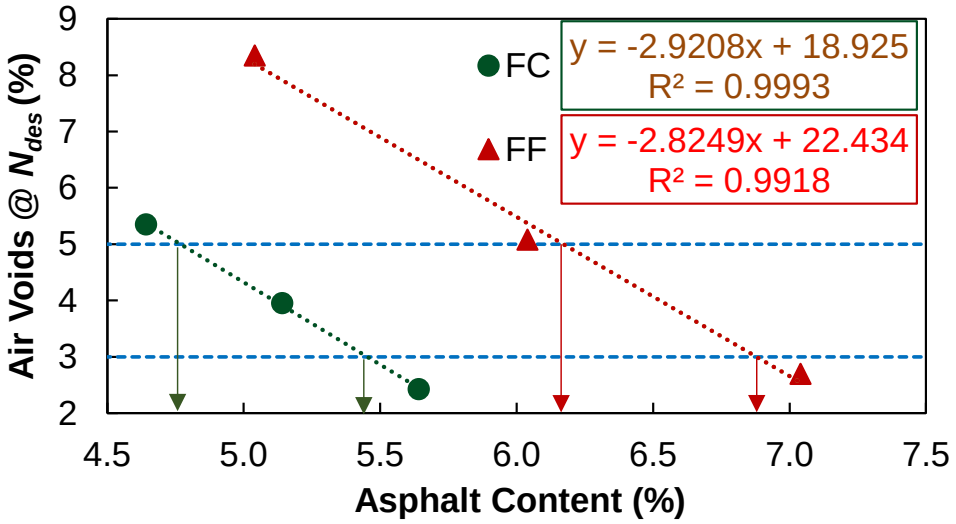

Figure 32: Example of AC-AV Relationship.

The procedure described above was followed to develop the two gradations for the three mixtures. The finalized stockpile proportions and developed gradations are presented in Table 10 to Table 12 and Figure 33 to Figure 35.

Table 10: RS9.5C Stockpiles Proportions of the Two Selected Gradations and JMF Gradations.

Gradation	Stockpile Proportion (%)		
	#78M	WS	RAP
FF	12	53	35
JMF	29	41	30
FC	35	30	35

Table 11: RI19.0C Stockpiles Proportions of the Two Selected Gradations and JMF Gradations.

Gradation	Stockpile Proportion (%)				
	#78M	#67	Sand	RAP	RAS
FF	23	20	29	25	3
JMF	25	22	25	25	3
FC	27	24	21	25	3

Table 12: RB25.0C Stockpiles Proportions of the Two Selected Gradations and JMF Gradations.

Gradation	Stockpile Proportion (%)				
	#467M	#78M	Sand	RAP	RAS
FF	10	34	28	25	3
JMF	22	29	21	25	3
FC	32	25	15	25	3

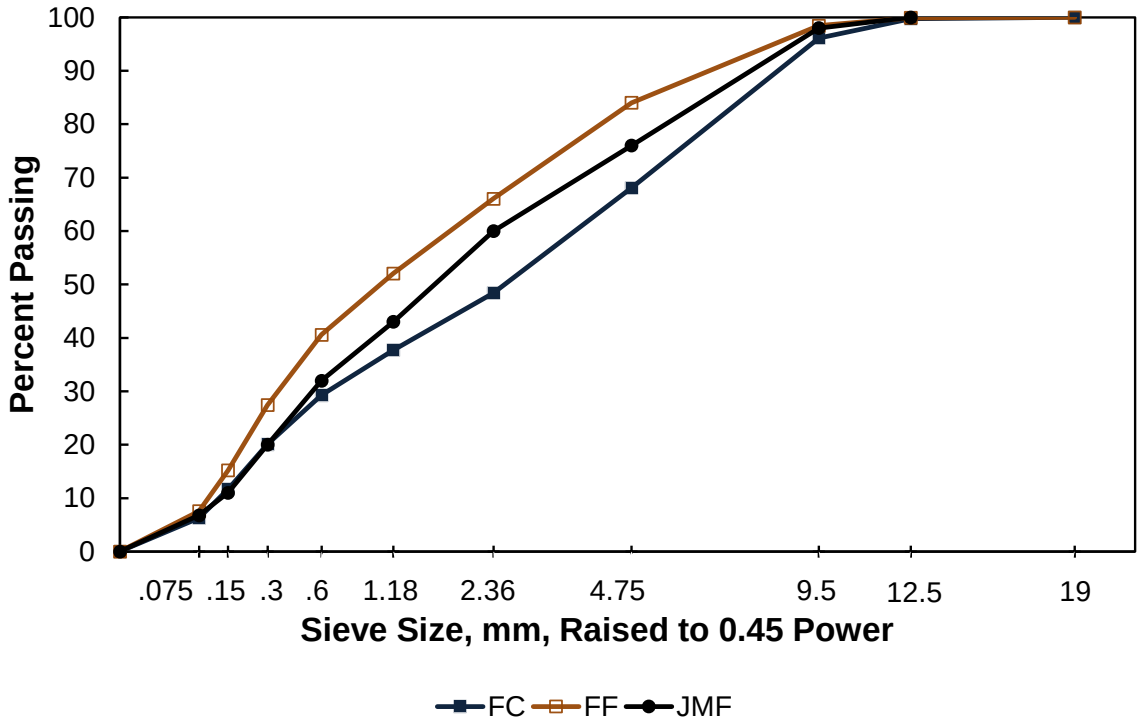

Figure 33: RS9.5C Two Gradations.

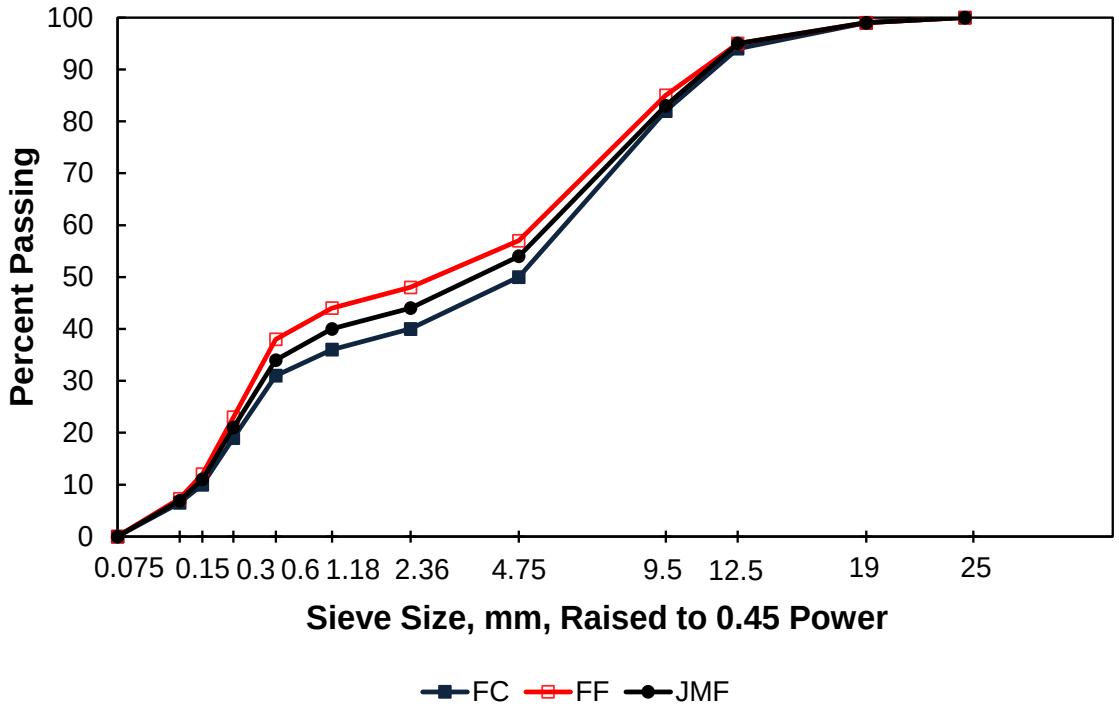

Figure 34: RI19.0C Two Gradations.

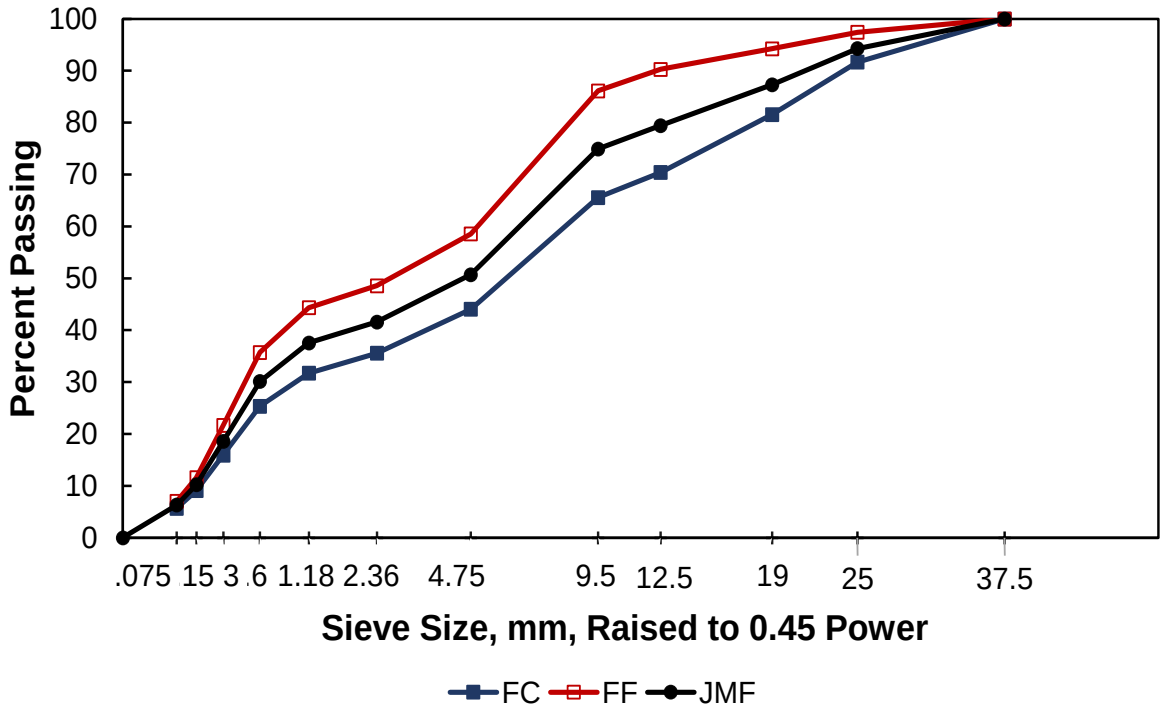

Figure 35: RB25.0C Two Gradations.

After the two gradations were finalized, two different volumetric conditions for each gradation are selected, which form the volumetric conditions for the selected four corners. It is important that the selected four corners capture the effects of changes in mixture factors on performance clearly. This point explains why the four corners selected for different mixtures are different. It also explains why the selected four corners for the fatigue side can be different from the ones selected for the rutting side. The volumetric domain for the fatigue and rutting sides are presented in Figure 36 to Figure 40. As can be seen from Figure 40, the RB25.0 PMLC samples are slightly out of the four corners volumetric range. To have most of the PMLC samples within that range, the researchers needed to replace the 57 corners with 59 corners. However, using the 59 corners would result in increased difficulty in compacting stable specimens as well as an increase in end failures during cyclic fatigue tests. For these reasons, 57 corners were used. It should be noted that the IVR function of RSI for the base mixture was not calibrated considering that rutting would not be of concern when dealing with a base mixture, especially with the pavement design used in this study that has 76.2 mm (3 inches) and 101.6 mm (4 inches) of surface and intermediate mixtures on top of the base layer.

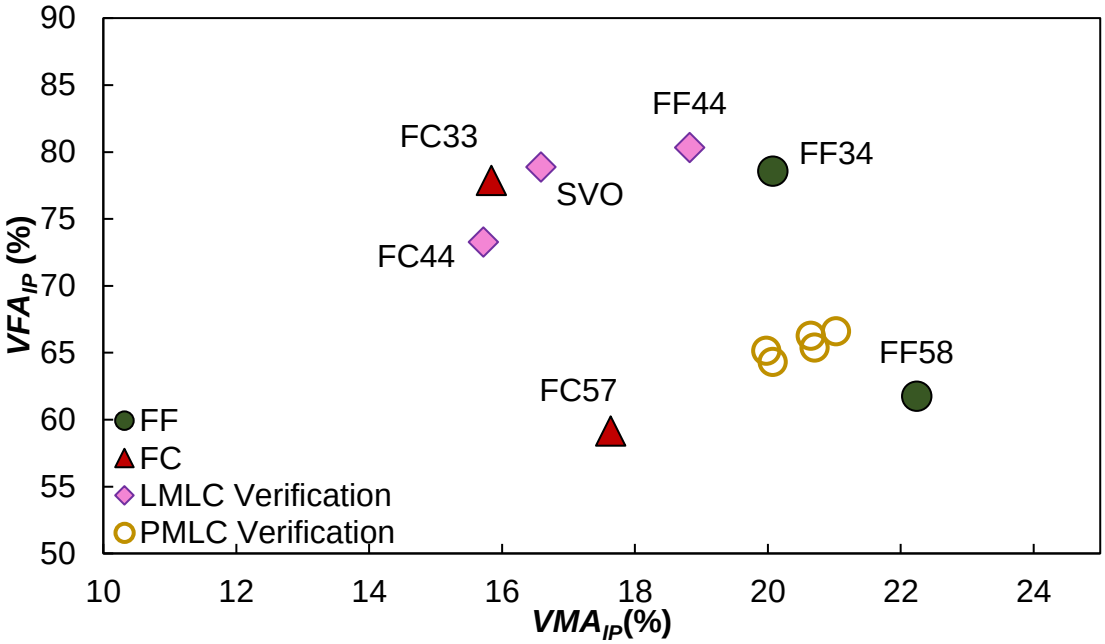

Figure 36: RS9.5C Fatigue Volumetric Domain.

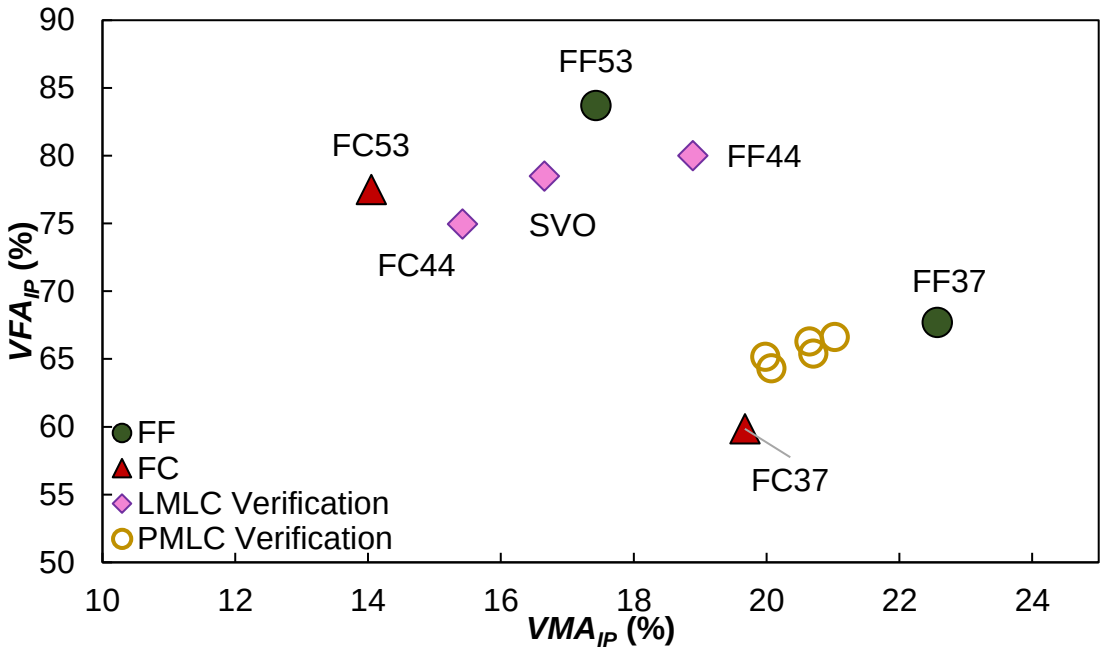

Figure 37: RS9.5C Rutting Volumetric Domain.

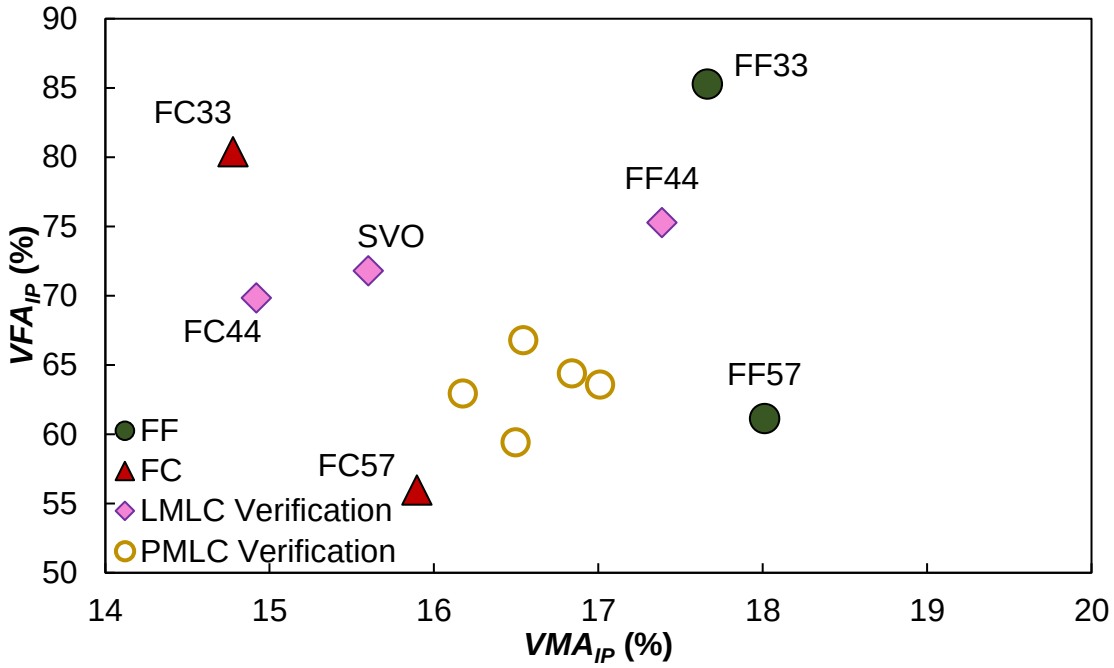

Figure 38: RI19.0C Fatigue Volumetric Domain.

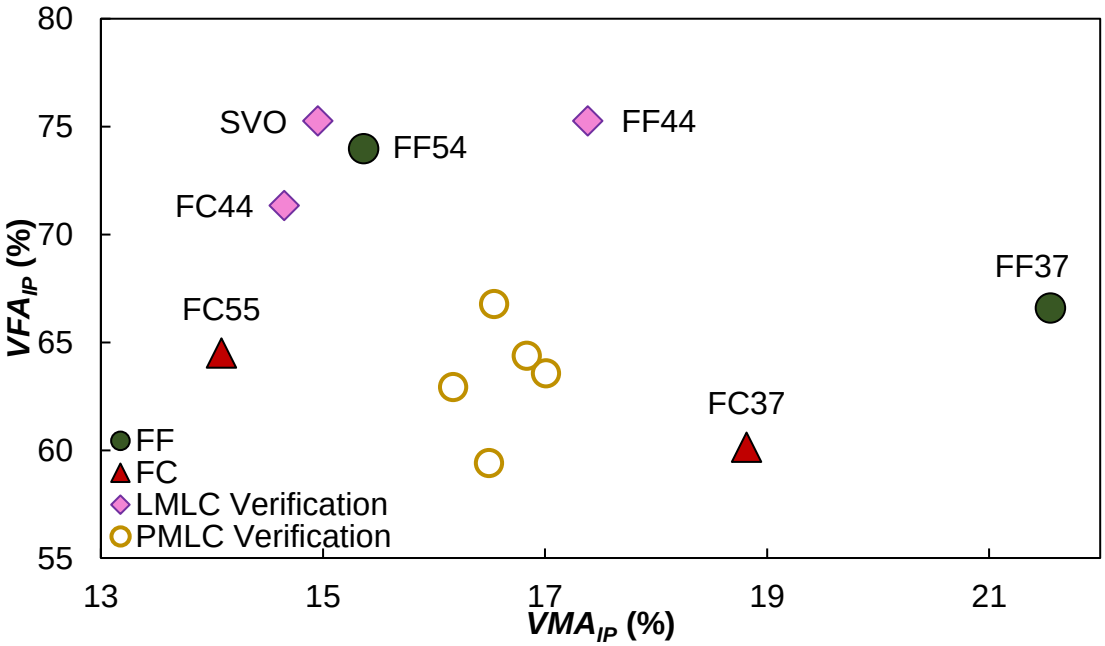

Figure 39: RI19.0C Rutting Volumetric Domain.

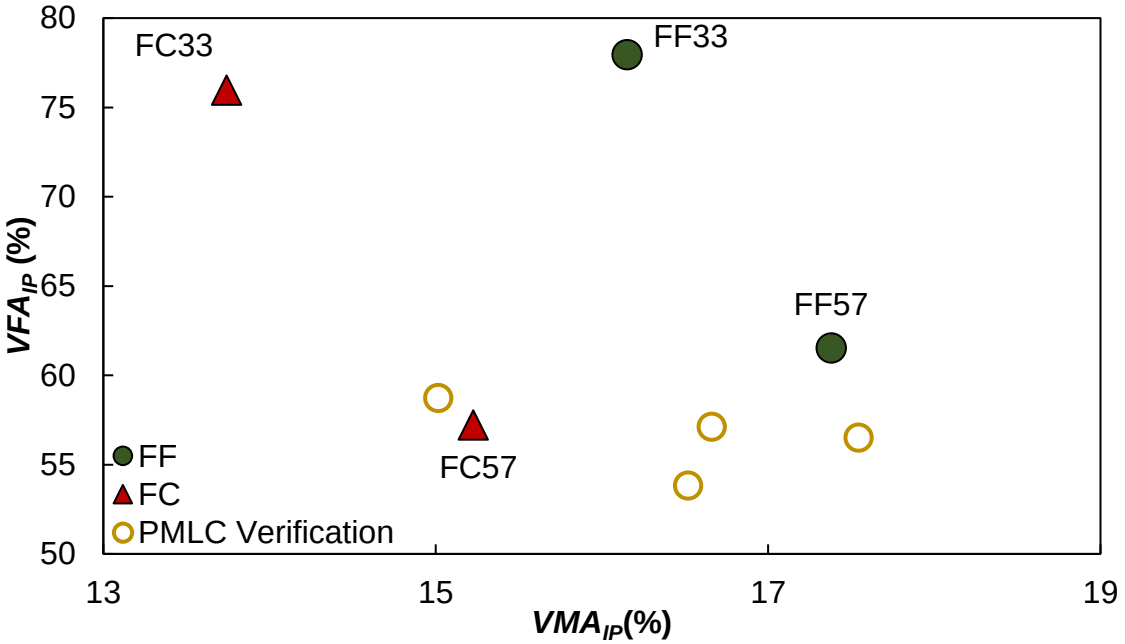

Figure 40: RB25.0C Fatigue Volumetric Domain.

Four corners and two points performance testing

RS9.5C mixture

Performance testing was performed on the volumetric conditions illustrated in the previous figures for the surface mixture. This testing included dynamic modulus, cyclic fatigue, and stress sweep rutting as well as Hamburg Wheel Tracker. For IDEAL-CT, the testing needed to be performed while maintaining a constant air void content. So, it was performed on the two points, namely, 44 conditions of the two gradations.

As can be seen from Figure 41, the coarser gradation exhibited a higher dynamic modulus compared to the fine gradation. This observation aligns with the fact that the coarser gradation had a lower binder content compared to the finer gradation when comparing the same volumetric condition. Within the same gradation, it can be seen that the corner 33 had a higher dynamic modulus followed by the 44 condition. The 57 corner showed the lowest dynamic modulus, in agreement with having the highest air void content and lowest binder content compared to the other points within the same gradation.

In general, it is expected that the cracking performance of 33 corners is better than that of 57 corners because of higher binder content and lower air void content of the 33 corners. However, the average D^R values shown in Figure 42 did not follow this expectation. The D^R values from 57 and 58 corners for FC and FF gradations are not that different the D^R values from 33 and 34 corners, respectively. The S_{app} values for the four corners and two points are presented in Figure 43. The trend in the S_{app} values aligns with the expected behavior, that is, S_{app} decreases as the volumetric condition moves from 33 to 44 to 57 conditions in the FC gradation and from 34 to 44 to 58 conditions in the FF gradation. The different trends in D^R and S_{app} demonstrate the

strength of the S_{app} index in representing the cracking performance of asphalt mixtures at different volumetric conditions.

Figure 42 and Figure 43 show that for the same volumetric condition, the fine gradation showed higher D^R and S_{app} values compared to the coarser gradation, indicating that fine gradations might be a better gradation than coarse gradations in terms of cracking performance. However, this observation needs to be balanced with the trend in RSI values that are shown in Figure 44, which indicate that coarse gradations are better in the rutting performance than fine gradations.

Overall, the D^R values ranged from 0.43 to 0.48 for the coarse gradation and from 0.52 to 0.56 for the fine gradation. The S_{app} values varied from 10.7 to 11.7 to 13.1 for the coarse gradation and from 19.3 to 25.9 for the fine gradation. The RSI values presented in Figure 44 show a superior rutting resistance of the coarse gradation compared to the fine gradation. Also, the RSI values were higher for the 37 corners followed by the 44 and 53 corners, respectively, which aligns with the expected trend due to the high binder content and high air void content of the 37 corners compared to the 53 corners. Note that the 37 corners' RSI values are significantly higher than those from the other corners. The 37 condition may not be a common volumetric condition in the field; that is, the compaction is easier with the high binder content, which results in low air void content. The same argument can be applied to the 53 condition, that is, it is difficult to compact the low binder content mixture down to 3% air voids. These observations may suggest the change in the four corners for the rutting evaluation to the ones that are more commonly observed in the field.

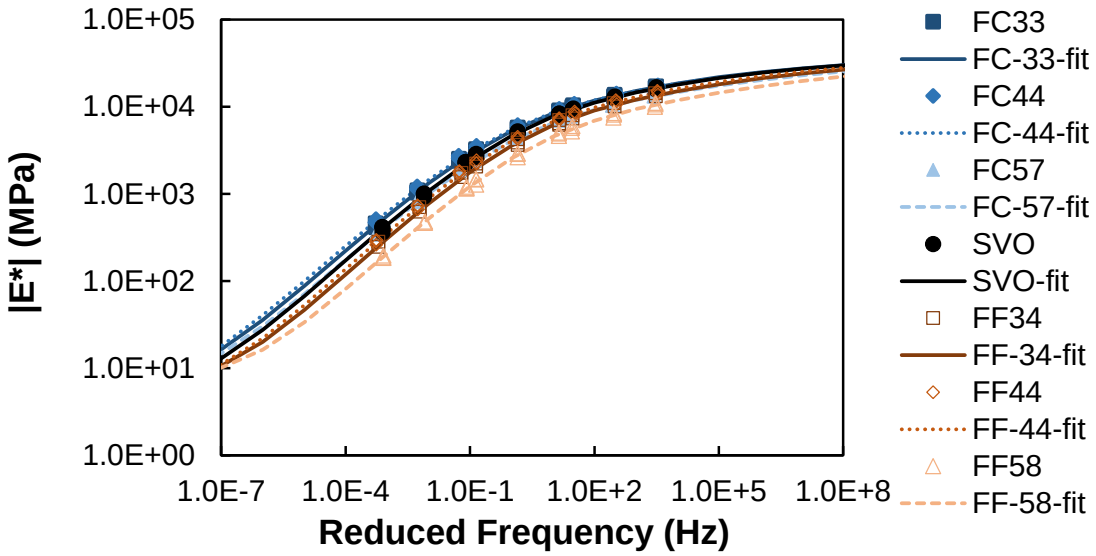

Figure 41: Four Corners and Two Points Dynamic Modulus Mastercurve.

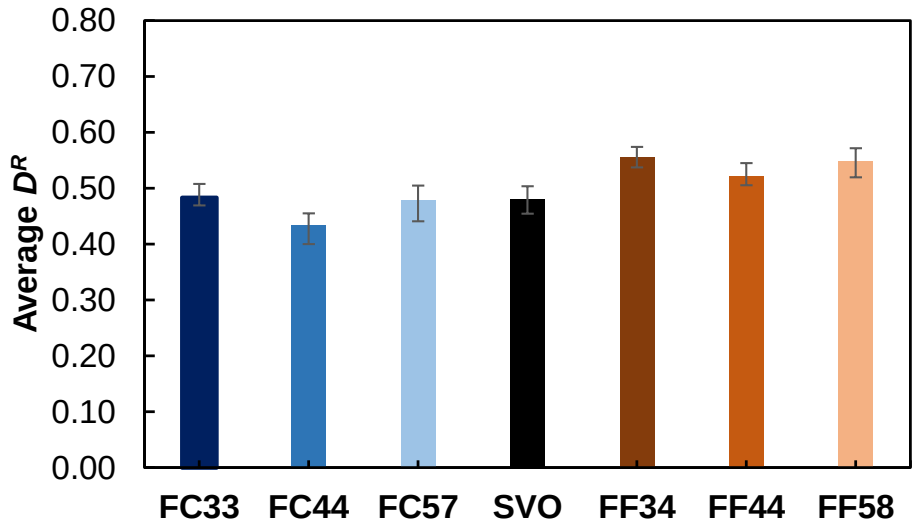

Figure 42: Four Corners and Two Points Average D^R Values.

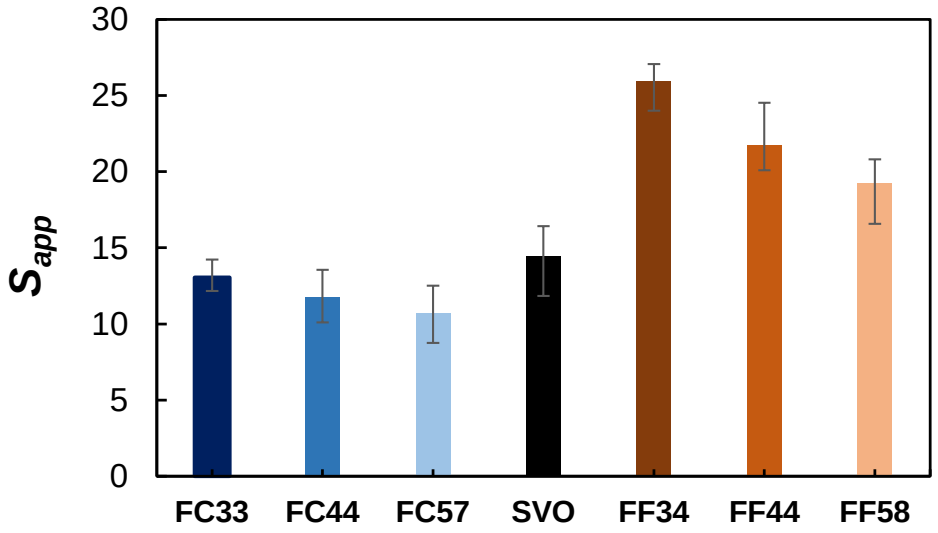

Figure 43: Four Corners and Two Points S_{app} Values.

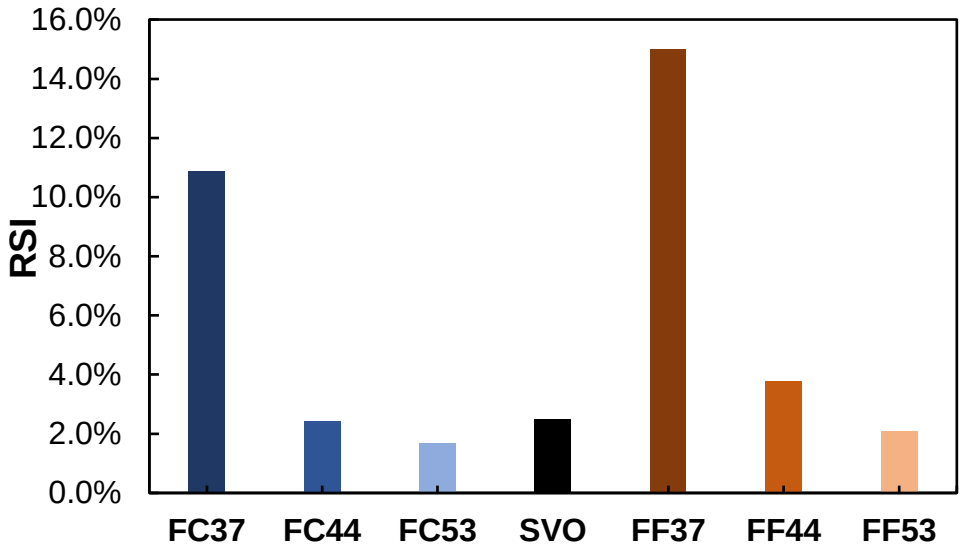

Figure 44: Four Corners and Two Points RSI Values.

The results from the IDEAL-CT and HWT tests are shown in Figure 45 and Figure 46, respectively. Both tests use the 62 mm tall specimens. The 53 corners were not possible to compact for the HWT specimens due to the specimen height and low target air void content. Consequently, HWT specimens were prepared for the 37 and 44 conditions only. Even for the FC44 condition, the specimen height had to be changed from 62 mm to 63 mm to achieve the target air void content of 4%. Similar to HWT testing, the coarse gradation specimen height for the IDEAL-CT testing needed to be increased from 62 mm to 63 mm. This change in height was accounted for when calculating the CT-Index. The fine gradation specimen height was kept at 62 mm.

In Figure 45, the FF44 condition has a CT-Index of 27.6, while the FC44 condition has a CT-Index of 12.9. The SVO's CT-Index was observed to be 23.4. These results align with the expected trend, i.e., cracking resistance increases as the gradation becomes finer. The rutting data shown in Figure 46 follow the expected trend, that is, rutting resistance increases as the binder content and air void content decrease and the gradation becomes coarser. It should be noted that the finer gradation showed significant stripping.

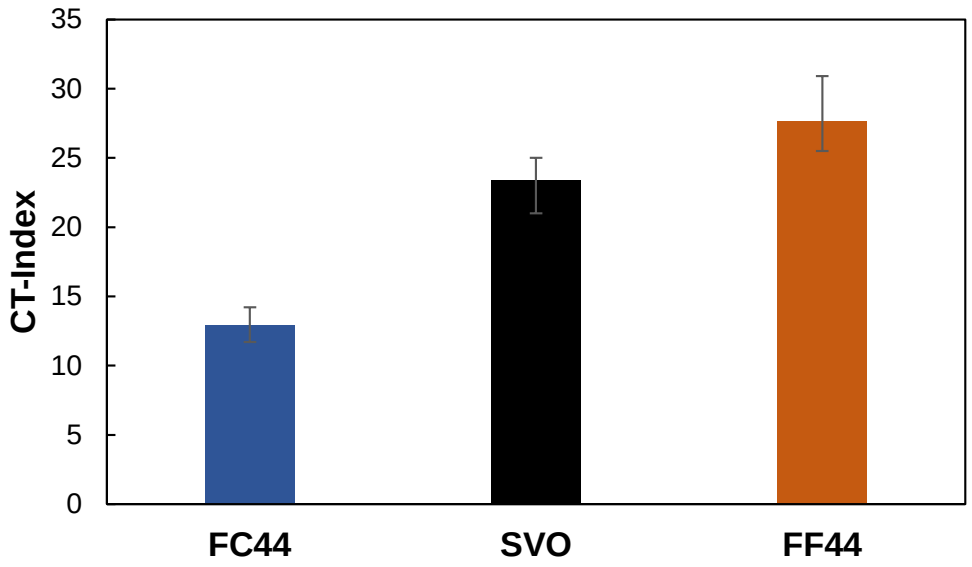

Figure 45: The Two Points CT-Index Values.

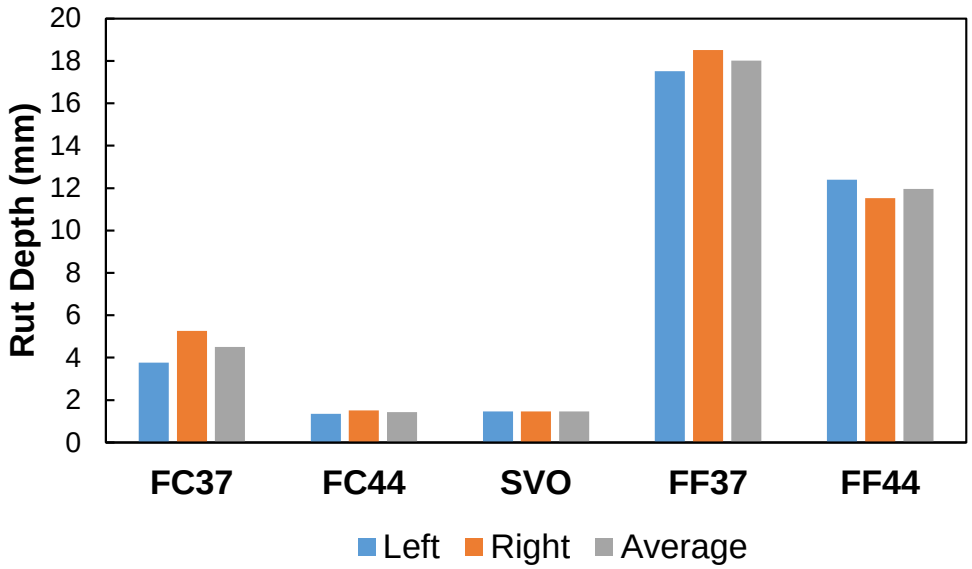

Figure 46: Four Corners and Two Points HWT Rut Depth.

RI19.0C mixture

Similar to the surface mixture, BMD+ and BMD performance testing was performed on the selected volumetric conditions. However, one difference is that the IDEAL-CT testing was performed on the 47 points to leverage the results in predicting the CT-Index of the PMLC samples.

The dynamic modulus mastercurves for the four corners and two points are presented in Figure 47. The coarser gradation points showed similar moduli; however, at lower reduced frequencies the 33 condition exhibited a slightly higher modulus. The same observation can be made for the finer gradation, with more pronounced differences observed at lower frequencies. It can be also noticed that the SVO mastercurve aligns better with the coarse gradation compared to the fine gradation.

The D^R values presented in Figure 48 show lower values for the coarse gradation compared to the fine gradation. Also, while the FC44 condition did not show a higher value compared to the FC57 corner, it can be noticed that the D^R value decreased systemically for the fine gradation when transitioning from the FF33 to FF44 to FF58 conditions. Overall, the D^R values ranged from 0.51 to 0.59 for the coarse gradation, and from 0.64 to 0.70 for the fine gradation. The same D^R trend was observed for the S_{app} values in Figure 49, as the fine gradation exhibited better fatigue resistance compared to the coarse gradation. The S_{app} values ranged from 11.7 to 18.2 for the coarse gradation, and from 19.5 to 32 for the fine gradation.

The RSI values presented in Figure 50 were higher for the 37 conditions followed by the 44 and 57 conditions, respectively. It can be also observed that the 37 condition's RSI values were significantly higher than the other corners due to the high binder and air void content. The coarser gradation exhibited lower RSI values compared to the fine gradation. The RSI values ranged from 1.20% to 4.40% for the coarse gradation and from 1.80% to 6.40% for the fine gradation.

As mentioned earlier, IDEAL-CT testing was performed on the 47 condition to leverage the results in predicting the PMLC CT-Index values. The presented CT-Index values in Figure 51 shows that the FC47 condition had a CT-Index of 30.1 and a value of 48.9 for the FF47 condition, with a CT-Index of 31.2 for the SVO condition. HWT rut depth presented in Figure 52 show higher rut depths and sensitivity for the finer gradation, as the rut depth did not change significantly from FC44 to FC55. This observation is mainly a result of the higher binder content associated with the fine gradation. HWT rut depth ranged from 1.15 mm to 3.94 mm for the coarse gradation and from 2.00 to 5.37 for the fine gradation.

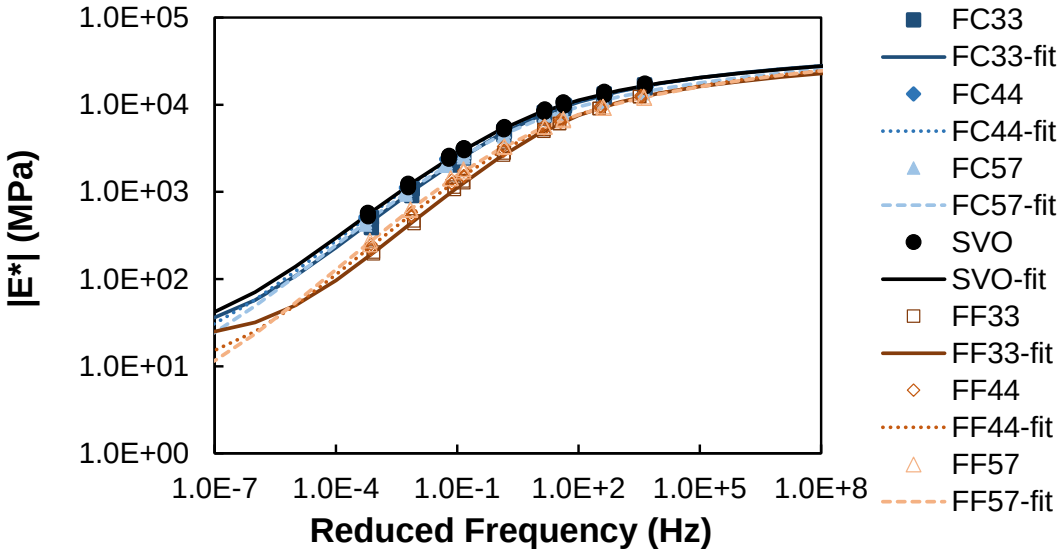

Figure 47: Four Corners and Two Points Dynamic Modulus Mastercurve.

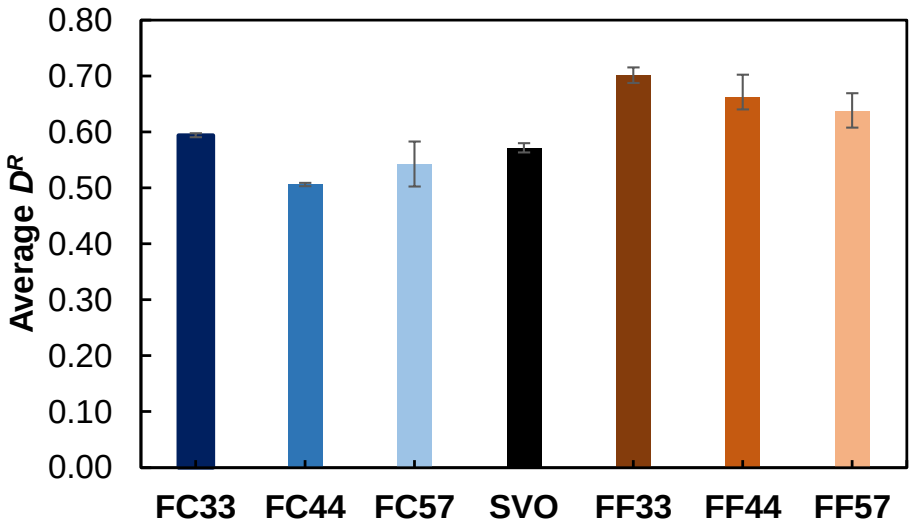

Figure 48: Four Corners and Two Points Average D^R Values.

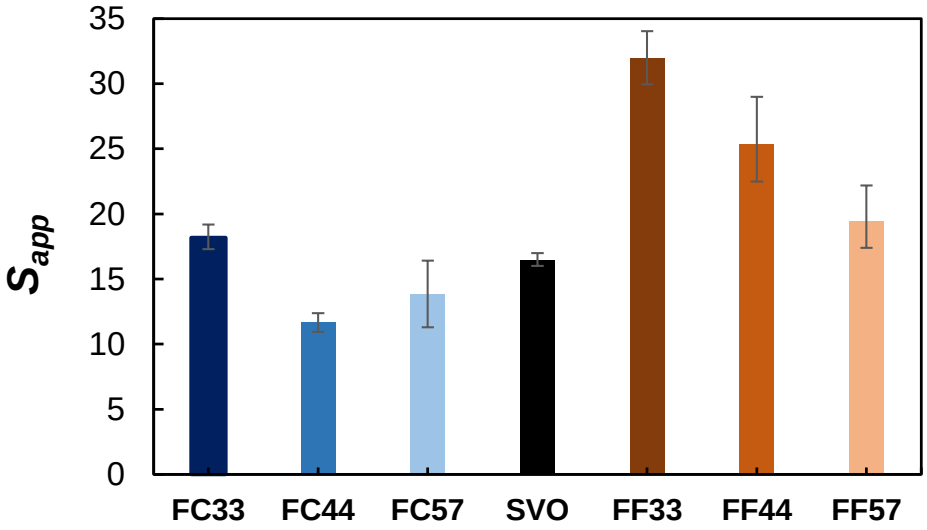

Figure 49: Four Corners and Two Points S_{app} Values.

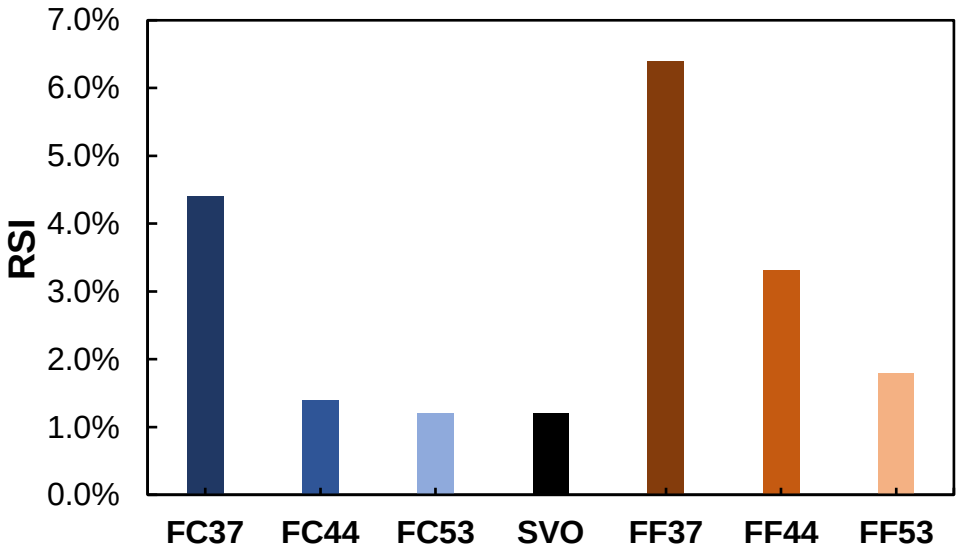

Figure 50: Four Corners and Two Points RSI Values.

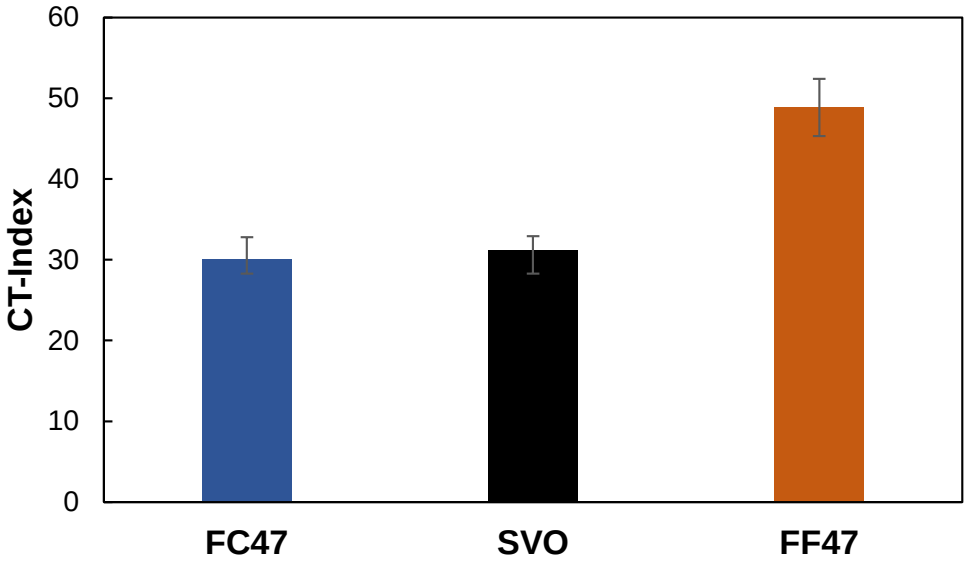

Figure 51: The Two Points CT-Index Values.

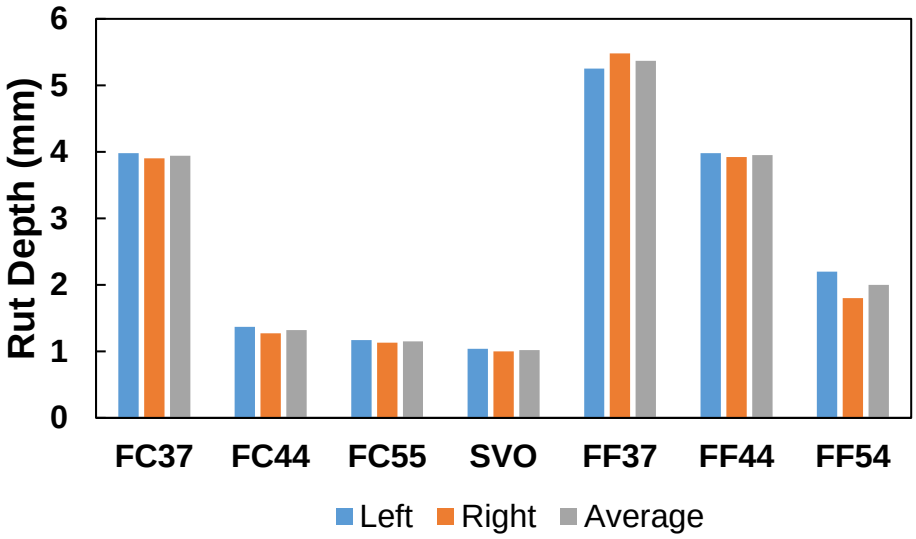

Figure 52: Four Corners and Two Points HWT Rut Depth.

RB25.0C mixture

For the base mixture, the main focus was the fatigue performance and not the rutting. Consequently, only dynamic modulus and cyclic fatigue tests were performed on the four corners. The dynamic modulus mastercurves show that the coarse gradation exhibited higher stiffness compared to the fine gradation. However, no significant difference was noticed between the 33 and 57 conditions for both gradations. In Figure 54, a higher variability in D^R values was observed for the coarse gradation. FC33 showed a higher D^R value compared to FF33, but FF57 showed almost the same D^R value compared to FC57. The S_{app} values shown in Figure 55 present

the effect of binder content, air void content, and gradation on cracking performance better than the D^R values in Figure 54. The fine gradation showed higher values compared to the coarse gradation as the S_{app} values ranged from 8.7 to 11.4 for the coarse gradation and from 11.6 to 21.1 for the fine gradation. The 33 conditions with higher binder content and low air void content showed higher S_{app} values compared to the 57 conditions regardless of the gradation type.

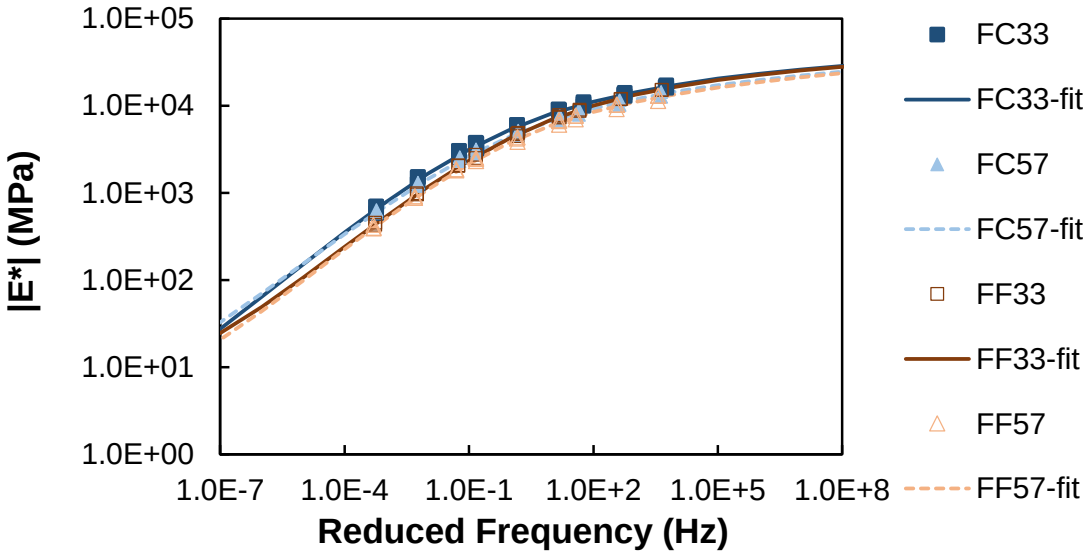

Figure 53: Four Corners Dynamic Modulus Mastercurve.

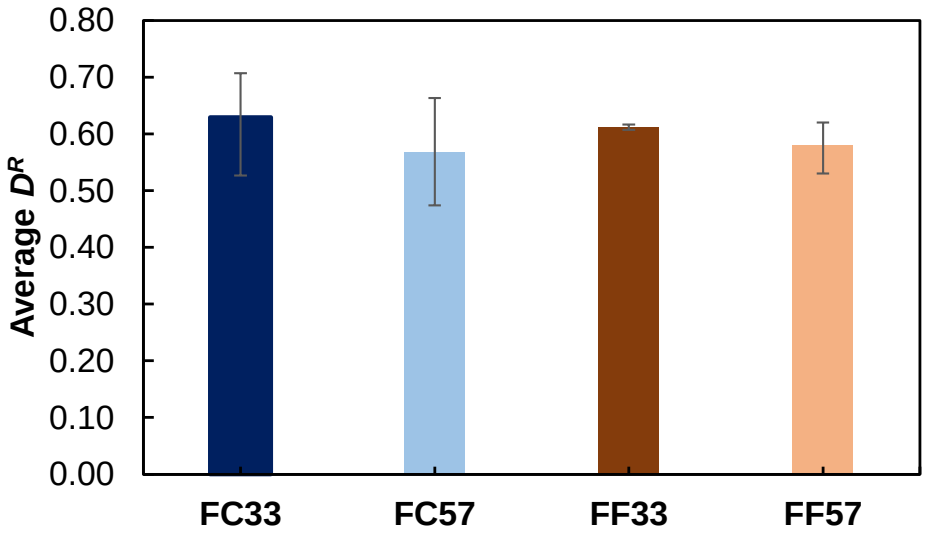

Figure 54: Four Corners Average D^R Values.

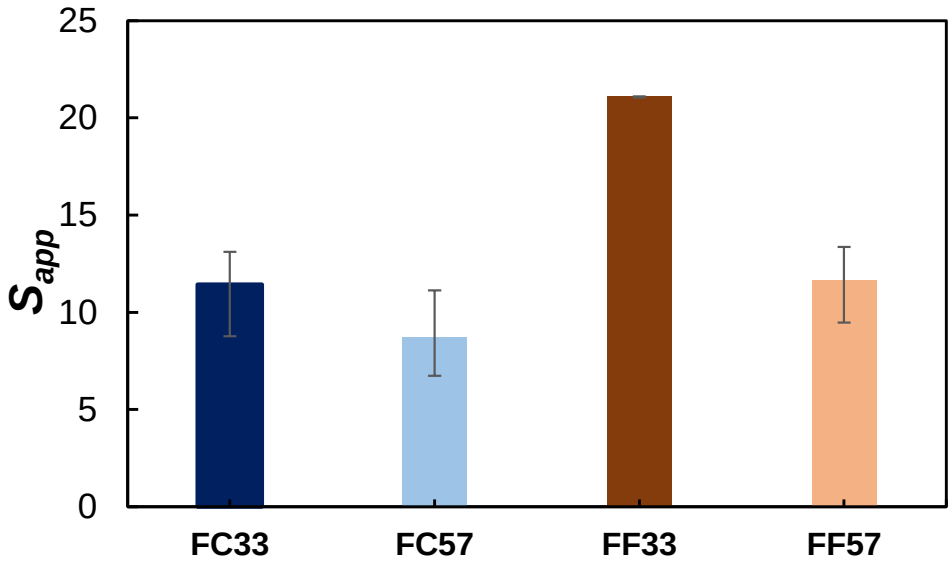

Figure 55: Four Corners S_{app} Values.

Index-volumetric relationships calibration and predictions

After performance testing was performed on the selected four corners for the three mixtures, the research team calibrated the IVR functions using the measured performance indexes as well as the determined volumetrics. The IVR functions are shown in Equations (9) to (12).

$$S_{app} = \alpha_1 \times VMA_{IP} + \alpha_2 \times VFA_{IP} + \alpha_3 \quad (9)$$

$$CT\text{-}Index = \gamma_1 \times AC_{eff.} + \gamma_2 \quad (10)$$

$$\text{RSI} = \beta_1 \times VMA_{IP} + \beta_2 \times VFA_{IP} + \beta_3 \quad (11)$$

$$\text{HWT-Rut Depth} = \delta_1 \times VMA_{IP} + \delta_2 \times VFA_{IP} + \delta_3 \quad (12)$$

Where VMA_{IP} is voids in mineral aggregates determined using in-place air void content, VFA_{IP} is voids filled with asphalt determined using in-place air void content, $\alpha's, \beta's, \gamma's, and\ \delta's$ are fitting coefficients for the performance indexes.

As can be seen from the equations, all the performance indexes IVR functions were calibrated with VMA_{IP} and VFA_{IP}, with the exception of CT-Index function which was calibrated using the effective binder content. The research team decided to use the effective binder content for calibrating the CT-Index IVR function based on evaluating the prediction accuracy of three different parameters: VMA_{IP}, asphalt binder content, and effective binder content. The research team used RI19.0C mixture results to perform this evaluation and the results are presented in Figure 56. As can be seen, effective binder content showed the best prediction accuracy out of the three evaluated parameters. Consequently, effective binder content was used to calibrate the CT-Index IVR function for RI19.0C as well as RS9.5C. The fitting coefficients for all the performance indexes IVR functions are presented in Table 13 to Table 16.

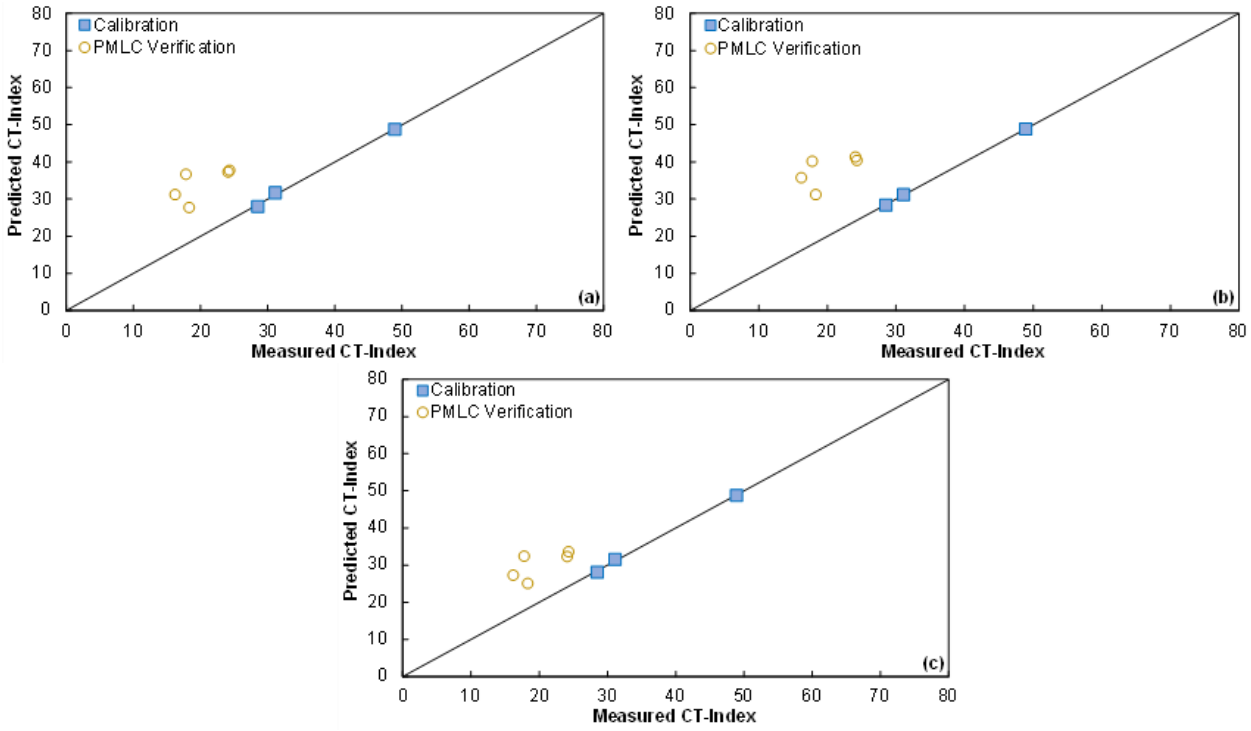

Figure 56: Evaluating CT-Index IVR Function Calibration Parameter: *(a)* Using VMA_{IP}; *(b)* Using Asphalt Binder Content; and *(c)* Using Effective Binder Content.

Table 13: S_{app} IVR Fitting Coefficients.

	α_1	α_2	α_3
RS9.5C	2.201	0.490	-58.474
RI19.0C	3.751	0.393	-69.698
RB25.0C	2.068	0.507	-53.824

Table 14: CT-Index IVR Fitting Coefficients.

	γ_1	γ_2
RS9.5C	8.433	-25.688
RI19.0C	16.300	-45.130

Table 15: RSI IVR Fitting Coefficients.

	β_1	β_2	β_3
RS9.5C	1.269	-0.244	1.625
RI19.0C	0.697	-0.013	-7.846

Table 16: HWT Rut Depth IVR Fitting Coefficients.

	δ_1	δ_2	δ_3

RS9.5C	2.842	0.489	-79.821
RI19.0C	0.607	0.011	-8.107

The research team predicted the performance of the calibration and verification samples using the developed IVR functions. The four corners represent the calibration points, and the other LMLC mixtures as well as the PMLC samples represent the verification points. The measured and predicted values as well as the percent error in predictions for the RS9.5C mixture are presented in Table 17 to Table 21. Percent error is defined as the absolute difference between measured and predicted values divided by the measured index value. The presented S_{app} predictions show that the IVR function was able to predict the measured performance with resonable accuracy. In Table 17, the calculated average percent error of the predictions varied from 9.7% for the calibration points to 7.2% for the LMLC verification points and 11.0% for the PMLC verification points. Figure 57 shows the measured and predicted S_{app} values along the line of equality (LOE). It can be observed that the calibration points as well as LMLC verification points were located along the LOE. Although the PMLC samples exhibited a slightly higher average prediction error compared to the LMLC verification samples, the accuracy level was reasonable.

For the IDEAL-CT test, the researchers used the 44 conditions as well as the SVO condition for the IVR calibration. The calibration average percent error was 11.7% in Table 18. Figure 58 shows the LOE plot of the calibration data. The CT-Index for the PMLC samples was not predicted since they were tested with a target air void content of 7% and there were not enough materials left to calibrate the CT-Index IVR function using the 47 conditions.

The presented RSI predictions in Table 19 show that while the numerical difference between measured and predicted values was not significant, the percent error was much higher for the conditions where the RSI index numerical value was small. This observation is supported by the results shown in Figure 59, which show that the predictions were reasonably accurate compared to the measured values. Overall, the average percent error in RSI values predictions varied from 36.5% for the calibration points to 42.8% for the LMLC verification points and 16.2% for the PMLC verification points.

HWT rut depth predictions presented in Table 20 show that while the average percent error for the calibration points was relatively low with a value of 21.5%, it was significantly higher for the SVO point and PMLC samples, with values of 304.9% and 98.6%, respectively. This overprediction can be seen clearly in Figure 60. The research team believed that the reason for this high variability was due to the combined effect of rutting and stripping that was captured by the testing of calibration points as well as the high rutting values observed for the 37 conditions, which yielded an overpredictve IVR function. To verify this hypothesis, a dry HWT testing was performed on the four corners to calibrate another HWT rut depth IVR function. The testing was also performed on four PMLC samples to evaluate the prediction accuracy of the newly calibrated IVR function. As can be noticed from Table 21, the average percent error was reduced from 21.5% to 2.1% for the calibration conditions, and from 98.6% to 24.7% for the PMLC samples. The LOE plot is shown in Figure 61. These results support the reasoning mentioned above for why the original HWT rut depth IVR function yielded high errors in predictions.

Table 17: RS9.5C S_{app} IVR Predictions.

Condition	Purpose	S_{app}				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF34	Calibration	20.1	78.6	25.9	24.3	6.5
FF58	Calibration	22.2	61.8	19.3	20.8	7.8
FC33	Calibration	15.8	77.9	13.1	14.6	11.8
FC57	Calibration	17.6	59.2	10.7	9.4	12.6
Four Corners % Error Average						9.7
FF44	LMLC Verification	18.8	80.3	21.7	22.4	2.9
SVO	LMLC Verification	16.6	78.9	14.4	16.7	15.9
FC44	LMLC Verification	15.7	73.3	11.7	12.1	2.7
LMLC Verification % Error Average						7.2
PMLC-S1	PMLC Verification	20.1	64.3	20.8	17.3	17.0
PMLC-S2	PMLC Verification	20.6	66.3	21.8	19.5	10.8
PMLC-S3	PMLC Verification	21.0	66.6	17.0	20.5	20.5
PMLC-S4	PMLC Verification	20.0	65.1	18.3	17.5	4.5
PMLC-S5	PMLC Verification	20.7	65.4	19.7	19.2	2.5
PMLC Verification % Error Average						11.0

Table 18: RS9.5C CT-Index IVR Predictions.

Condition	Purpose	CT-Index			
		AC_{eff}	Measured	Predicted	% Error
FF44	Calibration	6.46	27.6	28.8	4.4
SVO	LMLC Verification	5.43	23.4	20.1	14.3
FC44	Calibration	4.83	12.9	15.0	16.4
Three Points % Error Average					11.7

Table 19: RS9.5C RSI IVR Predictions.

Condition	Purpose	RSI				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF37	Calibration	22.6	67.7	15.0	13.8	8.3
FF53	Calibration	17.4	83.7	2.1	3.3	60.1
FC37	Calibration	19.7	59.8	10.9	12.0	10.4
FC53	Calibration	14.0	77.5	1.7	0.6	67.3
Four Corners % Error Average						36.5
FF44	LMLC Verification	18.9	80.0	3.8	6.1	62.0
SVO	LMLC Verification	16.7	78.5	2.5	3.6	46.4
FC44	LMLC Verification	15.4	75.0	2.4	2.9	19.8
LMLC Verification % Error Average						42.8
PMLC-S1	PMLC Verification	20.2	63.4	10.8	11.8	9.8
PMLC-S2	PMLC Verification	20.3	64.0	14.6	11.8	19.6
PMLC-S3	PMLC Verification	20.6	63.1	14.6	12.4	15.0

PMLC-S4	PMLC Verification	19.7	61.9	9.0	11.5	28.6
PMLC-S5	PMLC Verification	19.9	62.3	12.7	11.7	7.9
PMLC Verification % Error Average						16.2

Table 20: RS9.5C HWT Rut Depth (Wet) IVR Predictions.

Condition	Purpose	HWT Rut Depth (Wet)				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF37	Calibration	22.2	69.3	18.02	17.0	5.4
FF44	Calibration	19.0	79.6	11.96	13.0	8.5
FC37	Calibration	19.6	60.1	4.51	5.3	16.8
FC44	Calibration	15.6	73.7	1.44	0.6	55.3
Four Corners % Error Average						21.5
SVO	LMLC Verification	17.0	76.7	1.46	5.9	304.9
PMLC-S1	PMLC Verification	19.9	65.1	4.20	8.5	102.0
PMLC-S2	PMLC Verification	20.7	66.2	6.64	11.3	69.5
PMLC-S3	PMLC Verification	21.0	66.8	5.16	12.4	141.2
PMLC-S4	PMLC Verification	20.0	64.9	5.37	8.8	64.5
PMLC-S5	PMLC Verification	20.5	66.4	4.98	10.8	116.0
PMLC Verification % Error Average						98.6

Table 21: RS9.5C HWT Rut Depth (Dry) IVR Predictions.

Condition	Purpose	HWT Rut Depth (Dry)				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF37	Calibration	22.2	69.3	0.41	0.42	1.7
FF44	Calibration	19.0	79.6	0.38	0.37	1.9
FC37	Calibration	19.6	60.1	0.26	0.25	2.1
FC44	Calibration	15.6	73.7	0.20	0.21	2.9
Four Corners % Error Average						2.1
PMLC-S2	PMLC Verification	20.7	66.2	0.36	0.34	6.0
PMLC-S3	PMLC Verification	21.0	66.8	0.28	0.35	26.6
PMLC-S4	PMLC Verification	20.0	64.9	0.19	0.31	65.2
PMLC-S5	PMLC Verification	20.5	66.4	0.34	0.33	0.9
PMLC Verification % Error Average						24.7

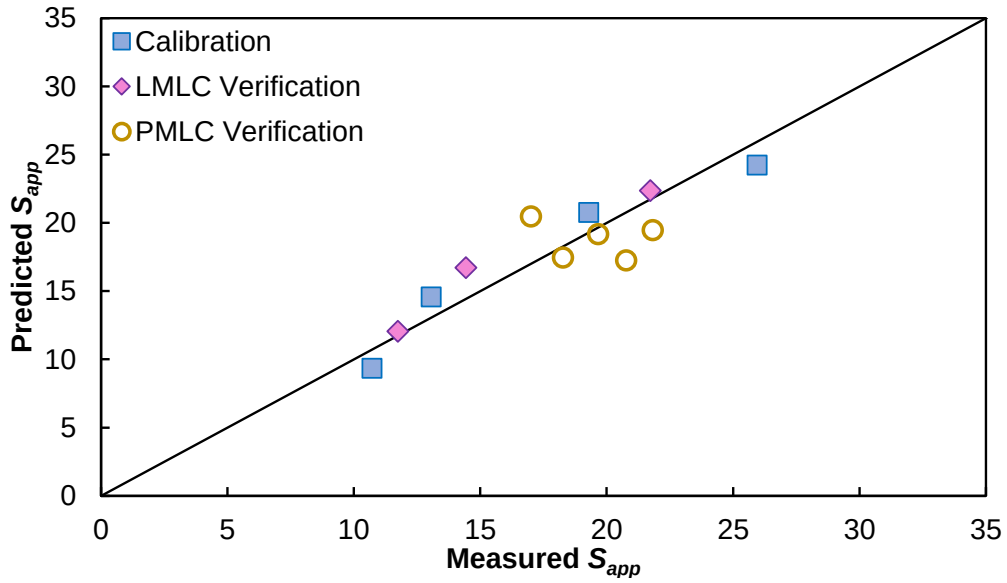

Figure 57: RS9.5C Measured and Predicted S_{app} Values.

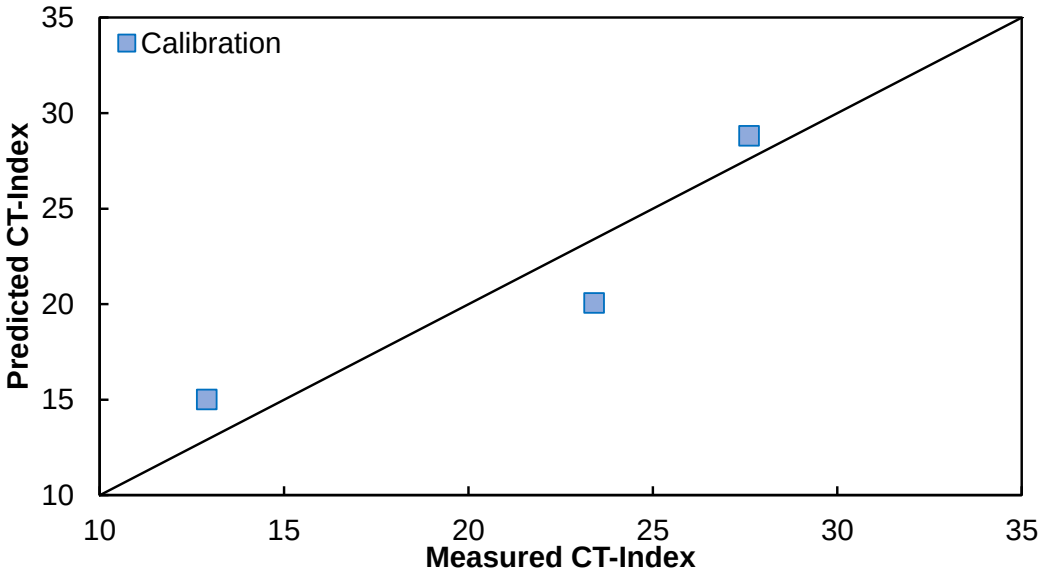

Figure 58: RS9.5C Measured and Predicted CT-Index Values.

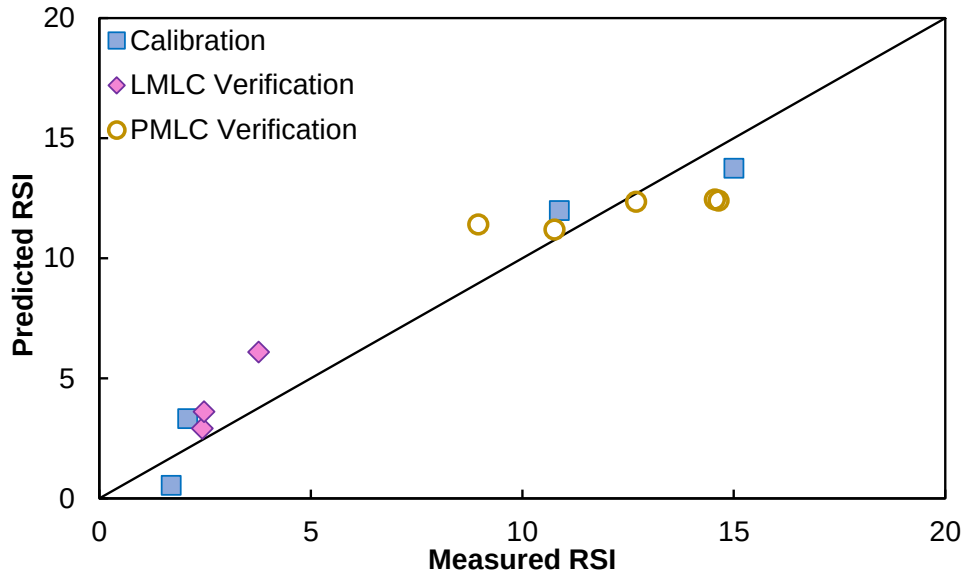

Figure 59: RS9.5C Measured and Predicted RSI Values.

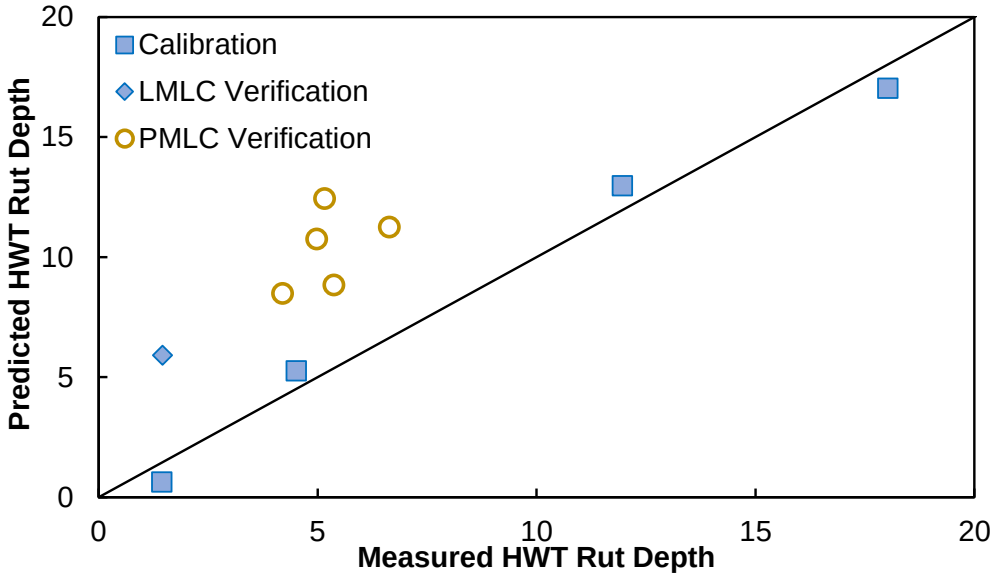

Figure 60: RS9.5C Measured and Predicted HWT Rut Depth (Wet) Values.

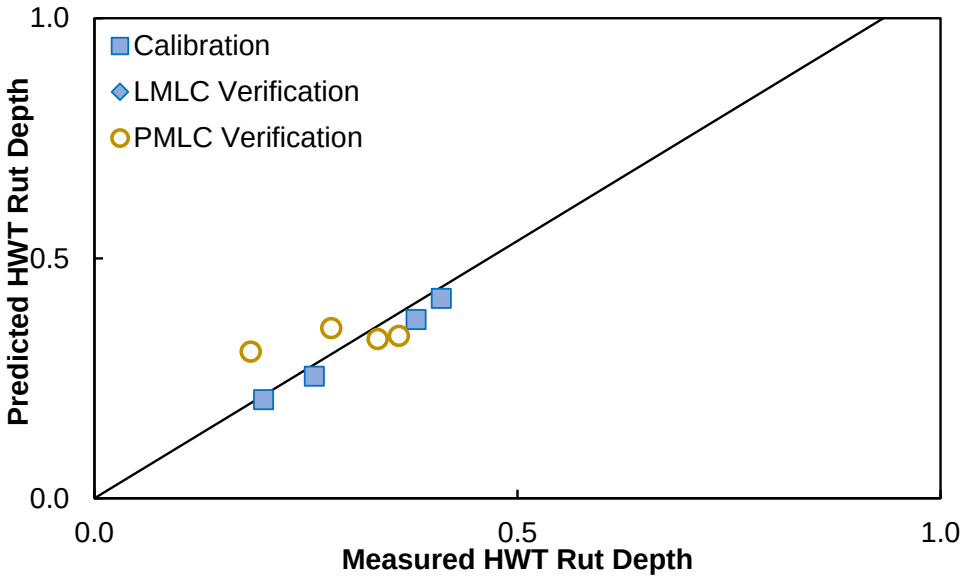

Figure 61: RS9.5C Measured and Predicted HWT Rut Depth (Dry) Values.

The measured and predicted index values using the calibrated IVR functions for the RI19.0C mixture are presented in Table 22 to Table 25 and Figure 62 to Figure 65. The presented S_{app} predictions show that the IVR function was able to predict the measured performance with resonable accuracy. The calculated average percent error of the predictions varied from 11.7% for the calibration conditions to 9.1% for the LMLC verification conditions and 12.8% for the PMLC verification conditions. The accuracy of predictions can be observed also in Figure 62.

The average percent error for the CT-Index predictions ranged from 0.94% for the calibration conditions to 52.1% for the PMLC verification conditions. It can be noticed from Figure 63 that while the PMLC samples CT-Index did not fall within the determined CT-Index of the calibration conditions, the accuracy of the predictions was reasonable and able to capture the effect of binder content.

The RSI predictions average percent error varied from 3.9% for the calibration conditions to 12.5% for the LMLC verification conditions and 48.3% for the PMLC verification conditions. The HWT rut depth predictions average percent error varied from 0.9% for the calibration conditions to 33.9% for the LMLC verification conditions and 54.7% for the PMLC verification conditions. Overall, all the results presented in Figure 62 to Figure 65 highlight the reasonableness of the IVR prediction accuracy.

Table 22: RI19.0C S_{app} IVR Predictions.

Condition	Purpose	S_{app}				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF33	Calibration	17.7	85.3	32.0	30.1	5.9
FF57	Calibration	18.0	61.1	19.4	21.9	12.9
FC33	Calibration	14.8	80.4	15.6	17.3	11.2
FC57	Calibration	15.9	56.0	14.3	11.9	16.5
Four Corners % Error Average						11.7

FF44	LMLC Verification	17.4	75.3	25.4	25.5	1.1
SVO	LMLC Verification	16.7	15.5	1.5	16.7	15.5
FC44	LMLC Verification	14.9	69.8	11.0	13.6	24.9
LMLC Verification % Error Average						9.1
PMLC-I1	PMLC Verification	16.5	66.8	16.9	18.6	10.2
PMLC-I2	PMLC Verification	16.5	59.4	15.9	15.5	2.2
PMLC-I3	PMLC Verification	16.2	62.9	13.7	15.7	14.8
PMLC-I4	PMLC Verification	16.8	64.4	16.2	18.8	16.0
PMLC-I5	PMLC Verification	17.0	63.6	15.8	19.1	21.0
PMLC Verification % Error Average						12.8

Table 23: RI19.0C CT-Index IVR Predictions.

Condition	Purpose	CT-Index			
		AC_{eff}	Measured	Predicted	% Error
FF47	Calibration	5.76	48.9	48.8	0.15
SVO	Calibration	4.70	31.1	31.5	1.41
FC47	Calibration	4.50	28.5	28.1	1.28
Three Points % Error Average					0.94
PMLC-I1	PMLC Verification	4.83	24.3	33.6	38.3
PMLC-I2	PMLC Verification	4.31	18.3	25.1	37.1
PMLC-I3	PMLC Verification	4.44	16.2	27.3	68.6
PMLC-I4	PMLC Verification	4.76	17.8	32.4	82.2
PMLC-I5	PMLC Verification	4.75	24.1	32.3	34.2
PMLC Verification % Error Average					52.1

Table 24: RI19.0C RSI IVR Predictions.

Condition	Purpose	RSI				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF37	Calibration	21.6	66.6	6.4	6.2	1.3
FF54	Calibration	15.4	74.0	1.8	1.9	3.8
FC37	Calibration	18.8	60.1	4.4	4.5	2.6
FC55	Calibration	14.1	64.5	1.2	1.1	8.0
Four Corners % Error Average						3.9
FF44	LMLC Verification	17.4	75.3	3.31	3.30	0.3
SVO	LMLC Verification	15.0	75.3	1.20	1.61	34.0
FC44	LMLC Verification	14.7	71.3	1.40	1.45	3.3
LMLC Verification % Error Average						12.5
PMLC-I1	PMLC Verification	16.4	67.1	2.13	2.79	28.2
PMLC-I2	PMLC Verification	16.1	61.4	1.72	2.48	48.3
PMLC-I3	PMLC Verification	15.9	64.2	1.55	2.39	55.9
PMLC-I4	PMLC Verification	16.6	65.6	1.75	2.90	63.5
PMLC-I5	PMLC Verification	17.0	63.6	2.19	3.22	45.7

PMLC Verification % Error Average	48.3

Table 25: RI19.0C HWT Rut Depth IVR Predictions.

Condition	Purpose	HWT Rut Depth				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF37	Calibration	21.0	69.0	5.37	5.35	0.4
FF54	Calibration	15.4	74.0	2.00	2.02	1.0
FC37	Calibration	18.8	60.1	3.94	3.96	0.5
FC55	Calibration	14.1	64.5	1.16	1.14	1.7
Four Corners % Error Average						0.9
FF44	LMLC Verification	17.5	74.8	3.95	3.30	16.5
SVO	LMLC Verification	15.0	75.3	1.02	1.78	74.5
FC44	LMLC Verification	14.5	72.4	1.32	1.46	10.6
LMLC Verification % Error Average						33.9
PMLC-I1	PMLC Verification	16.9	64.9	1.95	2.85	46.2
PMLC-I2	PMLC Verification	16.6	58.9	1.74	2.59	48.9
PMLC-I3	PMLC Verification	16.3	62.5	1.31	2.43	85.5
PMLC-I4	PMLC Verification	16.8	64.3	1.63	2.80	71.8
PMLC-I5	PMLC Verification	17.4	61.6	2.59	3.14	21.2
PMLC Verification % Error Average						54.7

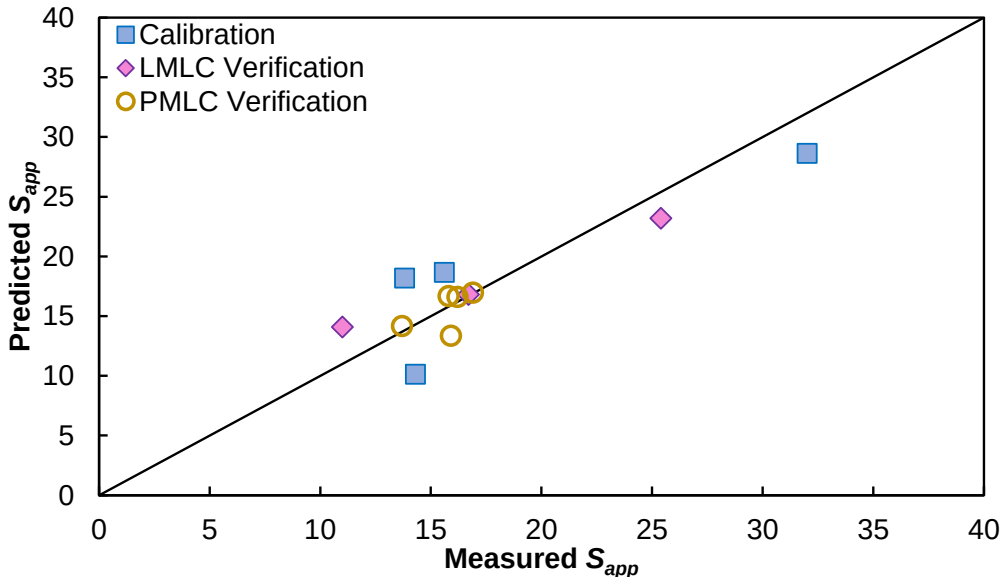

Figure 62: RI19.0C Measured and Predicted S_{app} Values.

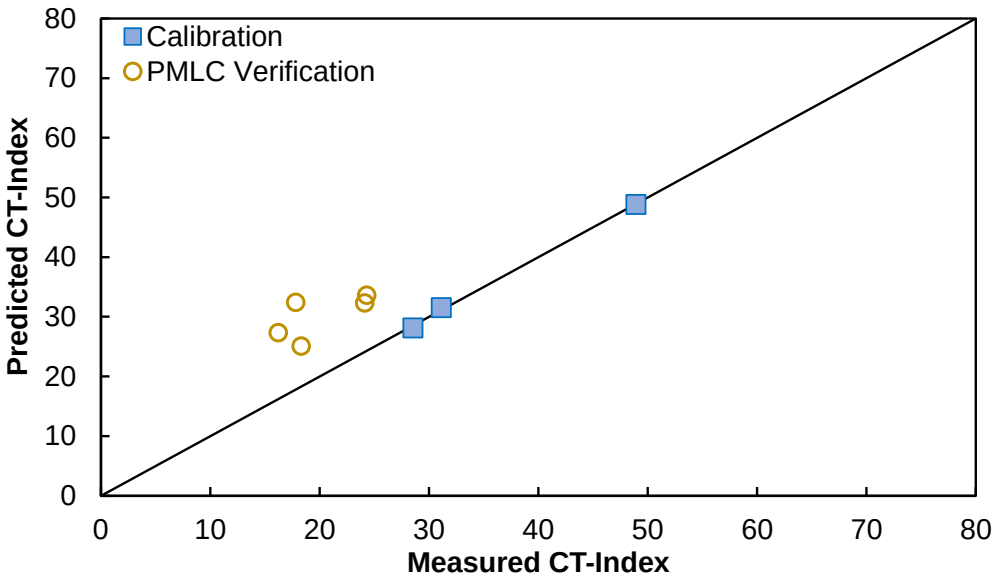

Figure 63: RI19.0C Measured and Predicted CT-Index Values.

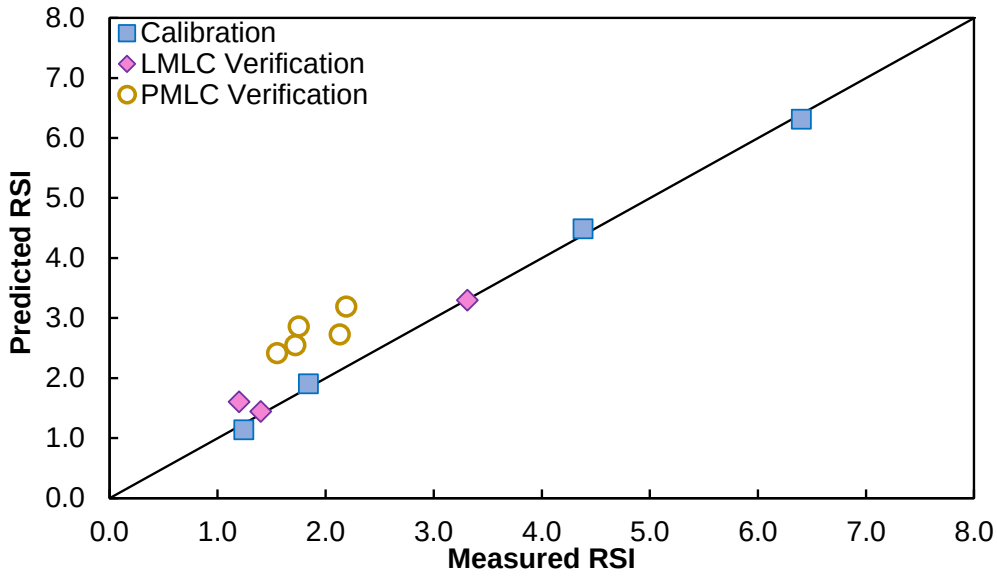

Figure 64: RI19.0C Measured and Predicted RSI Values.

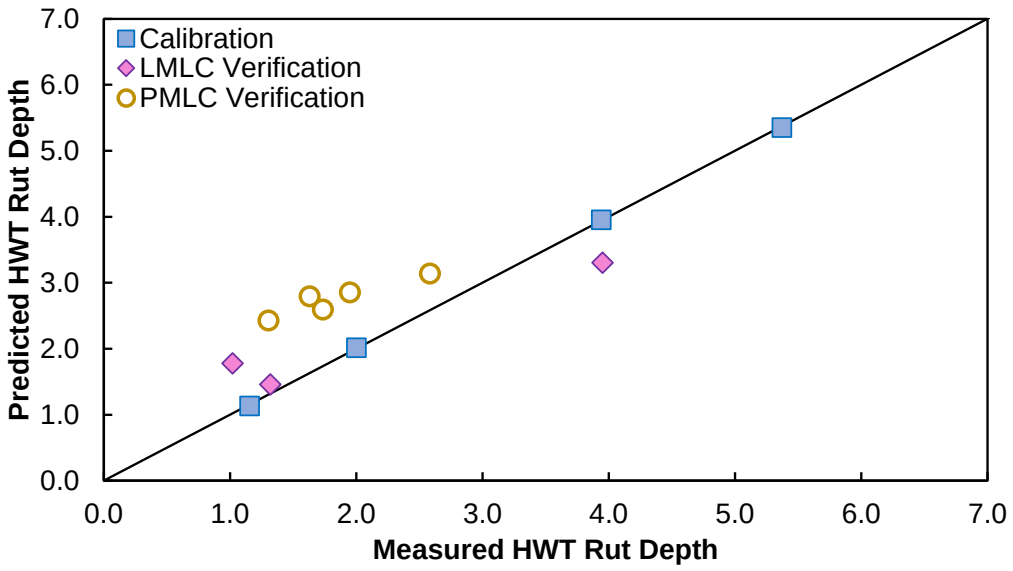

Figure 65: RI19.0C Measured and Predicted HWT Rut Depth Values.

RB25.0C measured and predicted S_{app} values are presented in Table 26 and Figure 66. The average percent error ranged from 15.3% for the calibration conditions to 32.0% for the PMLC verification conditions. As can be noticed from Figure 66, all the calibration points were located along the line of equality, and the S_{app} values for the PMLC samples were predicted with a reasonable accuracy except for PMLC-B1 where a higher variation was observed.

Table 26: RB25.0C S_{app} IVR Predictions.

Condition	Purpose	S_{app}				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF33	Calibration	16.2	78.0	21.1	19.1	9.2
FF57	Calibration	19.4	53.7	11.6	13.6	16.9
FC33	Calibration	13.7	76.0	11.4	13.1	15.1
FC57	Calibration	17.5	48.5	8.7	6.9	20.0
Four Corners % Error Average						15.3
PMLC-B1	PMLC Verification	16.7	57.1	17.9	9.6	46.3
PMLC-B2	PMLC Verification	16.5	53.8	10.4	7.6	26.3
PMLC-B3	PMLC Verification	17.5	56.5	8.5	11.1	31.5
PMLC-B4	PMLC Verification	15.0	58.7	9.2	7.0	23.8
PMLC Verification % Error Average						32.0

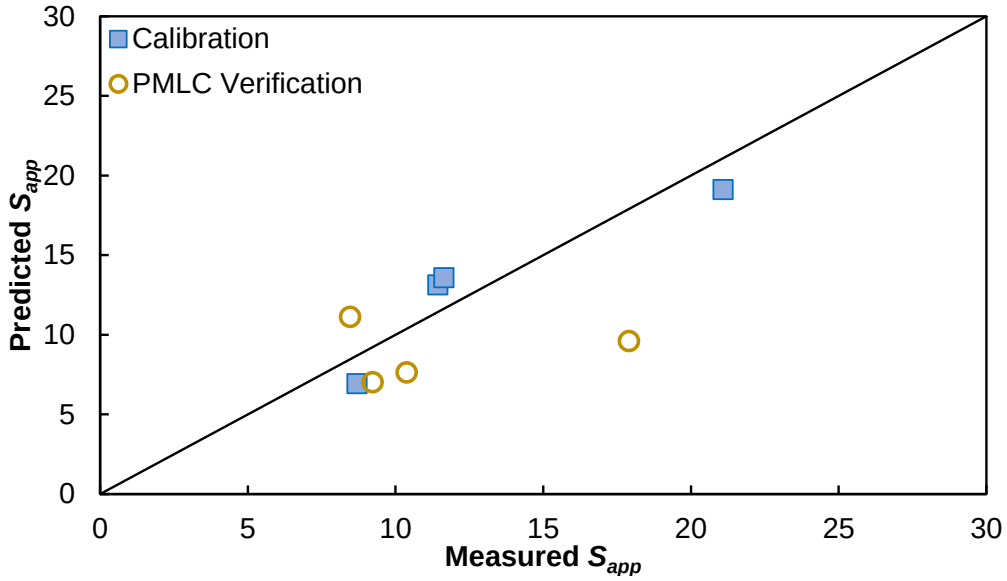

Figure 66: RB25.0C Measured and Predicted S_{app} Values.

An important observation to be made in all the IVR verification results is that the measured performance of PMLC samples shows more variability than predicted. Since actual binder content and air voids are accounted for through VMA_{IP} and VFA_{IP} in the IVR prediction but a single G_{sb} value is used for all PMLC samples, it is likely that this uniform G_{sb} assumption contributes to the lower spread in predictions compared to measurements.

CHAPTER 5. CALIBRATION OF PERFORMANCE-VOLUMETRICS RELATIONSHIPS

Performance-volumetrics relationship (PVR) is the relationship between volumetrics and performance and is based on testing mixtures under four different volumetric conditions, namely, the four corners. The PVR concept has been applied and verified in several PRS shadow projects, including NCDOT Project HWY-2017-29. To develop the PVR function, FlexPAVE simulations are required.

FlexPAVE is a software program that integrates VECD theory to account for the influence of loading rates and temperature variations on pavement responses and failure mechanisms. The software enables the modeling of pavement structures composed of asphalt concrete (AC) and unbound materials. Users can assign specific material properties to each AC layer by incorporating output files generated by FlexMAT. Additionally, project-specific factors, including location, and traffic conditions can be specified. Climatic conditions are determined through the Enhanced Integrated Climatic Model (EICM) based on the project's geographical location. The primary output of FlexPAVE simulations includes predictions of pavement performance, expressed as damage percentage (% damage) and rut depth (cm) over the pavement's design life.

To determine the aging inputs for the simulations, the research team used the level 3 procedure described in the NCHRP 9-54 final report (Kim et al., 2021). The Level 3 procedure consists of determining the virgin and recycled materials properties separately using Equations (13) and (14) shown below followed by the use of Equations (15) and (16) to determine the mechanistic properties of a mixture that contain recycled materials. It should be noted that FlexMAT provides the calculated parameters by providing the high-temperature performance grade for the virgin binder and recycled materials as well as the replacement ratio.

$$log|G^*|_0 = \frac{6.552}{1 + \left(\frac{HPG}{107.167}\right)^{-4.068}} \quad (13)$$

$$M = \frac{0.94}{1 + \left(\frac{HPG}{82.68}\right)^{7.96}} \quad (14)$$

$$log|G^*|_{0,Blend} = (1 - ABR) \times log|G^*|_{0,Binder} + ABR \times log|G^*|_{0,RAP} \quad (15)$$

$$RM_{Blend} = (1 - X_{RAP})RM_{Binder} + X_{RAP}RM_{RAP} \quad (16)$$

where

$|G^*|_0$ = short-term aged binder shear modulus at 64°C, 10 rad/s (kPa),
M = parameter related to fast reaction reactive material to be used in the kinetics model,
HPG = asphalt binder high temperature performance grade,
ABR = asphalt binder replacement,
RM = reactive material,
X_{RAP} = mass fraction of the RAP, and
$(1 - X_{RAP})$ = mass fraction of the binder.

For the RS9.5C mixture, the inputs necessary for the aging model were established through testing of both the virgin binder and the binder extracted from the sampled RAP. For the intermediate and base mixtures, only the virgin binder was tested directly. However, for the recycled materials, the research team utilized data from the NCSU database for materials obtained from the same contractor (Castorena et al., 2022). The determined aging model coefficients are presented in Table 27.

Table 27: Aging Model Coefficients.

Mixture		*c*	$log\|G^*\|_{STA}$	*M*
RS9.5C	FF34	1.710	0.888	0.644
	FF58	1.710	0.888	0.629
	FC33	1.710	0.888	0.611
	FC57	1.710	0.888	0.587
	SVO	1.710	0.888	0.638
RI19.0C		1.710	1.532	0.577
RB25.0C		1.710	1.612	0.552

Considering that the surface mixtures were sampled from a different project compared to the intermediate and base mixtures, the research team used the same typical structure used in the intermediate and base mixtures construction project. The structure consists of 76.2 mm (3 inches), 101.6 mm (4 inches), and 114.3 mm (4.5 inches) for the surface, intermediate, and base mixtures, respectively. FlexPAVE simulations were run using the volumetric properties and performance test results from the four corners of the surface mixture to calibrate the PVR functions. Note that the material properties, including the dynamic modulus mastercurve and the S-VECD properties of the intermediate and base mixtures, were kept the same among different simulations. That is, only the surface mixture properties were changed to develop the PVRs.

Damage and damage factor contours at the end of seasons 30, 60, and 80 are presented in Figure 67 to Figure 70 for the four corners, respectively. In FlexPAVE, the damage factor is calculated using Equation (17). A damage factor of one implies the failure of the material, i.e., a crack. In all the damage factor contours, it can be seen that bottom-up cracking is more prevalent than top-down cracking. Also, the difference among the damage contours from the four volumetric conditions is deemed insignificant.

$$Damage\ Factor = \frac{(1-C)_{accumulated}}{N_{accumulated}} \times \frac{1}{D^R} \quad (17)$$

where

$(1\text{-}C)_{accumulated}$ = the accumulated loss in pseudo stiffness up to the current point in time, and
$N_{accumulated}$ = the cumulative number of cycles up to the current point.

In order to quantify the amount of damage, FlexPAVE introduces '% Damage' parameter using Equation (18). In this equation, the subscript *i* refers to each of the nodal points within the given

pavement cross-section, M is the total number of nodal points, and A_i is the area represented by nodal point i within the finite element mesh.

$$\% \, Damage = \frac{\sum_1^M (Damage \; Factor)_i \times A_i}{\sum_1^M A_i} \times 100 \quad (18)$$

Three different % Damage values are calculated in FlexPAVE: total damage, top damage, and bottom damage. The total damage uses the damage factors in the entire reference area, which is 1.55 m (61 inches) wide and the combined asphalt layer thickness high. The top damage uses the top-third thickness of the combined asphalt layers in calculating the % Damage, whereas the bottom damage uses the lower two-thirds of the combined asphalt layers.

Figure 71 to Figure 73 present % Damage for total damage, top damage, and bottom damage, respectively. The total and bottom damage outcomes in Figure 71 and Figure 73 show negligible difference in % Damage among the four volumetric conditions. The difference in % Damage can be seen in the top damage in Figure 72, FF58 exhibiting the highest % Damage, followed by FF34, FF57, and FC33, in that order. The reason for the negligible difference in the top and bottom damage is that the material properties in the intermediate and base layers were the same in all the simulations. The effect of the difference in cracking performance of the different mixtures in the surface layer was lessened when the total reference area was used in calculating the total damage. Also, different material properties in the surface layer did not affect the cracking performance in the intermediate and base layers, resulting in identical % Damage for the bottom damage. These findings highlight the significant role of the intermediate and base mixtures within the pavement structure, emphasizing the need to optimize their design alongside the surface mixture to achieve substantial improvements in overall pavement performance.

Note that the top third of the pavement used in the top damage calculations includes the materials in the top 97.28 mm (3.83 inches), whereas the surface layer thickness used in the simulations was 76.2 mm (3 inches). To investigate the differences in performance between the four corners more accurately, the researchers calculated the top damage in the surface layer only. The results for the top damage in the surface layer only are presented in Figure 74. Although the damage difference among the four corners did not drastically increase from what is seen in Figure 72, the results still show better differentiation in performance among the four corners compared to Figure 72.

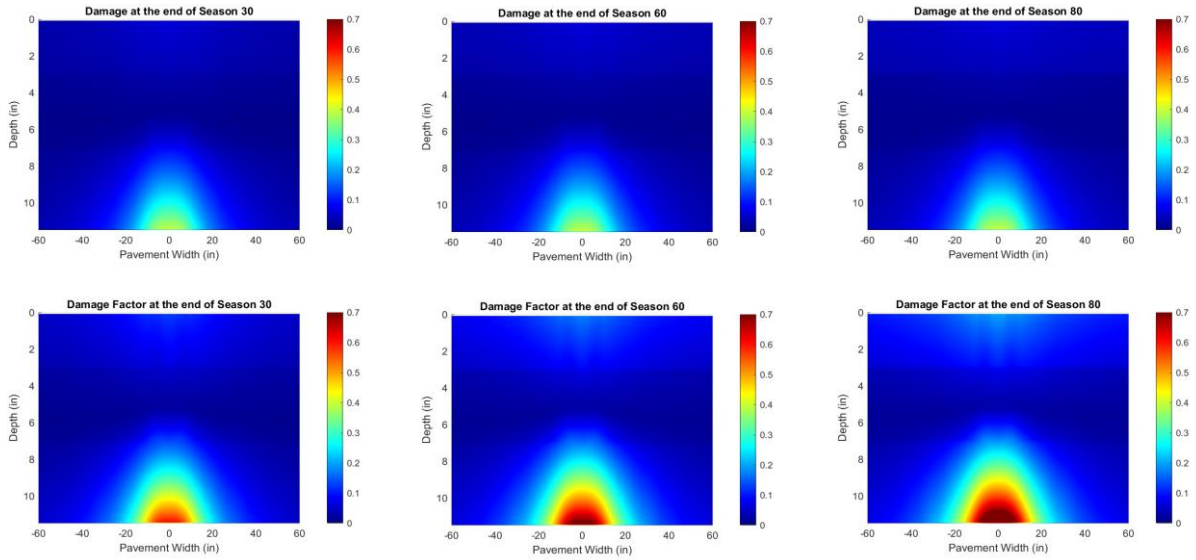

Figure 67: Damage and Damage Factor at the End of Season 30, 60, and 80 for FF34 Corner of RS9.5C Mixture.

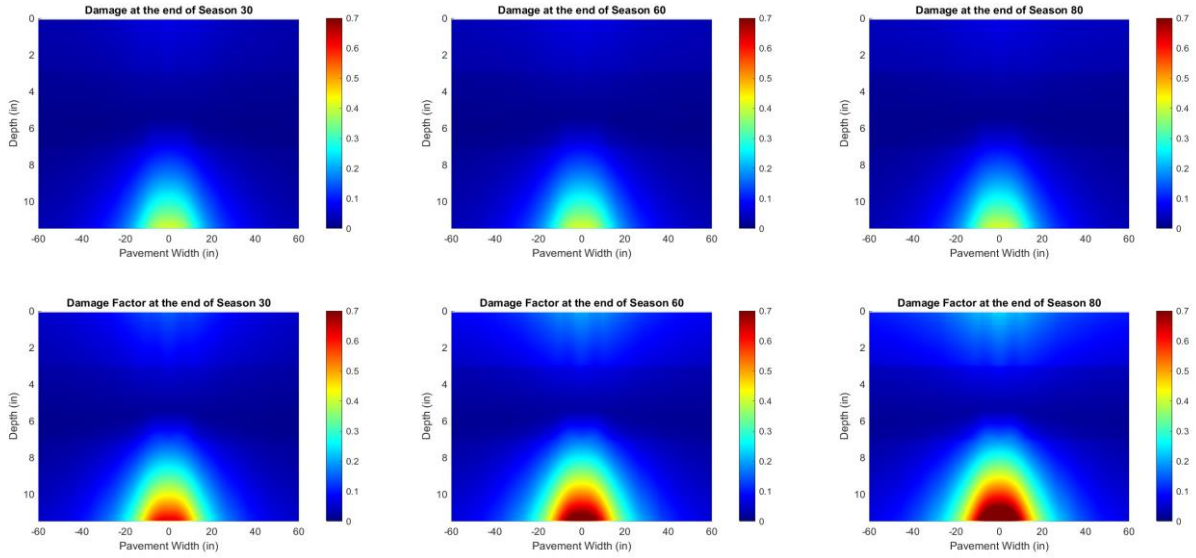

Figure 68: Damage and Damage Factor at the End of Season 30, 60, and 80 for FF58 Corner of RS9.5C Mixture.

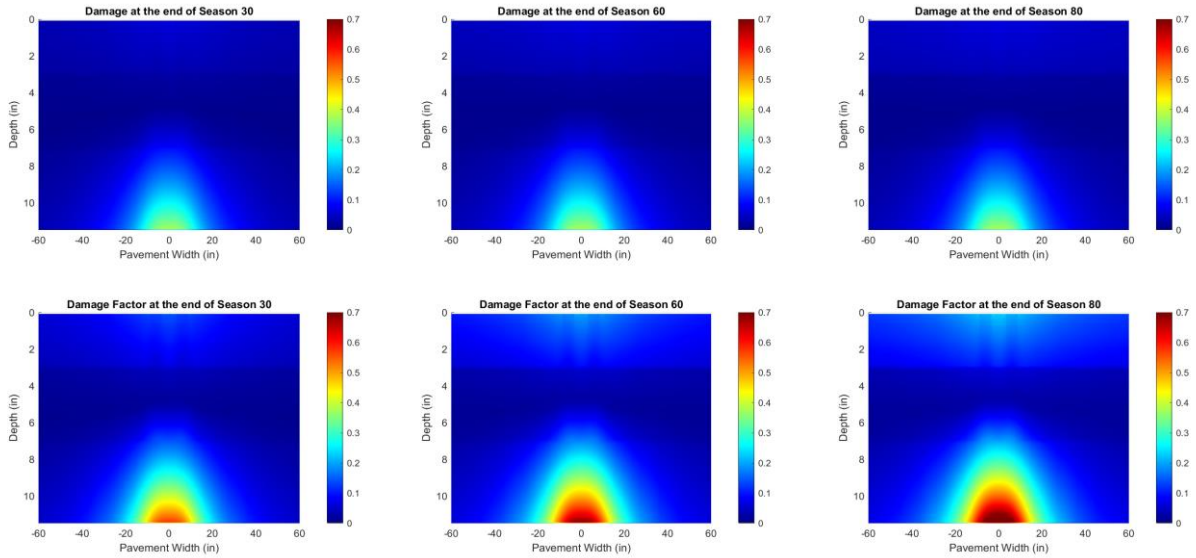

Figure 69: Damage and Damage Factor at the End of Season 30, 60, and 80 for FC33 Corner of RS9.5C Mixture.

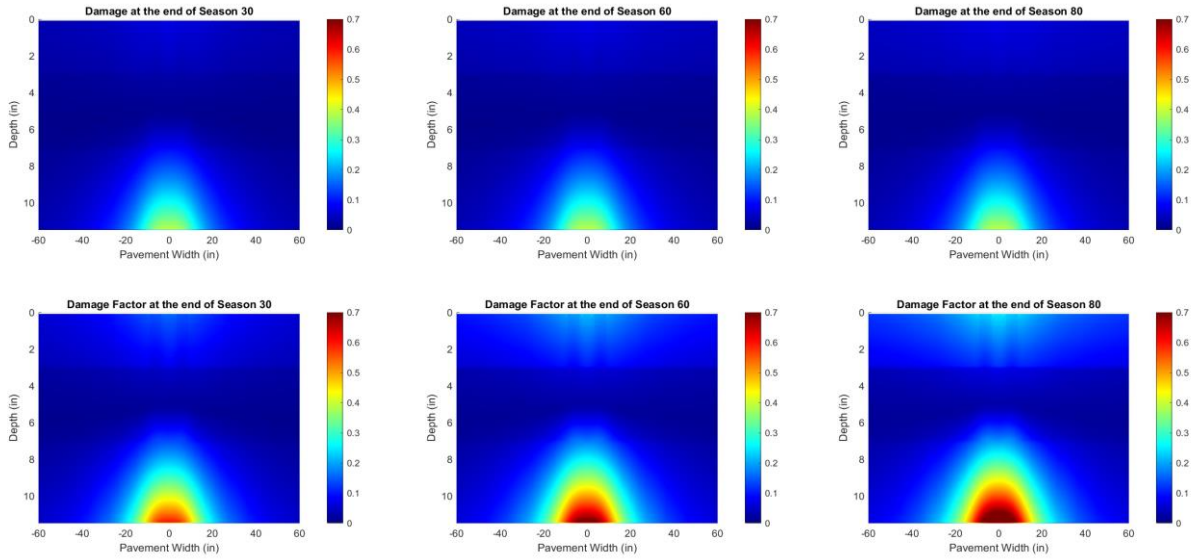

Figure 70: Damage and Damage Factor at the End of Season 30, 60, and 80 for FC57 Corner of RS9.5C Mixture.

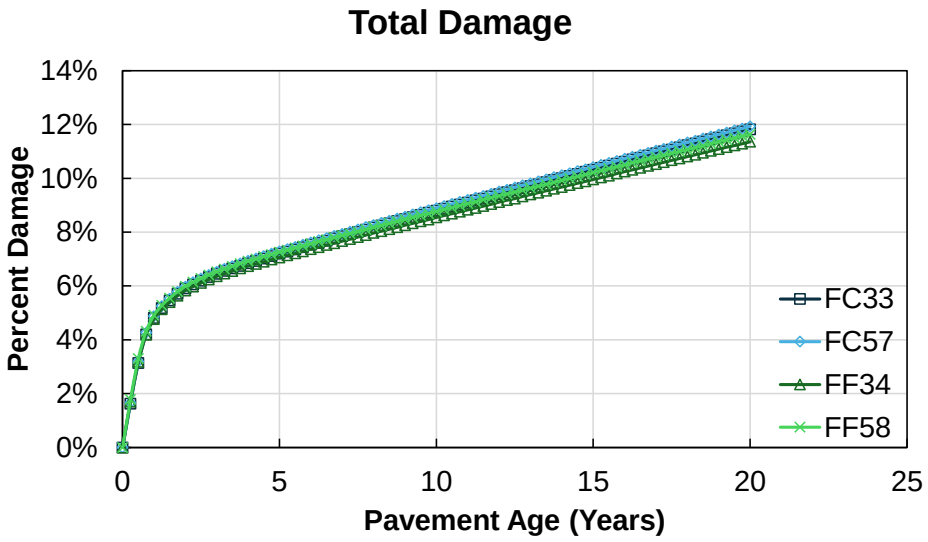

Figure 71: Total Damage Progression for RS9.5C Four Corners.

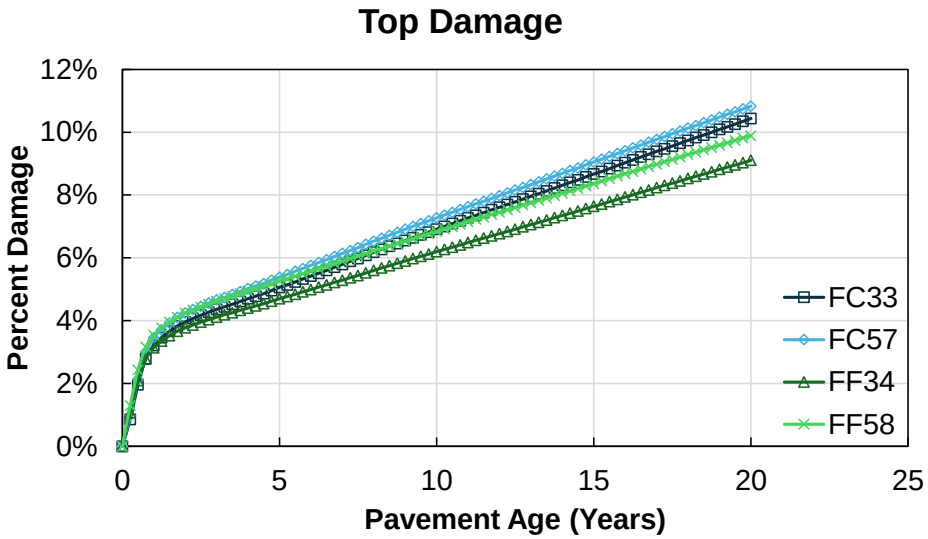

Figure 72: Top Damage Progression for RS9.5C Four Corners.

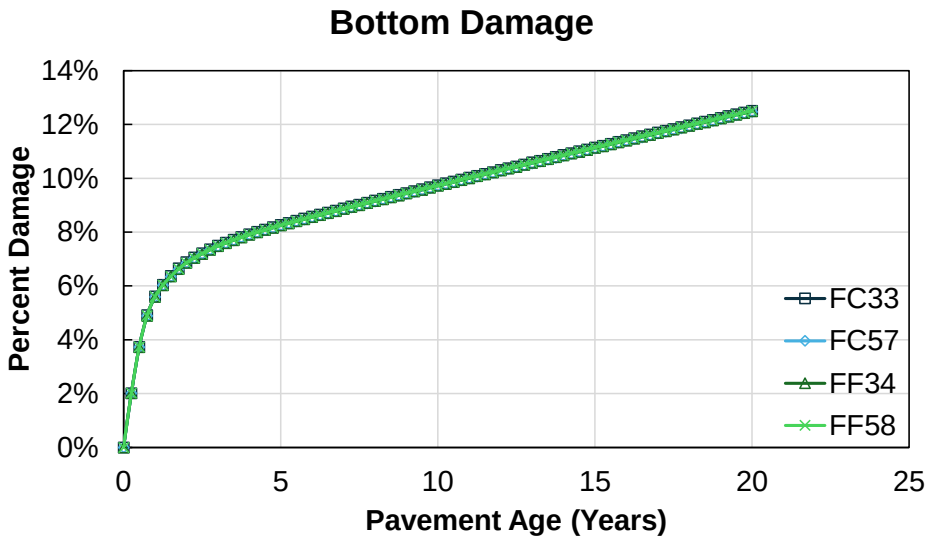

Figure 73: Bottom Damage Progression for RS9.5C Four Corners.

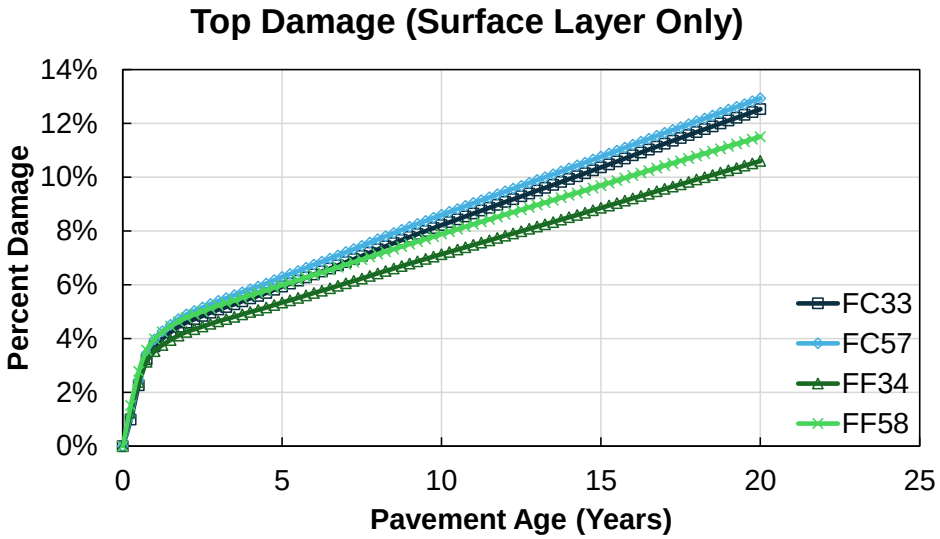

Figure 74: Top Damage Progression for RS9.5C Four Corners (Surface Layer Only).

Once the simulations were conducted, the results were used to calibrate the PVR functions for the total damage and top damage shown in Equations (19) and (20), respectively. The % Damage from the FlexPAVE simulations and the % Damage predicted from the developed PVRs are presented in Table 28 and Table 29 for the total and top damages, respectively. Average % error was 0.5% and 1.6% for the total damage and top damage prediction, respectively, indicating reasonable accuracy of the developed PVR functions.

$$\%\ Damage\ (Total) = -0.076VMA_{IP} - 0.019VFA_{IP} + 14.441 \quad (19)$$

$$\%\ Damage\ (Surface\ Layer\ Only) = -0.346VMA_{IP} - 0.074VFA_{IP} + 23.536 \quad (20)$$

Table 28: RS9.5C Four Corners PVR Function Calibration Results (Total Damage).

Condition	Purpose	*%Damage (Total)*				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF34	Calibration	20.1	78.6	11.36	11.42	0.6
FF58	Calibration	22.2	61.8	11.64	11.58	0.5
FC33	Calibration	15.8	77.9	11.82	11.76	0.5
FC57	Calibration	17.6	59.2	11.93	11.98	0.4
Four Corners % Error Average						0.5

Table 29: RS9.5C Four Corners PVR Function Calibration Results (Top Damage in Surface Layer Only).

Condition	Purpose	*%Top Damage (Surface Layer Only)*				
		VMA_{IP}	VFA_{IP}	Measured	Predicted	% Error
FF34	Calibration	20.1	78.6	10.60	10.81	1.9
FF58	Calibration	22.2	61.8	11.50	11.32	1.6
FC33	Calibration	15.8	77.9	12.53	12.34	1.5
FC57	Calibration	17.6	59.2	12.93	13.10	1.3
Four Corners % Error Average						1.6

CHAPTER 6. EVALUATION OF CONSTRUCTION VARIABILITY AND QUANTIFYING THE IMPACT OF QMS LIMITS ON ASPHALT MIXTURES' PERFORMANCE

Evaluation of construction variability

In CHAPTER 4, the IVR functions have been calibrated and the accuracy of predictions for different performance indexes has been demonstrated. The following step was to use the calibrated IVR functions to predict the performance of the three mixtures based on real QC/QA data obtained through the help of NCDOT personnel. The data obtained for the surface mixture contained asphalt content as well as the in-place density of the obtained mixture samples. These two parameters were used to determine VMA_{IP} and VFA_{IP} and then predict performance. As mentioned in CHAPTER 4, for the CT-Index, the effective binder content was used instead of VMA_{IP} and VFA_{IP}. However, for the intermediate and base mixture, only in-place density data were available to the research team. Consequently, the research team randomly selected values for the binder content based on the mean and standard deviation of the obtained PMLC samples. Then the performance was predicted similarly to the surface mixture.

Figure 75 shows the variability in binder content for the surface mixture. The measured binder content fluctuates slightly around the target value, with occasional small deviations. The overall trend remains stable, and the measured values consistently stay within the upper and lower acceptable limits. The construction variability represented by the in-place density of the same mixture is presented in Figure 76. A moderate fluctuation is observed in the data. Almost all the data points fall within the acceptable range, with the exception of a single data point. Construction variabilities of the intermediate and base mixtures are presented in Figure 77 and Figure 78, respectively. A higher level of variability and number of data points falling outside of the acceptable range can be observed in these figures compared to the surface mixture.

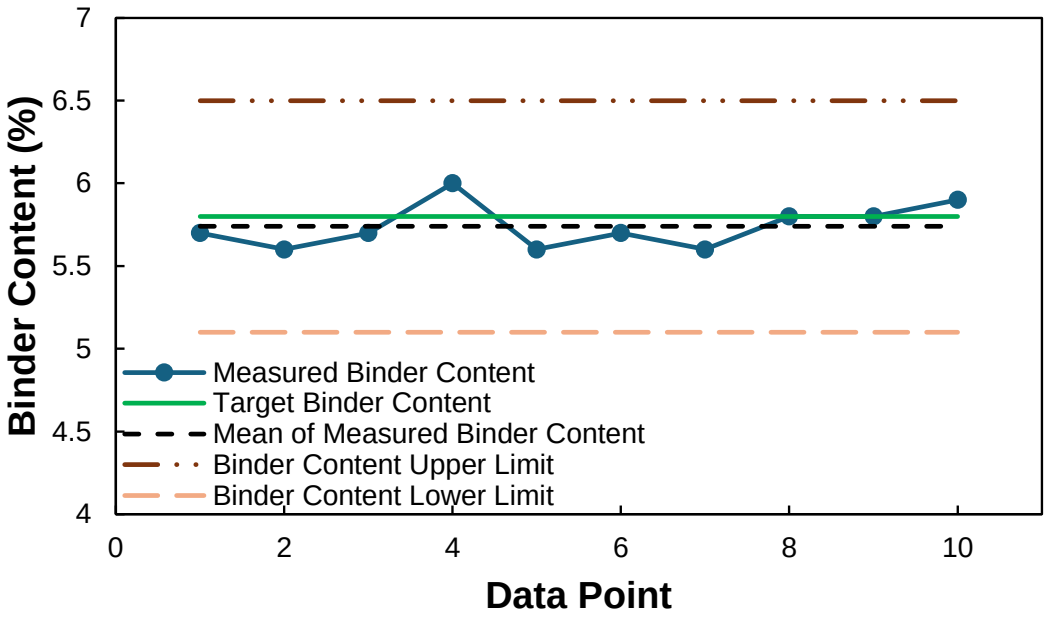

Figure 75: RS9.5C Binder Content Variability.

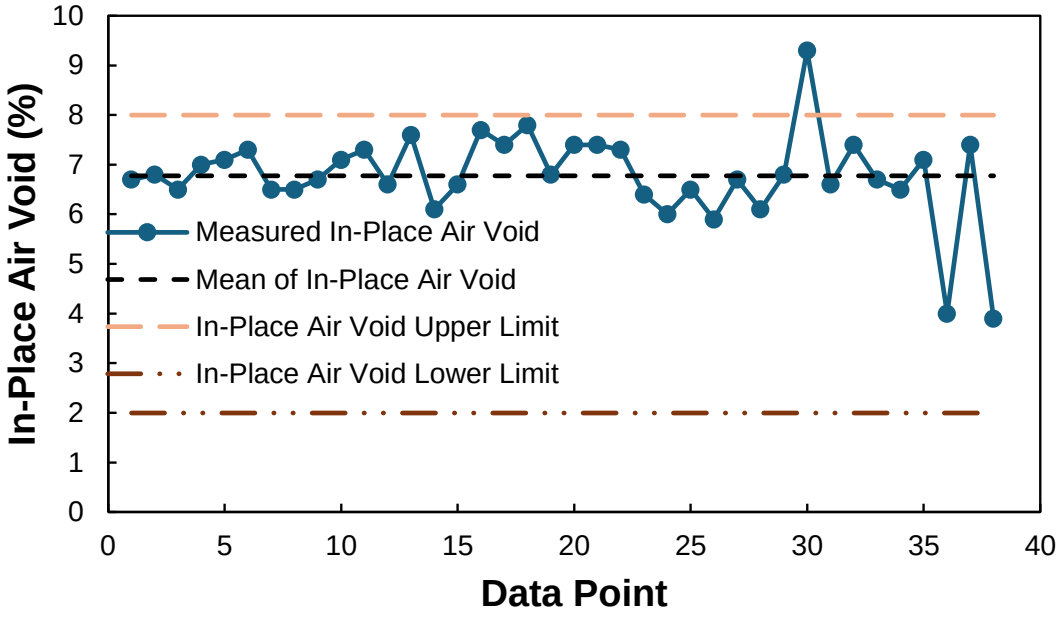

Figure 76: RS9.5C In-Place Air Void Content Variability.

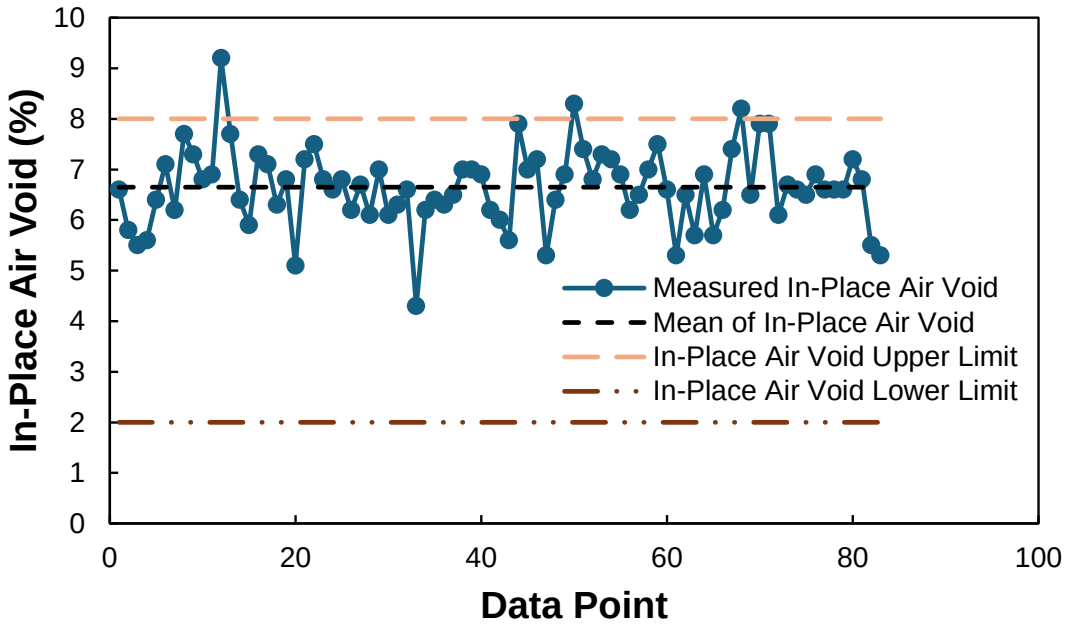

Figure 77: RI19.0C In-Place Air Void Content Variability.

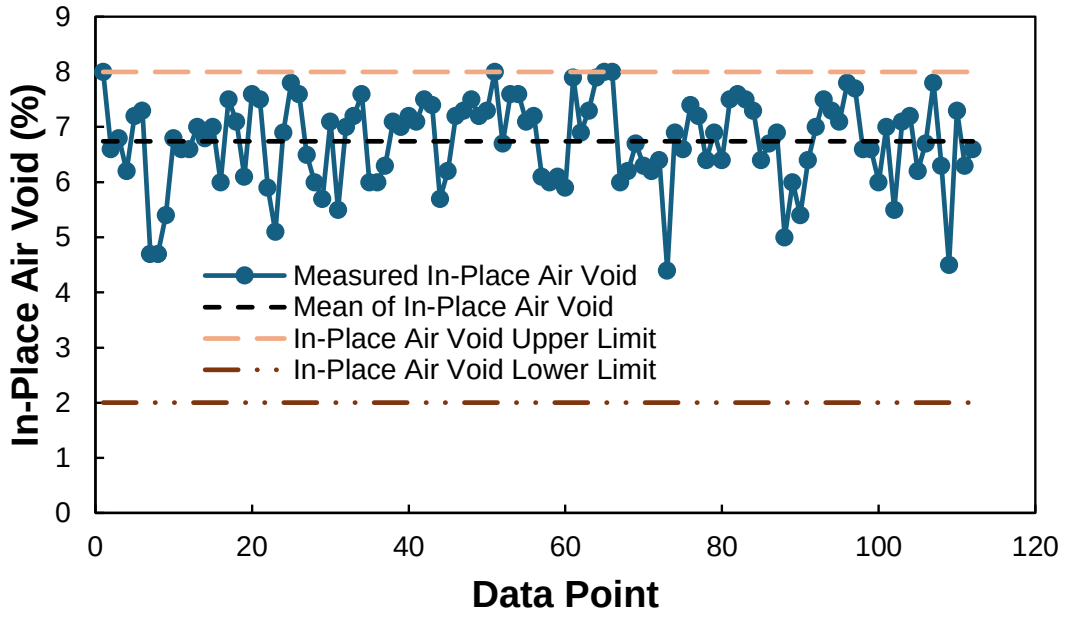

Figure 78: RB25.0C In-Place Air Void Content Variability.

Binder content and in-place air void content data points that had matching dates for the surface mixture were used to determine VMA_{IP} and VFA_{IP}. Then, the calibrated IVR functions were used to predict S_{app} and RSI values. The predicted index values are presented in Figure 79. The predicted S_{app} values ranged from 14.9 to 18.7 with an average of 16.8, while the RSI values

ranged from 8.6% to 12.0% with an average of 10.3%. As mentioned earlier, binder content data was not obtained for the intermediate and base mixture. Consequently, data points for the binder content were randomly selected based on the mean and standard deviation of the PMLC samples by assuming a normal distribution for the binder content data. The S_{app} and RSI values for the intermediate mixture are shown in Figure 80. The predicted S_{app} values ranged from 7.5 to 21.6 with an average of 13.7, and the predicted RSI values ranged from 0.2% to 4.9% with an average of 2.6%. The predicted S_{app} results for the base mixture presented in Figure 81 show that the predicted S_{app} values ranged from 5.9 to 18.4 with an average of 10.0.

Table 30 presents the S_{app} and RSI thresholds for different traffic tiers. For RS9.5C, all data points remained within the Standard traffic tier for both fatigue and rutting performance. For RI19.0C, all data points met the Standard traffic tier based on S_{app} values except for one outlier. However, traffic tier classifications based on RSI values varied, with 10% of data in the Standard tier, 64% in the Heavy tier, 18% in the Very Heavy tier, and 8% in the Extremely Heavy tier. This variation shows that while fatigue resistance remained consistent, rutting resistance varied significantly. For RB25.0C, 23% of data fell below the Standard traffic tier, while the remaining 77% met the Standard classification based on S_{app} values. This observation highlights the influence of production variability on mixture performance, potentially leading to premature fatigue cracking in some cases.

These results emphasize how variations in binder content and air voids during production and construction can significantly impact asphalt mixture performance, shifting its expected traffic tier classification and influencing pavement durability.

Table 30: S_{app} and RSI Thresholds.

Traffic (Million ESALs)	Limits		Tier	Designation
	S_{app}	RSI		
Less than 10	$S_{app} > 8$	RSI < 12	Standard	S
Between 10 and 30	$S_{app} > 24$	RSI < 4	Heavy	H
Greater than 30	$S_{app} > 30$	RSI < 2	Very Heavy	V
Greater than 30 and Slow Traffic	$S_{app} > 36$	RSI < 1	Extremely Heavy	E

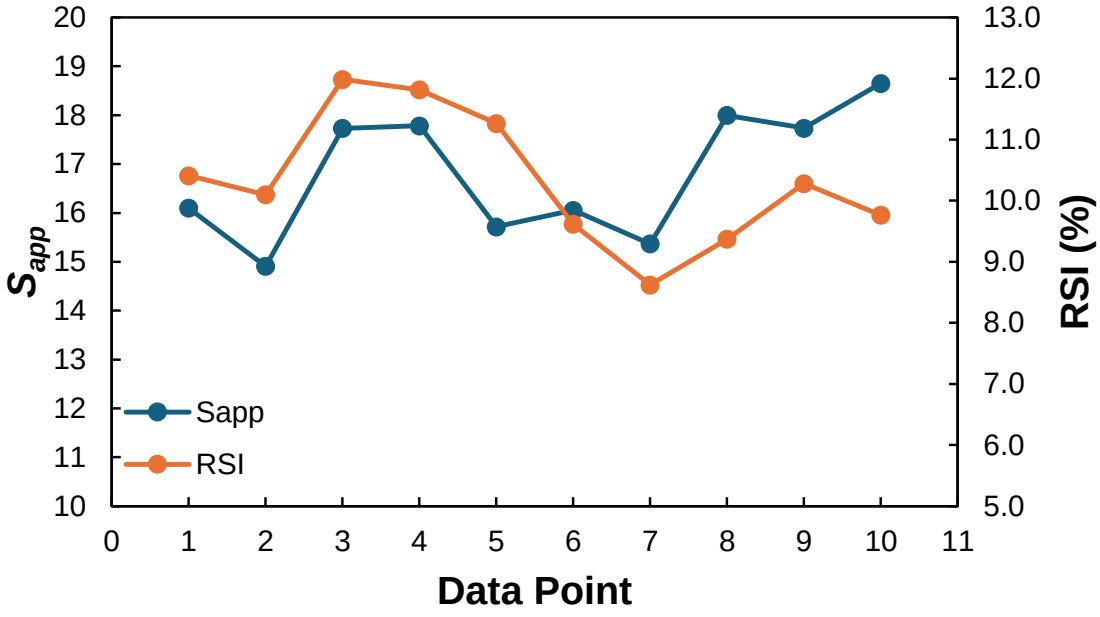

Figure 79: RS9.5C Predicted S_{app} and RSI Values Based on QC/QA Data.

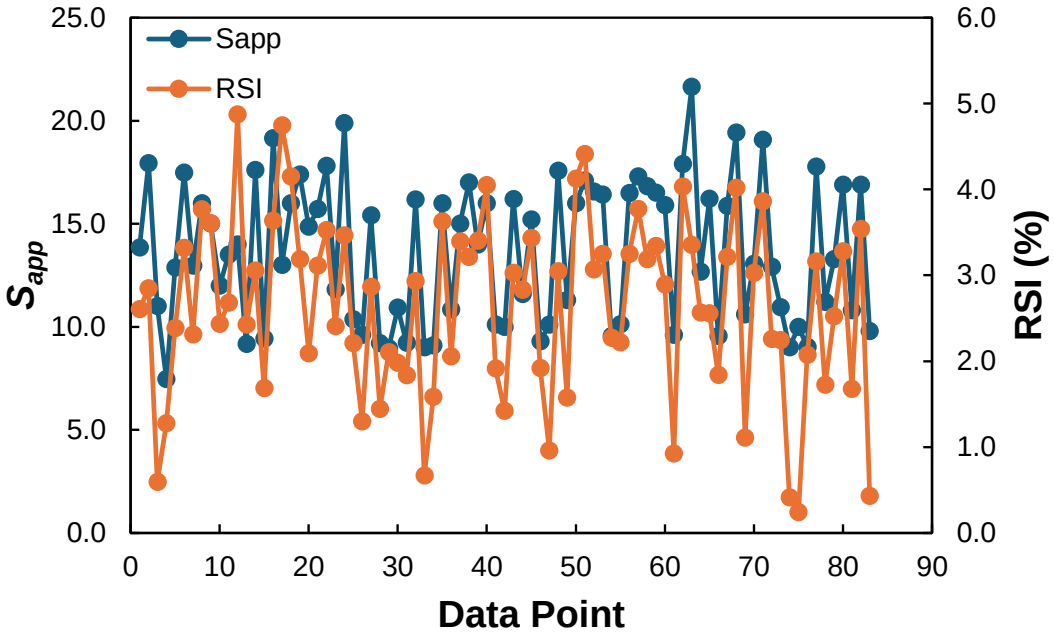

Figure 80: RI19.0C Predicted Sapp and RSI Values Based on QC/QA Data.

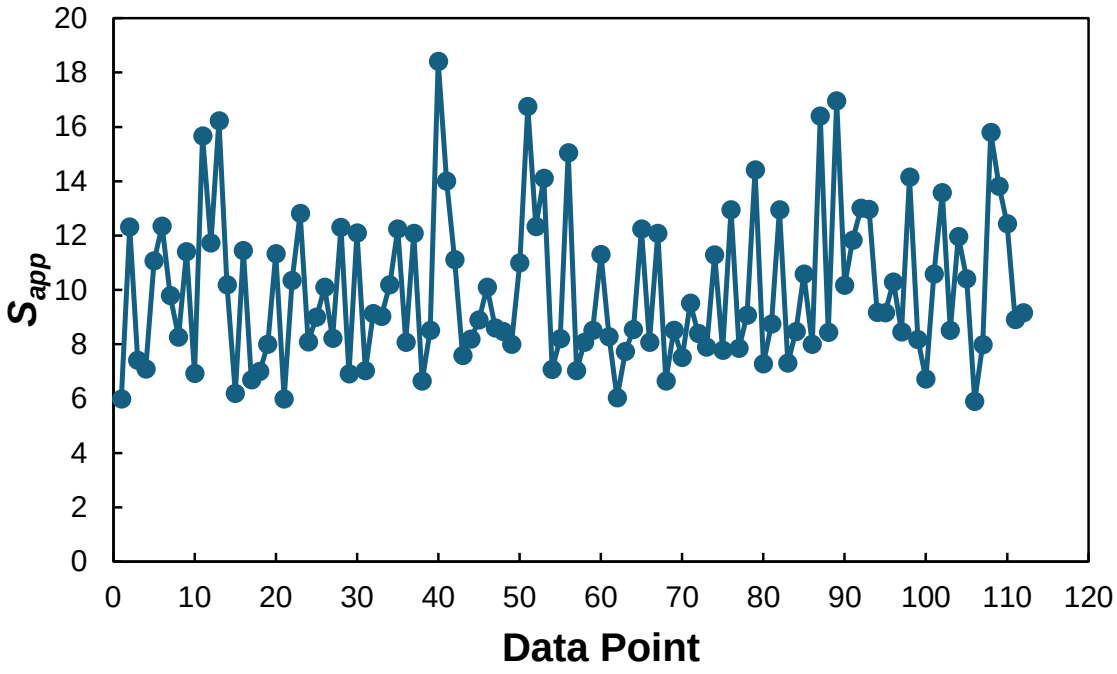

Figure 81: RB25.0C Predicted S_{app} Values Based on QC/QA Data.

Quantifying the impact of QMS limits on performance

As can be noticed from the QC/QA data presented in the previous section, the data does not cover the full range of allowable variability in binder content and in-place density that is provided in the QMS manual. To evaluate the impact of QMS limits on performance, the research team suggests the following framework:

- Step 1: Develop a distribution for binder content using Monte Carlo simulations by assuming a normal distribution. The mean value would be the target binder content for the mixture, and the standard deviation shall be determined such that the resulting distribution covers the full allowable range described in the QMS manual, namely, target binder content for the mixture +/- 0.7%.

- Step 2: Develop a distribution for in-place air void content using Monte Carlo simulations by assuming a normal distribution. The mean value would be the average in-place air void content of the mixture determined from QC/QA data, and the standard deviation shall be determined such that the resulting distribution's lowest value does not exceed the maximum air void content described in the QMS manual, namely, 8%.

- Step 3: Calculate VMA_{IP} and VFA_{IP} based on the developed distributions for the binder content and in-place air void content. It should be noted that for the rutting side, in-place density can be used as is considering that rutting is more significant in the pavement's early life. However, for the fatigue side, it is recommended to first estimate the in-place air void content after two years, since fatigue is not a concern in the first couple years of pavement's life. This estimation was done using a regression equation provided in the FHWA report titled "Hot-Mix Asphalt Performance-Related Specification Based on Viscoelastoplastic Continuum Damage Models". The regression equation correlates constructed air void content with air void content after 2 years. The equation is provided below.

$$\%AV_{2\ years} = 0.7883 \times constructed\ \%AV - 0.9687 \quad (21)$$

- Step 4: Predict S_{app} and RSI values using the calibrated IVR functions as well as the developed distributions in the steps above.

- Step 5: Fit a suitable distribution for the predicted S_{app} and RSI values. The fitted distributions can be used to determine the 95% confidence interval (CI) upper and lower limits (UL and LL) as well as the confidence interval range.

By following the steps described above, the research team quantified the impact of the binder content and air void content limits on fatigue and rutting performance for the three mixtures. The results are presented in Figure 82 to Figure 86. As can be seen from the figures, a normal distribution was suitable to represent the predicted S_{app} values, however, a normal three-mixture distribution was more suitable to represent the predicted RSI values. A normal three-mixture distribution (also called a three-component Gaussian mixture model) is a probabilistic model that represents data as a combination of three distinct normal distributions (Bishop et al., 2006). Table 31 summarizes key statistics for RS9.5C, RI19.0C, and RB25.0C. For RS9.5C, S_{app} has a

mean of 14.5 with a standard deviation of 1.54, and a 95% confidence interval range of 6.0, which represents a +/- 21% variation from the mean. The RSI parameter has a mean of 4.7, a standard deviation of 1.11, and a 95% CI range of 4.4, which represents a +/- 46% variation from the mean. For RI19.0C, S_{app} has a mean of 14.1, a standard deviation of 2.28, and a 95% CI range of 8.9, while RSI has a mean of 1.9, a standard deviation of 0.50, and a 95% CI range of 1.9. The RI19.0C results represent a +/- 32% and +/-51% variation from the S_{app} and RSI means, respectively. Lastly, RB25.0C shows a mean S_{app} of 12.0 with a standard deviation of 1.50 and a 95% CI range of 6.6, while RSI has a mean of 2.0, a standard deviation of 0.68, and a 95% CI range of 6.6, which represents a +/- 28% variation from the mean.

Based on the 95% confidence intervals for S_{app} and RSI across the three mixtures, the analysis shows that the entire 95% CI for S_{app} falls within the Standard traffic tier for all three mixtures. For rutting resistance based on RSI, the 95% CI for RS9.5C spans from Standard to Heavy. Meanwhile, the 95% CI for RI19.0C extends from Heavy to Extremely Heavy. These results highlight how the allowable binder content and air void content limits in QMS specifications influence traffic designation for asphalt mixtures.

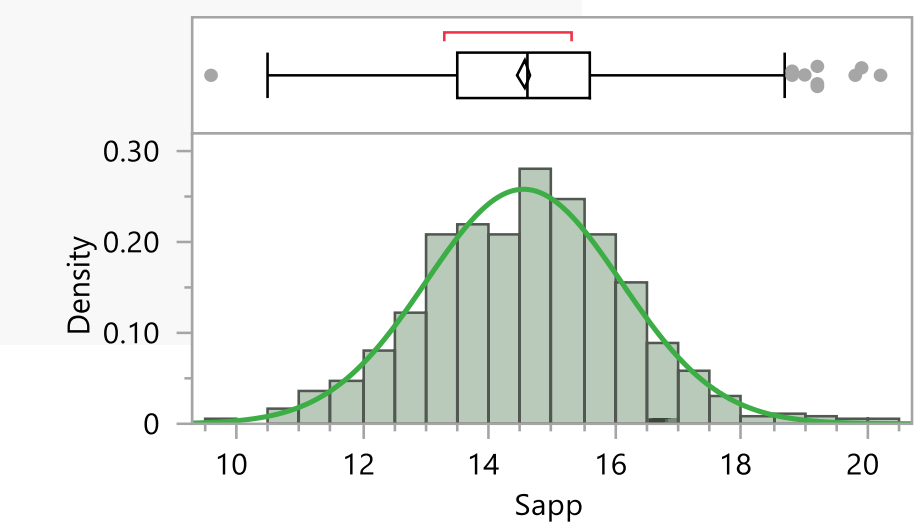

Figure 82: RS9.5C S_{app} Values Distribution Based on the QMS Limits.

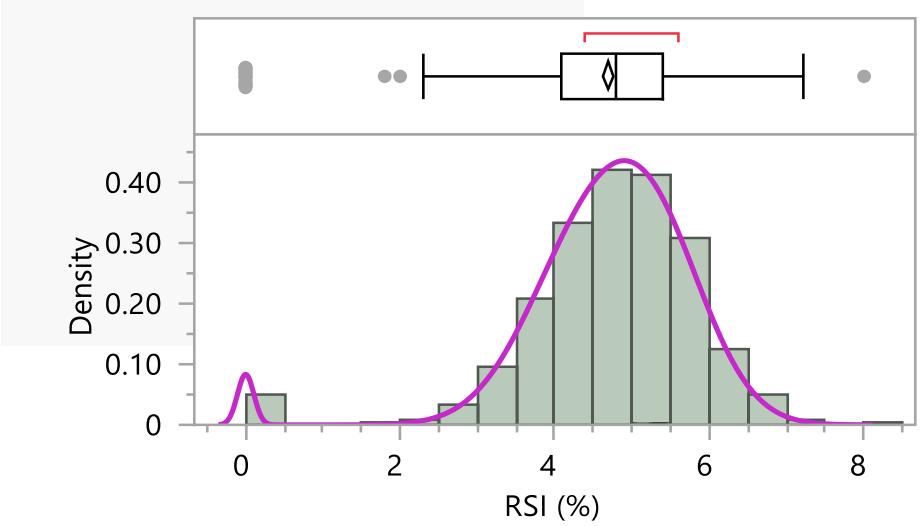

Figure 83: RS9.5C RSI Values Distribution Based on the QMS Limits.

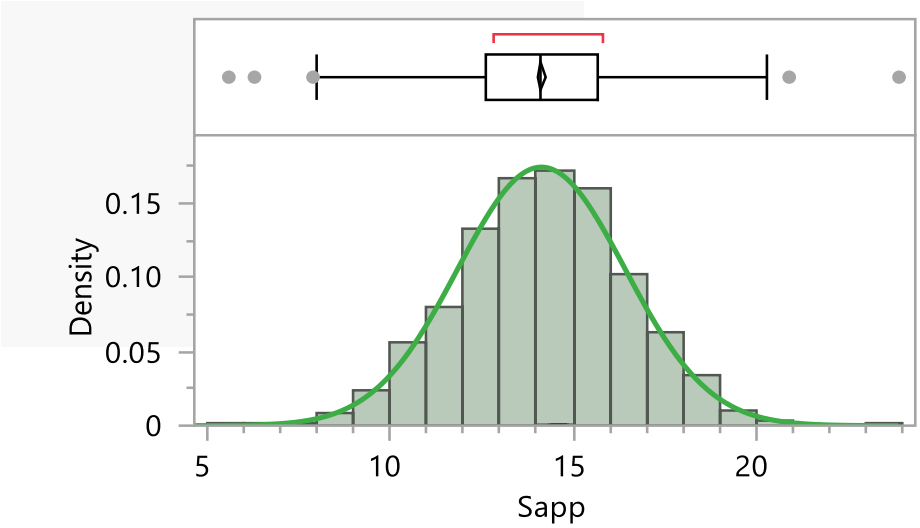

Figure 84: RI19.0C S_{app} Values Distribution Based on the QMS Limits.

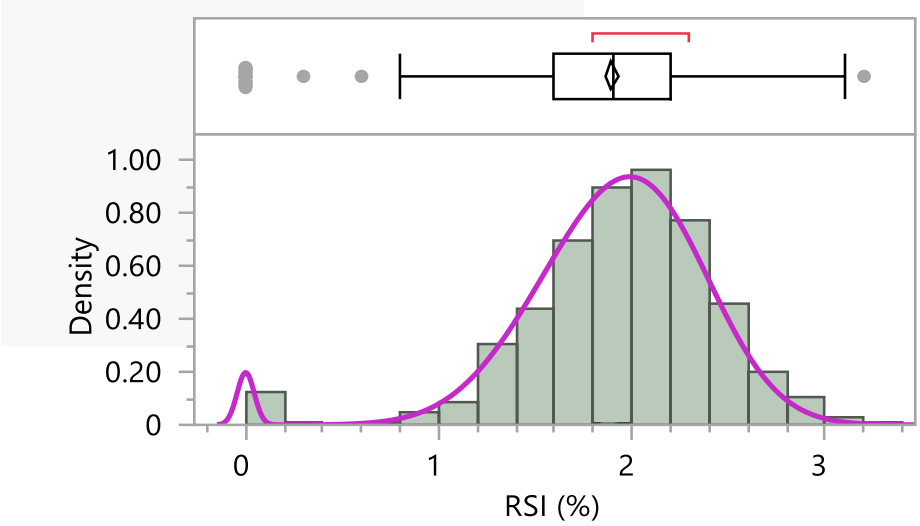

Figure 85: RI19.0C RSI Values Distribution Based on the QMS Limits.

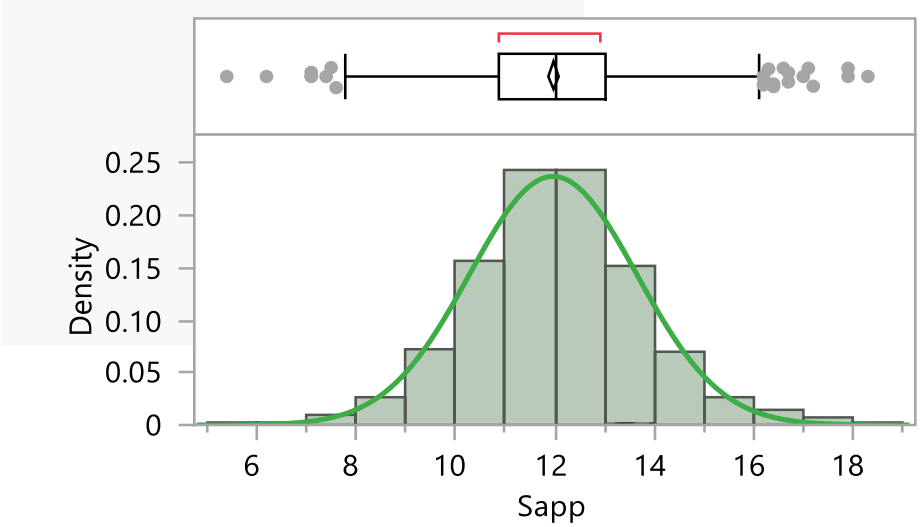

Figure 86: RB25.0C S_{app} Values Distribution Based on the QMS Limits.

Table 31: Key Statistics Summary for the Three Mixtures Based on the QMS Limits.

	RS9.5C		RI19.0C		RB25.0C
	S_{app}	RSI	S_{app}	RSI	S_{app}
Mean	14.5	4.7	14.1	1.9	12.0
Standard Deviation	1.54	1.11	2.28	0.50	1.68
95% CI UL	17.6	6.9	18.6	2.9	15.3
95% CI LL	11.5	2.5	9.7	0.9	8.7
95% CI Range	6.0	4.4	8.9	1.9	6.6

The framework presented above can help in translating the QMS limits into direct impact on fatigue and rutting performance. While three mixtures are not enough to develop recommendations to keep or change the current QMS limits, they provide an example of how this framework can be utilized with a large enough representative sample of North Carolina mixtures to evaluate the impact of QMS limits on performance. Once a more representative sample is used, the results may shed light on how the limits should be changed, ensuring more accurate and performance-based mix design criteria.

CHAPTER 7. CONCLUSIONS AND FUTURE RESEARCH RECOMMENDATIONS

The conclusions that can be drawn from the work presented in this report are as follows.

- The performance test results obtained for the construction samples generally showed good agreement with changes in binder content and air void content. However, the results also demonstrated the complexities involved in predicting performance solely based on AQCs when binder content and air voids are changing simultaneously.

- Short-term aging duration for LMLC samples was investigated, and the results showed that using two hours for short-term aging yielded a generally better agreement to the PMLC sample results compared to four hours. This finding is an important one when the performance of plant mixes is to be predicted using the performance test results from the laboratory-fabricated mixtures.

- The calibrated IVR functions demonstrated reasonable accuracy in predicting fatigue and rutting performance of LMLC and PMLC verification samples.

- The performed FlexPAVE simulations highlighted the importance of optimizing the performance of the intermediate and base mixture along the surface mixture in order to yield a pronounced improvement in the asphalt pavement performance.

- QC/QA data for the three mixtures showed a higher variability in the intermediate and base mixtures data compared to the surface mixture. Considering the importance of the intermediate and base mixtures for the asphalt pavement structure performance, more care should be given to maintaining the consistency of production for these two mixtures.

- S_{app} and RSI values were predicted using QC/QA data, showing that RS9.5C remained within the Standard traffic tier, while RI19.0C met the Standard tier for S_{app} but RSI values varied significantly (10% Standard, 64% Heavy, 18% Very Heavy, 8% Extremely Heavy); for RB25.0C, S_{app} values showed that 23% fell below Standard, highlighting the impact of production variability on mixture performance and potential fatigue cracking risks.

- The research project provided a framework for evaluating the impact of QMS limits for binder content and in-place density on fatigue and rutting performance. The results of the analysis showed +/- 21%, +/- 32%, and +/- 28% change in S_{app} values based on 95% confidence interval ranges for the surface, intermediate, and base mixtures, respectively. The results also showed +/- 46%, and +/- 51% change in RSI values based on 95% confidence interval ranges for the surface, and intermediate mixtures, respectively.

- Based on the QMS limits, the 95% confidence intervals for S_{app} across all three mixtures remained within the Standard traffic tier, while RSI confidence intervals varied, with RS9.5C spanning Standard to Heavy and RI19.0C ranging from Heavy to Extremely Heavy. These results highlight how QMS limits for binder and air void content influence traffic designation for asphalt mixtures.

- The provided framework provides a systematic approach for evaluating the impact of QMS limits on fatigue and rutting performance. The framework can be used with a large enough representative sample of North Carolina standard mixtures to develop recommendations and amends to the current limits in a way that ensures improvement to asphalt mixtures performance and asphalt pavement durability.

The following topics are recommended to be investigated in future research:

- Short-term aging protocol optimization: Investigate short-term aging protocols for North Carolina asphalt mixtures to better replicate the performance of PMLC samples. This study should utilize a representative sample of North Carolina mixtures, including surface, intermediate, and base layers, that account for different traffic levels.

- Evaluation of the impact of the variability in QC/QA data and development of performance-based QMS limits: Analyze the typical variability in QC/QA data for asphalt mixtures for a large sample of contractors and develop performance-based acceptance limits that account for variations in mixture properties and their effect on asphalt mixtures' fatigue and rutting performance. Establish the performance-based QMS limits.

REFERENCES

1. Kim, Y. R., & Jeong, J. (2020). Implementation of Shadow Performance-Related Specifications for an Asphalt Paving Project.

2. American Association of State Highway and Transportation Officials. (2023). *AASHTO TP 132: Standard Method of Test for Determining the Dynamic Modulus for Asphalt Mixtures Using Small Specimens in the Asphalt Mixture Performance Tester (AMPT).* Washington, D.C.: AASHTO.

3. American Association of State Highway and Transportation Officials. (2023). *AASHTO T 400: Standard Method of Test for Determining the Damage Characteristic Curve and Failure Criterion Using the Asphalt Mixture Performance Tester (AMPT) Cyclic Fatigue Test.* Washington, D.C.: AASHTO.

4. Olard, F., & Di Benedetto, H. (2003). General "2S2P1D" model and relation between the linear viscoelastic behaviours of bituminous binders and mixes. *Road materials and pavement design*, *4*(2), 185-224.

5. American Association of State Highway and Transportation Officials. (2023). *AASHTO TP 134: Standard Method of Test for Stress Sweep Rutting (SSR) Test Using the Asphalt Mixture Performance Tester (AMPT).* Washington, D.C.: AASHTO.

6. ASTM International. (2019). *ASTM D8225-19: Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Test.* West Conshohocken, PA: ASTM International.

7. Wang, Y. and Y. R. Kim. (2017). Development of a Pseudo Strain Energy-Based Fatigue Failure Criterion for Asphalt Mixtures. *International Journal of Pavement Engineering*, DOI: 10.1080/10298436.2017.1394100.

8. Ghanbari, A., B. S. Underwood, and Y. R. Kim. (2020). Development of a Rutting Index Parameter Based on the Stress Sweep Rutting Test and Permanent Deformation Shift Model. *International Journal of Pavement Engineering*. DOI: 10.1080/10298436.2020.1748190.

9. ASTM International. (2019). *ASTM D8225-19: Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Test.* West Conshohocken, PA: ASTM International.

10. American Association of State Highway and Transportation Officials. (2019). *AASHTO T 324-19: Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Hot Mix Asphalt (HMA).* Washington, D.C.: AASHTO.

11. AASHTO T 27. (2023). *Sieve Analysis of Fine and Coarse Aggregates*. American Association of State Highway and Transportation Officials. Washington, D.C.

12. AASHTO T 19. (2022). *Standard Method of Test for Bulk Density ("Unit Weight") and Voids in Aggregate*. American Association of State Highway and Transportation Officials. Washington, D.C.

13. AASHTO T 84. (2022). *Standard Method of Test for Specific Gravity and Absorption of Fine Aggregate*. American Association of State Highway and Transportation Officials. Washington, D.C.

14. AASHTO T 85. (2022). *Standard Method of Test for Specific Gravity and Absorption of Coarse Aggregate*. American Association of State Highway and Transportation Officials. Washington, D.C.

15. AASHTO T 308. (2022). Standard Method of Test for Determining the Asphalt Binder Content of Asphalt Mixtures by the Ignition Method. American Association of State Highway and Transportation Officials. Washington, D.C.

16. AASHTO T 228. (2022). *Standard Method of Test for Specific Gravity of Semi-Solid Asphalt Materials*. American Association of State Highway and Transportation Officials. Washington, D.C.

17. ASHTO T 209. (2022). *Standard Method of Test for Theoretical Maximum Specific Gravity and Density of Asphalt Mixtures*. American Association of State Highway and Transportation Officials. Washington, D.C.

18. AASHTO T 319. (2022). *Standard Method of Test for Quantitative Extraction and Recovery of Asphalt Binder from Asphalt Mixtures*. American Association of State Highway and Transportation Officials. Washington, D.C.

19. AASHTO T 100. (2022). *Standard Method of Test for Specific Gravity of Soils*. American Association of State Highway and Transportation Officials. Washington, D.C.

20. Kim, Y. R., Castorena, C., Saleh, N. F., Braswell, E., Elwardany, M., & Rad, F. Y. (2021). *Long-term aging of asphalt mixtures for performance testing and prediction: Phase III results* (No. Project 09-54).

21. Castorena, C., Pape, S., Mocelin, D., Xue, L. G., Aparicio Alvis, M. C., & Ravichandran, M. (2019). *Improving the Design of RAP and RAS Mixtures*. NCDOT Report, 21, 2022.

22. Kim, Y. R., Guddati, M. N., Choi, Y. T., Kim, D., Norouzi, A., Wang, Y. D., ... & Underwood, B. S. (2022). *Hot-mix asphalt performance related specification based on viscoelastoplastic continuum damage (VEPCD) models* (No. FHWA-HRT-21-093). United States. Federal Highway Administration. Office of Research and Technology Services.

23. Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.