DEVELOPMENT OF A PERFORMANCE-RELATED ASPHALT MIX DESIGN SPECIFICATION FOR THE ILLINOIS TOLLWAY

Final Report
January 19, 2021

Principal Investigator
William G. Buttlar, Glen Barton Chair in Flexible Pavements
Civil and Environmental Engineering, University of Missouri

Co-Principal Investigators
Henry Brown, P.E.
Civil and Environmental Engineering, University of Missouri

James Meister, Research Engineer
Civil and Environmental Engineering, University of Missouri

Research Assistants
Behnam Jahangiri (PhD Candidate)
Punyaslok Rath (PhD Candidate)
Hamed Majidifard (PhD Candidate)
Loreto Urra (MS)

Civil and Environmental Engineering, University of Missouri

Authors
William Buttlar, Behnam Jahangiri, Punyaslok Rath, Hamed Majidifard, Loreto Urra, James Meister, Henry Brown

Sponsored by
Illinois State Toll Highway Authority

A report from
The University of Missouri-Columbia
E2509, Lafferre Hall, S Sixth Street,
Columbia, MO 65211.
Phone: 573-882-4375/ Email: muengrcee@missouri.edu

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
MU #		

4. Title and Subtitle	5. Report Date
DEVELOPMENT OF A PERFORMANCE-RELATED ASPHALT MIX DESIGN SPECIFICATION FOR THE ILLINOIS TOLLWAY	March 2020
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
William Buttlar, Behnam Jahangiri, Punyaslok Rath, Hamed Majidifard, Loreto Urra, James Meister, Henry Brown	
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
Department of Civil and Environmental Engineering University of Missouri E2509 Lafferre Hall Columbia, MO 65211	**11. Contract or Grant No.**
12. Sponsoring Organization Name and Address	**13. Type of Report and Period Covered**
Illinois State Toll Highway Authority 2700 Ogden Ave, Downers Grove, IL 60515	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

In the past, traditional asphalt mixtures involved relatively simple combinations of virgin asphalt binder and aggregates to meet performance requirements. However, modern, heterogeneous asphalt mixtures exhibit more complex behavior due to a proliferation of new ingredients and because of the interactions that subsequently occur. As a result, recent asphalt mixes require advanced performance tests to account for the effects of the added components, increases in traffic loads, and the environmental conditions that prevail. In this project, fourteen different mixtures produced in 2018 on mainline and shoulders sections across Tollway system in Chicago were selected to characterize performance testing trends in current Tollway mixtures and to study the ability of the different performance tests to predict pavement performance. To this end, performance tests were performed on the collected plant produced mixtures, and later, on selected field-cored sections. The latter included both good and poor performing sections, which were determined based on an extensive survey of the Illinois Tollway in May of 2019. Evaluation of Tollway asphalt surface pavement management data indicated excellent overall performance vs. time, with minor amounts of several cracking forms developing gradually over time. These included transverse cracking, usually associated with reflective cracking on the mainline and/or thermal and block cracking on shoulders, and longitudinal cracking (typically along the construction joint between lanes). Rutting, on the other hand, was not observed to be a significant form of distress on modern Tollway sections. The Disk-shaped Compact Tension (DC(T)) test was chosen to be retained in the PRS for the design of crack-resistant mixtures due to its high degree of correlation with field results and its best repeatability. A systematic approach was developed, which allowed different reliability levels to be addressed in the specification, along with a consensus step to take advantage of local practitioner experience. Tailored Hamburg rut depth thresholds were established based on lift position relative to the pavement surface. Experimental results led to a new approach where the Hamburg stripping inflection point is used in lieu of the TSR test as the first step in moisture sensitivity verification. If a mix is determined to have stripping potential after analyzing Hamburg results, the TSR test can be then be employed as a final determination of stripping potential. For SMA mixtures with rut depths less than or equal to 4.0 mm, it is recommended that the SIP computation be waived and the mix categorized as non-stripping.

17. Key Words	18. Distribution Statement
Specification, Performance, Lab Testing, Field Data	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	191	NA

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

ACKNOWLEDGMENTS

The authors would like to thank the Illinois State Toll Highway Authority for their assistance during the various stages of this research investigation. The authors would also like to thank our colleagues from S.T.A.T.E. Testing, LLC in Chicago for their assistance throughout the study, particularly with field sampling and for binder extraction, recovery, and testing of field-sampled materials. The authors would also like to express their gratitude to Applied Research Associates of Champaign, IL, for providing extensive pavement condition rating and performance data. We would also like to thank the entire Technical Review Panel (TRP) for their insightful comments and kind guidance provided throughout this project. A list of TRP members is provided in Chapter 1.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Traditional asphalt mixtures have generally involved relatively simple combinations of virgin asphalt binder and aggregates to meet load-bearing needs of the roads and surfaces. Accordingly, simple tests such as Marshall Stability and Flow were used in an effective manner for asphalt mixture screening and quality control purposes. In recent years, there has been a proliferation of asphalt ingredients available to designers, especially in the case of recycled materials, compaction aides, and mixture performance and/or sustainability promoting products. These include reclaimed asphalt pavement (RAP), recycled asphalt shingles (RAS), warm mix agents, antistripping agents, rejuvenators, ground tire rubber, and even waste plastic. These modern, heterogeneous asphalt mixtures exhibit more complex behavior as compared to earlier mixes containing fewer ingredients and predominantly virgin materials. As a result, recent asphalt mixes require advanced performance tests to account for these complexities, while factoring in traffic and environmental loads for the given mixture type being designed.

According to recent literature, mixture performance can be evaluated using various tests to mitigate different distress types such as cracking, rutting, and moisture damage. In this project, fourteen different mixtures produced in 2018 on mainline and shoulders sections across the Tollway system were selected to characterize performance testing trends in current Tollway mixtures and to study the ability of the different performance tests to predict pavement performance. To this end, performance tests such as the DC(T), I-FIT, IDEAL-CT, IDT, Hamburg, and TSR were conducted on the collected plant produced mixtures. The process of sample fabrication, ease of conditioning and testing, repeatability and ability to correctly rank various Tollway mix types was taken into consideration in selecting the appropriate performance tests to be used in the Tollway's mix design asphalt specification. The DC(T) test was found to possess the best correlation to field performance, and significantly outperformed the I-FIT test in terms of test repeatability. Both the I-FIT and IDEAL tests returned failing results for a number of SMA mixes, and dense-graded mixes with high recycling content, which have traditionally performed well on the Tollway. This provided additional motivation to retain the DC(T) test as the cracking test to be used in the Tollway's asphalt mixture design specification.

In addition to performance testing of plant-produced mixtures in 2018, various existing roads including good and bad performing sections were selected after a site visit in May of 2019. In this field investigation, the main distresses on the Tollway were identified. Also, several field cores were obtained from mainline and shoulder sections to evaluate the laboratory performance of existing asphalt mixtures across a range in-situ aging levels. Analyzing the available field performance data such as international roughness index (IRI), condition rating system (CRS), and rut depths and comparing them with laboratory testing results provided a robust data set to establish updated performance test thresholds for the Tollway mixture design specification.

The analysis presented in this study, in conjunction with field observations, led to the identification of various cracking types as the primary distresses observed on Tollway mainline and shoulder sections surfaced with asphalt. Rutting and stripping were not found on Tollway asphalt surfaces at the present time. The Disk-shaped Compact Tension (DC(T)) test was chosen to be retained in the performance related specification (PRS) for the design of crack-resistant mixtures due to its high degree of correlation with field results and best repeatability. The DC(T) fracture energy

values for different types of SMA and dense-graded mixtures included in this study were computed after taking field aging, production and test variability into account. A systematic approach was developed, which allowed different reliability levels to be addressed in the specification, along with a consensus step to take advantage of local practitioner experience. Similarly, for high-temperature performance, Hamburg thresholds for binder course mixtures were tailored for different mixture types and use cases. In some cases, by relaxing Hamburg requirements, designers have more leeway in building crack resistance into the mixture and/or to utilize higher amounts of recycled materials. Because stripping appears to be absent under the current Tollway mix design specification, only minor changes were recommended in the Hamburg stripping inflection point thresholds. For SMA mixtures, it was observed that low-rut-depth mixes were sometimes identified as having stripping potential in the Iowa method. However, similar mixtures have not exhibited stripping in the field. As a result, it is recommended that SMA mixtures with rut depths less than or equal to 4.0 mm after 20,000 passes should be characterized as non-stripping and do not need to be assessed for stripping potential using the Iowa method. Based on experimental results, it is also recommended to use the Hamburg test in lieu of the TSR stripping test for moisture sensitivity evaluation. In the event of failing results, the TSR test can be used as a secondary method to assess adequate moisture resistance.

Chapter 1

INTRODUCTION

1.1. Overview

Some eighty years ago, the Marshall stability and flow test and Hveem stabilometer and cohesiometer devices were developed to supplement asphalt binder purchase specifications and volumetrics-based mix design methods by providing 'tests on the mix.' In both cases, tests were developed to provide bookends on high and low temperature asphalt pavement performance, i.e., rutting and durability/cracking. These were necessarily very simple, empirical tests run at room temperature or higher, as it was difficult to test in the low in-service temperature range in that era, or to reliably measure fundamental material properties. In the late 80's and early 90's, the Strategic Highway Research Program (SHRP) undertook an ambitious program to radically improve asphalt binder purchase specifications, aggregate requirements, mixture compaction, and performance-based mixture tests with associated models. The SHRP program created 'Superpave' products such as the PG Binder specification, new collections of aggregate consensus and source property tests, a new standardized asphalt gyratory compactor, and provided minor changes and national standardization of mixture volumetric design principles and use of the AASHTO T-283 tensile strength ratio test to evaluate moisture damage.

After much painstaking debate early in SHRP, it was agreed that fundamental tests were the key to moving forward from the shortcomings of past, empirical, strength-of-materials, and torture test approaches. For instance, one cannot use binder penetration (PEN) values in a finite element model to relate binder properties to low-temperature cracking. On the other hand, creep stiffness from the bending beam rheometer is a fundamental measure and has been used in mixture and pavement models to develop Superpave binder specification limits. However, in the end, the advanced mixture tests and models that were developed and used to calibrate and validate the Superpave PG binder specification were deemed to be too complicated to serve as replacements for the Marshall stability and flow test or the Hveem mixture tests. For instance, two of the tests: the Superpave Shear Tester and Superpave IDT, were far too expensive, cumbersome, time consuming and variable to be practically used in mixture design.

More than 25 years have passed since the completion of the SHRP program, and although the asphalt community has produced PG-plus binder tests and has conducted extensive research to investigate simpler mixture performance tests (such as tests developed in NCHRP 1-37A for use with the M-E Pavement Design Guide), a national consensus on mixture performance tests to supplement volumetric mixture design still does not exist. In terms of permanent deformation, a few tests have gradually been adopted by agencies for routine mix design of heavy volume roadways, including the asphalt pavement analyzer, Hamburg wheel tracking test, and flow number test. Advances in applying fracture mechanics principles to asphalt concrete mixtures have led to the development of new, robust fracture tests, which show promise as tools that can accompany volumetric mix design to limit pavement cracking. Adding to this scene are a number of newly developed or re-vamped empirical mix cracking tests. Thus, there are a number of fundamental and empirical cracking tests being proposed at this time, which include the disk-shaped compact tension test (ASTM D7313), the Texas overlay test, several tests stemming from the semi-circular bend test geometry (fracture toughness, fracture energy, and flexibility index), and several revamped IDT-based tests, such as the NCAT IDT N_{flex} factor test. A brief review of the current literature is included in Appendix A. While a few agencies have investigated performance specifications for asphalt mixtures, none of them have led to the establishment of a rigorously calibrated performance-related specification (NCHRP 492 Synthesis). The Illinois Tollway was one of the first agencies to investigate, and to adopt, modern performance-related specifications for asphalt pavements. The Tollway now seeks to validate and to evolve their performance-based asphalt specification.

While the rutting test debate has settled down with the successful implementation of the Hamburg wheel tracking test in many states, a number of interconnected factors have undoubtedly clouded the national debate on mixture cracking performance tests. These complicating factors include:

1) the existence of many cracking forms, including single event low-temperature (thermal) cracking, thermal fatigue cracking, block cracking, top-down cracking, bottom-up (traditional) fatigue cracking, reflective cracking (which is further clouded by the important role of the underlying pavement in the development of this distress), and interface debonding and cracking;

2) the differing mechanisms behind each of these cracking forms;

3) the differing pavement configurations, climatic conditions, and traffic conditions nationwide, leading to differences in cracking types observed and subsequent selection of cracking performance tests;

4) the moving target resulting from constant changes in binder refining, recycled material types and amounts, use of additives such as warm-mix, antistrip, rejuvenators, and REOB, and;

5) adjustments to Superpave mix design volumetrics that have been proposed to address higher recycling amounts, such as lowering the target air voids, use of a binder availability factor for recycled asphalt shingles, and/or raising the required VMA or volume of effective binder for mixtures with higher binder replacement.

With the Tollway's emphasis on constructing and maintaining high volume expressways, high performing asphalt mixtures are needed to ensure durable, long-lasting pavements. The use of validated and improved performance-related specifications (PRS) could also lead to lower maintenance needs, lower mix costs, and reduced user delays. Asphalt mixture test methods in a practitioner-friendly PRS should be repeatable, straightforward, commercially available and sufficiently standardized. They should also reliably control the most critical distresses identified by the owner (Tollway). The types of critical distresses to be controlled may differ when developing PRS limits for surface mixes, binder course mixes, and shoulder mixes. In theory, a benefit arising from adoption of performance-related mixture specifications is the ability to provide additional flexibility to the mix designer by relaxing or removing any over-constrained method-based requirements (gradation bands, dust-to-asphalt ratio range). This should be simultaneously investigated to shorten the PRS development and implementation cycle, saving time and money.

On the basis of these needs, this research investigation involved a comprehensive literature review, extensive laboratory and field investigations, frequent consultation with Tollway and area practitioners, discussions with related agencies, analysis and ranking of tests based on numerous metrics, and the evaluation, evolution, and consensus-based tuning of the Tollway asphalt mixture PRS for a range of mixture types. Position in the pavement; i.e., surface vs. non-surface, and mainline vs. shoulder, was also considered.

1.2. Research Questions

The key research questions posed in the research study included:

- What are the most critical flexible pavement or asphalt overlay distresses to be controlled by a new performance-related mixture specification to be used by the Tollway?
- Based on the distresses to be controlled, which mixture performance tests and associated limits or ranges can be used to most reliably and effectively control these distresses over the design life in a cost-effective and practical manner?
- Has the Hamburg device successfully eliminated the need for AASHTO T-283 (TSR)?
- Which test devices have the versatility to be adapted and used to obtain other fundamental properties and performance measures for research/forensic purposes, and possibly for the development of future performance measures as materials, recycling, and pavement practices evolve?
- Can all of the critical cracking forms be controlled by a single test, or are multiple tests needed?
- If only one cracking test can be employed from a practical standpoint, should it be at intermediate temperature or low temperature?
- If a low temperature test is selected, would cracking at intermediate temperatures also be sufficiently controlled by the test, or would changes in mixture volumetrics be needed to fill the gap?
- If an intermediate temperature test is selected, how would low temperature cracking be controlled, such as by testing of the recovered binder for mixtures, particularly when higher recycling levels are used?
- Finally, can any mixture volumetric criteria be eliminated or relaxed in light of the mixture performance tests? In theory, the use of mixture performance tests should open the door for mix design flexibility and innovation, which might mean relaxing criteria such as dust proportion requirements, VTM target range, and VFA. To be truly performance-based, method-based specification aspects should in theory be eliminated. Practically, method-based specification aspects should be minimized and/or relaxed as appropriate.

The key tasks identified to address these research questions were:

- A comprehensive review of the literature and relevant project reports and a survey to identify the state-of-the-art for asphalt mixture performance-related testing methods, and recent adjustments to Superpave mixture volumetric targets aimed at ensuring sufficient binder content for mixture durability.

- Evaluating available performance tests used to predict or measure the resistance of asphalt mixtures to distresses such as rutting, low-temperature cracking, fatigue cracking, and moisture damage.
- Evaluation and calibration of the Tollway's performance-related asphalt mix design specification, including evaluation of candidate cracking tests and adjustment of specification limits for Tollway mix designs based on the testing of samples obtained from new and existing Tollway pavements, and by collecting and analyzing field performance data.
- Updating the Tollway with quarterly reports and frequent meetings of the technical review panel.
- Development of a comprehensive final report, to include calibrated updates to the performance-related specification used for the design of Illinois Tollway asphalt mixtures.

1.3. Research Approach/Detailed Work Plan

After consultation with the Technical Review Panel (TRP), a detailed research plan was developed, as follows.

1.3.1. Task 1: Literature Review

A thorough review of the available literature and ongoing project documentation was carried out by the research team to determine the current state-of-the-practice for asphalt mixtures testing and to identify those asphalt mixture performance tests that should be included in a performance-related specification for asphalt mixtures. Appendix A provides a preview of the available current literature in the area of asphalt mixture performance tests and specifications. Some of the subtasks performed within Task 1 included:

- A comprehensive search of journal papers, conference papers, technical reports, theses, agency and industry websites was conducted.
- Development and deployment of a brief, targeted survey administered to state highway, toll authorities and other related agencies in the US was carried out. The survey provided an updated snapshot of the asphalt tests and specifications being used, developed or considered by relate agencies, along with information regarding the objectives of the PRS, including: which distresses are being addressed, and how; what test procedures and limits are being used; what recent changes in Superpave mixture volumetric design have been used, and why; how the PRS is being developed and validated, implications of added cost for mixture design and asphalt bid prices and expected life extension and overall life cycle savings, and; lessons learned during PRS development and implementation.
- A compilation of the draft and final literature search and survey synopses, after synthesizing, analyzing, interpreting, and organizing the findings was produced.
- A meeting was held with the TRP to review findings and to discuss the approach for developing the first round of revised PRS recommendations. The desired scope of the study

moving forward was also discussed, for instance, to determine how much relative emphasis to place on SMA surface mixes, and dense-graded mixes for lower lifts and shoulders. It was agreed that a robust PRS should be equally applicable to SMA and dense-graded mixes, although volumetric and performance test limits would obviously differ between these mixes. This is due to differences in their composition, their usage (riding surface vs. structural or shoulder use), the key distresses to be controlled, the types and levels of recycled materials to be incorporated, and to differences in desired mixture economics.

Task 1 deliverables: Literature Review (provided in Appendix A).

1.3.2.Task 2: Draft Revised PRS

A draft, revised performance-related specification for Tollway SMA and dense-graded asphalt mixtures, complete with sampling, testing and mix volumetric target and range recommendations for implementation on near-term Tollway rehabilitation and reconstruction projects was to be developed in this task. The subtasks conducted and factors considered in the development of the draft revised specification included:

- Review of past experiences and availability of performance test results from previously designed and constructed Tollway asphalt mixture sections, and information garnered from the literature review and survey.
- Review of key distresses to be controlled in the PRS, including rutting, single event low-temperature cracking, thermal fatigue cracking, block cracking, fatigue cracking (top-down and traditional bottom up), reflective cracking, raveling, moisture damage, and interface debonding. Obviously, it was not anticipated that a practical, first-generation PRS will be able to directly address all of these common asphalt distress types. However, the intentional prioritization and strategic selection of a suite of tests and associated limits based on the identified priorities was targeted to arrive at an effective and efficient first-generation PRS. Directly or indirectly, the performance tests and mixture volumetric design changes proposed were to be chosen to mitigate or significantly deter many of the listed distresses.
- A narrowly-focused and aggressively-scheduled laboratory study was performed in the first half of the study, designed to answer some of the critical research questions, namely: the number and type of mixture cracking tests needed in the PRS; if and how adjustments to Superpave mixture volumetrics can be used along with a single selected mixture cracking performance test to effectively control multiple forms of pavement cracking, and; if low temperature binder tests on the recovered binder are needed as part of mix design in light of the selected mixture cracking performance test and adjusted volumetric targets. Tests performed (hereafter referred to as ‘Performance Test Suite’), after adjustments from Task 1 and input from the TRP included:
    - Hamburg wheel tracking (AASHTO T324)
    - Disk-shaped compact tension test (ASTM D7313-07)
    - IL-SCB, or ‘I-FIT’ (ITP-405)
    - Extraction and recovery and Superpave binder testing, including MSCR testing (AASHTO T350, M320) (presented in Appendix B)
    - Mixture volumetrics and TSR (AASHTO T166, T209, T283)
- Results were reviewed in conjunction with the Technical Review Panel (TRP), and a draft PRS was developed and fine-tuned through consensus with the TRP.

Task 2 deliverables: Recommended tests for use in the performance-related asphalt mixture design specification. The adjustment of test thresholds in the PRS was deferred until field performance results and statistical analysis of data from Task 3 was completed.

1.3.3. Task 3: Project Shadowing with PRS

Laboratory testing was performed on a large number of samples obtained from Tollway projects to ensure that the performance testing suite was suitably vetted for modern, heterogeneous recycled Tollway mixtures. Parallel testing on lab-produced specimens that relate to the corresponding field mix designs was also conducted. These activities included:

- Working closely with the mixture designer on Tollway asphalt projects to develop several candidate mix designs for selected mixtures on selected projects, to generate more laboratory data to support finalization of the PRS, and to open the door for the construction of long-term monitoring test sections (side-by-side comparisons of mixes designed to meet the candidate PRS, but with significantly different design approaches).

- Analyzing data to address any unanswered research questions, and to develop more data on inter-laboratory repeatability of the proposed performance tests. For instance, lab-aging to validate and/or calibrate the cracking performance test thresholds specified for short-term aged specimens, while providing the ability to control mixture performance to long-term aging levels experienced in the field.

- Documenting designs, field trials, and associated laboratory tests to maximize what will be learned from the sections developed with the new PRS as they perform under traffic and environmental loading in the years that follow. For instance, new machine learning tools show great promise in connecting the dots between mixture ingredients, predicted performance test results and predicted field performance. However, they require fairly extensive amounts of well-documented laboratory and field study data to ensure high accuracy.

- Analysis of results and in-depth discussions with the TRP were carried out to finalize the suggested changes to the Tollway PRS.

Task 3 deliverables: Documentation of mix designs, observed construction activities and other details of the selected field sections, and laboratory test results on the as-produced materials and any parallel, laboratory-prepared mix testing. A revised PRS, based on field data and consensus with the TRP.

1.3.4. Task 4: Final Report

Task 4 involved the compilation of this comprehensive final report.

1.4. Meetings with TRP

One of the distinctive features of this project was the close coordination between a diverse array of technical experts in academia (University of Missouri), agencies (Tollway, IDOT) and industry (including STATE Testing, LLC, local contractors, the Illinois Asphalt Pavement

Association, Seneca Petroleum, and Vulcan Industries). Several meetings with the TRP or ad-hoc TRP subcommittees were held, involving research progress reports, data review sessions, and field sampling and specification discussions.

Table 1-1 lists the meetings held during the conduct of this study. The TRP members, along with their affiliations, are provided in Table 1-2.

Table 1-1. Summary of technical meetings

No.	Date	Location	Participants	Description
1	10/4/2017	Lisle, IL	TRP-MU	Literature review
2	6/26/2018	Lisle, IL	TRP-MU	Updated literature review and sampling plan
3	11/5/2018	Lisle, IL	TRP-MU	Sampled mixtures and testing results for SMAs
4	2/15/2019	Zoom Call	Sub TRP-MU	Update on performance testing results- Comparing cracking tests results
5	3/7/2019	Lisle, IL	TRP-MU	Stripping performance tests. Performance-space diagram (DCT-Hamburg plot)
6	5/30/2019	Lisle, IL	Sub TRP-MU	Organizing the field visit - Field coring plan
7	8/5/2019	Lisle, IL	TRP-MU	Performance modeling results-Creep testing results - Picking up field cores
8	11/5/2019	Zoom Call	Sub TRP-MU	Field core testing results - Field performance data analysis - Spec development
9	12/9/2019	Lisle, IL	TRP-MU	Performance thresholds-Spec draft-Consensus

Table 1-2. Names/current affiliations of TRP members

Name		Company/ Organization
1.	Jay Behnke	S.T.A.T.E. Testing LLC
2.	Ross Bentsen	Quigg Engineering Inc.
3.	Kevin Burke	Ill. Asphalt Pavement Assn
4.	Richard Duval	FHWA
5.	Jay Gabrielson	Vulcan Materials
6.	Dan Gallagher	Gallagher Asphalt
7.	Dan Gancarz	Applied Research Associates
8.	Brian Hill	Illinois DOT
9.	Stephen Jones	Illinois DOT
10.	Steve Kennedy	Rock Road Companies
11.	John Lavallee	S.T.A.T.E. Testing LLC
12.	Jeff Kern	Open Road Paving, LLC
13.	Alicia Pitlik	Illinois Tollway
14.	Mike Schilke	Illinois DOT
15.	Don Sjogren	Seneca Petroleum
16.	William Vavrik	Applied Research Associates
17.	Cindy Williams	Illinois Tollway
18.	Richard Willis	NAPA

1.5. Organization of the Remainder of the Report

A comprehensive literature review was conducted and provided in Appendix A. The obtained insight from literature review along with consultation with the TRP assisted in the selection of projects to be shadowed in 2018, and in the selection of older projects for coring and collection of performance data. The references cited throughout the body of this report can be found in the reference list provided at the end Appendix A (chapters do not contain reference lists). The overall project organization is summarized as a flowchart in Figure 1-1.

Chapter two provides details regarding the collected plant produced asphalt mixtures selected for sampling in this study. Fourteen mixtures, produced at six asphalt plants were selected in consultation with the TRP. These were sampled during production in the summer of 2018, with details on the sampling and storage techniques used, mix designs, etc., provided in this chapter.

Chapter three provides the laboratory testing results from the selected 2018 Tollway plant-produced mixtures. After transferring the collected samples from the asphalt plants to the Missouri Asphalt Pavement and Innovation Lab (MAPIL), testing samples were fabricated and a suite of performance tests were conducted. The testing results were used to evaluate the pros and cons of the investigated performance tests, including their repeatability, practicality and their expected relationship to field performance. Detailed plots and a statistical analysis of data is presented in Chapter 3. Preliminary recommendations for performance tests to be used in the revised asphalt mixture PRS for the Tollway were generated from the results obtained in this data set.

Chapter four elaborates on a site visit conducted by the MU team in May of 2019 to examine the performance of selected mainline pavements and shoulders, including the sampled 2018 sections. In addition, the site visit provided the opportunity to visit older, existing sections with the goal of identifying good and poor performing sections with varying service lives. These sections were later cored to obtain additional performance testing samples, and the corresponding field performance data versus time was obtained via collaboration with Applied Research Associates, LLC.

Chapter five presents the results of the second phase of the laboratory testing, involving field core samples obtained in 2019. Results from field-aged cores were studied alongside test results on short-term aged samples such that the mixture PRS could be calibrated. This calibration enables the PRS to specify the use of short-term aged, laboratory-prepared specimens, along with suitably conservative property thresholds that take into account the expected property (and performance) effects of subsequent long-term aging in the field.

Chapter six presents the field performance data provided by ARA and analyzed by the research team. The distress and overall performance data were used to set final thresholds in the revised asphalt mixture PRS developed herein.

Chapter seven establishes the framework for the development of the recommended PRS threshold adjustments. Considering the selected performance tests and their thresholds from the previous chapters, this chapter describes the systematic process used for finalization of the performance specification. The framework utilizes a combination of laboratory and field investigation results, a straightforward statistical approach to conservatively account for test and sampling variability, and documents the consensus process used to incorporate practical considerations and experience from selected experts on the project TRP, which led to verification and/or final rounding of PRS thresholds.

Chapter eight summarizes the conclusions and findings of this research investigation, and provides recommendations for future research.

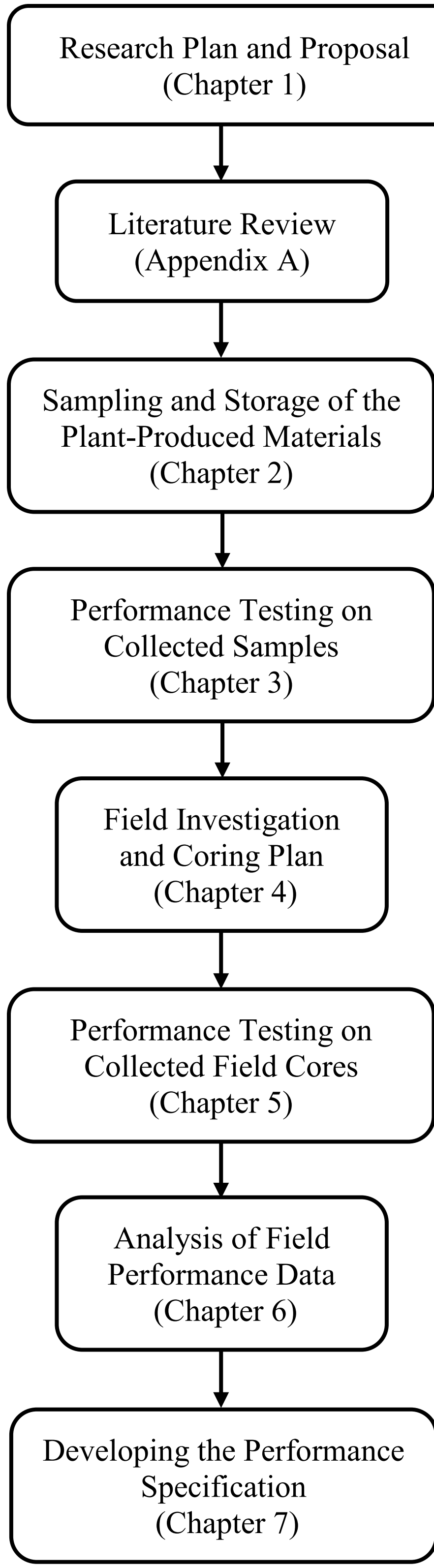


Figure 1-1. Flowchart for performance-based specification project

Chapter 2

MATERIAL SAMPLING AND PROCESSING

2.1. Overview

The selected plant-produced asphalt mixtures in this project were sampled per AASHTO T-168-03 across six asphalt plants in the Chicago area, as shown in Figure 2-1. Mixtures were sampled into uncoated, 5-gallon steel pails with tight-fitting lids. A representative from the Missouri Asphalt Pavement and Innovation Lab (MAPIL) of the University of Missouri-Columbia conducted all sampling with the assistance of local quality control (QC) lab staff. At the time of sampling, daily G_{mm}, and asphalt plant 6-min reports and cumulative mix tons were recorded in most cases. Selected materials were temporarily stored at the Tollway maintenance yard in Naperville, IL (Figure 2-2) for approximately one month before collection by MAPIL researchers.

(a) (b) (c)

Figure 2-1. Sampling plant produced mixtures from different asphalt plants. a) Mix 1836, Wm Ch; b) Mix 1845, Curran, and; c) Mix 1807, Geneva.

Figure 2-2. Temporarily storage of samples at Tollway maintenance yard in Naperville, IL

Table 2-1 provides details regarding the asphalt mixtures sampled from Chicago area plants producing Tollway asphalt mixtures in the summer of 2018. In total, six Illinois Tollway SMAs and eight dense graded mixtures were sampled from different asphalt plants in northern Illinois. Table 2-2 summarizes key asphalt mixture properties for each sampled section.

Table 2-1. Mixture sampling details

Date Sampled	No. of Buckets	Tollway Mix ID	Producer	Mix Type/Usage
9-Jul	10	90WMA1826	Plote, West Chicago	N70 9.5mm Surface
9-Jul	10	90WMA1803	Curran DeKalb	N50 19.0mm Binder - 3.0 voids
9-Jul	10	90WMA1840	Geneva North Aurora	N80 12.5 mm SMA Surface
9-Jul	10	90WMA1844	K-Five Romeoville	N80 12.5 mm SMA Friction Surface
13-Jul	10	90WMA1836	Wm Charles Airport	N80 12.5 mm SMA Surface
13-Jul	10	90WMA1845	Curran DeKalb	N80 12.5 mm SMA Friction Surface
16-Jul	10	90WMA1807	Geneva North Aurora	N50 19.0mm Binder - 3.0 voids
16-Jul	10	90WMA1828	Curran DeKalb	N50 4.75mm IL-4.75
18-Jul	10	90WMA1834	K-Five Romeoville	N70D 9.5mm Surface
18-Jul	10	90WMA1829	Geneva North Aurora	N50 4.75mm IL-4.75
18-Jul	10	90WMA1823	Rock Road Rochelle	N50 4.75mm IL-4.75
18-Jul	10	90WMA1835	Curran DeKalb	N80 12.5 mm SMA Friction Surface
2-Aug	10	90WMA1824	Rock Road Rochelle	N80 12.5 mm SMA Friction Surface
12-Sep	10	90WMA1818	Wm Charles Airport	N70E 9.5mm Surface

For the sake of simplicity, the common “90WMA” prefix in the Tollway Mix ID was omitted, and a shorter, four-digit sample ID was used throughout the report. The first four mixtures (1844, 1835, 1824, and 1845) are friction-surface-type SMAs, denoted as ‘Friction S.’ (used on highway curves and ramps), and the last two SMA mixtures (1836 and 1840) are regular SMA surfaces (used in lower trafficked, non-curved or tangent road alignments). The next three mixtures (1829, 1828, and 1823) are finer HMA mixtures (IL-4.75), which are used on the mainline below SMAs to promote pavement smoothness. Some engineers also believe that the IL-4.75 helps to reduce the rate of reflective cracking emanating upward from underlying Portland cement concrete joints and cracks. The three mixtures labeled as 1818, 1834, and 1826 represent surface shoulder materials (Shoulder S.). Finally, the last two sample IDs (1803 and 1807) represent shoulder binders, which appear below shoulder surface mixtures on the Tollway. As shown in the table, the design number of gyrations (N_{Design}) of all SMAs is 80, while the N_{Design} for shoulder surface mixtures is 70. IL-4.75 and shoulder binder mixtures used an N_{Design} level of 50 gyrations.

Table 2-2 also shows the binder system and reported modifiers used in each mix. Among the mixtures investigated, four of them, including 1844, 1824, 1836, and 1823 involved SBS-polymer-modified binder systems. Five mixtures (1835, 1845, 1840, 1829, and 1828) involved ground tire rubber (GTR), either by a terminal-blend wet process or by dry process. The 1835 mix utilized a relatively soft, neat binder (Superpave PG 46-34) combined with 10% engineered crumb rubber (ECR) by weight of binder (a dry-process GTR system). This mix also had the highest amount of recycled materials in any of the SMAs investigated (41.2% ABR), including 25.1% ABR by RAP and 16.1% ABR by RAS. Similar to 1835, the1845 mix also used PG 46-34 neat binder, which was later modified by 10.5% rubber by weight of the binder. The neat binder used in the 1840 mix was PG 58-28. This binder in this mix possessed 12.0% GTR, added to the binder via a terminal-blend, wet process. The binder used in dense graded shoulder mixtures involved neat (unmodified) Superpave binders.

The plan grade of binder used in Tollway SMAs and IL-4.75 mixes is PG 76-22. This implies that any extracted binder samples, which may include both a modifier (polymer or rubber), recycled binder components (usually RAP and RAS), and possibly rejuvenators and/or warm-mix and/or liquid antistrip additives, are expected to pass the performance grading criteria at 76℃ for the PG high temperature (PGHT) and -22℃ for the PG low temperature (PGLT). As for the shoulder mixtures, the plan grade is PG 64-22. The less stringent requirement on the PGHT of the shoulder plan grade is due to the lower traffic load that the shoulders experience throughout their service life. However, the plan PGLT requirement is the same for shoulder and mainline mixtures, as they experience the same low-temperature environmental conditions. Note also that the binder course mixtures on both shoulder and mainline sections undergo less critical low temperature and high temperature events, as they are thermally insulated and protected by the overlying surface mix. This should be considered when establishing PRS thresholds.

Aggregate gradations for the mixtures investigated are shown in Figure 2-3. It can be seen that the gradation of all SMAs investigated are quite similar, possessing a nominal maximum aggregate size (NMAS) of 12.5 mm. Likewise, the gradation of dense-graded mixtures within the groups, including IL-4.75, shoulder surface, and shoulder binder, are quite similar, within 4.5, 9.5, and 19 mm nominal maximum aggregate size (NMAS) groups, respectively. Despite the similarities in the aggregate gradations, it should be noted that the aggregate type used by each asphalt contractor can and does vary in the Chicagoland area. Therefore, the overall characteristics of the aggregate skeleton in each mix investigated herein should be viewed as unique.

Table 2-2. Details of mixture ingredients

Mix. ID	Mix Type	Base Binder	Plan Grade	ABR by RAP	ABR by RAS	NMAS
1844	N80 SMA Friction S.	SBS 70-28	76-22	10.8	16.0	12.5
1835	N80 SMA Friction S.	46-34 +10%ECR	76-22	25.1	16.1	12.5
1824	N80 SMA Friction S.	SBS 64-34	76-22	20.4	16.7	12.5
1845	N80 SMA Friction S.	46-34 +10.5%Lehigh	76-22	23.9	15.4	12.5
1836	N80 SMA Surface	SBS 64-34	76-22	16.2	16.3	12.5
1840	N80 SMA Surface	58-28 +12%GTR	76-22	15.9	9.8	12.5
1829	N50 Dense IL-4.75	58-28 +12%GTR	76-22	17.8	9.3	4.75
1828	N50 Dense IL-4.75	46-34 +10%ECR	76-22	35.3	9.2	4.75
1823	N50 Dense IL-4.75	SBS 64-34	76-22	24.1	14.2	4.75
1818	N70 Dense Shoulder S.	64-22	64-22	20.4	0.0	9.5
1834	N70 Dense Shoulder S.	58-28	64-22	20.0	0.0	9.5
1826	N70 Dense Shoulder S.	46-34	64-22	27.6	18.1	9.5
1807	N50 Dense Shoulder Binder	46-34	64-22	34.4	14.0	19.0
1803	N50 Dense Shoulder Binder	58-28	64-22	26.5	16.6	19.0

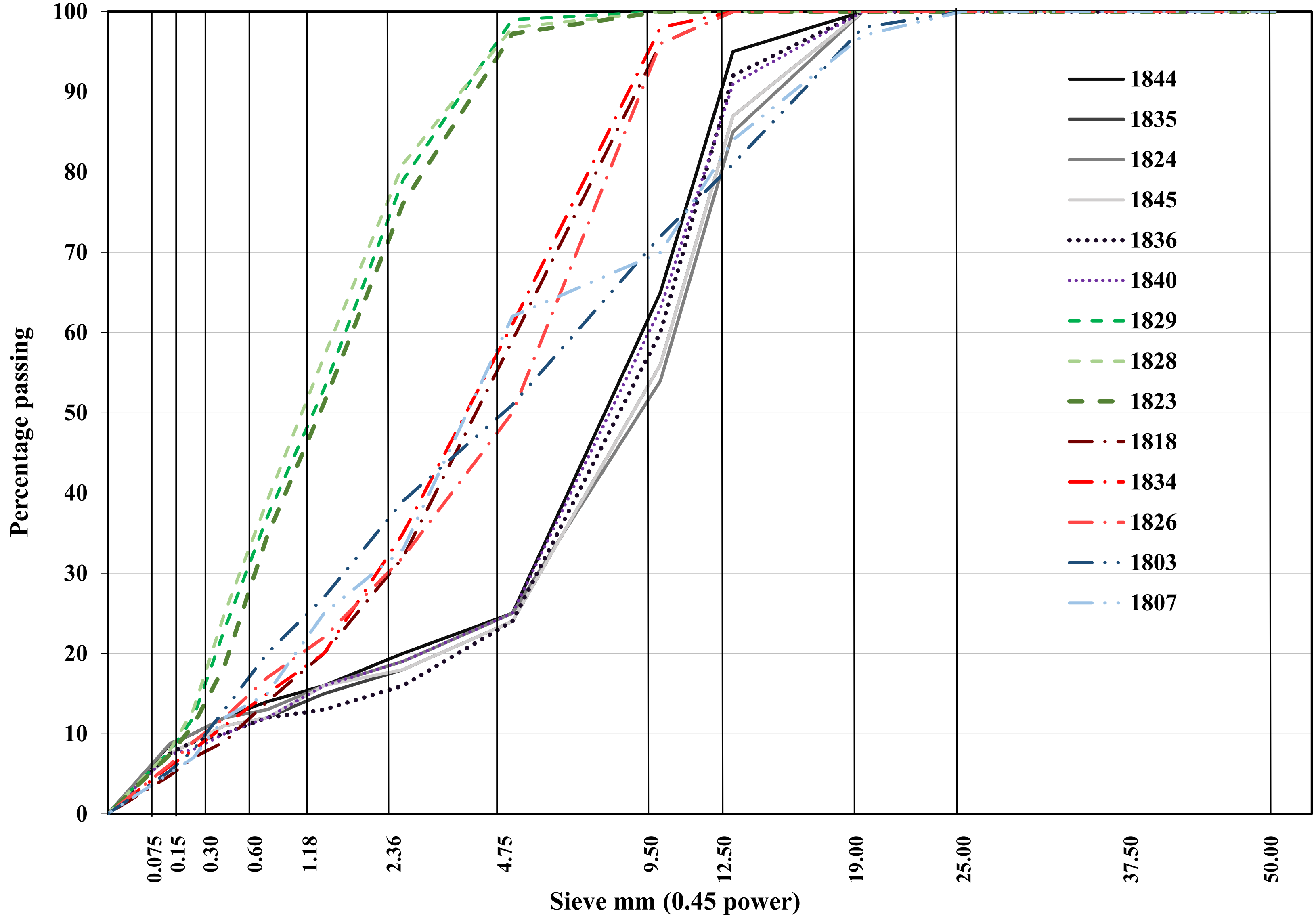


Figure 2-3. Aggregate gradations for the investigated plant-produced mixtures

Table 2-3 presents details regarding the sections that were paved in 2018, including Average Daily Traffic (ADT) and percent commercial vehicles. As indicated in Table 2-3 and shown on Figure 2-4, mixtures 1844 and 1834 were used to pave the mainline and shoulder, respectively, on route I-355 in starting from mile post (MP) 12 to 22. Mix1826 was used to pave the I-355 shoulder from MP 22 to 30. All other mixtures were used on I-88 between the indicated mile posts. The direction of the route is indicated in the 'Mile Post' column whenever the mixture appeared only in one direction. It is also worth mentioning that the 1845 mix was used on a small test section on the shoulder even though it was a friction-surface-type mix.

Table 2-3. Location of the paved section using the studied mixtures

Mix. ID	Mix Type	Route	Mile Post	Location	Traffic (ADT, and % Commercial Vehicles-CV)
1844	SMA Friction S.	I-355	12-22	Mainline	65,000 – 10% CV
1835	SMA Friction S.	I-88	93-103	Mainline	16,900 – 25% CV
1824	SMA Friction S.	I-88	EB 76-91	Mainline	10,600 – 25% CV
1845	SMA Friction S.	I-88	WB-105	Shoulder	16,900 – 25% CV
1836	SMA Surface	I-88	WB 76-91	Mainline	10,600 – 25% CV
1840	SMA Surface	I-88	103-113	Mainline	16,900 – 25% CV
1829	IL-4.75	I-88	103-113	Mainline	16,900 – 25% CV
1828	IL-4.75	I-88	92-103	Mainline	16,900 – 25% CV
1823	IL-4.75	I-88	WB 79-91	Mainline	10,600 – 25% CV
1818	Shoulder S.	I-88	EB 76-91	Shoulder	N.A.
1834	Shoulder S.	I-355	12-22	Shoulder	N.A.
1826	Shoulder S.	I-355	22-30	Shoulder	N.A.
1807	Shoulder Binder	I-88	103-113	Shoulder	N.A.
1803	Shoulder Binder	I-88	92-103	Shoulder	N.A.

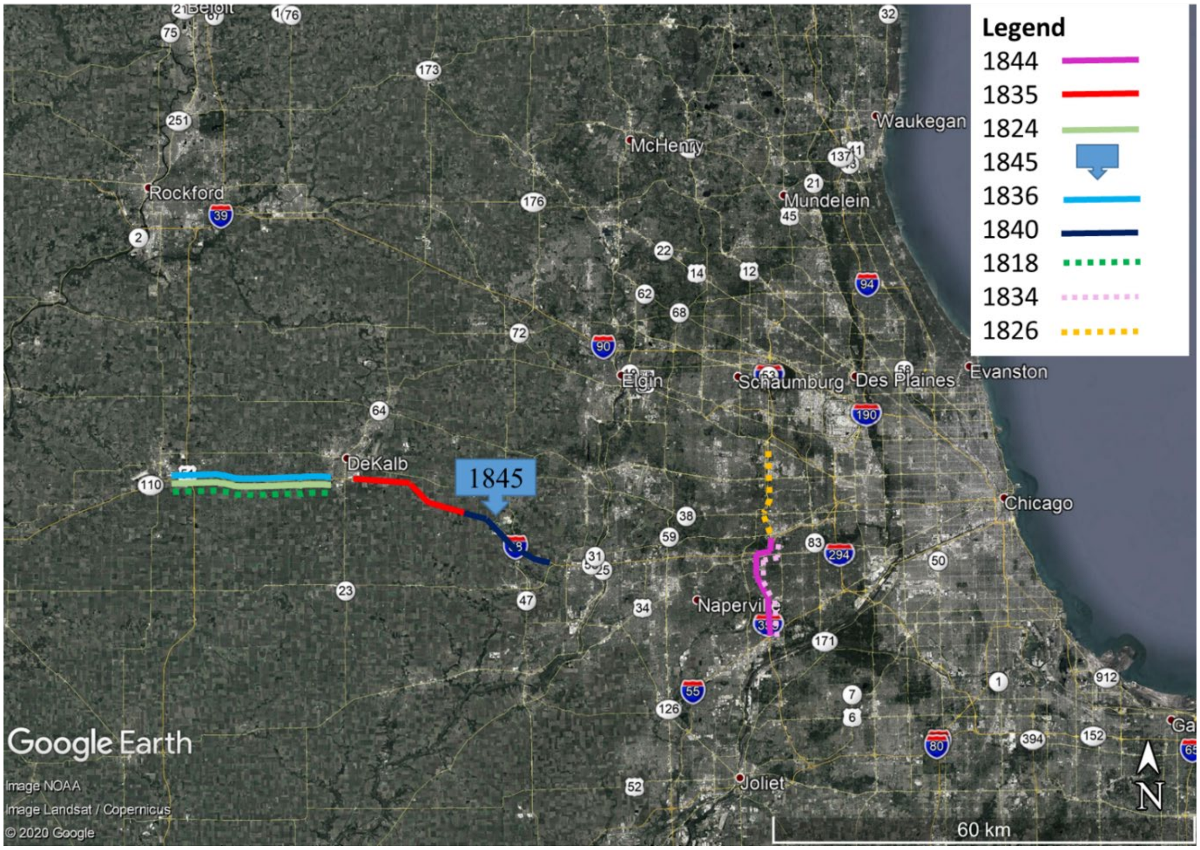


Figure 2-4. Location of paved roads using the sampled mixtures on Google Earth

2.1. Sample and Fabrication

The sampled plant-produced mixtures were brought back to MAPIL in 5-gallon steel pails. The plastic handles were removed and then pails were placed in a forced draft oven to heat the asphalt mixture to a workable consistency (~100 ℃). The heated mixture was then reduced to the gyratory sample mass following the quartering method in AASHTO R47 (see Figure 2-5). After reduction, two 1500 gr sets were collected in order to measure the maximum specific gravity (G_{mm}) of the mixtures as per AASHTO T 209 (see Figure 2-6). Although the G_{mm} was mentioned on the job mix formula (JMF) of each mix, it was attempted to verify it as the G_{mm} at the time of production might vary from the one on JMF. Figure 2-7 shows three different G_{mm} values obtained for each mixture. The blue bars are the G_{mm} values measured at MAPIL after reheating the buckets and collecting the G_{mm} samples. The orange bars are the G_{mm} measured at the night of mix production in the asphalt plant for quality control purposes, and the gray bars are the G_{mm} mentioned on the JMF sheets. It is also worth mentioning that the MAPIL measured values were used to measure the air void content of the gyratory compacted specimens. In order to avoid segregation during the sample production process, the heated asphalt mixture in the pans was transferred to a chute, as shown in Figure 2-8, and then was poured into the mold. A Pine GB2 Superpave gyratory compactor was used to compact the reheated samples and make cylindrical specimens.

Figure 2-5. Splitting the bucket of mixture as per AASHTO R47

Figure 2-6. Preparing samples for theoretical maximum specific gravity (G_{mm}) testing

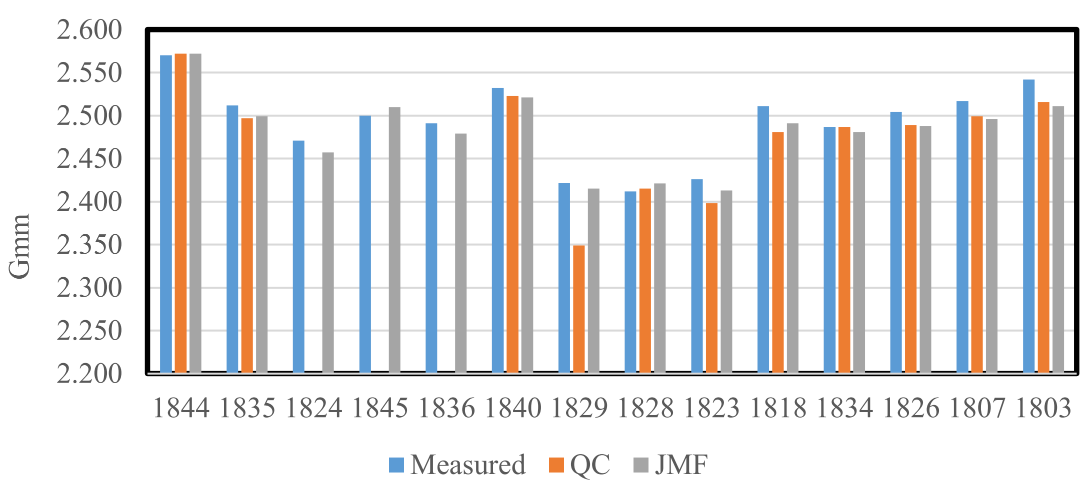


Figure 2-7. Comparing different G_{mm} values for the studied mixtures (Measured = Test performed in MAPIL; QC = Measured by quality control crew; JMF = job mix formula)

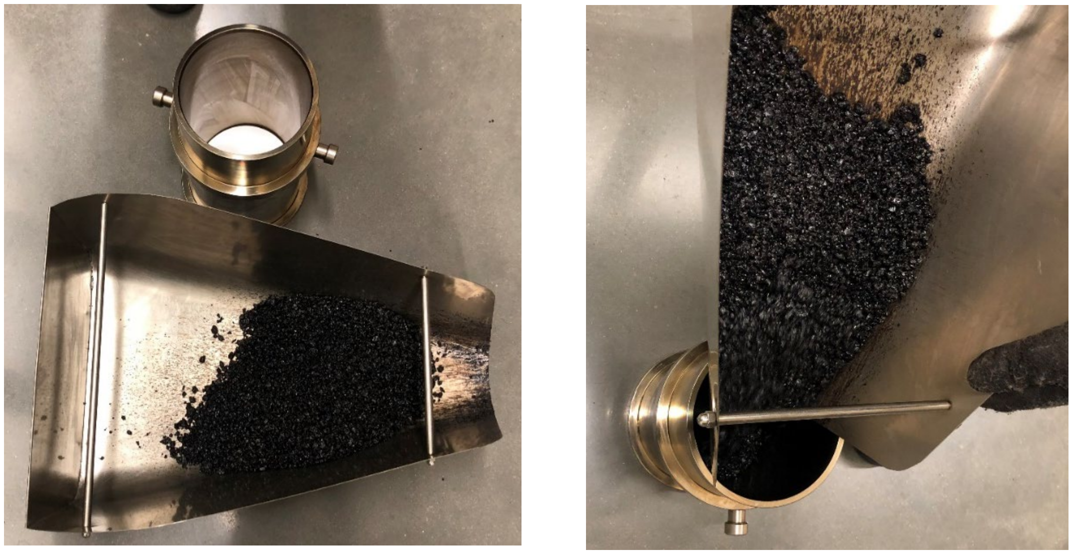


Figure 2-8. Transferring the heated mix to the mold

After splitting to desired mass, the asphalt mixture was heated to compaction temperature (155 and 143 ℃ for modified and unmodified mixes, respectively). All SMA testing samples were compacted to 6.0 % air voids while the target air void for dense graded mixtures was 7.0 %. For DC(T) and I-FIT samples, air voids were measured on the 50 mm slices before notching and coring for the DCT specimens, and before cutting the slice in half and notching in the case of I-FIT specimens. For Hamburg specimens, the original gyratory specimen (62 mm in height) was used for G_{mb} testing prior to cutting the flat face on one side. The TSR and IDEAL-CT tests were performed on 95 mm gyratory compacted samples.

Chapter 3

TESTING RESULTS FOR PLANT-PRODUCED MIXTURES

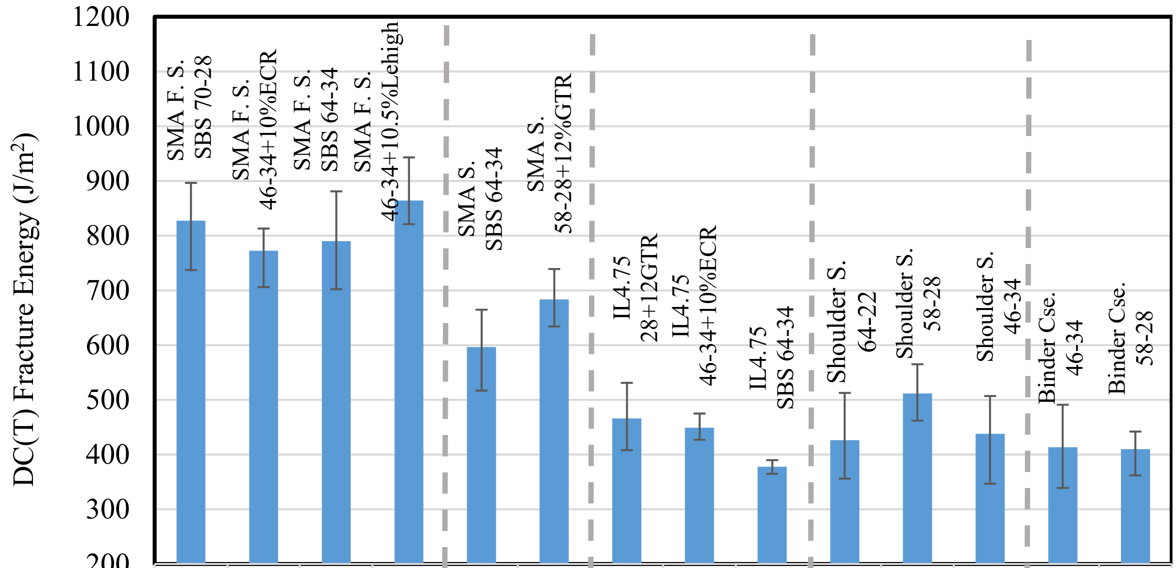
3.1. DC(T) Testing Results

The DC(T) test was developed to characterize the fracture behavior of asphalt concrete mixtures at low temperatures. The testing temperature is 10ºC warmer than the PG low temperature grade of the mixture, per (ASTM D7313-13). Thermal cracking in asphalt pavements can be considered as occurring in pure tensile opening or Mode I fracture, as the cracks propagate perpendicular to the direction of the thermal-induced stresses in the pavement, i.e., transverse to the direction of traffic (Wagoner et al. 2005). The fracture energy is computed as follows.

$$G_f = \frac{AREA}{B \cdot L} \quad [1]$$

where, G_f denotes fracture energy in J/m^2, AREA is the area under Load-CMOD$_{FIT}$ curve, until the terminal load of 0.1 kN is reached. B is specimen thickness in m, generally 0.050 m (except for field cores) and L is ligament length, usually around 0.083 m. The DC(T) test procedure used in this study includes conditioning of the fabricated specimen at the selected test temperature in a temperature-controlled chamber for a minimum of two hours. After the conditioning, the specimens are suspended on loading pins in the DC(T) machine. A portable Test Quip DC(T) device was used, which is housed at MAPIL. The test is performed at a constant crack mouth opening displacement (CMOD) rate, which is controlled by a CMOD clip-on gage mounted at the crack mouth. The CMOD rate specified in ASTM D7313-13 is 0.017 mm/s (1 mm/min). To begin the testing sequence, a seating load no greater than 0.2 kN (typically about 0.1 kN) is applied to 'seat' the specimen on the loading pins. Once a stable seating load is confirmed, the test is run at the specified CMOD rate the test is completed when the crack has propagated and the post-peak load level is reduced to 0.1 kN. The fracture energy can be obtained by measuring the area under the load-CMOD curve and dividing it by the fractured area (ligament length times thickness).

Figure 3-1 shows the DC(T) fracture (G_f) testing results at -12 ℃, using samples fabricated at MAPIL. The error bars provide the range of the values obtained for the three replicates tested in DC(T) fracture. In addition to the bars shown in the figure, the table attached to the figure provides the mix ID, the average fracture energy and also the ABR of each mix. Also, the type of the mix and the binder system are shown above each bar. The cracking resistance of the SMA friction surface (F. S.) mixes was expected to be the highest, followed by SMA surface mixes. Additionally, as the shoulder surface mixtures experience the same environmental conditions, they should ideally be designed with relatively high cracking resistance. The IL-4.75 and shoulder binder mixtures were expected to have lower cracking resistance as they are used in sublayers of the pavement. As shown in Figure 3-1, the expectations for relative crack resistance were found to be in close agreement with the measured DC(T) fracture energy results. This finding was among the early indications that the DC(T) test is a viable candidate for the control of cracking in Tollway asphalt mixtures, at least from the standpoint of mixture performance.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
Gf at -12 °C	828	772	790	864	596	684	466	449	378	427	512	438	414	410
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1
COV(%)	9.9	6.5	11.3	7.9	12.5	7.7	13.3	5.4	3.3	18.6	10.1	18.8	18.4	10.3

Figure 3-1. DC(T) fracture energy at -12 ℃ using MAPIL samples

As shown in the figure, all the SMA friction surface type mixtures were found to pass the 750 J/m^2 fracture energy threshold, which was specified by the Tollway in prior completion of this project (for instance, in the 2019 specification). The difference between the highest and lowest recorded fracture energy is less than 100 J/m^2 which implies that the resistance of the SMA Friction surface with respect to low temperature cracking is expected to be similar between sections. On the other hand, the SMA surface mixtures had significantly lower fracture energy than the SMA friction surface mixtures. The 1836 mix recorded the lowest fracture energy (596.5 J/m^2) and similar to 1840 which had 684 J/m^2, the sampled production mix did not pass the formerly practiced 700 J/m^2 limit for SMA surface mixtures in design. This may be possibly attributed to the effects of aging during sample storage (3 months on average) followed by reheating of the mix. More observations from the results are summarized as follows:

- The base binder system used in the 1824 and 1836 mixtures is the same (SBS 64-34). Although the 1824 mixture had higher amount of ABR, it has an additional 270 J/m^2 of fracture energy. This comparison reveals the importance of aggregate quality and its significant role in low temperature cracking resistance of the mix.

- Although the 1840 mix had the lowest amount of recycling, it was not found to pass the existing Tollway fracture energy criteria. Using a softer base binder on the low temperature side, adding a rejuvenator (recycling agent), and/or utilizing higher quality aggregate are strategies that could be used in the future to boost the fracture energy in this mixture.

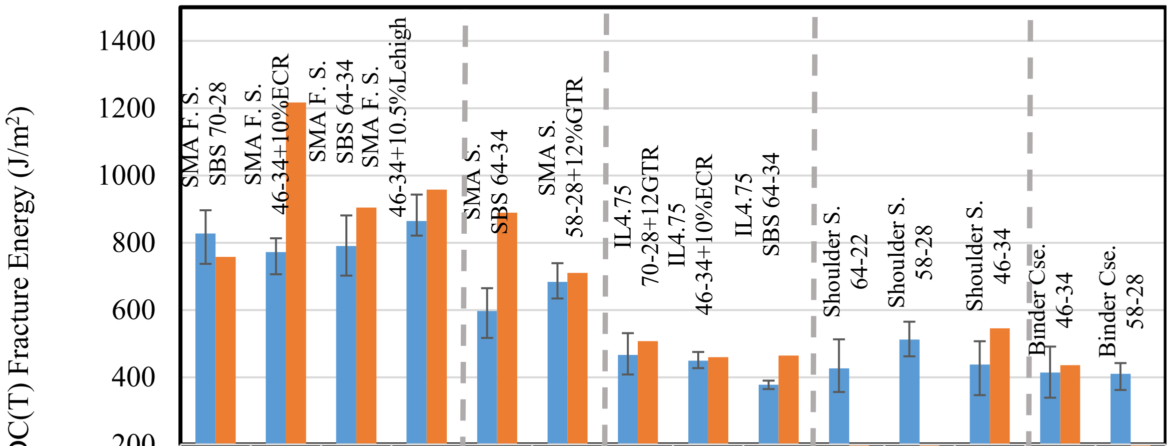
- Despite the high ABR (44.6%) incorporated in 1828 mix, using a softer binder system along with engineered crumb rubber (ECR) resulted in a relatively high fracture energy. In IL-4.75 mixtures, the 1823 mix with an SBS 64-34 binder system exhibited the lowest fracture energy.

- In the shoulder surface mix group, the 1826 mix benefited from the soft binder system and possessed a DC(T) fracture energy of 438 J/m^2 despite the high recycle content (ABR=45.7 %). Compared to 1818, the 1834 mix with a softer binder and similar recycle content performed notably better in terms of low temperature cracking. The softer binder system used in 1834 mix likely contributed to the additional 75 J/m^2 of fracture energy as compared to mix 1818. Differences in aggregate quality might also contribute to the difference in DC(T) fracture energy of these two mixtures.

- The shoulder binder mixtures including mix 1807 and 1803 yielded similar DC(T) fracture energy values, close to the 400 J/m^2 level specified by the Tollway at the time of their design. In the future, incorporating GTR could assist in raising the fracture energy of these mixtures, as was the case for IL-4.75s.

Figure 3-2 compares the DC(T) fracture energy measured at MAPIL using plant-produced lab compacted mixtures with the ones reported on the JMF. As shown, in most of the cases (all except 1844) the measured DC(T) energy at MAPIL is lower than the reported fracture energy. Storage and reheating of the plant-produced samples might have resulted in additional aging of the mixtures, which often leads to lower DC(T) fracture energy values (Buttlar et al., 2019).

Figure 3-3 presents a contour map that provides the pavement temperature at the surface produced by MAPIL researchers using LTPP bind software at a level of 98 % reliability. This temperature is used in order to determine the required PGLT for the binder. As shown in the figure, the required PGLT of the binder in the state of Illinois is in the range of -22 to -27 °C. The PGLT of -22 °C is mainly required in the southern part of Illinois while the -27 °C limit is suitable for the very northern part of the state. Therefore, the PGLT of plan grade of the binder in the upper parts of Illinois should be lower than -22 °C to reach 98 % reliability. As per ASTM D-7313, the DC(T) test is performed at 10 °C warmer than the binder grade. The Illinois Tollway currently uses a -22 °C plan low temperature grade, and thus testing at -12 °C used in the DC(T). Thus, the relatively high DC(T) thresholds specified by the Illinois Tollway reflect both the high project criticality of Tollway road surfaces, and also the fact that a slight adjustment has been factored in the specification based on the fact that Northern Illinois is somewhat colder than the assumed -22 °C PGLT used in the asphalt binder plan grades.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
Gf at -12 °C	828	772	790	864	596	684	466	449	378	427	512	438	414	410
Gf From Design	758	1217	904	958	889	710	507	460	464	N.A	N.A	545	436	N.A
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1

Figure 3-2. Comparing DC(T) fracture energy at -12 ℃: current study vs. JMF

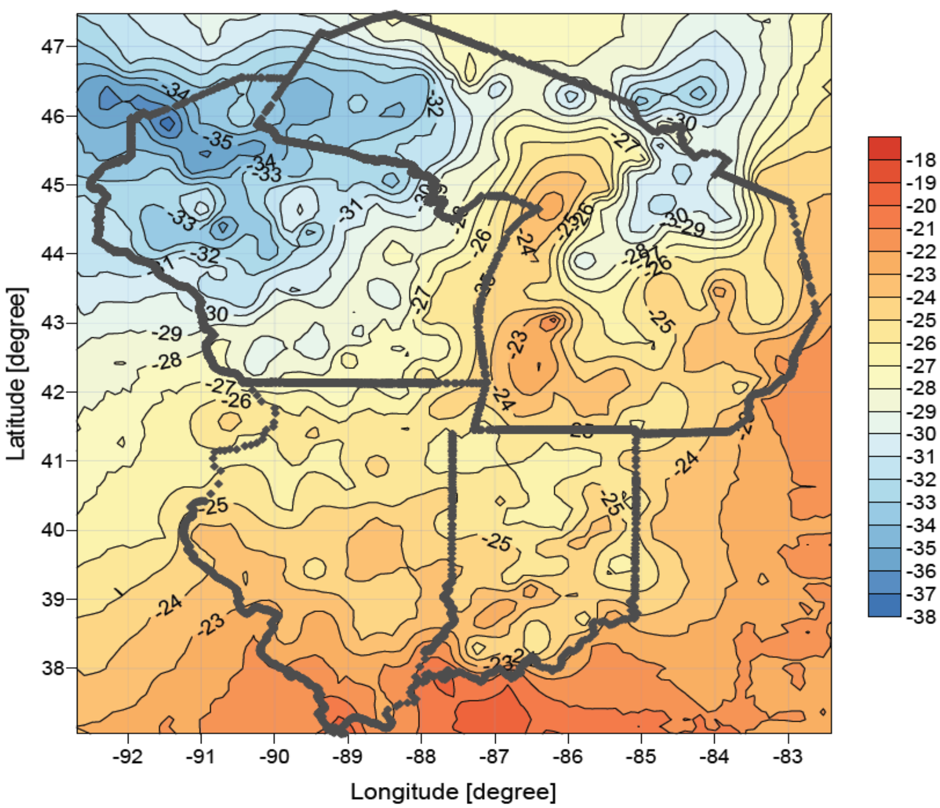
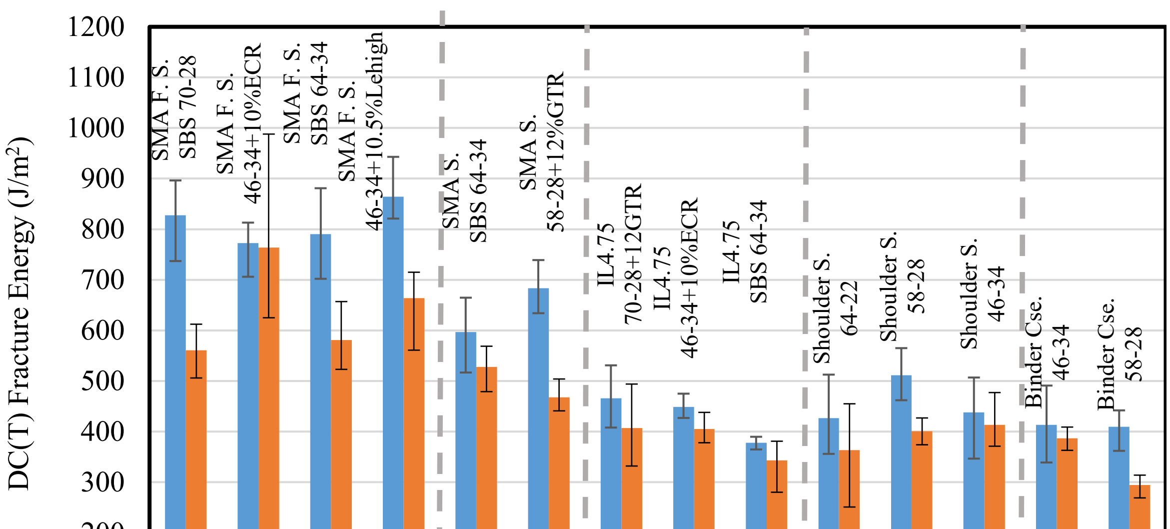


Figure 3-3. Pavement temperature at the surface in the northern part of the US (reliability=98%)

As testing temperature drops, asphalt binder becomes stiffer and as a result, the asphalt sample exhibits more brittle behavior and reduced fracture energy. Therefore, lower fracture energies are expected at -18 °C compared to -12 °C. As shown in Figure 3-4, the DC(T) fracture energy of all the mixtures at -18 °C are indeed lower than those obtained at -12 °C. Experiencing a drop of about 260 J/m^2, the 1824, which is an SBS modified mix, showed the highest temperature sensitivity in the SMA friction surface category. On the other hand, the 1835 mix, which is modified with a dry process Engineered Crumb Rubber (ECR-type GTR) showed almost the same fracture energy at -18 °C as compared to -12 °C. If the same DC(T) criteria were applied by Tollway at -18 °C, only the 1835 would pass. The 1840 mix in the SMA surface group experienced a significant drop in fracture energy after testing at -18 °C, which indicates that this mix may be highly temperature sensitive at low temperatures. Although the IL-4.75 and shoulder binder mixtures will not experience extreme low temperature events such as surface mixtures do, the DC(T) test at -18 °C was conducted nevertheless to evaluate their temperature sensitivity. Only the 1803 mix, which used a PG 58-28 base binder showed high sensitivity, where the DC(T) fracture dropped from 410 to 290 J/m^2 (a 28 % reduction in fracture energy).



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
Gf at -12 °C	828	772	790	864	596	684	466	449	378	427	512	438	414	410
Gf at -18 °C	560	764	581	664	528	468	407	405	343	363	401	413	387	294
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1

Figure 3-4. Comparing DC(T) fracture energy at -12 and -18 °C

3.2. I-FIT Testing Results

The flexibility index (FI) is an empirical index parameter that is computed as the total fracture energy divided by the absolute value of the slope of the post-peak softening curve. FI is proposed to provide a means to identify brittle mixtures that are prone to premature cracking, and was specifically developed to be sensitive to recycled material content (AASHTO TP124-16). The FI parameter is calculated as follows.

$$FI = \frac{G_f}{|m|}(0.01) \qquad [2]$$

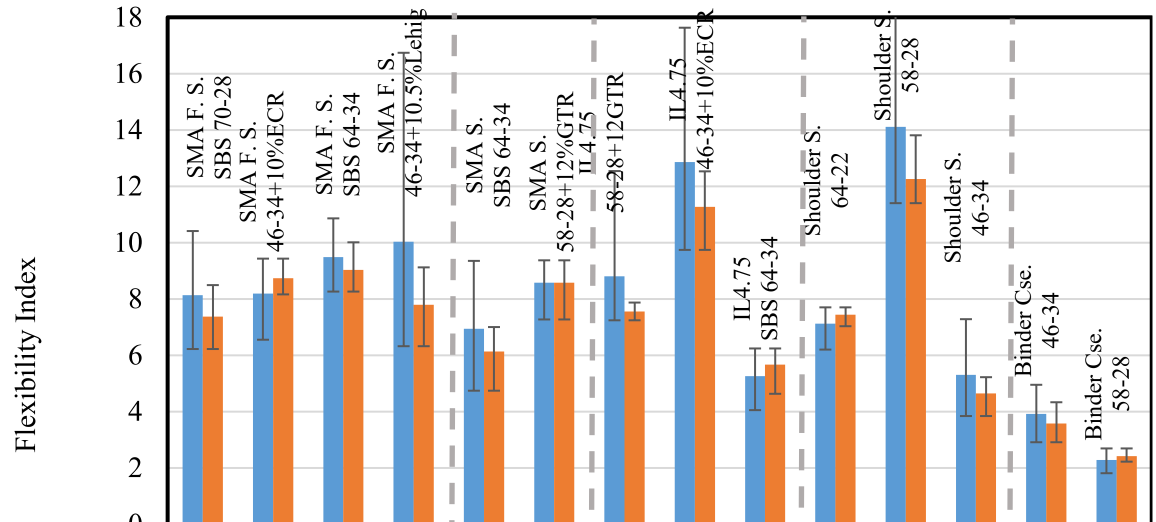
where G_f is computed in a similar manner as to the DC(T) test, and m represents the slope of the post-peak softening curve. There are numerous ways to estimate the slope of a curve resulting from a material test, which posed an inherent challenge from the perspective of test standardization in the development of tests such as the I-FIT. At present, to address this source of variability, the slope parameter is typically determined using a sophisticated software program available from the Illinois Center for Transportation (visit https://ict.illinois.edu/2016/07/01/i-fit-software-now-available-on-ict-website/). To fabricate samples, a notch is cut along the axis of symmetry of a semi-circular bend specimen to a depth of 15±1 mm. Test specimens are then conditioned in the environmental chamber at 25 ºC for 2 hrs. ± 10 min. After a contact load of 0.1 kN is reached, the test is carried out at a rate of 50 mm/min load line displacement (LLD). The test is considered to be complete when the load drops below 0.1 kN, which is identical to the DC(T) test termination definition. A sampling rate of 40 samples per second was used to collect the data during the test. A software named "SCB TestQuip LLC. V2.0.0rc4" was then applied to analyze the collected load-displacement data and calculated the FI parameter.

Figure 3-5 presents the results of I-FIT testing performed on the samples conditioned at 25 ºC. The blue bars represent the average of four replicates. As inferred from the large error bars, which display the upper and lower FIs obtained for each mix, the repeatability of the I-FIT test itself may be viewed as borderline (too high). In order to lower the FI variability, researchers in Illinois proposed that the replicate with the furthest FI from the average FI be removed, followed by a recalculation of the average of three remaining replicates (denoted herein as 'FI 3 Reps'). Although this approach is questionable from a statistical standpoint and may produce significant movement in the average (upwards or downwards), it clearly achieves the goal of reducing the reported variability in the averaged results. Figure 3-5 also compares these two averages (four replicates vs. three replicates).

Examining the FI trends, the cracking performance of the different groups of Tollway mixtures were not in close correspondence with expected, relative cracking performance trends. For example, two mixtures in the IL-4.75 group (1829 and 1828) possessed FIs that were higher than those of SMA friction surface mixes. In addition, 1834, which is a shoulder surface mix, yielded the highest FI amongst the studied mixtures. As will be shown later, the FI parameter is heavily dependent on aging. As these plant-produced mixtures have been reheated for sample preparation, the FIs values, including relative trends, might have been significantly affected.

Figure 3-6 shows the I-FIT testing results for specimens having varied levels of air voids. A straight line was fitted to data to quantify the sensitivity of the FI parameter to air voids. As the slopes suggests, air void content has a very significant effect on the FI. For example, the FI for the 1828 mix would increase by 2.4 for each percent increase in air void content. It is worth mentioning that the effect of air void on the I-FIT test was not the main goal of this study and the extra testing was done on the fabricated samples having air void levels outside of the acceptable range (6±0.5 for SMAs and 7±0.5 for dense graded mixes). Based on this relatively limited number of tested mixes, the maximum and minimum FI change per percent increase in air voids was found to be 5.1 and 2.4, respectively.

Figure 3-7 compares the average FI values from four replicates calculated through the TestQuip software with the outputs from the software provided by Illinois Center for Transportation (ICT). As seen, the difference between the FIs is not considerable and could be mainly attributed to the differences in curve fitting techniques including the slope (derivative) computation method used.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
FI (4 Reps)	8.1	8.2	9.5	10.0	6.9	8.6	8.8	12.9	5.3	7.1	14.1	5.3	3.9	2.3
FI (3 Reps)	7.4	8.7	9.0	7.8	6.1	8.6	7.6	11.3	5.7	7.4	12.3	4.6	3.6	2.4
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1
COV (4 Reps)%	21.9	14.8	12.3	46.1	27.3	13.3	28.5	26.3	20.7	9.6	27.4	27.2	23.1	19.4
COV(3 Reps)%	15.4	7.4	9.9	18.0	19.9	13.3	4.2	12.6	15.8	4.8	11.0	15.4	20.0	10.1

Figure 3-5. I-FIT testing results (four replicates vs. three replicates)

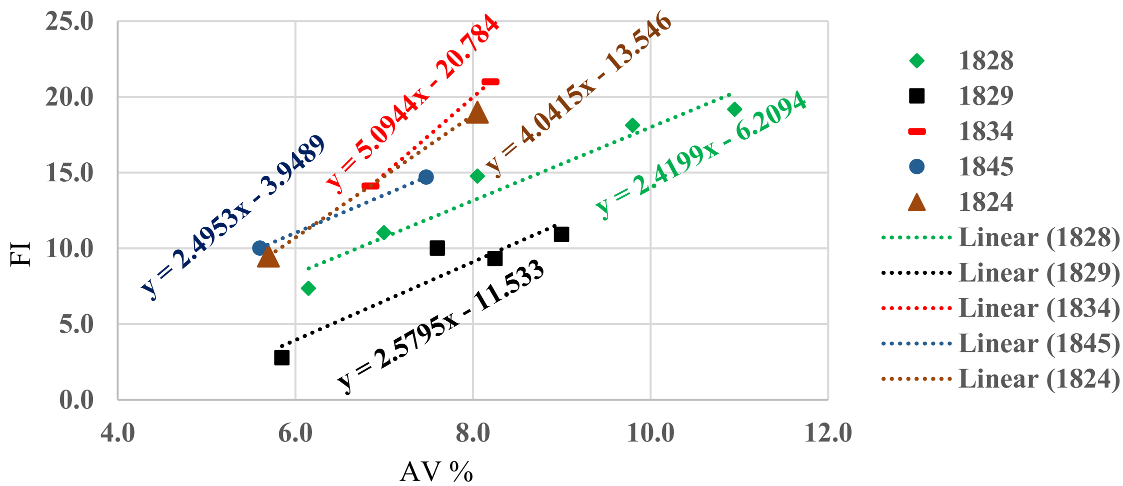
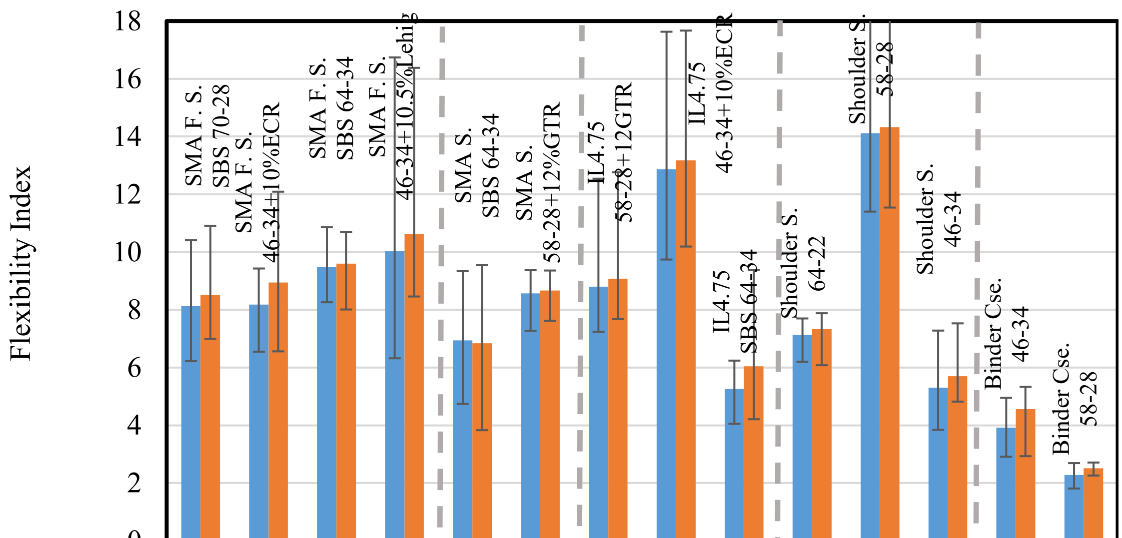


Figure 3-6. Effect of air voids on FI



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
FI (from TestQuip)	8.1	8.2	9.5	10.0	6.9	8.6	8.8	12.9	5.3	7.1	14.1	5.3	3.9	2.3
FI (from ICT)	8.5	8.9	9.6	10.6	6.8	8.7	9.1	13.2	6.0	7.3	14.3	5.7	4.6	2.5
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1
COV (TestQuip)%	21.9	14.8	12.3	46.1	27.3	13.3	28.5	26.3	20.7	9.6	27.4	27.2	23.1	19.4
COV (ICT)%	20.1	26.1	12.6	36.2	34.3	10.6	27.0	24.1	39.1	11.5	26.8	21.7	24.1	7.4

Figure 3-7. Comparing FIs calculated from Test Quip and ICT software

3.3. IDEAL-CT Testing Results

The IDEAL-CT test was developed to characterize the potential of cracking in asphalt mixes at room temperature. The test set-up is similar to the traditional indirect tensile strength test, but it is performed at 25°C under a constant loading rate of 50 mm/min until failure occurs (ASTM D 8225-19). The specimens are cylindrical with a diameter of 100 or 150 mm and a thickness of 38, 62, 75, or 95 mm, depending on the specification followed. The specimens do not require gluing, notching, drilling or additional cutting. The procedure of the test includes conditioning the specimens in a temperature-controlled chamber for a minimum of 2 hours at 25 ℃. After conditioning, the specimens were placed in a Test Quip™ load frame set up for the IDEAL test. A seating load of 0.1 kN was applied in order to make appropriate contact between the loading heads and the sample. The sample was then loaded under a displacement control mode of 50 mm/min while the loading level was collected by the device.

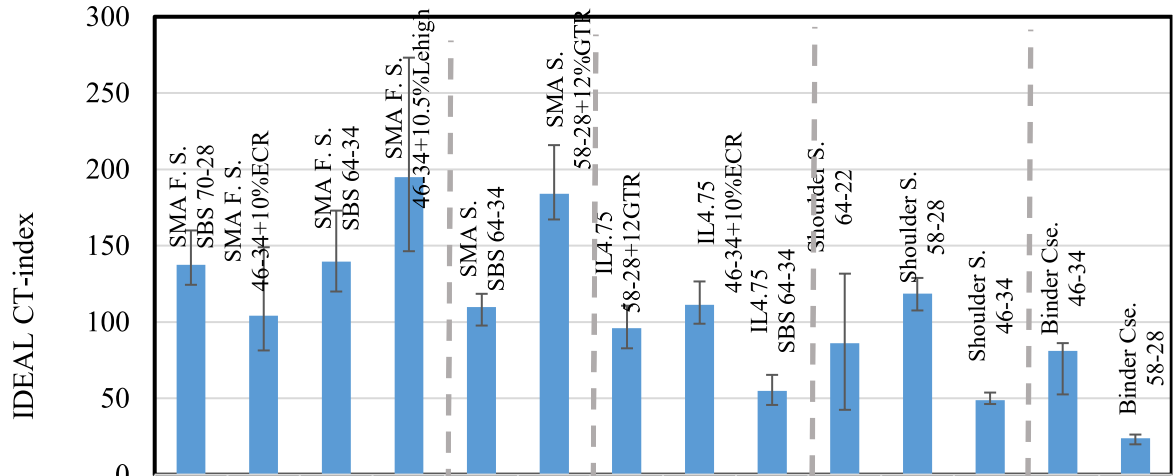
The cracking parameter for the IDEAL-CT is derived from the load vs. ram displacement curve. The larger the CT-index, the better the cracking resistance of the mixture according to the test developers. A minimum of CT-index for SMAs proposed is 145 while the recommended CT-index for Superpave dense graded mixes is 105 (Zhou, 2018). The CT index equation for a specimen of 62 mm thickness is as follows.

$$CT_{index} = \frac{G_f}{|m_{75}|} \times \left(\frac{l_{75}}{D}\right) \times \left(\frac{t}{62}\right) \quad [3]$$

where,

G_f=	Fracture energy (AREA under the curve normalized by the area fractured)
AREA=	Area under the load-displacement curve, until a terminal load of 0.1 kN is reached
m_{75}=	Modulus parameter (absolute value of the slope at 75% of peak load)
l_{75}=	Vertical displacement when the load is reduced to 75% of peak load
l_{75}/D=	Strain tolerance parameter (when load is reduced to 75% of peak load)
D=	Specimen diameter
t=	Specimen thickness

In this project, the IDEAL-CT test was performed on cylindrical samples compacted to 95 mm and conditioned for two hours in an environmental chamber. Three replicates were fabricated for each mix and tested to calculate the CT index. Similar to the previous testing figures, the error bars in Figure 3-8 shows the range of the calculated CTs for each mix. As shown, most of the SMA mixtures could not meet the threshold of 145, which is recommended for Texas. The dense graded mixtures had a difficult time reaching the minimum recommended CT-index threshold of 105. Given the fact that most of the SMA mixtures produced relatively high DC(T) fracture energy values and FI values in excess of 8.0, the CT-index thresholds recommended by developers for the Texas climate might be too stringent for Tollway mixtures. It is also worth mentioning that reheating the samples might have reduced the crack resistance of the mixtures due to excessive aging.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
CT-index	137	104	140	195	110	184	96	111	55	86	119	49	81	24
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1
COV (%)	14.3	37.3	20.8	35.2	8.0	15.0	14.6	12.6	18.1	51.8	9.0	8.8	29.3	14.7

Figure 3-8. Results from IDEAL-CT testing (95 mm sample thickness)

- Similar to the other two cracking tests (DC(T) and I-FIT), 1836 (SMA surface) recorded the lowest score in the IDEAL-CT test among SMA mixtures.
- The 1845 and 1840 mixtures, which are both GTR modified, recorded the best CT scores among the SMAs.
- Similar to the other two cracking tests (DC(T) and I-FIT), 1823 mix had the lowest CT score in the IL-4.75 group.
- The 1807 mix, which benefits from a softer binder as compared to 1803, had significantly better cracking performance based on IDEAL-CT results.

3.4. IDT Strength Testing Results

Following AASHTO T322-07, 2011, the Superpave Indirect Tensile Test (IDT) can be used to perform creep compliance and strength testing of asphalt mixtures. Field-cored or gyratory-compacted samples with heights ranging from 38 to 50 mm and diameters in the range of 150 ±9 mm are generally used. Three testing temperatures with 10 °C intervals are recommended for use, which are often taken as 0, -10, and -20 °C. Alternatively, temperatures can be selected to encompass the low-performance grade (PG) of the asphalt binder and can use a different temperature spacing, such as 0, -12, and -24 °C (Marasteanu et al., 2007). A creep test duration of 1000 seconds is generally required to ensure overlap between creep curves when constructing the master curve. Since the creep compliance should normally be characterized in the linear viscoelastic range, loading levels should be kept sufficiently low in order to retain this linearity. Therefore, a maximum deformation on the horizontal clip gage of 0.019 mm for 150 mm diameter samples is suggested to stay within the linear range. In addition, to circumvent the noise problems and drift inherent in sensors (displacement extensometers), a minimum deformation of 0.00125 mm at a 30-second loading time is recommended.

The IDT creep and strength tests were carried out using a Cooper universal testing machine (UTM) at MAPIL with the capacity of 100 kN. IDT creep and strength tests were performed following AASHTO T-322. To carry out the IDT creep test, three samples were conditioned at three different temperatures including 0, -12 and -24 °C. Each sample was kept at the testing temperature for 2 hours. The conditioned sample was then put into the IDT fixture. In order to compensate the temperature loss due to opening the chamber door and installing the extensometers, the sample was kept for another half an hour to reach the testing temperature. During this time, the response of extensometers was monitored to ensure temperature stabilization and the absence of sensor drift. Monitoring the response of the extensometer also helps in detecting potential problems with the attachment of the extensometers, which could affect the data. Next, a seating load of 0.1 kN was applied to the sample. The seating load fixes the sample position in the IDT fixture, ensures rapid creep loading without impact, and eliminates some of the slight nonlinearity exhibited at low load levels. In the test, the load level is rapidly increased as a steep slope-load function until the target creep load is reached, which may differ at each temperature. The closed-loop controls are tuned such that the creep load is attained in less than one second. The creep load was then maintained for 1000 seconds while displacements were recorded. Table 3-1 shows the testing parameters used in IDT creep test. Equation 4 presents the general equation used to convert load and deflection values to creep compliance (AASHTO T-322-17).

Table 3-1. Loading Properties in IDT Creep Test

Testing Temp. (C)	Chamber Temp. (C)	Seating Load (kN)	Ramp Time (s)	Creep Load (kN)	Creep Time (s)
0	-1.5	0.1	1	4	1000
-12	-14	0.1	1	8	1000
-24	-26	0.1	1	20	1000

$$D(t) = \frac{\Delta X_{tm,t} * D_{avg} * b_{avg}}{P_{avg} * GL} * C_{cmpl} \qquad [4]$$

where,

D(t)= Creep compliance as a function of time (1/GPa)
$\Delta X_{tm,t}$= Trimmed mean of normalized horizontal deformation (mm)
D_{avg}= Averaged diameter (mm)
b_{avg}= Averaged thickness (mm)
P_{avg}= Averaged applied load (kN)
GL = Guage length (mm)
C_{cmpl}= Creep compliance correction factor

After plotting the creep compliance curves at different temperatures versus time in a log-log space, the curves are shifted horizontally relative to the curve at the referenced temperature to construct a unique continuous curve, called the master curve. A power law function is then fitted to the master curve as shown in Eq. 5.

$$D(t) = D_0 + D_1 t^m \quad [5]$$

where D_0 and D_1 and m are model constant values and t denotes time.

The strength test is performed by applying an increasing load at a constant displacement rate until failure occurs in the specimen. Extensometers were removed prior to strength testing to avoid damage, as tensile strength was estimated using a simple 2D, plane-stress based solution.

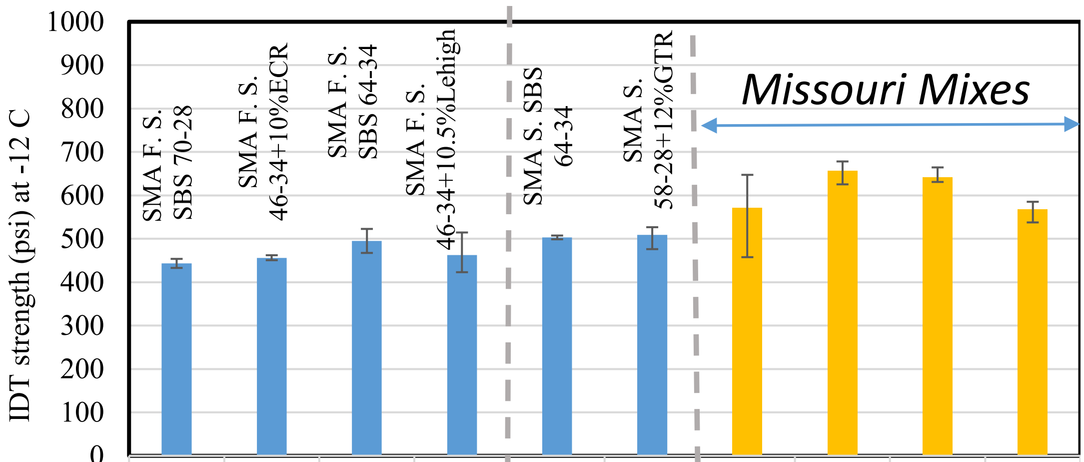
$$S_t = \frac{2P_{max}}{\pi * b * D} \quad [6]$$

where

S_t = Tensile strength (MPa)
P_{max}= Maximum recorded load (kN)
b= Sample thickness (mm)
D= Sample diameter (mm)

Figure 3-9 shows the IDT strength testing results performed on SMA slices with 50 mm thickness after being conditioned at -12 °C for 2 hours. As can be seen in this figure, the strength of the SMA mixtures are very similar to each other and there is no significant difference between the strengths. It is also worth mentioning that unlike the other cracking performance tests, the SMA surface mixes are exhibiting a slightly higher strength as compared to the SMA friction surface mixtures under the IDT strength test. The Tollway 2018 mixtures are also compared to selected dense-graded Missouri highway mixtures in this figure. It is interesting to note that the strength of the Missouri highway mixtures is higher than the Tollway mixtures. Additionally, the DC(T) fracture energy in the same Missouri highway mixtures was around 400 J/m^2, and measured creep compliance was relatively low, which indicates stiff and brittle behavior at low temperatures. This follows the general trend of high stiffness being related to high strength but low fracture resistance. This also follows the current thinking regarding the shortcomings in using simple tensile strength measurements as a parameter to rank low temperature cracking resistance (Buttlar et al., 2019).

Prior to IDT strength testing, the IDT creep test was conducted on the samples and the IDT creep compliance master curves were generated as shown in Figure 3-10. The creep compliance master curve, which is the output of the IDT creep test can be used to model the viscoelastic behavior of the mixtures and to predict the amount of low temperature cracking expected in the pavement.



	1844	1835	1824	1845	1836	1840	MO13_1	US63_1	US54_6	US54_1
Strength at -12C	443.3	455.6	495.1	462.5	503.2	509.1	571.0	656.2	641.7	568.0
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.0	35.0	31.0	0.0
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	0.0	0.0	0.0	33.0
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	17.0	35.0	31.0	33.0

Figure 3-9. IDT strength testing results at -12 ℃

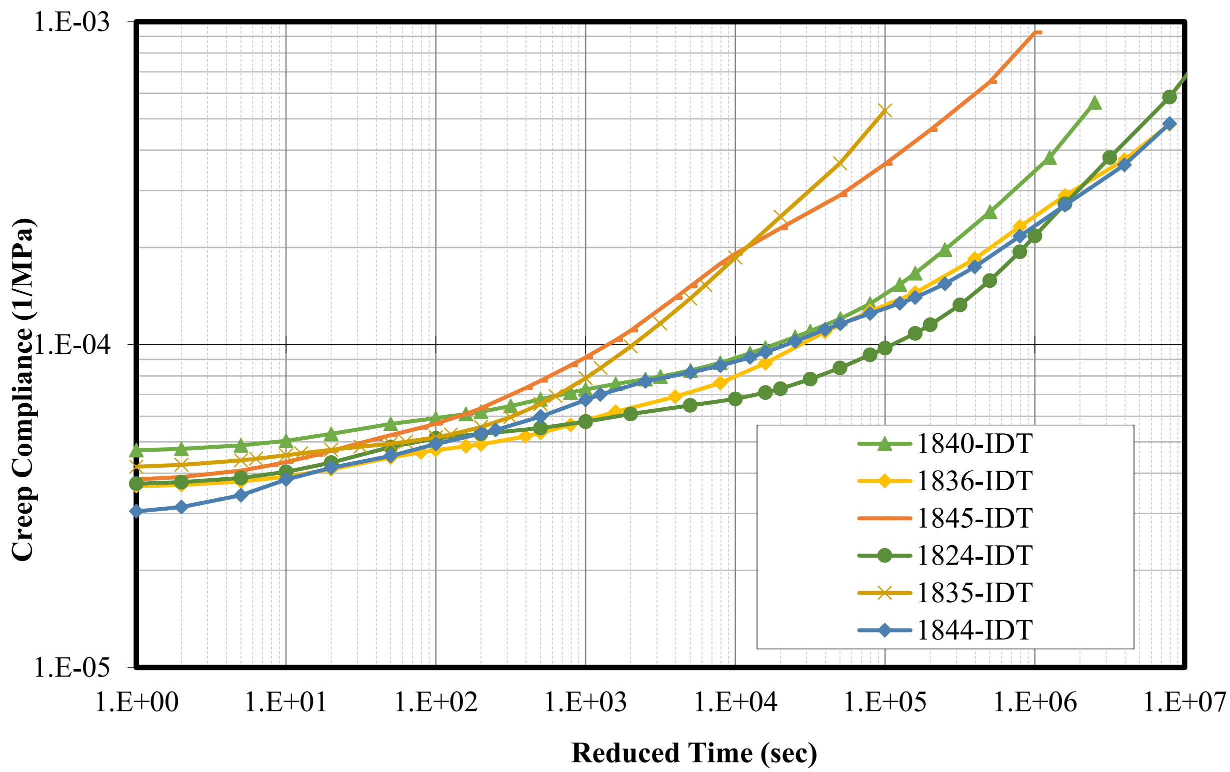


Figure 3-10. IDT creep compliance master curves (Reference temperature: -24 ℃)

3.5. HWTT Testing Results

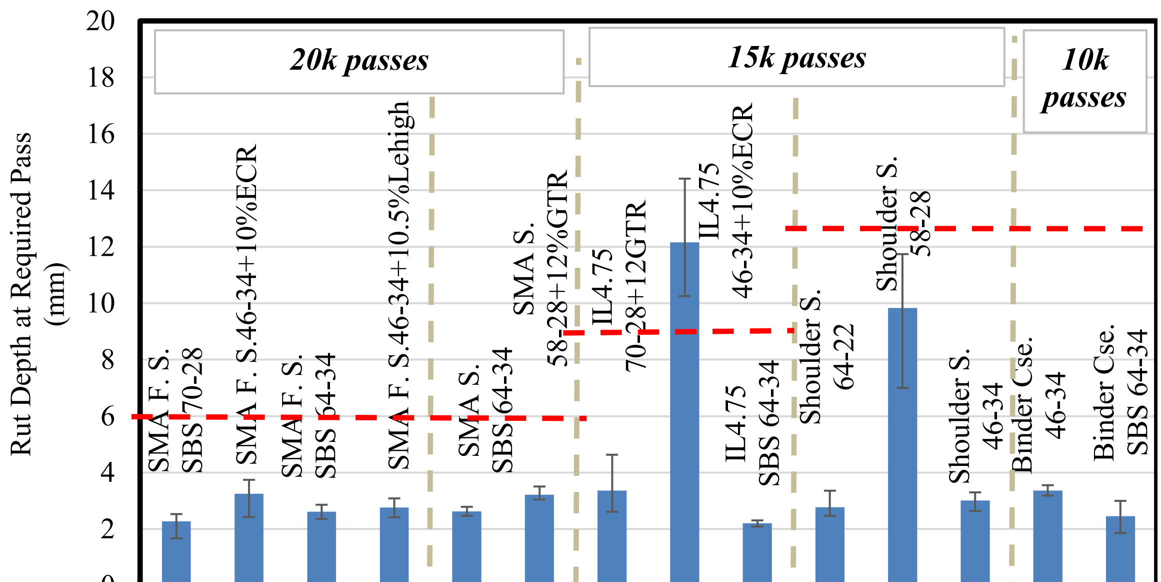
Permanent deformation (rutting) in asphalt pavement is a result of consolidation and shear flow caused by traffic loading in hot weather. This results in a gradual accumulation of volumetric and shear strains in the HMA layers. The measured deformation of different layers of flexible pavement revealed that the upper 100 mm (4 in.) serves the main portion of the pavement rut depth such that the asphalt layer accumulates up to 60 percent of total permanent deformation. Lack of shear strength of the asphalt layer to resist the repeated heavy static and moving loads results in downward movement of the surface and provides the potential for upheaval and microcracks along the rut edges. In addition to structural failure issues, safety concerns arise due to vehicle steering difficulties and the potential for increased hydroplaning. Wheel load tracking (WLT) tests are the most common performance tests used to control rutting potential in HMA mixes. The WLT methods simulate traffic by applying thousands of wheel load passes, simulating traffic loads on HMA specimen in an accelerated fashion at a selected temperature such as 50 ^{o}C.

The two most common WLT test devices are the Hamburg Wheel Tracking Test (HWTT) and the Asphalt Pavement Analyzer (APA) (formerly known as the Georgia-loaded wheel tester). The HWTT is performed in accordance to the AASHTO T324 standard. The vertical deformation of the specimen is recorded along with the number of wheel passes. In addition, conducting the test under water provides the opportunity to measure stripping potential. To this end, the concept of a stripping inflection point (SIP) has been defined and is currently used by state agencies in California, Wisconsin, Iowa, and Missouri. SIP is reported as the number of passes needed to reach the point at which the rutting vs. wheel pass curve displays a sudden increase in rut depth (inflection point in the curve). In this study, the Iowa method has been implemented to calculate the SIP as follows:

- Fit a 6th degree polynomial curve to the rut depth vs. wheel pass curve
- Take the first derivative of the fitted curve
- Determine the stripping line using the tangent at the point nearest to the end of the test where the minimum of the first derivative of the fitted curve occurs
- Determine the creep line using the tangent at the point where the second derivative of the fitted curve equals zero
- Intersect the creep and stripping lines - the wheel pass at which these two lines intersect is taken as the SIP

The Hamburg wheel tracking test was carried out in order to evaluate the rutting susceptibility of the mixtures. As mentioned before, the required number of wheel passes for Tollway SMAs is 20,000 and for shoulder surface mixtures is 15,000. Also, based on the current version of the Tollway asphalt mixture specification, the allowable rut depth at the required number of passes for SMA mixtures is 6 mm and 12.5 mm for shoulder mixtures. The measured rut depth under the Hamburg test along with the requirements for each mixture type is shown in Figure 3-11. From the figure, clearly Tollway SMAs have low rutting levels, as the maximum rut depth recorded was 3.3 mm in mix 1835. This means that the studied SMAs benefit from a robust aggregate structure and binder system, which is consistent with the observed resistance to permanent deformation of similar mixtures placed in the field over the past decade (see section 6).

As for the IL-4.75 mixtures, the 1828 mix recorded the highest rut depth (12.2 mm) under 15,000 wheel passes required for this category (see Figure 3-12). This mix was the only mix that could not meet the rutting requirements among all the Tollway mixtures. However, it is worth mentioning that the IL-4.75 mixtures are not placed on the surface of the pavement, and thus, they do not experience the same environmental and traffic conditions as the SMA mixtures. The lower number of wheel passes required for this category may reflect the fact that these mixtures do not undergo heavy traffic stresses. However, choosing a more appropriate testing temperature and/or adjusting the number of wheel passes to more directly account for the temperature difference between the surface and the binder course depth was addressed in this study, as documented later in this report. In addition, setting less stringent (more appropriate) Hamburg requirements for this category would result in more economic and/or allow more crack resistant mixtures to be designed in a simpler fashion.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
Rut Depth	2.3	3.3	2.6	2.8	2.6	3.2	3.4	12.2	2.2	2.8	9.8	3.0	3.4	2.5
ABR by RAP	10.8	25.0	20.4	23.9	16.2	15.9	17.8	35.3	24.1	20.4	20.0	27.6	34.4	26.5
ABR by RAS	16.0	16.1	16.7	15.4	16.3	9.8	9.3	9.2	14.2	0.0	0.0	18.1	14.0	16.6
Total ABR	26.8	41.2	37.1	39.3	32.5	25.7	27.0	44.6	38.3	20.4	20.0	45.7	48.4	43.1
COV (%)	6.6	21.4	11.4	8.0	9.3	4.0	21.9	14.1	6.2	15.4	18.1	8.2	7.6	13.9

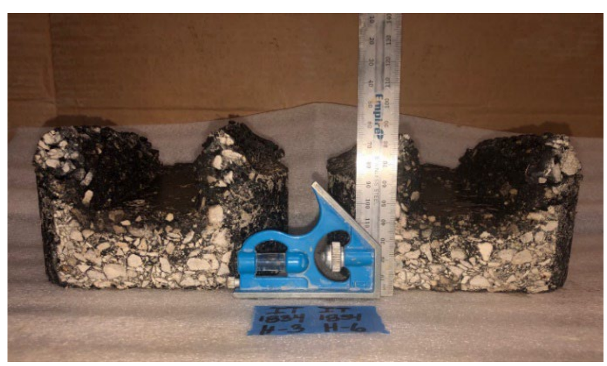
Figure 3-11. Hamburg testing results at required number of passes at 50 ℃

In the shoulder surface mix category, the 1818 and 1826 shoulder mixtures possessed low Hamburg rut depths as compared to the allowable. However, the 1834 mix (see Figure 3-12), which used a softer binder system as compared to 1818, had the highest rut depth (9.8 mm) among the tested mixtures. This shoulder mix is not designed for heavy traffic loads, and the higher Hamburg rut depth opens the door to obtain higher fracture energy due to the softer binder grade while employing a relatively economical mix design. Similar to the IL-4.75s, the number of required load pass is lower for shoulder binders due to the lower stress which is induced in the sublayers of the pavement. The 1807 mix had a slightly higher rut depth which is attributed to the softer binder as compared to the 1803 mix. As the maximum rut depth allowed for this category is the conventional 12.5 mm, in the future, these mixtures could benefit from a softer binder system – especially if new or more stringent fracture requirements are introduced for shoulder mixes as discussed later in this report.

As discussed earlier, the moisture damage potential of the mix can be evaluated through SIP determination. Higher SIP values indicate that the mix can tolerate more wheel passes prior to stripping. A minimum SIP threshold is thus specified for different mix types. Based on the Iowa method, a mix is first identified as potentially stripping when the ratio of the stripping line slope to the creep line slope exceeds 2.0. To illustrate the calculation of SIP, examples for four mixtures are provided in Figure 3-13.

The first example is shown in Figure 3-13-a, which represents the first replicate of the 1835 mix. Recall that in Figure 3-11, the maximum rut depth recorded for this mix was relatively low, at 3.3 mm. However, as Figure 3-13-a shows, the slope of the creep line is very low, such that the slope ratio did in fact exceed 2.0. Thus, the Iowa method determines this mix to have stripping potential, and the SIP was subsequently recorded to occur at 10,744 passes. However, the visual inspection did not show any de-bonding between aggregate and binder, which is generally expected in the case of actual stripping. This example implies that the mathematical process used in the SIP calculation might lead to misleading results, especially in cases where the deformation rate at the end of the densification phase of plastic deformation (i.e., the creep slope) is very low. This makes the denominator of the slope ratio very small and results in relatively higher slope ratios even if the stripping phase followed by densification phase is not problematic (i.e. stripping slope is relatively low). Figure 3-13-b (1845-Rep 1) presents another example of what appears to be an incorrect indication of a stripping-prone mix. However, the 1845 mix did not show visible stripping, where the rut depth recorded at the end of 20,000 passes was also very small (just over 2 mm). In addition to the visual inspection, performing other moisture damage tests such as AASHTO T-283 and the Texas boiling water test were be used to further evaluate the Iowa method based SIP parameter. Figure 3-13-c and d present the rut depths and slopes for the 1828 and 1834 Tollway mixtures, which incurred the highest rut depths among the studied mixtures.

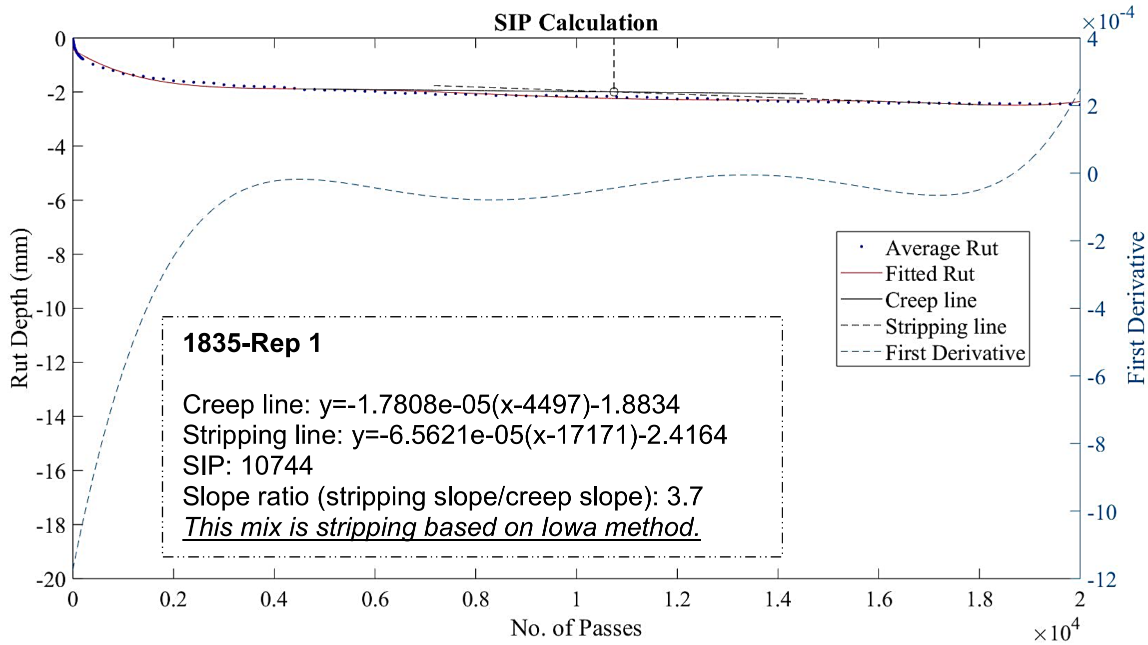
(a)



(b)

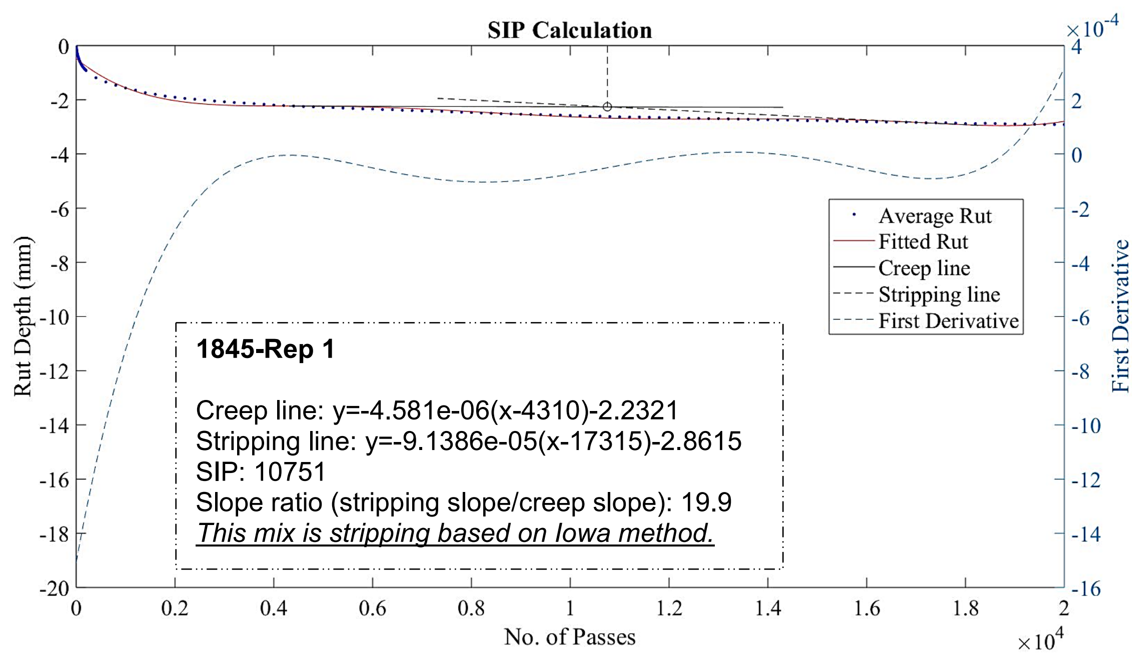
Figure 3-12. Tested samples after 20K pass at 50 ℃. a) 1828 mix, b) 1834 mix

Table 3-2 shows the average of creep and stripping slopes, and the SIP of the plant produced 2018 Tollway mixtures investigated. Of these, five mixtures including 1835, 1845, 1829, 1828, and 1834 had slope ratios over 2.0, indicating stripping potential according to the Iowa method. It is also worth mentioning that although SIP could be calculated for all the mixtures, as long as the slope ratio is lower than 2.0, the mix is not considered as stripping based on the Iowa method. Mixture 1828 was found to be a stripping prone mix, although borderline (9,861 < 10,000).

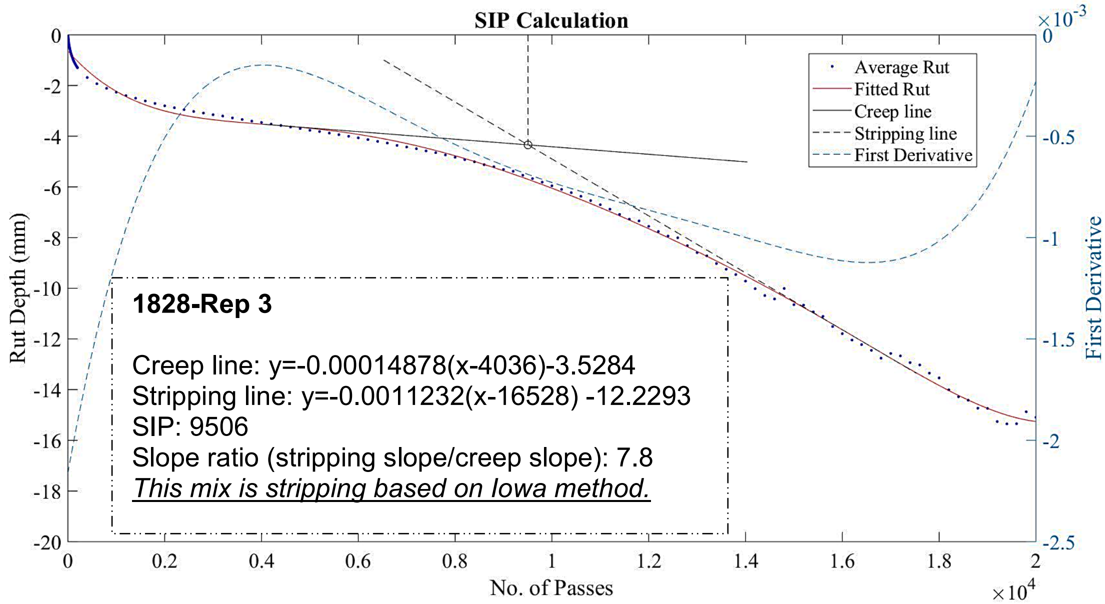


,

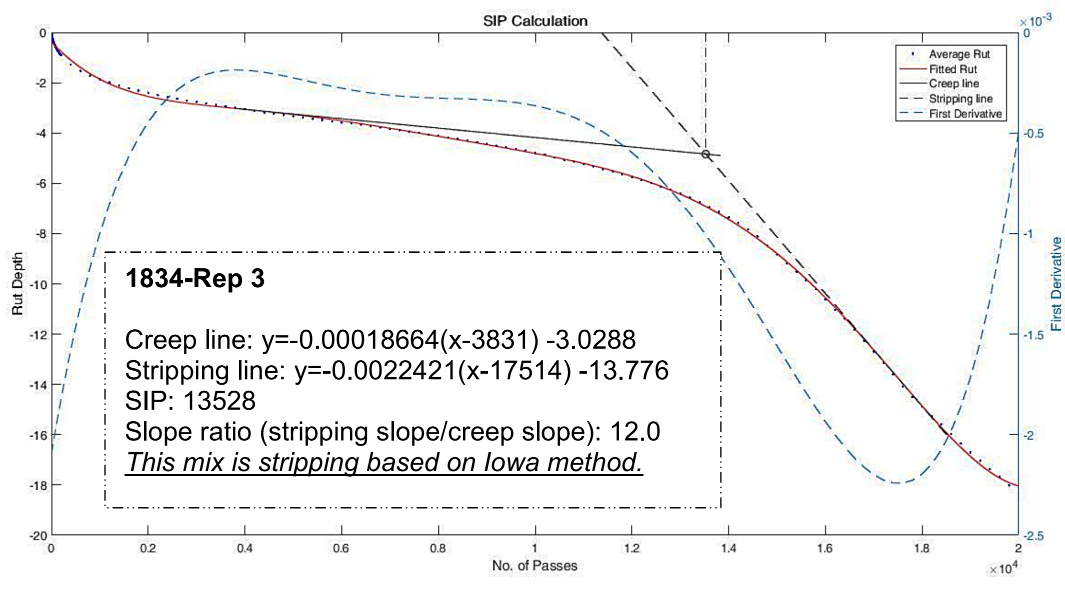
(a)



(b)



(c)



(d)

Figure 3-13. Examples of determination of SIP: a) 1835, b) 1845, c) 1828, d) 1834 mixtures

Table 3-2 also provides the minimum SIP required by the Tollway in their 2019 asphalt mixture specification for different mix categories. For all SMA categories, a minimum SIP of 15,000 passes is used. As highlighted, the average SIP measured on the plant-produced, reheated, lab-compacted specimens for mixes 1835 and 1845 are lower than the required minimum values and were identified as stripping mixtures. On the other hand, since the average SIP of the other three mixtures with slope ratios greater than 2.0 is higher than the required SIP, they are not tagged as stripping mixtures.

Having very low creep and stripping slopes can result in extremely low or even negative SIPs. This phenomenon especially occurs when most of the deformation occurs during the densification phase and both the creep and stripping lines bear similar slopes. Given the average slopes and SIP values obtained for mix 1836 in Table 3-2, it can be seen that the Iowa method resulted in an average SIP of 1711 passes; however, the slope ratio was lower than 2.0 and did not trigger the stripping detection. As shown in Figure 3-14, in one of the replicates of the 1836 mix, the almost parallel creep and stripping lines has shifted the intersection back to negative computed wheel passes at the SIP (obviously not possible). Clearly, the model fitting and numerical steps used in Iowa method for the SIP calculation often fails to work well for mixtures such as SMAs, which experience a negligible rut depth during the densification stage. In addition to negative SIPs, positive creep slope (upward deflection) can also be observed due to curve fitting issues and numerical calculation in the SIP determination. For this reason, it is recommended that for SMA mixtures with very low rut depths, for instance, for those with no greater than 4.0 mm of rutting at 20,000 passes, that the mixture be considered as 'non-stripping' without the need to compute the slope ratio and SIP value.

Table 3-2. Identifying the stripping mixtures based on SIP requirements

Mix. ID	Creep Slope	Stripping Slope	Slope Ratio	SIP	Min. SIP	Status
1844	3.99E-05	6.47E-05	1.6	13430	15000	OK
1835	4.29E-05	9.85E-05	2.8	13562	15000	Stripping
1824	5.62E-05	6.20E-05	1.1	14375	15000	OK
1845	1.98E-05	8.18E-05	10.5	10633	15000	Stripping
1836	6.37E-05	7.32E-05	1.5	1711	15000	OK
1840	5.39E-05	9.45E-05	1.8	12382	15000	OK
1829	6.95E-05	1.60E-04	2.3	12450	10000	OK
1828	2.19E-04	1.56E-03	7.1	9861	10000	Stripping
1823	4.48E-05	7.49E-05	1.7	12565	10000	OK
1818	6.51E-05	1.07E-04	1.6	13107	10000	OK
1834	2.43E-04	1.83E-03	7.5	13149	10000	OK
1826	6.16E-05	1.01E-04	1.7	13608	10000	OK
1807	6.62E-05	1.26E-04	1.9	12084	10000	OK
1803	7.68E-05	9.65E-05	1.3	12222	10000	OK

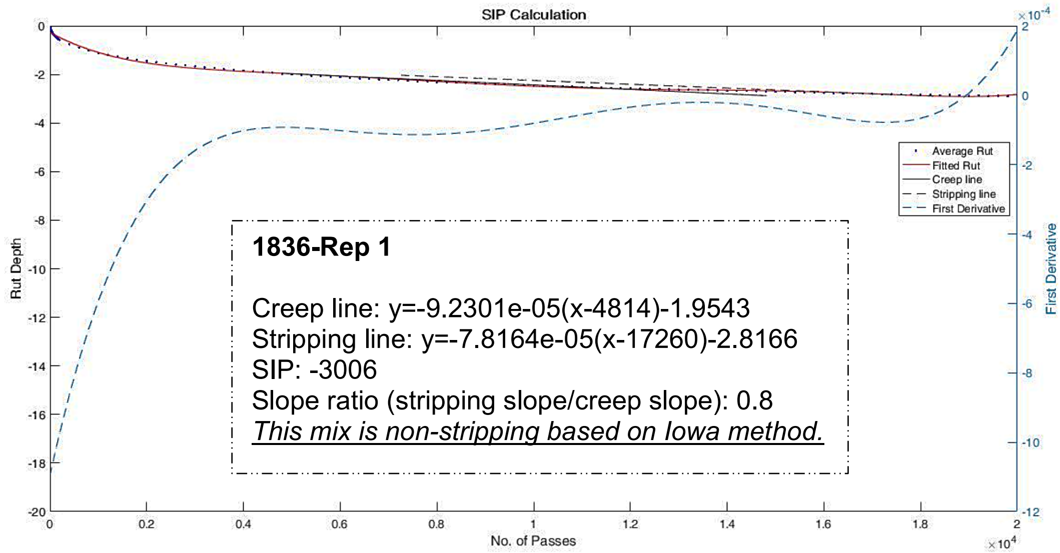
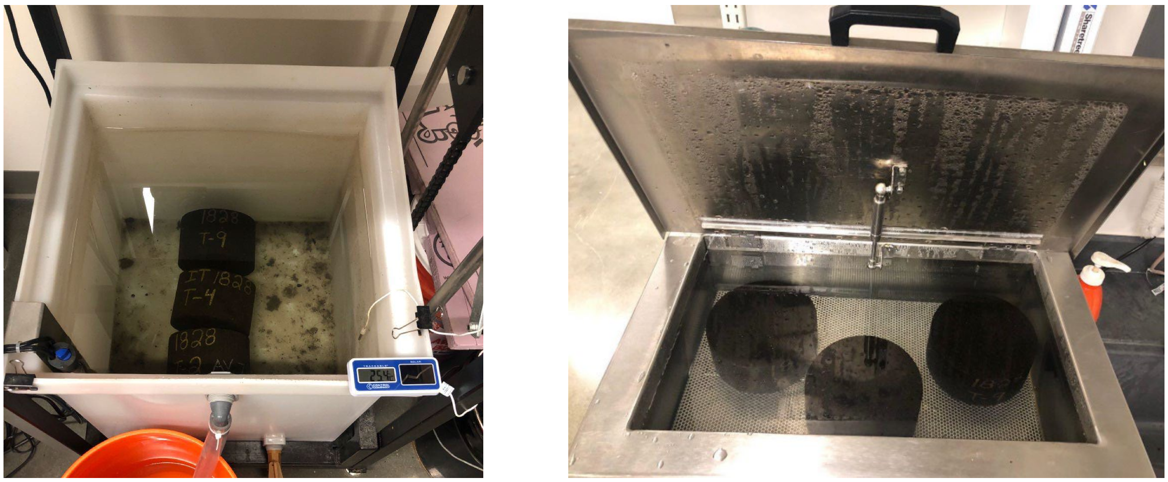


Figure 3-14. Parallel creep and stripping lines for 1836 mix.

3.6. TSR Testing Results

Prior to the introduction of the Hamburg test, the tensile strength ratio (TSR) test was the preferred method to evaluate moisture damage resistance of Tollway asphalt mixtures. The TSR test was conducted using the Illinois modified version (accelerated moisture conditioning) of the AASHTO T283 specification. In this method, the indirect tensile strength ratio of two subsets of samples is be measured and compared to the required TSR threshold. The first subset of samples included at least three replicates of gyratory compacted samples with 95 mm thickness. The air void content of the samples was kept within 6±0.5 % for SMAs and 7±0.5 % for dense-graded mixtures. After being compacted and cooled down at room temperature, the dry samples were conditioned in water bath at 25 ℃ (see Figure 3-15-a) for two hours. After conditioning, the samples were tested to measure the indirect tensile strength. Although the conditioning process is completed in water, this subset of samples is termed the dry subset.

The next subset of samples is called wet subset and includes at least three replicates with the same geometry and air void content as the dry subset. This subset is subjected to a vacuum saturation process, followed by soaking in warm water. To this end, specimens were placed in a vacuum container, supported a minimum of 25 mm (1 in.) above the container bottom by a perforated spacer. The container was then filled with potable water at room temperature so that the specimens had at least 25 mm (1 in.) of water above their surface. A vacuum of 13 to 67 kPa absolute pressure (10 to 26 in. Hg partial pressure) was applied for approximately 5 to 10 minutes. The vacuum pressure was then removed, and the specimen left submerged in water for a short time (approximately 5 to 10 min). The time required for some specimens to achieve the correct degree of saturation (between 70 and 80 percent) may in fact be less than 5 min. In addition, some specimens may require the use of an absolute pressure of greater than 67 kPa. After performing a first run of vacuum saturation trials, the saturation level is measured. If the degree of saturation is between 70 and 80 %, the sample will be ready for warm water conditioning. If the degree of saturation is less than 70 %, another period of saturation is needed. In case the saturation degree is over 80 %, the sample would not be representative and was discarded. After vacuum-saturating the samples, they were placed in warm water at 60 ℃ (see Figure 3-15-b) for 24 hours. After this warm conditioning prior to testing, the samples were placed in water at 25 ℃ for two hours. Finally, the IDT strength of the samples was measured. The minimum acceptable tensile strength is set at 60 psi for mixtures containing unmodified asphalt binders and 80 psi for mixtures containing modified asphalt binders.

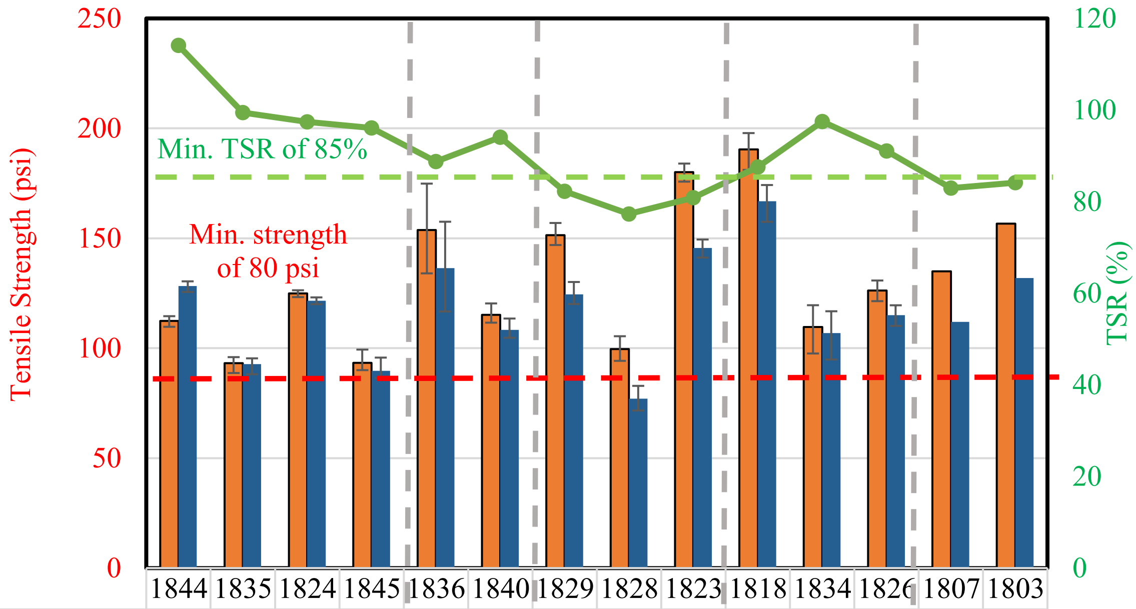


(a) (b)

Figure 3-15. TSR sample conditioning. a) Dry set at 25 ℃, b) Wet set at 60 ℃

Figure 3-16 presents the tensile strength of the sample subsets on the left axis in form of the bars (dry in orange and wet in dark blue colors). Also, the tensile strength ratio (TSR) is shown on the right axis as indicated by the green data points. Clearly, most of the mixtures possessed strength values in both dry and wet conditions in excess of 80 psi, indicating that the IDT strength minimums were easily met. The only mix which could not meet the minimum strength of 80 psi (set for modified mixtures) was 1828. Referring to the Hamburg testing results, the rut depth of this mix was also high, although its SIP was found to meet the requirement. The TSR calculated for this mix is 77.3 which is below the minimum TSR of 85% required in the Tollway specification. Therefore, the 1828 mix showed moisture damage potential in both the Hamburg and TSR tests. Fortunately, the IL-4.75 and shoulder binder mixes, which yielded TSR values less than 85%, are used in the sublayers of the pavement structure and are therefore insulated from the full intensity of stripping and freeze-thaw distress driving mechanisms present on the surface.

The TSR values for the 1835 and 1845 mixtures, which were marked as stripping mixtures by the Iowa method, were more than 96 % with their strength values greater than 90 psi in both wet and dry conditions. The rut depth of these mixtures in the Hamburg test was very low, but the slope ratios were high, and the mixes were detected as stripping by Iowa method. Therefore, these two performance tests do not completely match in terms of detecting the stripping. This finding supports the recommendation of waiving the Iowa SIP calculation and requirement for SMA mixtures with low total rut depths at 20,000 wheel passes. Based on these findings, it is also recommended that the Hamburg wheel track test be the primary test for moisture damage assessment. If the Hamburg test indicates stripping potential, the designer may opt to run the AASHTO T-283 procedure. If the TSR meets the required value, then the mixture can be rated as passing, and considered as non-stripping.



	1844	1835	1824	1845	1836	1840	1829	1828	1823	1818	1834	1826	1807	1803
Dry	112	93	125	93	154	115	151	100	180	190	110	126	134.9	156.7
Wet	128	93	122	90	136	108	125	77	146	167	107	115	112.0	131.9
COV (Dry) %	2.2	4.3	1.3	5.6	13.3	4.0	3.4	5.6	2.3	4.5	10.2	3.7	3.4	3.7
COV (Wet)%	7.1	1.8	3.3	5.0	1.1	2.7	3.9	5.5	7.2	3.2	5.5	4.4	1.2	8.1
TSR	114.1	99.4	97.4	96.1	88.7	94.0	82.3	77.3	80.8	87.6	97.5	91.1	83.0	84.2

Figure 3-16. Wet and dry strengths and TSR values

3.7. Boiling Water Test Results

In order to investigate the discrepancy between the Hamburg and TSR tests, a third test – the Texas boiling water test, was conducted per ASTM D3625-12. In this test, two, 250-gram samples of loose mixture were collected. After warming the samples to about 100 °C, one sample is placed in boiling water (100 °C) and the other sample is kept in water at room temperature for ten minutes. Then, the samples are carefully drained and visually compared with one other. If the bituminous coating of the aggregates conditioned in the boiling water was removed or observed to change in color, the mixture is identified as having stripping potential. In addition, residual material deposited on the wall and bottom of the water container can also help to indicate the separation of binder from aggregates and thus the potential for stripping in the evaluated mix. Figure 3-17 presents the pictures from nine conditioned loose mixtures in both boiling water and room temperature water and the binder residual on the wall and the bottom of the boiling water container. As Figure 3-17-b and Figure 3-17-d show, there is not a significant difference observed between the boiling water conditioned and room temperature water conditioned 1835 and 1845 samples. In addition, the residual remaining on the boiling water container is relatively negligible. As expected, the 1828 mix, which showed stripping potential though SIP and TSR parameter, had higher concentration of residual after being conditioned in boiling water for 10 minutes (Figure 3-17-g). However, due to the very fine particles present, it was difficult to detect any difference between the two differently conditioned samples from this mix with the naked eye. Based on these results and those of the previous sections, the boiling water test does not seem to be worth pursuing at the present time as a simple alternative to the Hamburg or TSR stripping tests. However, from a research perspective, the test provided a useful outside perspective when comparing Hamburg and TSR results.

a)1844 mix

b)1835 mix

c)1824 mix

d)1845 mix

e)1836 mix

f)1840 mix

g)1828 mix

h)1823 mix

h)1803 mix

Figure 3-17. Boiling water test samples and container residue for different mixtures

3.8. Performance Test Repeatability

In the previous section, the capability of various performance tests and associated parameters to mitigate different distresses such as cracking, rutting, and moisture damage has been presented. Error bars were provided on figures to graphically illustrate the variability of the tests with respect to their mean values. In this section, the coefficient of variation (COV) values are summarized and compared to assess the relative repeatability of the tests. Except for I-FIT test, where four replicate specimens were tested for each mix, all other tests involved three test replicates. The COV parameter allows scaling of the standard deviation with respect to the mean value of the result obtained. The COV is computed as the standard deviation of the replicate test results divided by the mean of the test results. Thus, the COV can be interpreted as the standard deviation of the test results expressed as a percentage of the mean. For instance, a 30% COV computation would estimate that 68% of test results would fall within +/- 30% of the true mean value. The dispersion can come from a number of sources, including variability in the material sampled, variability in procuring or producing the sample (gyratory compaction, coring and core procurement and handling/shipping/storage), variability in splitting samples, variability in fabricating the sample, human variability introduced during testing and possibly in data analysis, and finally, variability introduced by the testing device. By comparing similar materials evaluated with different testing approaches, one can obtain a general sense of the relative proportion of the COV that is attributable to the test device versus the inherent variability of the samples being tested. Lower COVs are generally associated with factors such as: fine-grained materials, homogeneous materials, samples and ligament areas larger than their representative volume element (RVE) dimensions, factory-produced materials, low-strain tests, modulus tests, and other highly controlled variables (temperature, loading rate, aging levels, specimen geometry). Higher COVs are generally associated with coarse-grade materials, heterogeneous materials, smaller samples tested below the RVE size, field produced materials, chaotic processes such as fracture or plastic shear flow in heterogeneous materials, and poorly or difficult to control variables. Clearly, our industry has its hands full when considering the realities of our material, our construction environment, and the desire to use simple test geometries and to test small samples with low number of replications when possible.

Table 3-3 and Table 3-4 show the averaged COVs and standard deviations (STDs) of the performance tests (or parameters) for different mix categories. It can be seen that the maximum COV of the DC(T) test for both -12 and -18 °C temperature is less than 17 %. After applying the procedure for omitting the furthest FI value from the average of the four replicates (as recommended by IDOT), the COV of the FI parameter using the three remaining FI values was significantly reduced (in the range of the DC(T) test). The COV of CT parameter is comparable to the FI (four replicate results). The IDT strengths in the TSR test (both wet and dry conditions) yielded the lowest COVs. The Hamburg rut depth COV never exceeded 15 % for the tested mix categories. By far, the SIP parameter had the highest variability when considering two mix categories - SMA surface (131.7%) and shoulder binder (38.6%). In the future, the low repeatability of the SIP parameter for the SMA mixtures will be mitigated by applyihg the prescreening method introduced in Chapter 7. The average COVs for the performance tests are summarized in Figure 3-18. It is worth noting that all of the performance tests were conducted at the Mizzou Asphalt Pavement and Innovation Laboratory (MAPIL). It would be helpful to investigate the repeatability of performance test results between independent labs in the future.

Tables 3-5 present more details regarding the variability associated with parameters obtained from the IFIT and IDEAL cracking tests. As shown, the average COV of the fracture energy (FE) calculated in IDEAL-CT test is 4.6 %, which is much lower than that of the I-FIT test. This is likely due to the larger sample size used, and thus, larger fracture process zone size relative to the inherent material RVE. Moreover, the post peak slope calculated through the IDEAL-CT test is generally less than half as compared to the I-FIT post peak slope. This is a factor explaining why the slope calculated by the IDEAL-CT method is less variable and therefore more reliable that that found in the IFIT. It was found that the COV of the final indices did not vary as much (20.9 vs. 22.7 %) in the case of the plant-produced mixtures.

Table 3-3. Test COV averages for different mix categories

Mix. Type	DC(T) at -12 °C	DC(T) at -18 °C	FI (4 Reps)	FI (3 Reps)	CT	Strength (Wet)	Strength (Dry)	Hamburg (Rutting)	SIP
SMA F. S.	8.9	15.1	23.8	12.7	26.9	4.3	3.3	11.9	14.1
SMA S.	10.1	7.8	20.3	16.6	11.5	1.9	8.6	6.6	131.7
IL-4.75	7.3	14.5	25.2	10.9	15.1	5.5	3.8	14.1	15.0
Shoulder S.	15.8	16.2	21.4	10.4	23.2	4.4	6.1	13.9	12.5
Binder Cse.	14.3	6.9	21.3	15.1	22.0	4.7	3.5	13.2	38.6

Table 3-4. Standard Deviation (STD) averages for different mix categories

Mix. Type	DC(T) at -12 °C	DC(T) at -18 °C	FI (4 Reps)	FI (3 Reps)	CT	Strength (Wet)	Strength (Dry)	Hamburg (Rutting)	SIP
SMA F. S.	71	102	2.2	1.0	39	4.9	3.3	0.3	3871
SMA S.	64	39	1.5	1.2	18	2.2	12.5	0.2	3132
IL-4.75	33	56	2.3	0.8	13	6.5	5.0	0.9	1770
Shoulder S.	71	62	2.0	0.9	20	5.4	8.1	0.8	1744
Binder Cse.	59	23	1.0	0,5	15	6.1	5.2	0.3	4711

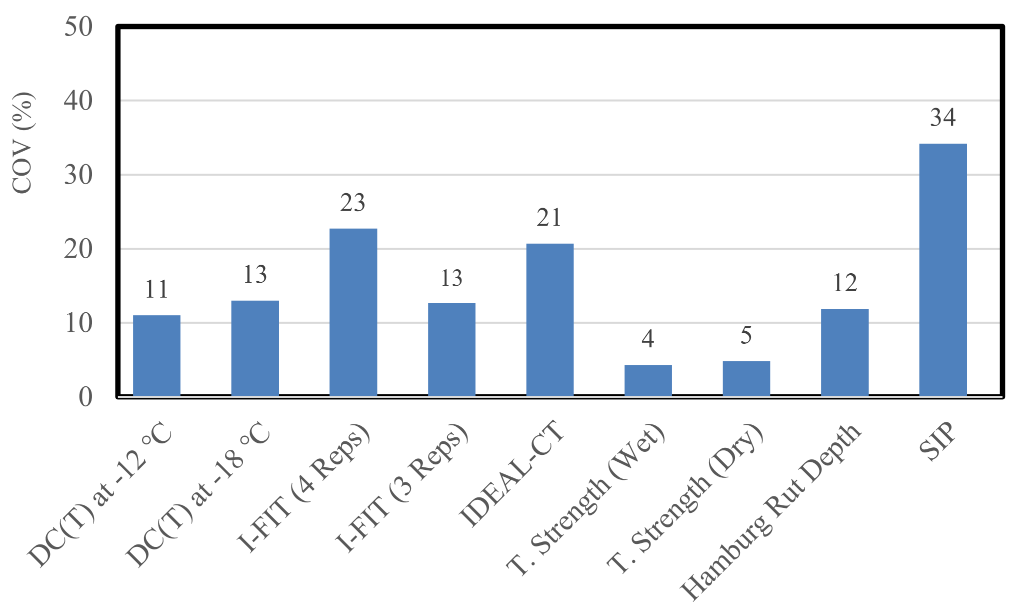


Figure 3-18. Average COV (%) of the tests

Table 3-5. COVs of different parameters obtained from IDEAL CT and I-FIT tests

	IDEAL-CT test				I-FIT test		
Mix. ID	**FE**	**Slope**	**L_{75}**	**CT**	**FE**	**Slope**	**FI**
1844	9.4	11.6	5.2	14.3	16.4	13.7	21.9
1835	4.5	16.8	13.2	37.3	23.1	22.0	14.8
1824	6.1	16.2	8.3	20.8	9.2	5.1	12.3
1845	6.8	13.5	13.4	35.2	18.9	23.0	46.1
1836	2.8	7.9	4.0	8.0	14.2	39.9	27.3
1840	6.4	5.2	6.2	15.0	10.0	24.1	13.3
1829	2.5	6.3	8.7	14.6	8.3	19.1	28.5
828	0.9	5.4	6.6	12.6	7.2	19.0	26.3
1823	1.9	9.2	7.3	18.1	41.1	50.9	20.7
1818	5.0	20.1	10.1	51.8	7.9	14.9	9.6
1834	4.1	10.7	2.6	9.0	5.1	24.1	27.4
1826	4.4	4.9	4.3	8.8	4.6	23.8	27.2
1807	4.5	16.2	6.9	32.5	8.8	22.6	23.1
1803	5.4	13.9	6.9	14.7	9.5	19.8	19.4
AVG	***4.6***	***11.3***	***7.4***	***20.9***	***13.2***	***23.0***	***22.7***

3.9. Performance-space Diagram

Figure 3-19 presents a useful x-y plotting form known as the 'performance space diagram,' or more specifically in this case, the Hamburg-DC(T) plot (Buttlar et al., 2016; Jahangiri et al., 2019). This plot allows the simultaneous evaluation of rutting and cracking behavior. Some useful trends that can often be observed when viewing data in this form are:

- The best overall performing mixtures will appear in the upper-right corner of the diagram (low rutting depth, high fracture energy). These can be considered as high 'total energy' mixtures; i.e., rut and crack (or damage) resistant. These are high toughness mixtures, and the best candidates for surfacing materials especially in demanding climates and for high traffic volumes.
- Mix variables that increase net total energy in the mix and thus 'move' mixtures in the direction of the upper-right corner of the plot include:
    - Higher quality binder (low temperature susceptibility, higher Useful Temperature Interval, or UTI), degree of polymer modification;
    - Higher quality aggregate (stronger, more angular, better bond with asphalt), and;

  - The presence of crack interceptors or rut mitigators, such as fibers, rubber particles, and even RAS (but only if properly used).
- Other salient features of the plot include:
  - Binders with different grades but similar UTI tend to move a mixture along a 'binder tradeoff axis', or roughly speaking, diagonal lines moving in the upwards-left or downwards-right directions, for stiffening and softening, respectively;
  - Pure stiffening elements, such as RAP, tend to move points upwards and to the left;
  - Pure softening elements, such as rejuvenators, tend to move points downwards and to the right;
  - Binders with higher UTI, where the grade bump is on the high temperature grade, tend to move points mainly upwards, but also slightly to the right due to the benefits of polymer in intercepting cracks, and;
  - Binders with higher UTI, where the grade bump is on the low temperature grade, tend to move points mainly to the right, but also slightly upwards, again, due to the benefits of polymer in intercepting cracks.
  - Data points that appear in the undesirable middle-to-lower-left portion of the plot are sometimes those that contain RAP and insufficient binder bumping, and possibly poor bond, where the RAP tended to cause lower DC(T) values, and the nature of the RAP-virgin material combination led to a moisture-susceptible mix with high Hamburg rut depth value.

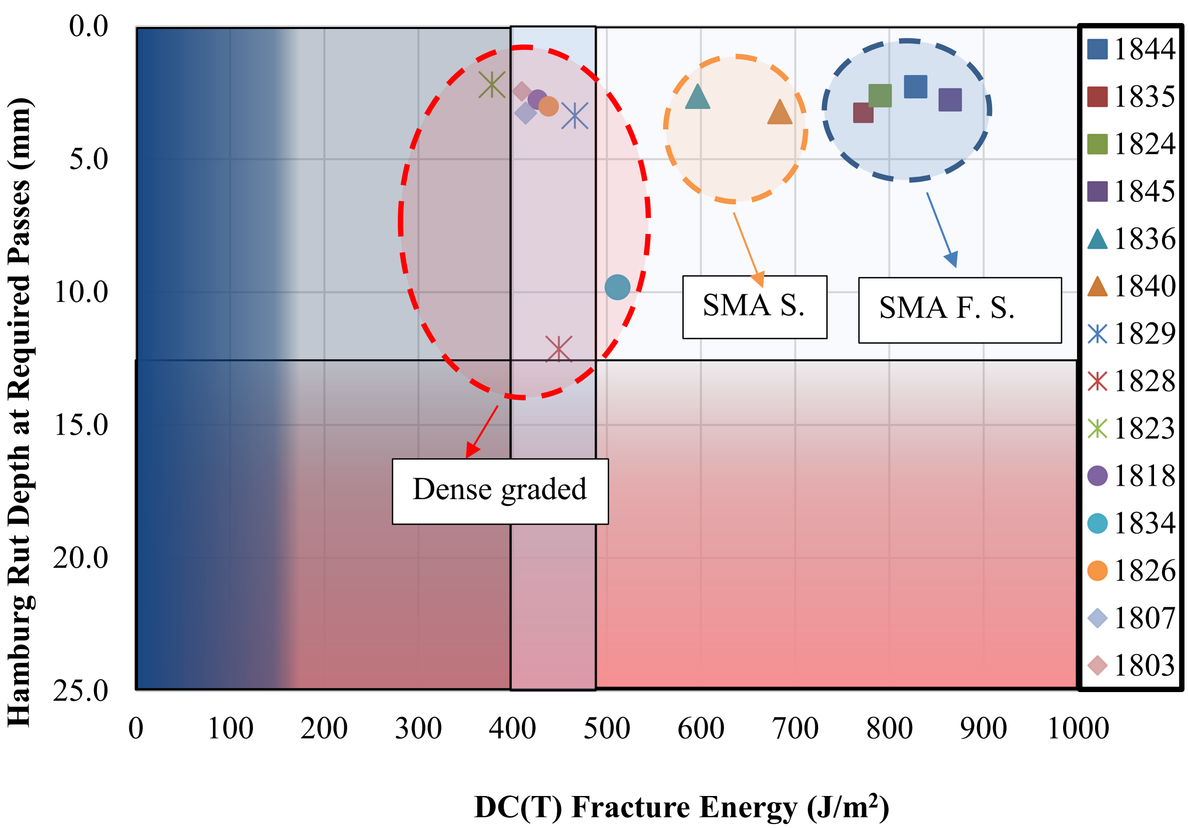


Figure 3-19. Hamburg-DC(T) performance-space diagram for 2018 mixtures

A number of interesting findings can be extracted from the results Tollway 2018 mixtures, including:

- The SMA friction surface mixtures (solid squares) are located on the upper right corner of the performance-space meaning that these mixtures have the highest fracture energy and very low rutting potential.
    - Given the fact that there is still room between the recorded rut depth and the 6 mm threshold, these mixtures could benefit from a softer binder system and/or rejuvenator to further improve the DC(T) fracture energy.

- The two SMA surface mixtures (solid triangles), including 1836 and 1840, are placed on the left-hand side of the SMA friction surface mixtures.
    - The horizontal alignment observed for SMA surface and SMA friction surface mixtures underlines the importance of aggregate quality in DC(T) fracture energy.
    - The rutting and cracking performance of the 1836 mix could be more balanced by employing strategies to soften the mix.

- The IL-4.75 mixtures (asterisks) studied in this project exhibited greatly varying behavior based on Hamburg-DC(T) plot.
    - 1823 mix showed an excellent resistance to rutting, although its DC(T) fracture energy was the lowest among the studied mixtures.
    - The GTR used in 1829 mix along with the fine aggregate structure resulted in a rut-resistant mix that also possessed reasonable cracking resistance.
    - Unlike the 1829 mix, the 1828 mix (the other rubber modified mix in this category) did not perform very well in the Hamburg test. Similar cracking resistance was measured. Lack of room for the swelling of the dry-processed rubber modification in the fine aggregate structure could have possibly led to the poor performance of this mix in the Hamburg test.

- The three shoulder mixtures (solid circles) performed similar to the other dense graded categories. Since the shoulder surface mixtures experience the same environmental conditions as the SMAs do, their DC(T) fracture energy should also be expected to be relatively high. That being said:
    - The 1834 mix recorded a higher DC(T) fracture energy than the other two mixes in this category. However, its rut depth in Hamburg test was high. However, the poor Hamburg performance may not be problematic in practice due to the low load intensity typically experienced on the shoulder.
    - The 1818 mix had a similar amount of recycled materials (~20 %) as compared to mix 1834. However, the base binder used in mix is one PG grade stiffer. In the future, to improve the fracture energy of the 1834 mix, a softer binder system should be considered.
    - The 1826 mix has the highest amount of recycling (~46 %) and accordingly used the softest binder system (PG 46-34) in this category. This mix design strategy seemed to pay off, as the mix is characterized as one of the better overall performers based on the performance-space diagram.
    - None of the shoulder mixtures tested in this project were modified with rubber or SBS. Incorporation of GTR, as a recycled material, might improve the

sustainability and also performance (especially cracking resistance) of these mixtures.

- The two shoulder binder mixtures including 1807 and 1803 (solid diamonds) performed very similarly in terms of their relatively low and high temperature test results.
    - As a result of high levels of recycled materials used (more than 43 % ABR), these mixtures are very stiff and showed negligible rutting in the Hamburg.
    - Similar to the shoulder surface mixtures, softer binders and GTR modification should be considered in future designs.

3.10. Long-term Aging

DC(T) testing has generally been performed on short-term aged laboratory mixes. However, the rheological properties of the binder continue to change during the service life of the pavement, resulting in higher cracking potential due to increased stiffening. The DC(T) was inherently calibrated to account for these differences during its development in the National Pooled Fund Study on Low Temperature Cracking (Pooled Fund Study #776). However, the calibration contained many sections from Minnesota, along with other participating states (mostly northern). This calibration had not been performed for the Illinois Tollway specification prior to this study. Thus, a targeted laboratory and field study was performed towards this end.

AASHTO R 30, which is the most commonly used method to simulate long term aging of the asphalt mixtures, suggests keeping the compacted samples at 85 °C for five days. However, given the highly variable climatic conditions across the US, this description likely does not closely simulate the environmental conditions in Chicago area. In addition, oven aging of the compacted samples could result in non-uniform aging, distortion, and change in air voids of the testing samples. Recent studies have suggested that loose mix oven aging at 95 °C may be the most promising long-term aging method to simulate field aging for asphalt mixtures, at least for research purposes. For example, NCHRP Report 781 generated aging duration maps for mixtures aged in a forced-draft oven at 95 °C (see Figure 3-20) for the U.S. Three field age targets (4 years, 8 years, and 16 years) were selected for the purpose of matching field aging effects at three depths (6 mm, 20 mm, and 50 mm) with oven aged results at 95 °C. The recommended aging protocol was developed by means of a series of laboratory experiments on field cores and asphalt binders along with a system to select the aging index properties (AIPs) that were integrated with pavement aging models.

3.10.1. Methodology for Tollway DC(T) Calibration to Account for Long-Term Aging

The limited literature available suggested using a 15% increase in DC(T) fracture energy thresholds on short-term aged specimens during mix design to account for the eventual fracture energy loss expected during long-term aging (Braham et al., 2009). However, this fracture energy reduction was recommended for one specific type of hot mix asphalt (HMA) and did not cover the various mix types used by the Tollway. To validate the recommended value and also to establish the aging characteristics for all of the Tollway mix types, University of Missouri (UM) researchers attempted to apply the NCHRP aging protocol on the loose mixtures to simulate eight years of aging on Tollway pavements placed in the Chicagoland area. In addition to DC(T) fracture energy, the sensitivity of FI parameter to aging was also studied.

Table 3-6 shows the duration of oven aging needed to simulate different field aging times. For this project, the surface mixtures were aged for 6 days in a forced-draft oven at 95 °C (see Figure 3-21) to account for eight years of in-situ aging. It should be noted that the plant-produced mixtures were sampled in mid-2018 and were kept in the storage facility until late 2019. However, any steric (thixotropic) hardening occurred would likely be reversed by sample heating and stirring.

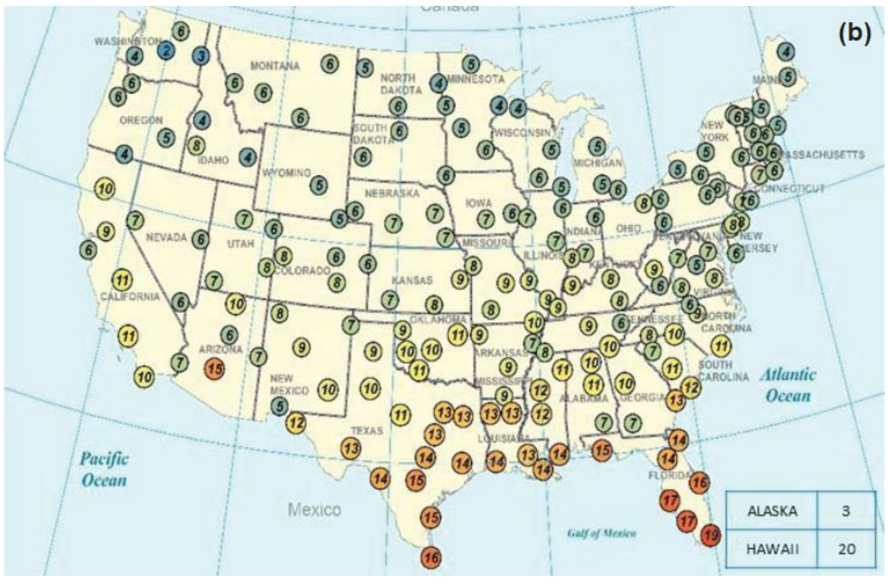


Figure 3-20. Required oven aging durations at 95°C to match level of field aging 6 mm below pavement surface for 8 years of field aging (Kim et al., 2018)

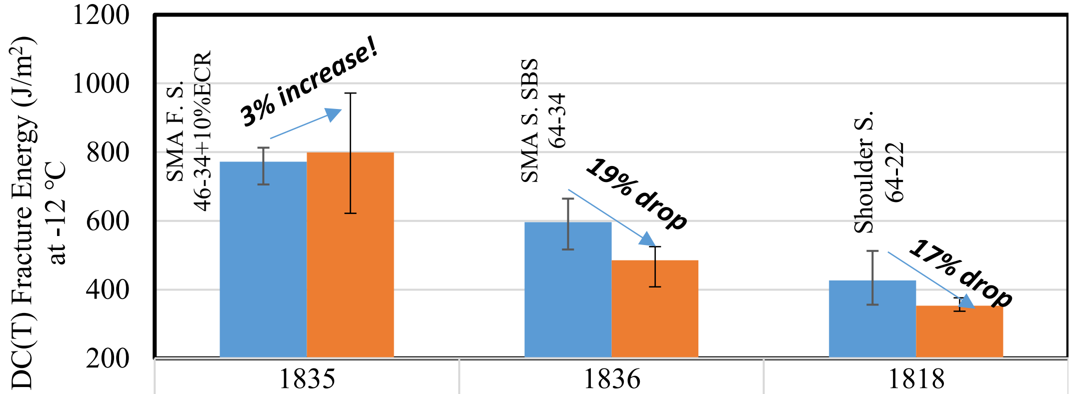
Table 3-6. Illinois oven aging duration based per NCHRP-871 “(6 mm below the surface) (Elwardanya et al., 2018):

Field Aging Time	**Oven Aging Time at 95 ℃**
4 years	3 days
8 years	***6 days***
16 years	12 days

Figure 3-21. Aging the Tollway 2018 plant produced mixtures in oven at 95 ℃ for 6 days

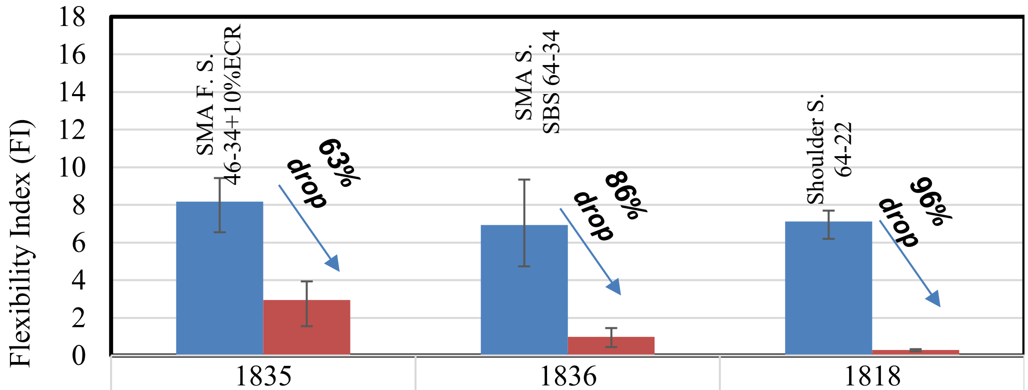
3.10.2. Testing Results

- Surprisingly, mix 1835 which is an SMA friction surface mix modified by dry-process GTR did not experience a drop in DC(T) fracture. This mix also had the smallest drop in FI score (63% drop) among the tested mixtures. Possibly the combination of GTR (containing carbon black, an antioxidant) along with a high amount of pre-aged recycled materials (ABR=41.2%) in this mix led to the relatively stable aging behavior.
- Mix 1836, which is an SBS-modified SMA surface mix, was measured to have a 19% decrease in fracture energy. However, the FI score underwent a decrease of 86% upon this aging. The DC(T) fracture energy 'bump' inherently considered in the current Tollway specification is thought to be around 15%, which is very close to aging effect measured in this mix.
- Mix 1818, which is an unmodified shoulder surface mix, experienced 17% drop in DC(T) fracture energy after aging. This mix had the highest drop in FI score (96%). It is worth mentioning that the peak load in I-FIT exceeded 10 kN for this mix and a snap-back (very brittle) behavior was observed for this mix such that the slope of Load-Deflection curve could not be properly calculated.



	1835	1836	1818
Gf (STA)	772	596	427
Gf (LTA)	799	485	353
ABR by RAP	25.0	16.2	20.4
ABR by RAS	16.1	16.3	0.0
Total ABR	41.2	32.5	20.4

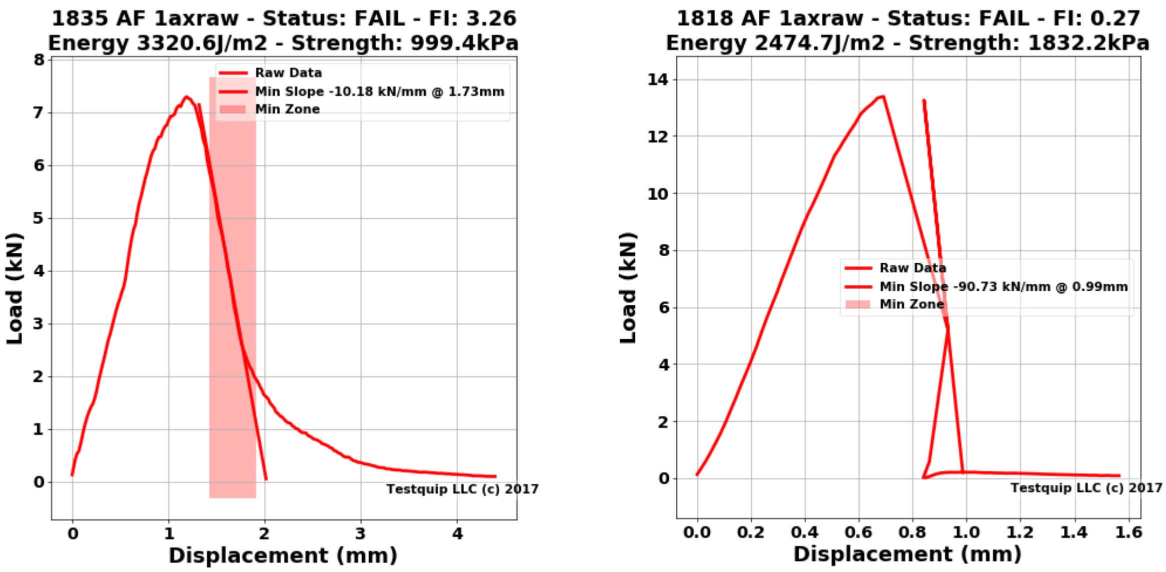
Figure 3-22. Comparing the DC(T) fracture energies of short-term aged (STA) with long-term aged (LTA) samples



	1835	1836	1818
FI (STA)	8.2	6.9	7.1
FI (LTA)	3.0	1.0	0.3
ABR by RAP	25.0	16.2	20.4
ABR by RAS	16.1	16.3	0.0
Total ABR	41.2	32.5	20.4

Figure 3-23. Comparing the FI of short-term aged (STA) with long-term aged (LTA) samples

Some of the tested aged samples (e.g. the 1835 mix) showed normal load-displacement curves (Figure 3-24a) while there were mixtures such as 1818 that had straight-line (due to a fast moving crack that outpaced the data acquisition rate used in the test) and snap-back shaped load-displacement plots (Figure 3-24b) with peak loads in excess of 10 kN. This behavior led to a very low FI value, less than 1.0, and in some cases, nearly zero. These specimens exhibited very brittle behavior in the I-FIT test, with snap-back type softening curves and very few data points following the peak load, indicating a very brittle failure. Analysis of data sets with very steep post-softening curves is not adequately described in the test specification, and requires analyst judgement. These mixes often possess the highest variability between test replicates. These observations underscore the difficulty in using the I-FIT test with respect to mix specification calibration on long-term aged materials.



(a) (b)

Figure 3-24. Load-displacement response for aged mixtures under I-FIT testing: a)1835 b)1818

Chapter 4

SITE VISIT AND CORING PLAN

4.1. Overview

Development of a truly performance-based specification requires comprehensive laboratory testing combined with extensive field performance data. Besides overall condition, details regarding the type, extent and severity of individual distresses should be considered when trying to control rutting, cracking and moisture damage. This can be achieved through site visits (visual inspection), automated data collection vehicles or preferably both. Both types of data were used herein to assist in updating the Tollway's asphalt mix design performance test thresholds.

The plant-produced asphalt mixtures studied in this project were used to pave different sections in the Tollway road system during the summer of 2018. The sampled mixtures were used to fabricate testing samples and evaluate the efficiency of the performance tests and to assess the expected performance in these mixtures. Various performance tests were carried out, and the results were presented in the previous chapter. In order to observe the service quality of the mixtures in-situ, the MU team had a two-day site visit from May 30th to May 31st, 2019. After finalizing a location of the targeted sections and the milepost ranges in a meeting with the TRP subcommittee, and the condition of the sections was visually observed. Although the visited sections did not age considerably, they experienced a record-breaking winter and severe cooling events at the beginning of 2019 that provided an opportunity to reveal any poor performing mixtures in terms of low temperature cracking.

In addition to the 2018 overlaid sections, other good and bad performing sections were located and observed as follows. Most of these selected sections were already studied in previous projects, and their laboratory performance data either on field cores or plant produced mixtures were available. The list of the projects from which the sections were selected as follows.

- Illinois Tollway I-88 Ground Tire Rubber Test Sections: Laboratory Mix Designs and Performance Testing- Report Published in 2017
- Laboratory Investigation of Illinois Tollway Stone Matrix Asphalt Mixtures with Varied Levels of Asphalt Binder Replacement- Report Published in 2016
- Characterization of Hot Mix Asphalt Containing Post-Consumer Recycled Asphalt Shingles and Fractionated Reclaimed Asphalt Pavement- Report Published in 2010

4.2. 2018 Overlaid Sections

- Route 355: The 1844 mix on mainline and the 1834 mix on shoulder
    - Mile post range: 12-22

   - Observation on mainline: frequent fat spots, generally a result of the production issue, not an over-asphalted mix- Some reflective cracks (see Figure 4-1)
   - Observations on shoulder: Infrequent reflective cracking-Open longitudinal joint between outer lane and shoulder- Occasional bumps

(a) (b)

Figure 4-1. Pictures from I-355: mix 1844, mile post: 17 and 18 on mainline

- Route I-355: The 1826 mix on shoulder
   - Mile post range: 22-30
   - Observation: No noticeable distress

- Route I-88: The 1840 and 1829 mixes on mainline, the 1807 mix on Shoulder
   - Mile post range: 123-103
   - Observations: Very low number of transverse (low temperature) cracks (see Figure 4-2)- Some periodic hairline cracks on the shoulder

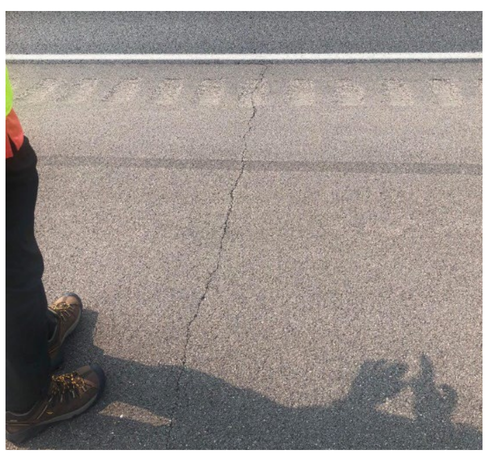


Figure 4-2. Picture from I-88: mix 1840 on mainline and mix 1807 on shoulder

- Route I-88: The 1835 and 1828 mixes on mainline (WB), the 1845 mix on shoulder
   - Mile post range: 103-93

- Observations: Very low number of transverse (low temperature) cracks (see Figure 4-3)

Figure 4-3. Picture from I-88 at mile post 101 (WB)

- Route I-88: The 1836 and 1823 mix on mainline (WB)
  - Mile post range: 91-76
  - Observations: Transverse cracks ~ 100 ft spacing (see Figure 4-4)

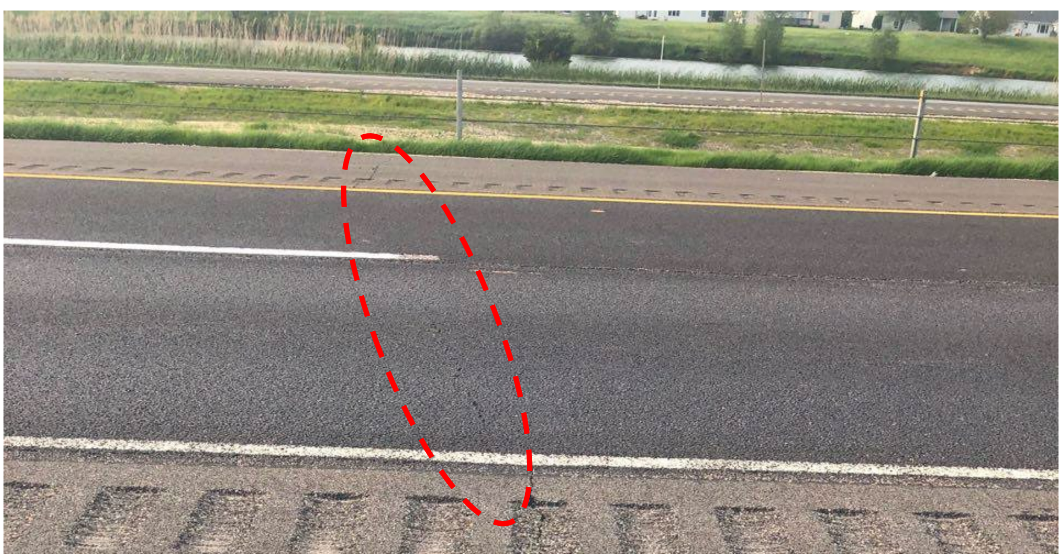


Figure 4-4. Picture from I-88 at mile post 77 (WB)

4.3. Ground-Tire Rubber Test Sections (Report Published in 2017)

The Illinois Tollway constructed test sections for three Ground Tire Rubber (GTR) asphalt modifier technologies on the Reagan Memorial Tollway (I-88) in April 2016. Apart from estimating the performance characteristics of the new GTR technologies, the study also examined the effect of softer virgin binder and an increased amount of reclaimed asphalt on mix performance properties. Accordingly, the GTR technologies were incorporated into SMA mixes with 33% asphalt binder replacement (ABR) using a 'standard' base or virgin binder (PG 58-28) and a softer base binder (PG 46-34). A third design was also used, where the softer base binder was combined with an increased asphalt binder replacement (ABR) percentage (PG 46-34 with 47% ABR), obtained by increasing the content of recycled asphalt shingles (RAS).

- Route I-88: The 1636 mix (Elastiko PG 46-34 High ABR) on passing lane (EB)
    - Mile post: 61.0
    - Observations: Some thermal cracks were observed- Cores were taken from the mainline. Some cracks stopped once they reached the rubber modified mix on the inside lane (see Figure 4-5-a)

Figure 4-5. Pictures from I-88 at mile post 61 (EB shoulder and mainline)

- Route I-88: The 1631 mix (Evoflex PG 46-34 High ABR) on outside shoulder (EB)
    - Mile post 65.9
    - Observations: SMA shoulder- Many cracks in asphalt (Figure 4-6).

Figure 4-6. Picture from I-88 at mile post 65.9-High ABR rubber modified asphalt (RMA) mix (EB shoulder and mainline)

4.4. SMA Study (Published in 2016)

In order to maximize the environmental and economic benefits of RAP, RAS, and GTR, innovative pavement agencies and mix designers tend to utilize these recycled products in various combinations to reduce virgin asphalt and aggregate content to the maximum extent possible, leading to significant cost savings and enhanced sustainability. In general, SMA surface mixtures containing high percentages of asphalt binder replacement (ABR) from RAP/RAS would be more susceptible to thermal and block cracking as compared to virgin asphalt mixtures, unless specific measures are taken to counterbalance the recycled materials with a softer virgin binder base grade and/or through the use of a rejuvenating-type modifier. Such countermeasures have been taken in the design of Tollway high-traffic, stone matrix asphalt mixtures; however, the design of theses mixtures pre-dated the existence of modern low temperature mixture cracking tests. In addition, the Illinois Tollway made an early move to a lower design voids target in an effort to enhance mixture durability when recycled materials are used. The primary objectives of this study were to evaluate the low temperature characteristics and expected performance of cores obtained from seven Tollway projects constructed between 2008 to 2012 using stone-mastic asphalt (SMA) mixtures with varying ABR levels and virgin materials.

- Route I-294: Mix G (PG 70-28 SBS), overlaid in 2012
    - Mile post range: 25-27
    - Observations: Some potholes- Reflective cracking (some were skewed)- Fat spots- Rough ride (see Figure 4-7)
    - Heavy traffic load (see Figure 4-8)

Figure 4-7. Picture from I-294, mix G, mile post range: 25-27

Figure 4-8. Heavy Truck Traffic on I-294, near accident site

- Route I-90: Mix A (on WB) and mix B (on EB)
    - Mile post range: 2-15
    - Observations: Mostly longitudinal and joint cracks

(a) (b)

Figure 4-9. Pictures from I-90 route: a) Mix A (Gravel), WB, MP:3 ¾, b) Mix B (Diabase), EB, MP: 7 ¼

4.5.RAS Test Section on I90-Shoulder (Published in 2012)

In the summer of 2009, a field demonstration project was conducted by the Illinois Tollway on the Jane Addams Memorial Tollway (I-90). Eight mix designs containing zero or five percent RAS and varying percentages of FRAP were developed and placed in the pavement shoulder. With more transportation agencies studying the options of adding RAS or using higher amounts of RAP through fractionation, the Tollway became interested in adopting these techniques in their construction specifications. The objective of this new research was to determine how replacing five percent of the FRAP in these new mixes with five percent post-consumer RAS would affect the performance of asphalt pavements. Figure 4-10 provides sample images from the shoulder and Figure 4-11 presents the properties of the mixtures used on the shoulder with the mile markers and description for each section superimposed on the plan view.

(a) (b)

Figure 4-10. Pictures from I-90, WB shoulder, a) MP:4 ¼, b) MP: 5 1/2

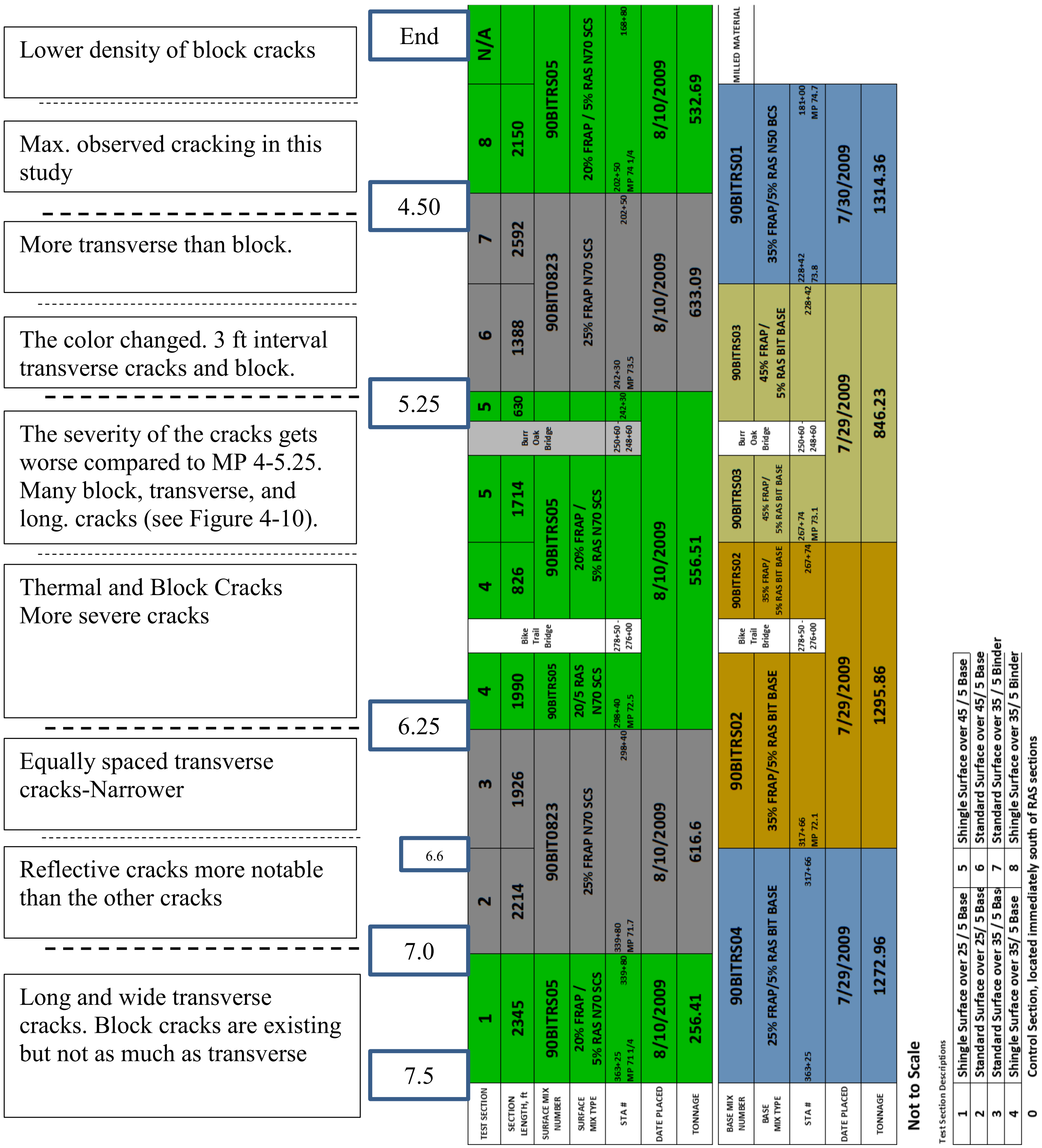


Figure 4-11. Plan view and notes on the 2011 study section

4.6. Other Sections

- Route I-90: SMA mix, overlaid in 2018, GTR modified
    - Mile post range: 16.5-17.9 EB, and I-90 west to I-39 ramp
    - Observations: High density block cracking on the SMA mix (observable when walking, less noticeable when driving), ride is still reasonably good

(a) (b)

Figure 4-12. Pictures of a) I-90 west to I-39 ramp and b) I-90 MP 17.0

Based on these observations, a list of good and poor performing sections was prepared (see Table 4-1). In this list, sections from mainlines (SMAs) and shoulders (dense graded) with different levels of age ranging from three to eleven years of service life were selected. Obtaining field cores and testing the laboratory performance of these sections were the next phase carried out in this study. This table also shows the location and number of cores obtained from each section along with a short description of the distresses observed on each section.

4.7. Cored Sections and Mixture Properties

Previously, the selected sections for the field core investigation were introduced. These sections were selected to cover the wide range of the mixture types that Tollway used on both mainline and shoulders. Also, considering the good- and poor performing sections could help with the specification calibration. It is also worth mentioning that these sections cover a wide range of service life and most of the sections have experienced at least eight years and many cold events. Wang Engineering collected 51 full-depth cores and 81 partial-depth cores at various locations along I-88, I-90 and I-294 shown in Table 4-1 in July and August 2019 (see Figure 4-13). The cores were obtained using a coring machine equipped with a 6.0-inch diameter core barrel. Figure 4-14 and Figure 4-15 show the field cores in the storage Wang's facility and the transferring of these cores to MAPIL using a truck, respectively.

Table 4-1. List of the selected sections for coring and their distress description

Route	***Lane***	***Dir.***	***Mile Post Range***	***Total No. of Cores***	*No. of Mid-Depth**	*No. of Surface layer***	***Description***
I-88	Inside lane-Tangents	EB	45.00-55.10	**12**	3	9	On rubblized JCP
I-88	Inside lane-Tangents	EB	61.30-60.10	**12**	3	9	Control-SBS- located between the rubber sections
I-90	Mainline-Tangents	WB	15.00-2.00	**12**	3	9	Gravel- Aged SMA- construction joints- crack sealants
I-90	Mainline	EB	15.00-2.00	**12**	3	9	Diabase- Aged SMA
I-90	I-90 Ramps	EB	17.80-16.50	**12**	3	9	High density block cracking
I-294	Mainline	NB	30.50-36.50	**12**	3	9	Quartzite mix- Reflective cracking- Rough ride-Placed on jointed concrete
I-88	Shoulder	EB	45.00-55.10	**12**	3	9	Visually good performing
I-88	Shoulder	EB	55.10-60.00	**12**	3	9	Poor performing-transverse cracks, low severe block cracks.
I-90	Shoulder	WB	7.50-7.00	**6**	3	3	Poor performing-transverse and block cracks
I-90	Shoulder	WB	6.60 6.25	**6**	3	3	Poor Performing- severe transverse and block cracks
I-90	Shoulder	WB	6.25-5.25	**6**	3	3	Poor performing-3 ft. interval transverse cracks and block cracking
I-90	Shoulder	WB	5.25-4.50	**6**	3	3	Poor performing- more transverse cracks than block
I-90	Shoulder	EB	9.50-10.50	**6**	3	3	Good performing
I-90	Shoulder	WB	9.50-10.50	**6**	3	3	Poor performing-transverse cracks

*Mid-depth core: 6" deep- **Surface layer core: 3-4" deep

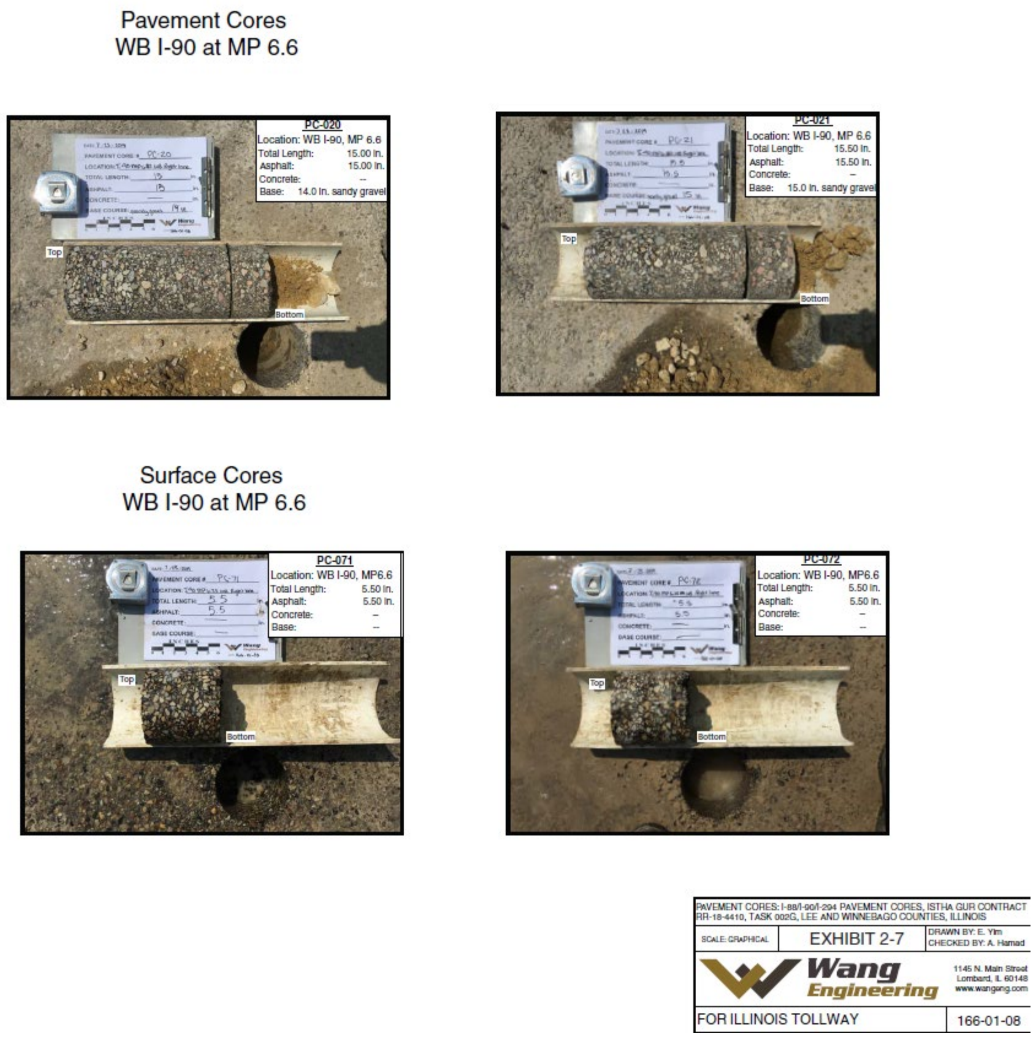


Figure 4-13. Example of the field core pictures and details from Wang Engineering Report

Figure 4-14. Field cores located in storage facility

Figure 4-15. Transferring the field cores from storage facility to MAPIL

The details about those cored sections including the location of the core samples, mixture properties on first two lifts of the cores, and also the overlaying year will be discussed in this section. Table 4-2 shows the properties of the cores' top lift. These cores were obtained from I-88, I-90, and I-294. The first six mixtures are SMAs including both SMA friction surface and SMA surface. The next six mixtures are shoulder sections located on I-88 and I-90. The description of each column is provided below.

- "Location" column, which will be used as the label of the sections later, includes the route and also a number that represents the mile post range and is unique for each section.

- "Direction" column specifies the traffic direction on the cored section. For instance, EB stands for East Bound.

- "MP Range" represents the mile post range on the corresponding section.

- "Mix ID" is the label of the mixture mentioned on the Job Mix Formula sheet.

- "Year" shows the overly time (year) of the section.

- "Base Binder" is the binder system including the PG grade and the modification.

- "Mix Type" can be SMA surface, SMA friction surface, or N70 dense graded mix.

- "NMAS" and "ABR"s are the nominal maximum aggregate size and asphalt binder replacement by RAP or RAS, respectively.

As mentioned in the previous chapter, the performance-based specification covers various types of mixtures located within different levels of the pavement structure. Therefore, in addition to the top lifts, it was attempted to study the bottom lifts of the field cores and evaluate their performance. The test results for the bottom lifts will be used to calibrate the spec for the corresponding mixture types. The details of bottom lifts of the field cores were collected from JMF and presented in Table 4-3. The bottom lifts studied in this project are from seven sections. The first three mixtures are the SMAs used on the bottom lift, and the next four sections are N50 dense graded shoulder binders. The headings used in this table are similar to the ones used in Table 4-2.

Table 4-2. Mixture properties for the cored sections (Top Lift)

No.	Location	Direction	MP Range	Mix. ID	Year	Base Binder	Mix. Type	NMAS	ABR by RAP	ABR by RAS	Total ABR
1	**I88-47**	EB	45-55.1	90WMA1649	2016	SBS 70-28	SMA Surface	12.5	11.6	19.8	31.4
2	**I88-60.5**	EB	60.1-61.3	90WMA1528	2015	SBS 70-28	SMA Friction S.	12.5	14.8	20.7	35.5
3	**I90-6.6**	WB	2.0-15.0	90BIT0941	2009	76-22+ GTR	SMA Surface	12.5	13.9	0	13.9
4	**I90-6.0**	EB	2.0-15.0	90BIT0851	2008	70-28+ GTR	SMA Friction S.	12.5	16.3	0	16.3
5	**I90-17.8**	EB	16.5-17.9	90BIT0859	2008	76-28+ GTR	SMA Friction S.	19.0	16.0	0	16.0
6	**I294-34**	NB	30.5-36.5	90BIT1218	2012	SBS 70-28	SMA Friction S.	19.0	15.5	16.2	31.7
7	**I88-52**	EB	45.0-55.1	90WMA1531	2015	58-28	N70D Surface	9.5	19.1	19.6	38.6
8	**I88-57**	EB	55.1-60.0	90WMA1450	2014	58-28	N70D Surface	9.5	22.8	17.8	40.7
9	**I90-7.25**	WB	7.0-7.5	90BITRS05	2009	58-22	N70D Surface	9.5	16.7	20.1	36.8
10	**I90-5.12**	WB	4.0-5.25	90BIT0823	2009	58-22	N70D Surface	9.5	24.4	0.0	24.4
11	**I90-10E**	EB	9.9-10.1	90BIT0842	2008	58-22	N70D Surface	9.5	24.0	0	24.0
12	**I90-10W**	WB	9.9-10.1	90BIT0819	2008	58-22	N70D Surface	9.5	16.2	0	16.2

Table 4-3. Mixture properties for the cored sections (Bottom lift)

No.	Location	Direction	MP Range	Mix. No	Year	Mix. Type	Base Binder	NMAS	ABR by RAP	ABR by RAS	Total ABR
1	**I90-6.6**	WB	2.0-15.0	90BIT0941	2009	SMA Surface/Binder	76-22+ GTR	12.5	13.9	0	13.9
2	**I90-6.0**	EB	2.0-15.0	90BIT0831	2008	SMA Binder	76-22+ GTR	12.5	15.3	0	15.3
3	**I294-34**	NB	30.5-36.5	90BIT1216	2012	SMA Binder	SBS 70-28	12.5	17.1	19.2	36.3
4	**I90-7.25**	WB	7.0-7.5	90BITRS04	2009	N50 Binder	58-22	19.0	21.7	22.4	44.0
5	**I90-6.06**	WB	6.0-6.12	90BITRS02	2009	N50 Binder	58-22	19.0	32.9	24.2	57.1
6	**I90-5.12**	WB	5.0-5.25	90BITRS03	2009	N50 Binder	58-22	19.0	42.1	23.7	65.8
7	**I90-4.75**	WB	4.5-5.0	90BITRS01	2009	N50 Binder	58-22	19.0	31.2	23.7	54.9

4.8. Summary of Field Observations

In this two-day site visit, the MU team observed numerous sections both on mainline segments and shoulders. These sections were mainly paved using the asphalt mixtures that were previously tested through different projects. The most common distress observed on the surface of the roads was transverse cracking which calls for extra attention to selecting the appropriate cracking test and then setting the thresholds for the test output. A summary of the field observations is categorized based on the projects as follows.

<u>Sections in Performance-Based Specification Project (2018 mixtures)</u>

- Mixtures used in the 2018 study on the mainline did not have considerable transverse cracking. That being said, some reflective cracks were observed after the record-cold winter of 2018-2019.
- Mix 1845 which was used on the shoulder (Lehigh rubber mix) has begun to show thermal cracking after the 2018-2019 harsh winter.

<u>Sections in Rubber Study</u>

- All the mainline sections constructed in 2016 were performing very well (only a few, isolated thermal cracks were observed).
- The control sections (SBS) in-between the rubber sections (and west of them) exhibited more thermal cracks as compared to the GTR mainline sections.
- The dense-graded mix shoulders had frequent cracking.
- The SMA mix used on the shoulder (Evoflex RMA) showed extensive transverse cracking.

<u>Sections in SMA Study</u>

- The 2012 I-294 section now has many visible distresses, and is starting to ride rough. It should be mentioned that this mix has been placed on jointed concrete pavement.
- Heavy % trucks were observed.

<u>Sections in RAS Study (Shoulders)</u>

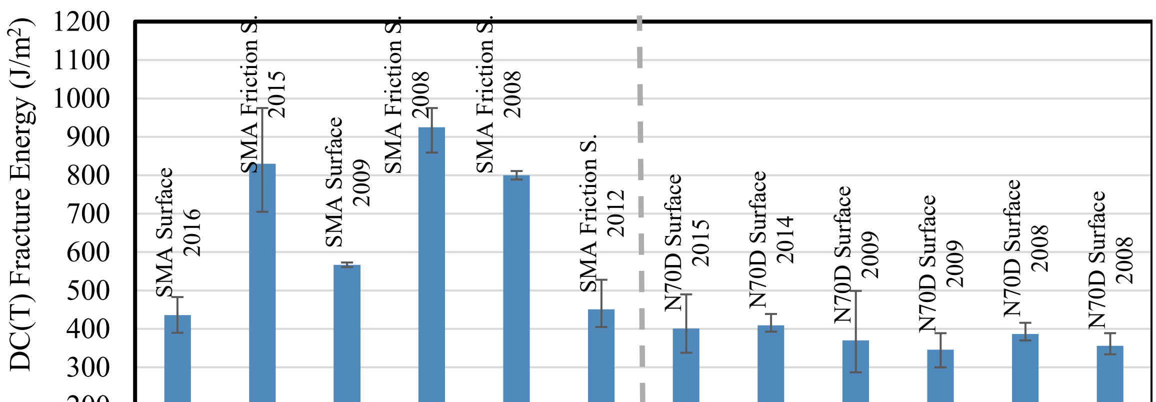
- The shoulders with RAP and RAS had many cracks, thermal and block.

Chapter 5

FIELD CORE TESTING RESULTS

5.1. DC(T) Testing Results

DC(T) samples were fabricated using the top lift of the collected field cores. The thickness of the top lift was at least 50 mm (2 inches) for all the sections, which made it possible to cut the DC(T) samples into 50 mm slices. Three replicates were tested for each section and the average of DC(T) fracture energies was calculated. Figure 5-1 shows the fracture energies tested at -12 °C. The error bars shown for each section covers the range of the obtained from testing the replicates. Also, mixture type and year of overlay are indicated for each section. The table attached to the figure provides the amount of recycled materials used in each mixture, including the ABR by RAP and RAS and total ABR. The details of the mixture ingredients such as the NMAS, binder system, and modification are provided in Table 4-2. The tested mixtures are divided into two categories, namely, SMAs and dense graded mixtures, using a gray dashed line. As expected, the DC(T) fracture energies of the SMAs are higher than the dense graded mixtures. Also noted were:

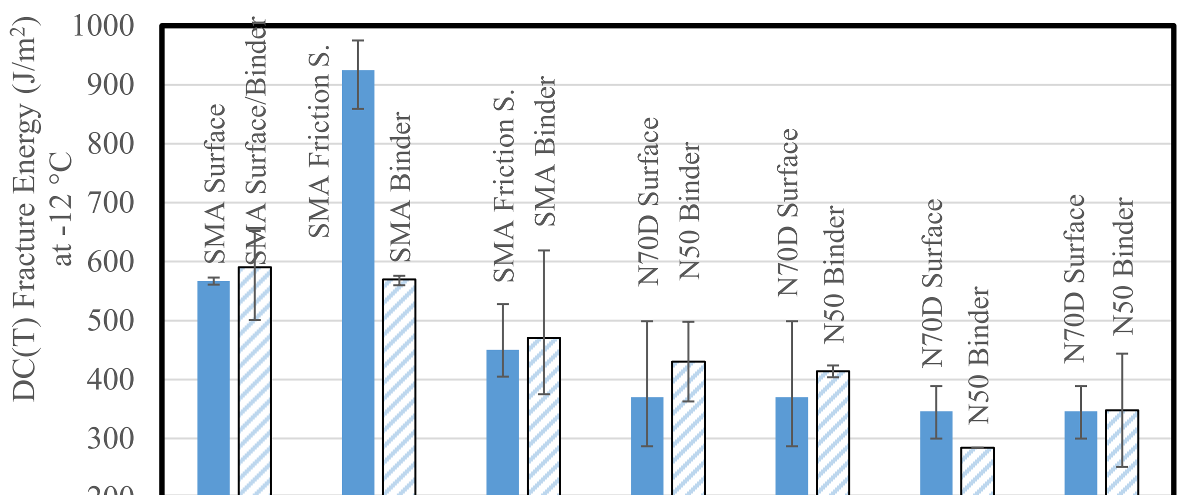


	I88-47	I88-60.5	I90-6.6	I90-6.0	I90-17.8	I294-34	I88-52	I88-57	I90-7.25	I90-5.12	I90-10E	I90-10W
Gf at -12 °C	436	830	567	925	800	451	401	409	370	346	387	356
ABR by RAP	11.6	14.8	13.9	16.3	16.0	15.5	19.1	22.8	16.7	24.4	24.0	16.2
ABR by RAS	19.8	20.7	0	0	0	16.2	19.6	17.8	20.1	0.0	0	0
Total ABR	31.4	35.5	13.9	16.3	16.0	31.7	38.6	40.7	36.8	24.4	24.0	16.2
Year	2016	2015	2009	2008	2008	2012	2015	2014	2009	2009	2008	2008
COV(%)	10.7	16.4	1.5	6.4	1.9	14.9	19.7	6.4	30.5	11.2	6.5	8.1

Figure 5-1. DC(T) testing results at -12 °C for the top lift of the field cores

- I88-47 and I88-60.5 both used SBS 70-28 binder systems. However, benefiting from higher quality aggregates, I88-60.5 yielded a significantly higher fracture energy (436 v. 830 J/m^2) although this section has aged one year longer than I88-47.
- The combination of higher aggregate quality and also a softer binder system (PG 70-28 GTR) used in the I90-6 mix (SMA friction surface) resulted in higher DC(T) fracture energy as compared to I90-6.6.
- Referring to the distress summary for I90-6.0 and I90-6.6 listed in Table 4-1, the section with lower fracture energy (I90-6.6) started to show transverse cracking while I90-6 with higher fracture energy did not.
- Although the I90-17.8 section experienced block cracking on its surface, the mixture performed well in the DC(T) test with a fracture energy value of 800 J/m^2. We believe this is due to the nature of aggregate in this mix, perhaps combined with mix volumetrics.
- The I294-34 mix was placed on jointed concrete pavement and has experienced significant reflective cracking. The fracture energy of this mix was low (451 J/m^2).
- The fracture energy of all the shoulder mixtures including poor and good performing sections was around 400 J/m^2. This indicates that a long-term aged fracture energy level of 400 J/m^2 may be borderline with respect to ensuring adequate resistance to environmentally-based cracking in the Chicagoland area.

Figure 5-2 shows the DC(T) fracture energies for two lifts of the studied sections. Generally, no significant difference was observed in fracture energy for two different lifts of the same section. Although the mix used in the bottom lift may not be as crack resistant as the first lift, the environmental and traffic loading conditions that the first lift experiences are more severe than the bottom lift. Given the fact that the studied mixtures have aged for many years in-situ, the higher crack resistance could be balanced with harsher environmental conditions (i.e. cooling cycles their severity) such that the difference between fracture energies is not considerable.



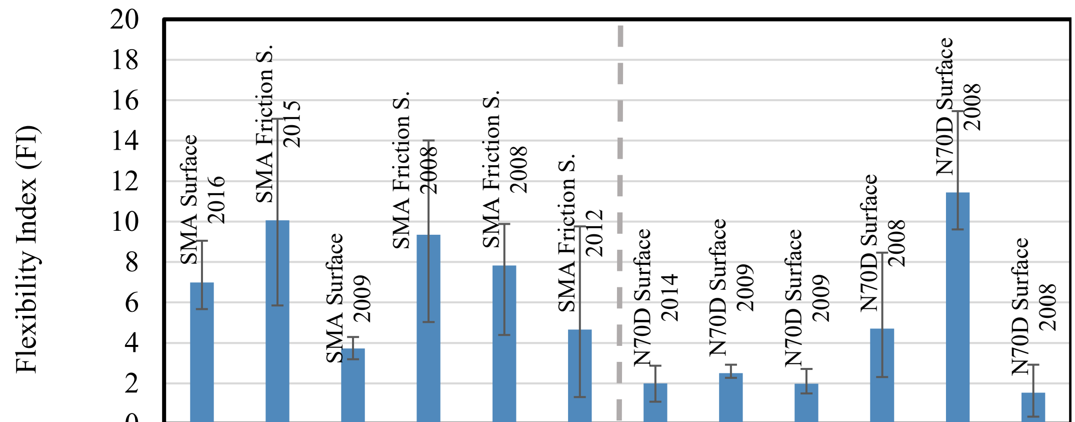
	I90-6.6	I90-6.0	I294-34	I90-7.25	I90-6.06	I90-5.12	I90-4.75
Top Lift	567	925	451	370	370	346	346
Bottom Lift	590	570	470	431	414	285	348
Const. Year	2009	2008	2012	2009	2009	2009	2009

Figure 5-2. Comparing fracture energies for top and bottom lifts of the field cores

5.2. I-FIT Testing Results

I-FIT testing was performed on the sliced top lift of the field cores. Four replicates were prepared and tested and the average of the FI parameter for twelve sections are presented in Figure 5-3. Similar to the previous figure, the error bars show the range of the FIs calculated for the four replicates. Although the FIs for the SMAs are generally higher than those of dense graded mixtures (as expected), the I90-10E section, as a shoulder mix had the highest FI among the tested sections. I90-5.12 was another section that has an FI of 4.7 which is comparable to the FI of the SMAs. Among the six SMA mixtures, I88-60.5, which is a four-year-old SMA friction surface mix with an SBS 70-28 recoded the highest FI. As the I-FIT test is very sensitive to aging (based on the aging results in Chapter 3), it is expected that the FI of this section will drop significantly as it reaches the age of the other sections. The I90-6.6 mix, which performed well in DC(T) test and was ranked as the third-best SMA, yielded the poorest performance in I-FIT test among the SMA, with an FI of 3.7.

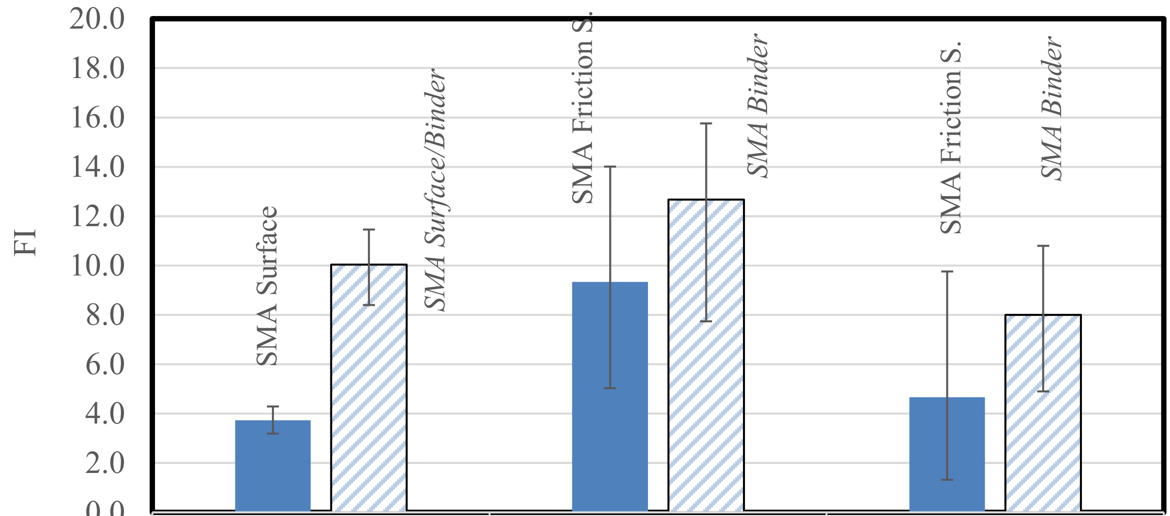
Comparing I90-7.25 and I90-5.12, which have the same binder system but different combinations of recycled materials, illustrates how the FI is sensitive to RAS binder. As a result, the FI of I90-5.12 is significantly higher than that of I90-7.25, although I90-5.12 is one year older than I90-7.25.



	I88-47	I88-60.5	I90-6.6	I90-6.0	I90-17.8	I294-34	I88-52	I88-57	I90-7.25	I90-5.12	I90-10E	I90-10W
FI	7.0	10.1	3.7	9.3	7.8	4.7	2.0	2.5	2.0	4.7	11.4	1.5
ABR by RAP	11.6	14.8	13.9	16.3	16.0	15.5	19.1	22.8	16.7	24.4	24.0	16.2
ABR by RAS	19.8	20.7	0	0	0	16.2	19.6	17.8	20.1	0.0	0	0
Total ABR	31.4	35.5	13.9	16.3	16.0	31.7	38.6	40.7	36.8	24.4	24.0	16.2
Year	2016	2015	2009	2008	2008	2012	2015	2014	2009	2009	2008	2008
COV(%)	22.6	45.6	13.0	39.7	30.4	79.2	37.0	11.5	26.5	57.0	24.0	75.2

Figure 5-3. I-FIT testing results for the top lift of the field cores

Figure 5-4 compares the FIs for two different lifts of the field cores collected from three sections. As shown in the figure, all of these three sections are SMAs. Unlike the DC(T) fracture energy, there is a significant difference between the FI of different lifts. As mentioned before, I-FIT tests is very sensitive to aging and the bottom lift of field cores had less aging than the top lift. This was especially reflected in I90-6.6 where the same mix was used in both lifts but the FI of the bottom lift was 170 % higher than the top lift. For the other two sections, I90-6.0 and I294-34, although the top lift is a SMA friction surface mix which benefits from very high quality aggregates, the FI of the bottom lift, which are SMA binders, are higher than the top lifts.



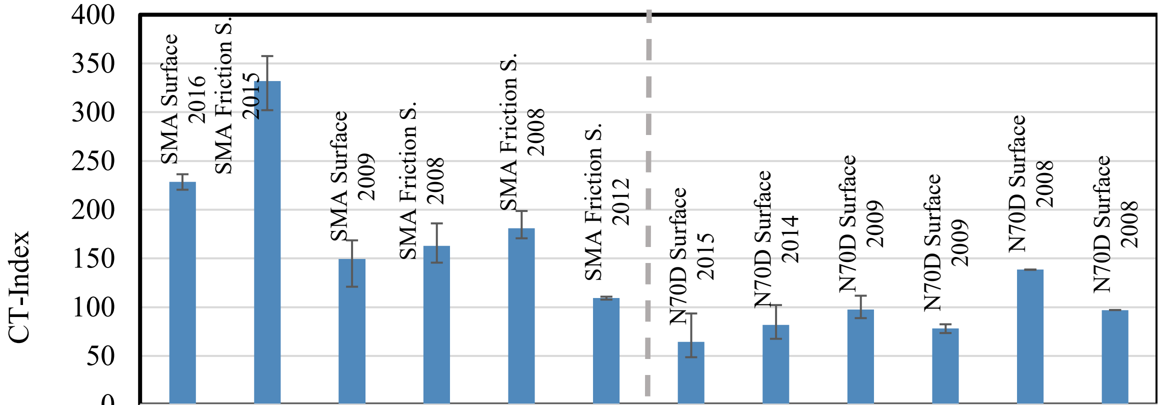
	I90-6.6	I90-6.0	I294-34
Top Lift	3.7	9.3	4.7
Bottom Lift	10.0	12.7	8.0
Const. Year	2009	2008	2012

Figure 5-4. Comparing the FIs for top and bottom lifts of the field cores

5.3. IDEAL-CT Testing Results

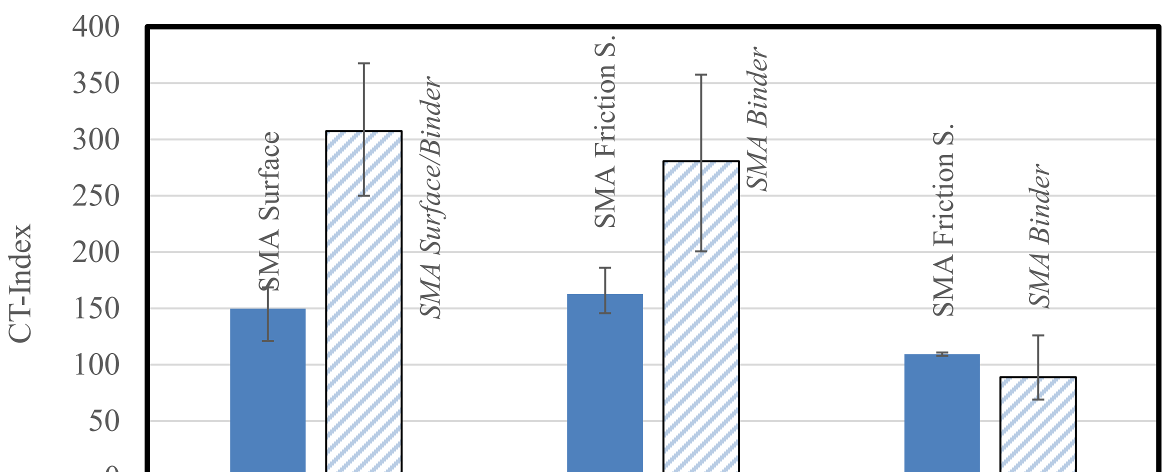
The IDEAL-CT test was the third cracking test which was carried out on the field cores. The averages of the CT index using three replicates for different sections are presented in Figure 5-5. It should be noted that the recently published IDEAL-CT specification (ASTM D8228-19) calls for 62 mm as the thickness of the testing samples. However, the thickness of the lifts was not enough to meet that requirement and 50 mm slices were tested. Similar to I-FIT test results, the I88-60.5 and I90-10E have the highest CT score among the SMAs and dense graded mixtures, respectively. That being said, the SMA friction surface mixtures did not outperform the SMA surface mixtures. The IDEAL displayed a considerably lower COV as compared to I-FIT. Similar to DC(T) test results, the shoulder mixtures performed very similarly in this test, where the CT score ranged from 64 to 139.

As shown in Figure 5-6, the CT scores of the bottom lifts for two sections (I60-6.6 and I90-6.0) are higher than their corresponding top lift. However, the I294-34 section recorded a higher CT index for its top lift.



	I88-47	I88-60.5	I90-6.6	I90-6.0	I90-17.8	I294-34	I88-52	I88-57	I90-7.25	I90-5.12	I90-10E	I90-10W
CT	228	332	150	163	181	109	64	82	98	78	139	97
ABR by RAP	11.6	14.8	13.9	16.3	16.0	15.5	19.1	22.8	16.7	24.4	24.0	16.2
ABR by RAS	19.8	20.7	0	0	0	16.2	19.6	17.8	20.1	0.0	0	0
Total ABR	31.4	35.5	13.9	16.3	16.0	31.7	38.6	40.7	36.8	24.4	24.0	16.2
Year	2016	2015	2009	2008	2008	2012	2015	2014	2009	2009	2008	2008
COV(%)	4.9	8.4	16.9	12.8	8.5	1.9	39.3	22.2	12.6	5.8	0.0	0.0

Figure 5-5. IDEAL-CT testing results for the top lift of the field cores



	I90-6.6	I90-6.0	I294-34
Top Lift	150	163	109
Bottom Lift	307	281	89
Const. Year	2009	2008	2012

Figure 5-6. Comparing the IDEAL-CTs for top and bottom lifts of the field cores

5.4. Repeatability of the Cracking Tests

Table 5-1 presents the COV and standard deviation (STD) of different cracking tests for different mixture types tested in this study. As mentioned before, the repeatability of a performance test should be a key consideration in selecting an appropriate test for specification development. As shown in the table, the DC(T) test has the lowest COV for all tests across each mix category. Note that the I-FIT and IDEAL-CT tests were not performed on the shoulder binder lift mixes so that enough materials could be retained to enable Hamburg testing.

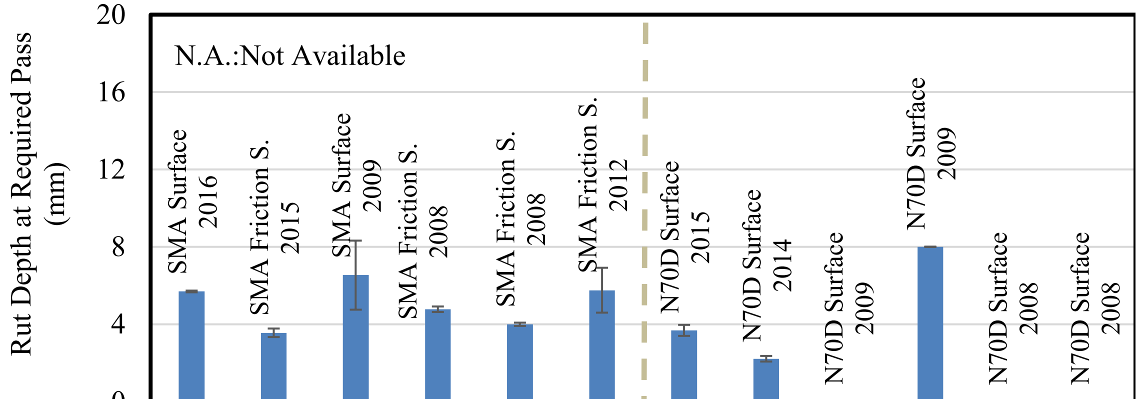
Table 5-1. Variability of the cracking performance test results for field cores

Cracking Test:	DC(T) at -12 °C		FI (4 Reps)		FI (3 Reps, Using Trimmed Mean)		CT	
Mix. Type	COV	STD	COV	STD	COV	STD	COV	STD
SMA F. S.	9.9	70	48.7	2.9	36.5	1.8	7.9	17
SMA S.	6.1	27	17.8	1.0	12.6	0.7	10.9	18
SMA B.	14.2	73	25.3	2.6	11.0	1.6	27.8	57
Shoulder S.	13.7	52	38.5	1.4	26.5	0.6	20	11
Shoulder B.	16.2	62	NA	NA	NA	NA	NA	NA

5.5. Hamburg Testing Results

Based on the Hamburg test specification and testing fixtures, 62 mm thickness samples are needed. In order to fit the samples into the Hamburg fixtures, concrete slices were fabricated at a thickness of about 12 mm and placed in the fixtures as vertical shims (see Figure 5-7). Figure 5-8 presents the rut depths measured in the Hamburg test on asphalt cores. The rut depth of the first six mixtures (SMAs) were recorded at 20,000 passes while for shoulder mixtures, 15,000 wheel passes were used. As already seen in the plant-produced mixtures in Chapter 3, rutting is not a concern for SMAs. The only SMA mix with a rut depth higher than 6 mm was I90-6.6 which had the lowest amount of recycled materials (ABR=13.9%). Similarly, the I90-5.12 section, which had the lowest ABR among the three tested shoulder mixtures with an ABR of 24.4%, recorded the highest rut depth (8.0 mm). It is also worth mentioning that the testing samples obtained from the field cores for the rest of the shoulder mixtures, including I90-7.25, I90-10E, and I90-10W, were used for the cracking tests, as cracking was the main distress on the shoulders. The relatively low rut levels on field cores are probably due to the age-hardening of the mixes, but also indicate the proper performance characteristics in Tollway mixtures, such as resistance to stripping or aggregate degradation.

Figure 5-7. Hamburg samples with 12mm PCC shims placed below asphalt specimens

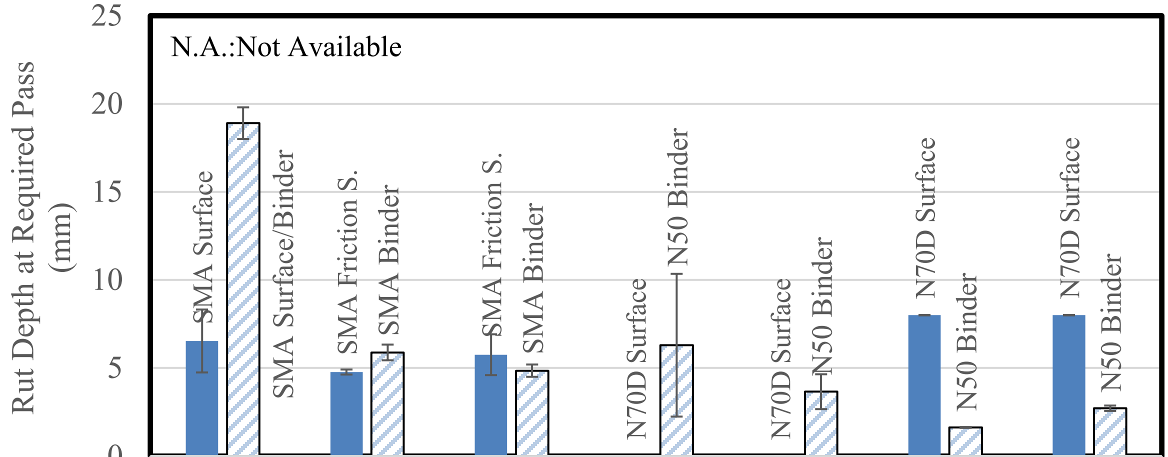


	I88-47	I88-60.5	I90-6.6	I90-6.0	I90-17.8	I294-34	I88-52	I88-57	I90-7.25	I90-5.12	I90-10E	I90-10W
Rut Depth (mm)	5.7	3.6	6.5	4.8	4.0	5.8	3.7	2.2	N.A.	8.0	N.A.	N.A.
ABR by RAP	11.6	14.8	13.9	16.3	16.0	15.5	19.1	22.8	16.7	24.4	24.0	16.2
ABR by RAS	19.8	20.7	0	0	0	16.2	19.6	17.8	20.1	0.0	0	0
Total ABR	31.4	35.5	13.9	16.3	16.0	31.7	38.6	40.7	36.8	24.4	24.0	16.2

Figure 5-8. Hamburg testing results for the top lift of the field cores

Figure 5-9 compares the rut depths from testing of surface vs. top binder course lifts in the Hamburg on cores. For the SMA binders and shoulder binders the required number of Hamburg passes is 20,000 and 10,000, respectively. The bottom lift of the I90-6.6 section, which used the same mix as the top lift, did not perform well in the Hamburg test at 50 °C. It is not clear why this mix experienced poor scoring in the Hamburg. However, due to its position in the pavement (shoulder, binder), rutting is not expected to be a concern.

The results obtained from field cores were instrumental in validating and calibrating DC(T) and Hamburg thresholds in the Tollway asphalt performance specification, as described in Chapter 7.



	I90-6.6	I90-6.0	I294-34	I90-7.25	I90-6.06	I90-5.12	I90-4.75
Top Lift	7	5	6	N.A.	N.A.	8	8
Bottom Lift	18.9	5.9	4.9	6.3	3.7	1.6	2.7
Const. Year	2009	2008	2012	2009	2009	2009	2009

Figure 5-9. Comparing the rut depths for top and bottom lifts of the field cores

Chapter 6

FIELD PERFORMANCE DATA AND ANALYSIS

6.1. Overview

Pavements serve as an important and critical part of a nation's infrastructure, and it is essential to preserve its functioning to maintain national development and prosperity. Pavements, like all other infrastructure assets, deteriorate over time and thus require routine maintenance activities to be conducted by transportation agencies in order to avoid any loss of serviceability. The first step towards planning a pavement maintenance activity is to be aware of the pavement condition, and that is achieved through systematic Pavement Condition Surveys (PCSs).

PCSs refer to activities that quantify the pavement serviceability and its physical condition and are mainly comprised of three aspects: data collection, condition rating, and quality management. The data collection, which is mostly semi-automated or automated, provides a measure of the distresses prevalent in an existing pavement section. The data might also include other details about the pavement construction, such as length and width of the section, location of underlying structures, and details of last conducted preservation or maintenance activity. The condition rating is usually index- or scale-based to quantify the condition of a pavement section. Various systems for condition rating exist, and adoption of a particular system depends on available resources and familiarity with the said rating system. Finally, based on which pavement section falls below the set condition rating thresholds, adequate maintenance treatments are applied to retain a certain minimum serviceability. Condition rating data collected over time could provide an overall performance of any particular section and could provide an objective basis for selecting future maintenance techniques, affecting the short- and long-term budget planning of a transportation agency.

In this chapter, the field performance data collected by Applied Research Associates (ARA) are presented based on further processing and analysis by the research team. These data consist of condition rating survey (CRS) results including the severity of the observed asphalt pavement distresses, International Roughness Index (IRI), and rut depth collected from the mainline sections. Since all of the studied mainline sections are located in Northern Illinois, it is assumed that they experienced the same environmental conditions and their low-temperature cracking performance can be compared. The asphalt mixtures used on the studied mainlines have been tested in the lab, and the results were presented in Chapter 5. The objective of the field performance data analysis is to establish a link between the field performance and laboratory testing results. The link will ultimately be used to determine the thresholds and calibrate the performance specification.

6.2. Condition Rating Survey (CRS)

As mentioned before, a historical record of pavement condition rating allows for a) the proper planning of maintenance activities to be undertaken, b) the adequate allocation of funds to maintain a minimum amount of serviceability in the existing pavement network, and c) the ability to predict future requirements for maintenance leading to adoption of relevant preservation techniques. The Condition Rating System (CRS), used in this study, is an index between 1 and 9, representing a failed and a new pavement condition respectively.

6.2.1. CRS Trends in Service Life

Figure 6-1 shows the CRS measurements for six different mainline sections that were already introduced in Chapter 4 and evaluated for performance as described in Chapter 5. Different zones ranging from "Excellent" to "Poor/Failed" were superimposed on the figure to more clearly identify the condition of the sections based in CRS values. The CRS of each section is plotted in different years. Where the data was not available, a dashed line was used to extrapolate the CRS values, especially at the early years of the sections' service life. As seen in Table 4-2, the I88-47 section is the newest section (overlaid in 2016) and has the highest CRS score. Also, noted were:

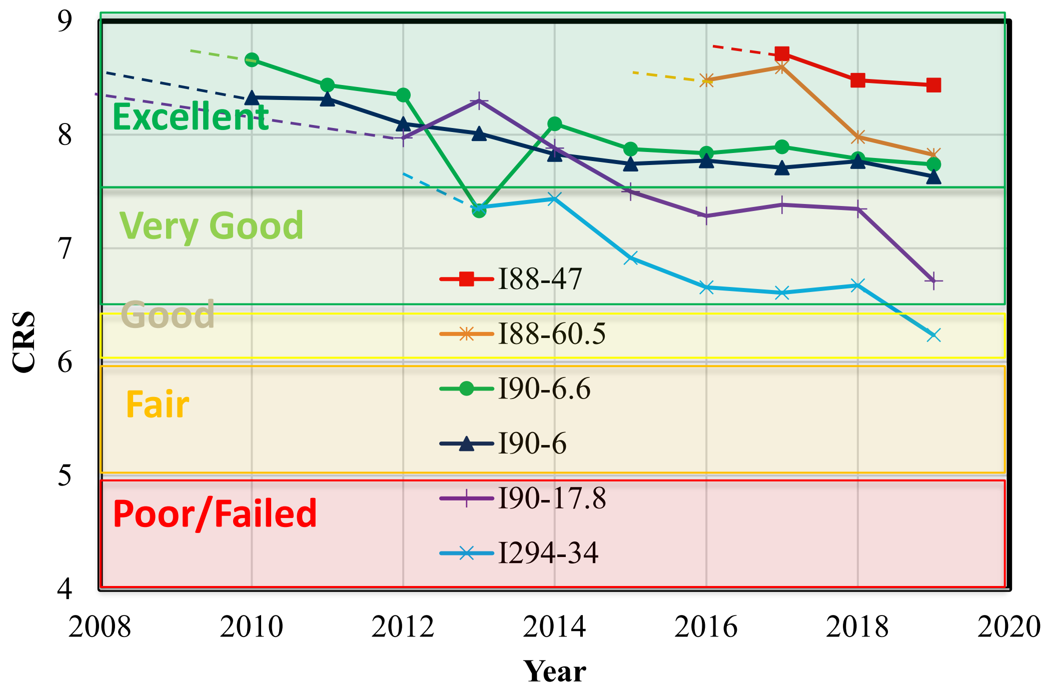
- The lowest CRS is recorded for the I294-34 section which is placed on a jointed concrete pavement and undergoes heavy traffic loads as discussed in Chapter 4 (refer to Figure 4-8). In addition, I294-34 is the only section with an existing CRS value below 6.5, and the section entered the "good" condition based on CRS score.

- The I90-6.6 and I90-6.0 sections show similar trend and CRS values at different years and both hold CRS values in "Excellent" condition after more than 10 years of service life.

- Although the DC(T) fracture energy of I90-6.6 (567 J/m^2) was much lower than that of I90-6.0 (925 J/m^2), it was high enough to maintain the CRS values similar to I90-6.0.

- I90-17.8 showed block cracks on the surface as shown in Figure 4-12. Accordingly, the CRS decreased at a high rate especially in 2018 and 2019 such that the CRS went below the "Very Good" condition and into the "Good" zone.

- Despite the fact that I88-60.5 section is in its early stages of life, the CRS deterioration rate was high such that this section is approaching the "Very Good" CRS zone.



--- *Dashed lines show extrapolated CRS to the year of last overlay*

Figure 6-1. Comparing the CRS values as a function of year

6.2.2. Distress Type and Severity

The presented CRS values were determined based on the type, extent, and severity of different distresses. The CRS system has a pavement distress guide which can be used to characterize the distress identification and coding. This guide contains distresses for both concrete and asphalt pavements. Some of the important distresses which were frequently observed on the studied sections are presented in Table 6-1. As noticed, these specific distresses are all cracking type distresses which highlight the importance of this mode of deterioration in the Tollway system. As shown, "centerline deterioration", "longitudinal/center of lane cracking", "transverse cracking/joint reflection cracking", and "block cracking" are denoted as "S", "Q", "O", and "M", respectively. Each distress has different severity levels that can be identified using the digit after the distress code. For instance, S4 is used to characterize "centerline deterioration" that is "frequent". It should be noted that the "S" distress is mainly due to the construction and is referred to as cold joint. Although this study does not attempt to mitigate this crack, improving the construction methods for paving patterns are expected to address it. Centerline cracks (Q) are developed mostly due to traffic load and could form block cracks after joining the transverse cracks. Finally, the transverse and reflective cracks (O) can be formed due to cooling cycles and propagation of the cracks from underneath layers, respectively.

Table 6-1. Examples of CRS distress characterization and coding

Distress Type	Severity Levels
Centerline Deterioration	S1 – Tight cracking with little or no spalling.
	S2 – Cracking with low to medium spalling.
	S3 – Infrequent: Cracks are open with medium to severe spalling.
	S4 – Frequent: Cracks are open with medium to severe spalling
Longitudinal/Center of Lane Cracking	Q1 – Beginning Stage: Cracks are tight (width is less than or equal to ¼”) with little or no spalling.
	Q2 – Infrequent: Cracks are between ¼” and ½” and may have minor spalling.
	Q3 – Frequent: Cracks are between ¼” and ½” and may have minor spalling.
	Q4 – Infrequent: One or more of the following conditions exist: Cracks are greater than ½” in width Cracks have severe spalling Major maintenance activity has been performed on the crack
Transverse Cracking /Joint Reflection Cracks	O1 – Beginning Stage: Hairline cracks at any frequency.
	O2 – Infrequent: Cracks are open and less than or equal to ¼” in width and may have low to moderate levels of associated distress.
	O3 – Frequent: Cracks are open and less than or equal to ¼” in width and may have low to moderate levels of associated distress.
	O4 – Infrequent: Cracks are greater than ¼” in width and may have moderate to severe levels of associated distress.
Block Cracking	M1 – Low level: Hairline cracks with none or only a few interconnecting cracks. Cracks are not spalled.
	M2 – Medium level: Further development of interconnecting cracks into a pattern. Cracks may be lightly spalled.
	M3 – High level – Infrequent: Cracks have progressed so that the pieces are well defined and/or spalled at the edges.

6.2.3. Analysis of the Distress Data

The studied sections have different length and the CRS system provides the field performance data for subsections (typically) with a length of one mile. For each of these subsections, the field performance data such as CRS, IRI, and rut depth are provided. In addition to the aforementioned performance indices and parameters, each subsection determines the observed distresses and their severity. Figure 6-2 shows an example (section I88-60.5) which is divided into two one-mile subsections (60.0 to 61.0 and 61.0 to 62.0). This figure also shows the number of lanes and the type of pavement. These data are organized and presented based on the year. As seen in this figure, the first subsection had three distresses including O2, Q1, and S2 in 2019. However, the severity of “O” distress was 1 (O1) on the second subsection in 2019, which indicates that the transverse/reflective cracks had relatively lower severity within the second subsection. A weighted average (based on the length of the subsection) is implemented to calculate the average severity of distresses. In this example, the average severity of O distress is calculated as 1.5 in 2019 as the two subsections of I88-60.5 had equal length (1 mile). This section was overlaid in 2015 and did not show any distresses from 2016 to 2017. There were some subsections (especially long ones such as I90-6.0 and I294-34) paved with concrete pavement as concrete subsections are commonly used next to bridges. Those subsections were excluded from the average severity calculation. The averaged severity of the important distresses was calculated for all of the studied sections in their service lives and will be discussed later. It should be mentioned that the only section which showed block cracking ("M" type distress) according to the ARA data is the I90-17.8 section. This section started to develop the "M1" crack in 2019. As no other sections developed block cracking, the average severity for "M" is not presented.

		Station (mi)		No of	New Pvmt	Distresses					Avg. Severity		
	Route	From	To	Lanes	Type	1	2	3	4	5	S	Q	O
2019	EWEB	60.00	61.00	2	HMAC	O2	Q1	S2			2	1	1.5
	EWEB	61.00	62.00	2	HMAC	O1	Q1	S2					
2018	EWEB	60.00	61.00	2	HMAC	Q1	S2				2	1	0.5
	EWEB	61.00	62.00	2	HMAC	S2	Q1	O1					
2017	EWEB	60.00	61.00	2	HMAC						0	0	0
	EWEB	61.00	62.00	2	HMAC								
2016	EWEB	60.000	61.000	2	HMAC						0	0	0
	EWEB	61.000	62.000	2	HMAC								

Figure 6-2. Example of analyzed field performance data for I88-60.5

The average severity of the centerline deterioration distress (S) is calculated for each section and presented in Figure 6-3. As shown in the figure, the severity of all the sections reached at least one after three years of service life. The I88-60.5 and I294-34 showed the highest rate of the distress development and reached the average of 2 at the fourth year of their service life. The calculated average severity of this distress for different sections suggests that the centerline deterioration stops growing after five years of service life. However, maintenance strategies should be applied in order to prevent the development of joint deterioration as surface water can penetrate through the cold joint and affect the structural capacity of the sublayers.

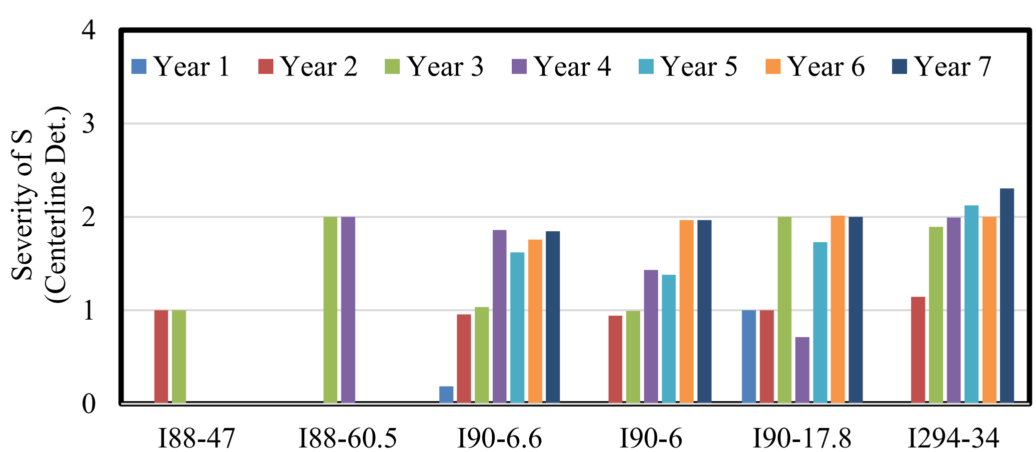


Figure 6-3. Average severity of centerline deterioration distress (S) as a function of service life

As shown in Figure 6-4, Section I88-47 did not have significant longitudinal/center lane cracking after three years of life. On the other hand, the severity of this distress reached Q1 on section I88-60.5 after four years in service. Similar to the CRS and "S" distress, the I90-6.6 and I90-6.0 performed alike in terms of "Q" distress and did not grow considerable longitudinal/center lane cracks. I90-17.8 was the worst performing section in "Q" distress based on CRS data. This is aligned with the block cracks observed on this section. The I294-34 section developed some "Q" cracks after seven years of service life. Given the high traffic load that this section carries, it was expected to observe "Q" cracks on this section.

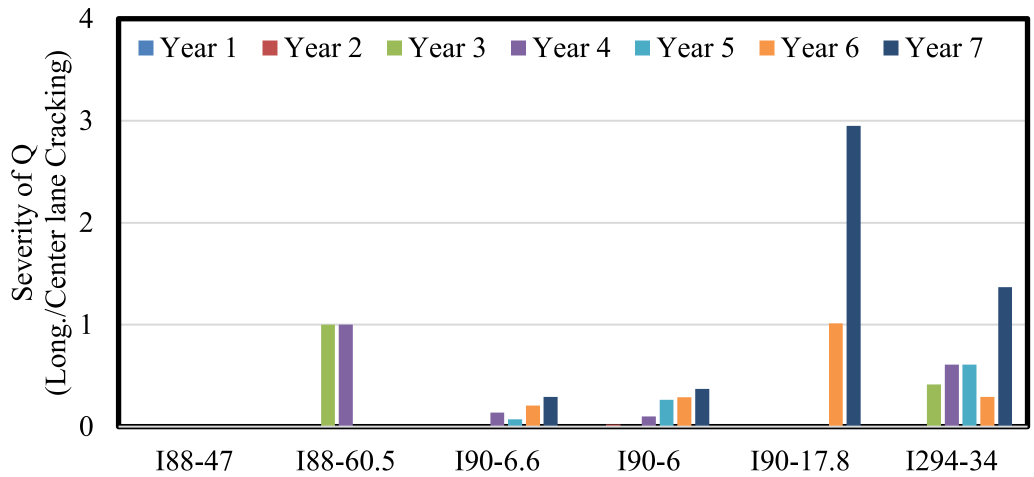


Figure 6-4. Average severity of longitudinal/center lane cracking distress (Q) as a function of service life

Figure 6-5 shows the average severity of "O" distress in the studied sections. As expected, the I294-34 section developed the highest severity of "O" distress. As mentioned before, the asphalt layer is placed on jointed concrete and the cracks observed on the surface (see Figure 4-7) are reflected from the concrete joints. Similar to the previous performance indices and distresses, I90-6.6 and I90-6.0 are performing comparably. This similarity will be used later to set the criteria on the DC(T) fracture energy of the mainline mixtures. Section I90-17.8 was performing very well in terms of "Q" distress but suddenly started to grow the transverse cracks at year six. It appears that the I88-60.5 section has high potential for the transverse cracking and does not benefit from a crack resistant mix. That being said, this mix performed very well in all of the cracking tests and was one of the two best performers in the DC(T), I-FIT, and IDEAL-CT tests. It should be mentioned that this section was the shortest among the studied sections (less than two miles) and there might be difficulties associated with a comprehensive survey due to this short length. Another possible reason for this discrepancy between laboratory and field performances is the quality of construction. According to Table 4-1, this short section is paved using SBS modified asphalt mixture and is located between the rubber modified test sections and might have been placed on rubblized concrete pavement.

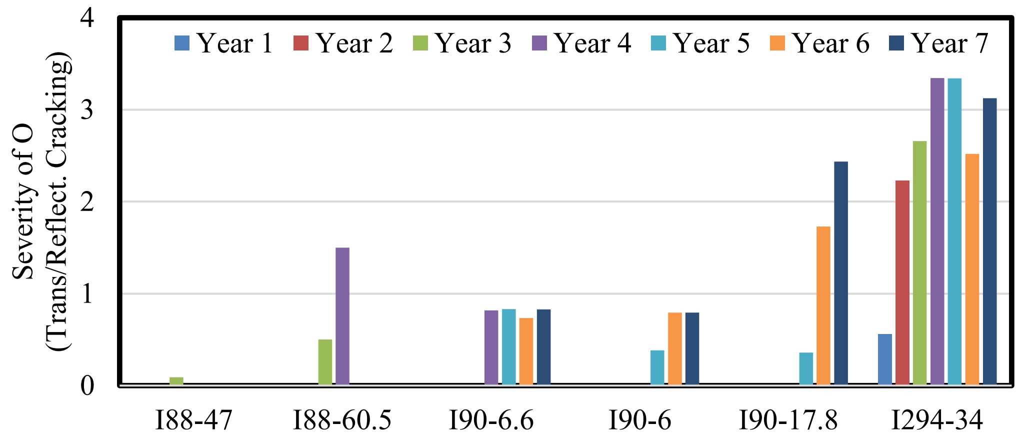


Figure 6-5. Average severity of transverse/reflective cracking distress (O) vs. service life

As discussed in one of the TRP meetings, transverse and block cracking might be improperly characterized in current CRS databases. For instance, a longitudinal crack might form first, followed by a transverse crack, with the two joining shortly thereafter (see Figure 6-6). As time progresses, a clear block crack pattern form. However, once the longitudinal and transverse crack are codified in early performance assessments, this typically biases future rating assessments. This would lead to an over-estimation of longitudinal and transverse crack estimates, and an underestimate of block cracking. The implication of this miscategorization is driven by the fact that the deduction factor for block cracking is lower than that for transverse cracking. Therefore, one would err on the conservative side by considering these cracks as separate transverse and longitudinal cracks from a condition rating standpoint, and in fact, we found this to be a common practice. In the future, research could be focused on delineating block cracking from other cracking forms in order to arrive at even more precise performance test threshold recommendations.

Figure 6-6. Longitudinal cracks that join the transverse cracks - I90-6.6 section

6.3. International Roughness Index (IRI)

Pavement roughness is generally defined as an expression of irregularities in the pavement surface that adversely affect the ride quality of a vehicle (and thus the user). Roughness is an important pavement characteristic because it affects not only ride quality but also vehicle delay costs, fuel consumption, and maintenance costs. Roughness is also referred to as "smoothness" although both terms refer to the same pavement qualities. IRI is a standardized measure of the reaction of a vehicle to roadway profile and roadway roughness as expressed in "inches per mile". Generally, higher IRI values represent rougher roads and vice versa. IRI is used to define a characteristic of the longitudinal profile of a traveled wheel track and constitutes a standardized roughness measurement. The commonly recommended units are meters per kilometer (m/km) or millimeters per meter (mm/m). Figure 6-7 shows the IRI values for the six studied sections by year. As shown in the figure, the IRI trends are very similar to those of CRS presented in Figure 6-1.

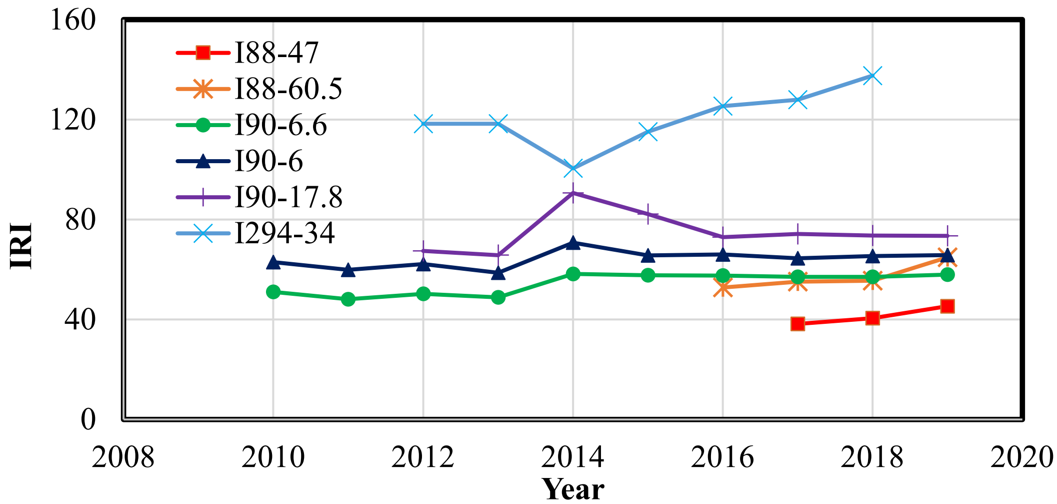


Figure 6-7. Comparison of IRI values vs. year in service

The I294-34 section recorded the highest roughness, which accumulated at a high rate. It is noted that the block cracks observed on I90-17.8 did not affect the ride quality and the IRI did not increase considerably even in recent years as compared to CRS values. The I90-6.0 and I90-6.6 sections performed well and had smooth ride quality. The other two sections, I88-47 and I88-60.5, did not have considerable aging and are at early stages of service life.

6.4. Rut Depth

The permanent deformation measured on the surface of the studied sections is presented in Figure 6-8. It is noted that the measured rut depths sometimes exhibit a decrease in rut depth in specific years. These unexpected rebounds are likely data anomalies, especially since they are such low values. That notwithstanding, and somewhat coincidently, the final rut depths measured in 2019 are similar to the rut depths measured in the lab using the Hamburg test. Referring to Figure 5-8, the I90-6.6 mix had the highest rut depth (6.6 mm) at 20,000 passes under Hamburg wheels which is in accordance with the field performance. Also, I88-60.5 was the best field performer in terms of rutting and accordingly recorded the lowest rut depth (3.6 mm) in Hamburg test. These correlations and comparable field and laboratory rut depths show that the Hamburg test was able to mitigate the rutting distress, and the requirements already set for this test (20,000 at 50 °C) for mainline sections appear to be quite conservative.

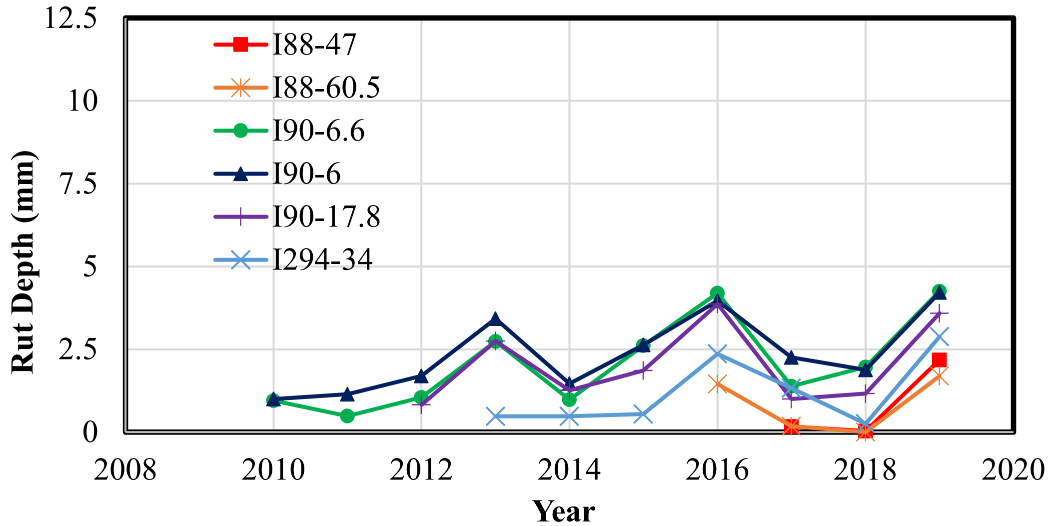


Figure 6-8. Comparison of rut depth values vs. year in service

Chapter 7

SPECIFICATION DEVELOPMENT

7.1. Overview

This chapter presents the data and methods used to validate and calibrate the Illinois Tollway asphalt mixture design specification. Based on observations from the site visit (Chapter 4) along with the field performance data (Chapter 6), various forms of cracking occur more frequently on Tollway pavements as compared to rutting. That notwithstanding, a systematic process was developed and deployed to validate, and to calibrate the Tollway's asphalt mix performance test thresholds as they relate to cracking, rutting and moisture damage. Adding and/or consolidating mix type categories was also addressed in an effort to match the recommended specification with current and future mix design practices. The concept of using the Hamburg as the primary screening method to evaluate moisture damage is also discussed.

7.2. Performance Test Use in Asphalt Mix Design Specifications

As presented earlier in this report, it is recommended that the Tollway retain its existing asphalt mix design performance tests in its asphalt mix design specification, namely the DC(T) fracture energy and Hamburg wheel tracking tests. Before reviewing/adjusting specification limits, a brief review of the testing parameters used in the existing specification and their link to performance is presented.

7.2.1. DC(T) as the Cracking Test

The Illinois Tollway has considerable experience in using the DC(T) test as part of mix design and material characterization. In this section, the data leading to the recommendation to retain the DC(T) test in the Tollway's mix design specification are reviewed. The ability to closely correlate a performance test criterion (or multiple criteria) to field performance should be a key consideration in selecting a performance test for a given distress category. Figure 7-1 shows an existing correlation between transverse cracking and fracture energy (colloquially referred to as "the bubble plot"), using data collected from field sections in various northern states in the US such as Minnesota, Missouri, and Illinois. As shown, there is a clear trend between DC(T) fracture energy and transverse cracking. The data in this plot has a rectangular hyperbola shape; mixtures with fracture energy above certain threshold have low-to-medium transverse cracking, while mixtures with low fracture energy values tend to have thermal cracking levels that 'bubble upwards' with age. As fracture energy of the asphalt mixture drops, the observed transverse cracking in the field increases (and the data dispersion, which is a factor in design reliability), with a sharp upward trend in the curve in the range of 400 J/m^2. The steep upward tick in the curve is likely related to the delineation between more brittle and more ductile binder systems. This is binder, aggregate, mix design, recycling and age dependent. Mixtures that behave as brittle when tested in the DC(T) at 10 ºC warmer than the PG low temperature plan grade (whether or not this happens may depend on the age of the mixture) are found to have low fracture energy values. Mixtures with fracture energies in excess of 600 J/m^2 are clearly found to have very low transverse cracking in this data set, and probably will not develop significant thermal cracking even in later stages of their service life (higher reliability design).

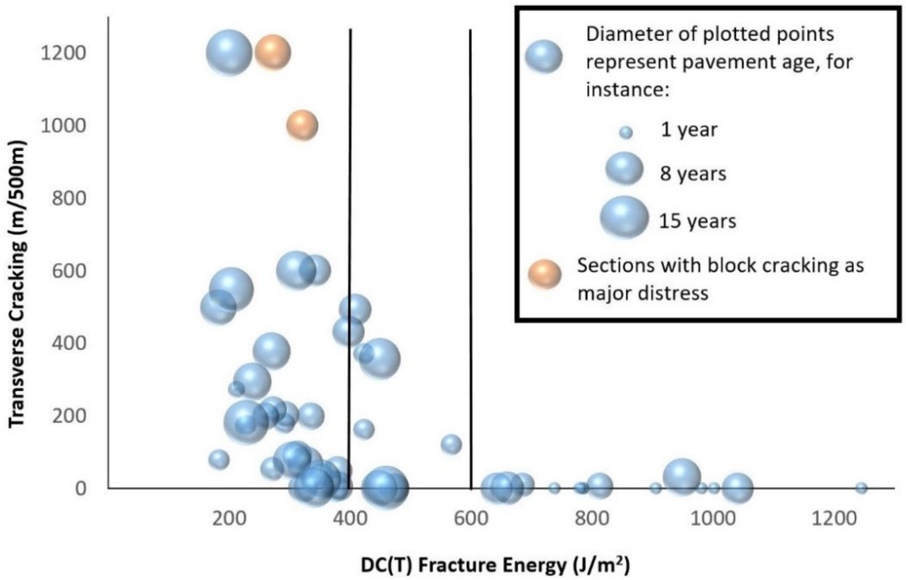


Figure 7-1. Transverse cracking vs. DC(T) fracture energy (Buttlar et al., 2018)

As discussed in sections 3.8 and 5.4, the repeatability of the test is another key factor that should be considered when a performance test is being selected. As shown in Table 3-3 and Table 5-1, the DC(T) test produced low COV's for both plant-produced and field-cored samples, lower than either the IDEAL-CT or I-FIT tests. Lower COV values allow laboratories and transportation authorities to make more confident decisions, especially when borderline results are obtained.

Another key consideration in performance test selection its ability to characterize and score (or to rank) asphalt mixes based on clear performance expectations. In other words, an expensive mix (such as SMAs) containing high quality aggregates and modified binder system along with premium volumetrics (e.g. high VMA) had led to nearly two decades of outstanding rut and cracking resistance in the field for the Illinois Tollway. Therefore, SMAs are expected to attain better cracking scores than dense-graded mixes, those containing unmodified binders, and mixes with lower aggregate quality or volumetric requirements (binder/shoulder mixes). In addition to meeting this expectation better than the I-FIT or IDEAL tests, the DC(T) was found to logically capture the effects of different mix ingredients. For example, although a similar binder system was used in both the 1824 and 1836 mixes, mix 1824 was categorized as an SMA friction surface mix due to its higher aggregate quality. This difference has been reflected in DC(T) fracture energy, where mix 1824 had almost 200 J/m^2 higher fracture energy than 1836 mix at -12 °C. On the other hand, the I-FIT and IDEAL tests sometimes scored dense-graded mixtures higher than SMA mixes. Finally, the DC(T) is more stable and predictable with respect to sample air void levels and mixture aging level, rendering it easier to calibrate based on testing on field cores.

7.2.2. DC(T) Spec Calibration Process

Following the recommendation to retain the DC(T) as the cracking test in the Tollway mix design specification, Figure 7-2 presents the approach developed by the research team to calibrate the specification for various mix types. In the first box, the threshold for DC(T) fracture energy must be chosen for different mix types based on field performance observations and testing results obtained on field cores. For example, by comparing poor performing and good performing sections and their corresponding DC(T) fracture energy results, baseline thresholds can be established. These can also be compared to the bubble plot which includes data from other studies (Figure 7-1) to ensure that recommendations are in range with broader national trends. However, it is acknowledged that the Tollway thresholds should be set towards the most stringent extreme of national thresholds because: (1) at DC(T) test temperature of -12 ºC is desired by local practitioners, but this is more than 10 ºC warmer than the PG low temperature for Chicagoland for a 98% reliability level and therefore somewhat unconservative (suggesting that fracture energy thresholds should be adjusted upwards to account for the warmer test temperature used), and; (2) a very high reliability should be used in Tollway pavement material specifications, considering the very high traffic levels, high speeds, and in consideration of the high user delay costs associated with construction and maintenance activities on the Tollway.

The core samples tested and used to define the DC(T) threshold have been long-term aged in the field. However, in order to circumvent the impracticalities associated with long-term aging of mixtures in the lab prior during mix design, the effect of the aging must be calibrated into the specification limits for tests carried out on short-term aged specimens. In the next box shown in Figure 7-2, standard deviations associated with DC(T) testing will be used to account for test variability as a means to instill a high degree of reliability into the thresholds. SMA friction surfaces should have the highest reliability levels, as they are used on the surface of the mainline pavement in high traffic load sections and/or on the curves to provide skid resistance. Therefore, SMA friction surface type mixes have been assigned the highest level of reliability, as discussed later. The reliability approach developed is also intended to cover the uncertainties associated with field performance data collection and evaluation, and variabilities associated with test scores. Finally, comments and recommendations from experts serving on the project TRP were used, as shown in the last box (consensus step), to capture practical limitations which, for instance, can help avoid high bid prices for certain mix types based on limitations in locally available materials with respect to reaching certain DC(T) thresholds for certain mix types. Consensus adjustment can also create more uniform and logical spreads between thresholds assigned to various mix types in the specification.

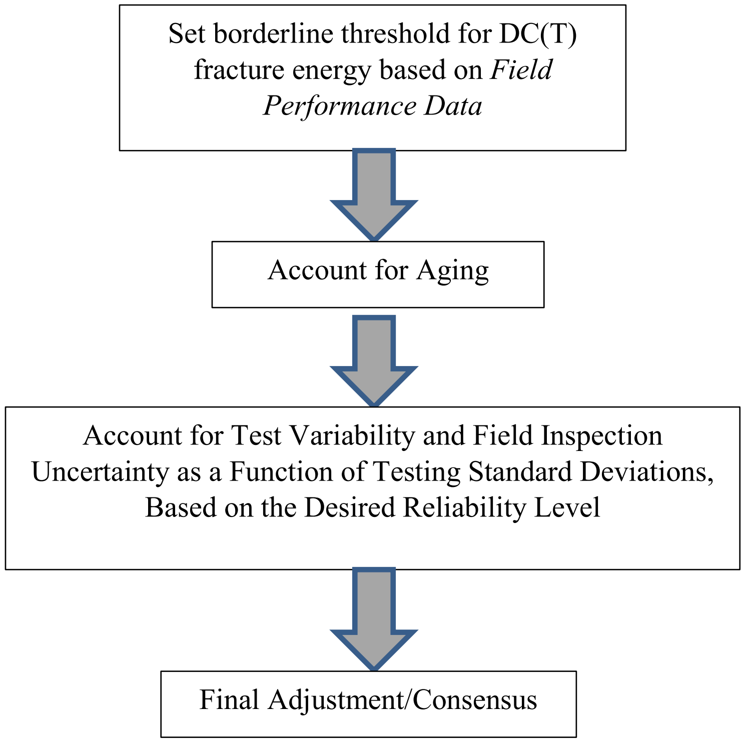


Figure 7-2. Steps in DC(T) spec development

7.2.3. Hamburg Test in Mix Design Specification

The Hamburg wheel tracking test has been used by many agencies across the U.S. to evaluate both rutting and stripping potential of the asphalt mixtures. The test has the capability to be performed at different temperatures as the water tank temperature is adjustable. Often, the test is performed at 50 °C and up to 20,000 wheel passes are used, based on traffic level. Limits such as 12.5 mm rut depth or lower are established, based on traffic level. Given the fact that the traffic load on the Tollway road facilities is relatively high (with AADT values up to 66,000 with 10% commercial vehicles), perhaps more than 20,000 wheel passes should be used. Instead, for practical reasons, the number of wheel passes is kept to 20,000 for SMAs and the maximum allowable rut depth is decreased to 6.0 mm to increase mix reliability (since this level is easily met with high quality aggregates, especially in higher ABR mixes). Following this approach, the Tollway has not experienced rutting issues on mainline pavement in the era of Hamburg use. For the binder mixtures (both shoulders and mainlines) and shoulder surface mixtures, lower wheel pass levels and higher rut levels are allowed in the specification. As no evidence of stripping has been observed, it is suggested to maintain SIP requirements. However, it is recommended that the mix be declared as non-stripping for mixes with Hamburg rut depths lower than 4.0 mm, to avoid erroneous slope ratio and SIP values that sometimes occur in very stiff mixes.

7.3. Effect of Depth on Pavement Response

The Tollway performance specification covers not only surface mixtures, but also binder course mixtures that are used in both mainline and shoulder layers. Therefore, the loading and environmental conditions for binder course layers need to be considered and factored into the test criteria. In this section, the effect of depth for low temperature (cracking) and high temperature (rutting) performance and its implications on adjustment of PRS thresholds will be discussed.

7.3.1. Low Temperature

Asphalt pavements experience the most extreme cold temperatures on the pavement surface during cold winter nights; temperatures in binder courses never reach these extreme levels. Figure 7-3 presents a pavement temperature analysis for very cold, 48-hour duration on a section located in Frazier, Minnesota which was investigated during the SHRP project in 1993. As shown, the difference in temperature at the top and a point 2 in. deep in the pavement is almost 4 °C at the lowest temperature peak. The difference in temperature extremes, and accordingly, the lower temperature gradients (cooling rates) in binder courses leads to lower tensile stress in these deeper layers of the pavement. Based on typical viscoelastic properties at low temperatures and using a convolution integration, the tensile stress induced at different depths has been calculated as shown in Figure 7-4. This analysis reveals that there is approximately 40 % difference between the stress levels at the surface and 2 in. deep in the pavement (reduction from 500 to 300 psi) during critical conditions. This significant difference in material response provides motivation to develop less stringent performance test requirements for binder course mixes to allow economical designs. To the end, an analysis of Illinois temperature data, and simple methods to apply the results to the adjustment of specification values were developed.

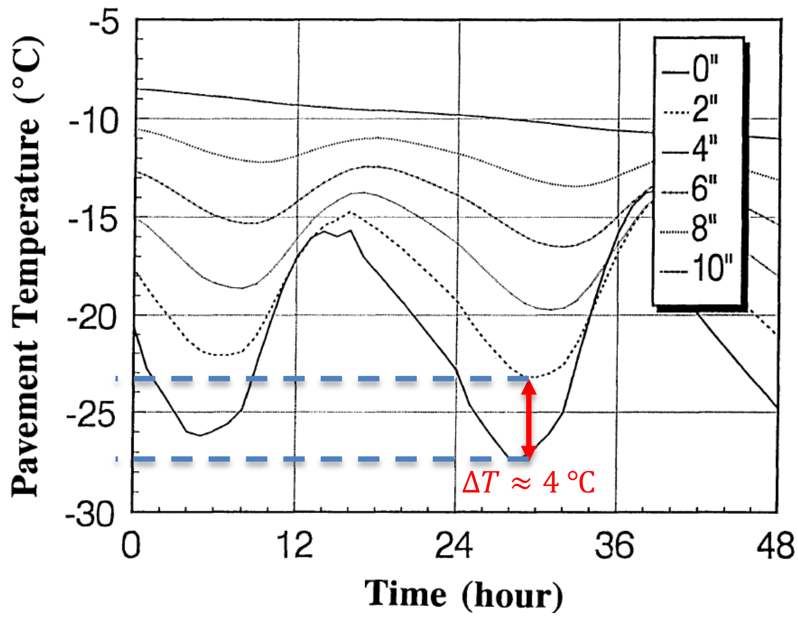


Figure 7-3. Effect of depth on the layer temperature as a function of depth (SHRP A357 Report, 1993- Location: Frazier, Minnesota)

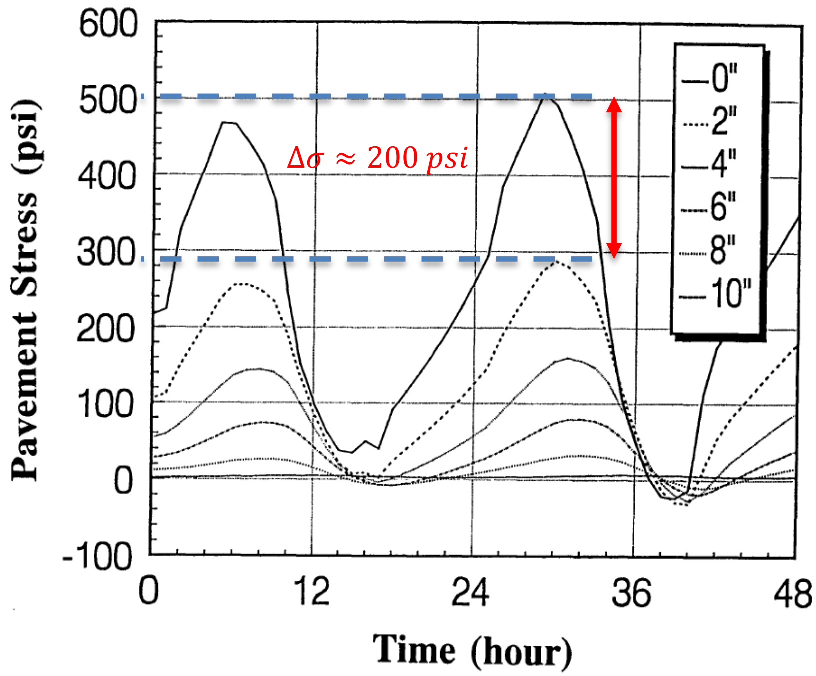


Figure 7-4. Stress relaxation due to depth (SHRP A357 Report, 1993- Location: Frazier, MN)

Using LTPP bind software, the pavement temperature data for the Chicago area at different pavement depths and reliability levels were extracted as presented in Figure 7-5. The station used for the pavement low temperature analysis is very close to the sections that were cored on I88 route (Station Name: Rochelle, ID: IL7354, MP=76 on I88). This provides the chance to investigate the temperature conditions of the sections whose field performance were studied. Considering 98 % reliability, the temperature of the pavement was determined as -27.2 °C while it was calculated to be -24.3 °C at a level of 50 mm of depth in the pavement. Following the temperature analysis shown in Figure 7-3, there was a 3 °C difference between the temperature determined at the surface and at depth of 50 mm (~ 2 in.). Scaling based the assumption of linear viscoelastic behavior, a 30% drop in thermal-induced stress is expected. Although a more rigorous viscoelastic analysis of specific creep data obtained on Tollway pavements would yield higher accuracy, this estimate was instead used for brevity and considering the need to calibrate the model to account for other uncertainties that exist in practical mix design.

Cleary, data from the LTPPBind software suggest that a lower testing temperature should be used for DC(T) testing in the Chicagoland area. The DC(T) test temperature of -12 °C corresponds to a low pavement temperature of -22 °C, as normally the DC(T) test is performed at 10 °C warmer than the pavement temperature. Based on the determined pavement temperature at the 98 % reliability level (-27.2 °C), the DC(T) test should be performed at -17.2 °C. However, as the Tollway has been previously conducting the DC(T) test at -12 °C, this temperature has been retained, and the temperature difference will be accounted for as part of the calibration of the performance specification.

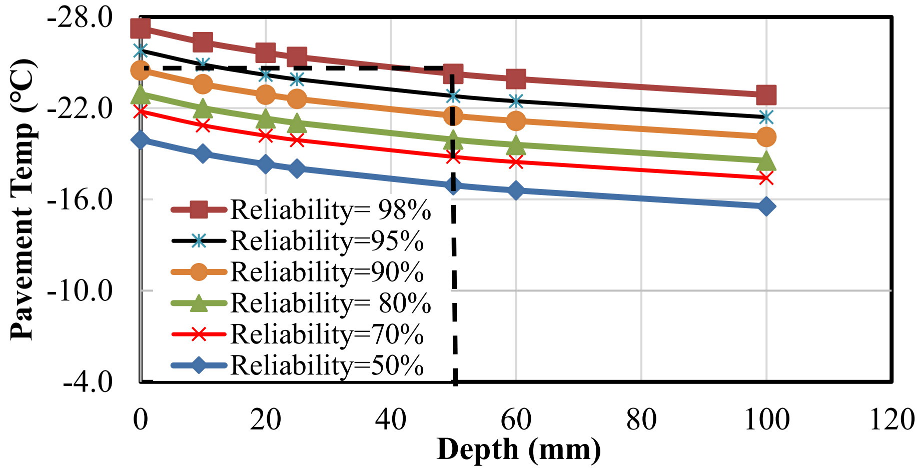


Figure 7-5. LTPP bind software outputs for pavement temperature in winter as a function of depth and reliability in Northern Illinois

7.3.2. High Temperature

The LTPPBind software was used to extract the pavement temperature data during summertime in the Chicago area. A weather station entitled Lake Villa (ID: IL4837), which has a similar latitude as the I90-6.6 and I90-6.0 sections, was selected. As shown in Figure 7-6, the temperature difference between the surface and 50 mm depth in the pavement was determined to be 6.7°C (52.5-45.8=6.7 °C). This is clearly more significant compared to the low temperature differences computed. Although the Hamburg test is normally performed at 50 °C, the environmental conditions are less stringent for the sublayers in terms of pavement high temperature. Therefore, a less stringent criteria (or lower number of passes) should be considered for the binder course mixtures.

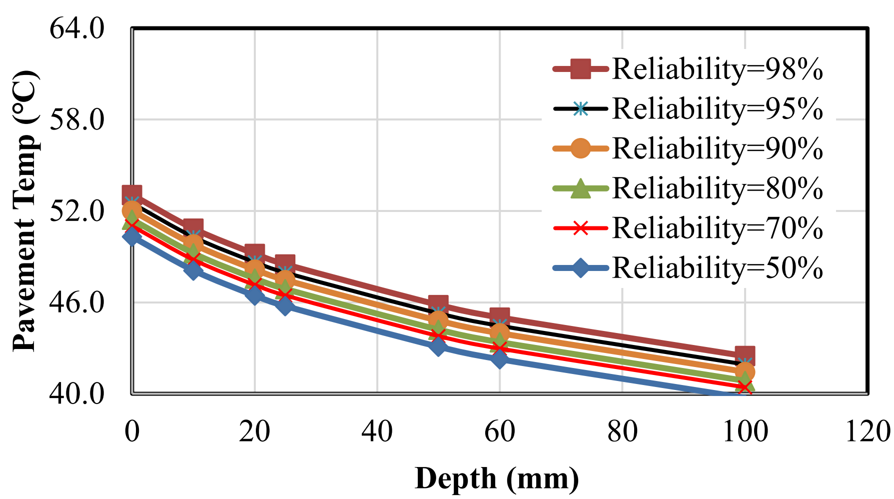
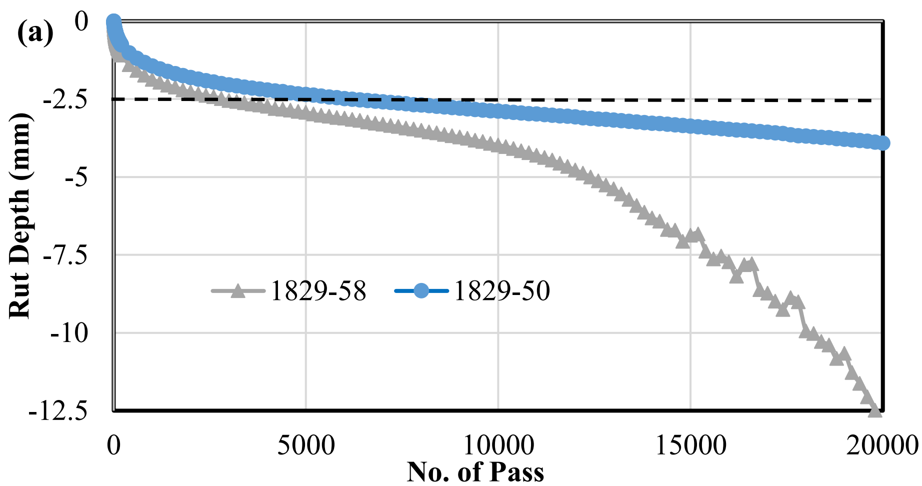


Figure 7-6. LTPP bind software outputs for pavement temperature in summer as a function of depth and reliability in Northern Illinois

To investigate the effect of temperature on the rutting performance of the asphalt mixtures, Hamburg tests at temperatures other than 50 °C were performed on the plant-produced mixtures. Figure 7-7 shows the rut depth as a function of wheel passes at 50 and 58 °C. Although the pavement temperature is not expected to reach 58 °C, as discussed above and shown in Figure 7-6, this testing temperature was selected such that a higher rut depth could be measured. Also, the 8 °C difference between these two testing temperatures is in the range of the 6.7 °C temperature difference determined using the LTPP bind temperature data. The difference between the observed rut depths at these two temperatures will be used to identify the number of passes that could be reduced from the requirements for the binder course mixtures. In essence, a wheel-pass-to-temperature superposition principle has been established.

To this end, a rut depth of 2.5 mm was selected as the reference point. This rut depth was chosen because in most cases at 50 °C, the rut depth recorded by the test sample will attain this level just after the densification phase. The difference in the number of passes to reach this level of rut depth at 50 and 58 °C was then determined to account for the less severe environmental conditions in the subsequent layers. Table 7-1 summarizes the number of wheel passes to reach 2.5 mm for different mixtures at 50 and 58 °C. The average of this wheel pass difference was then calculated as 4,680 passes. After rounding, 5,000 is proposed to be used to reduce the number of required wheel passes for IL-4.75 mixtures to account for the lower temperature present at that depth in the pavement.



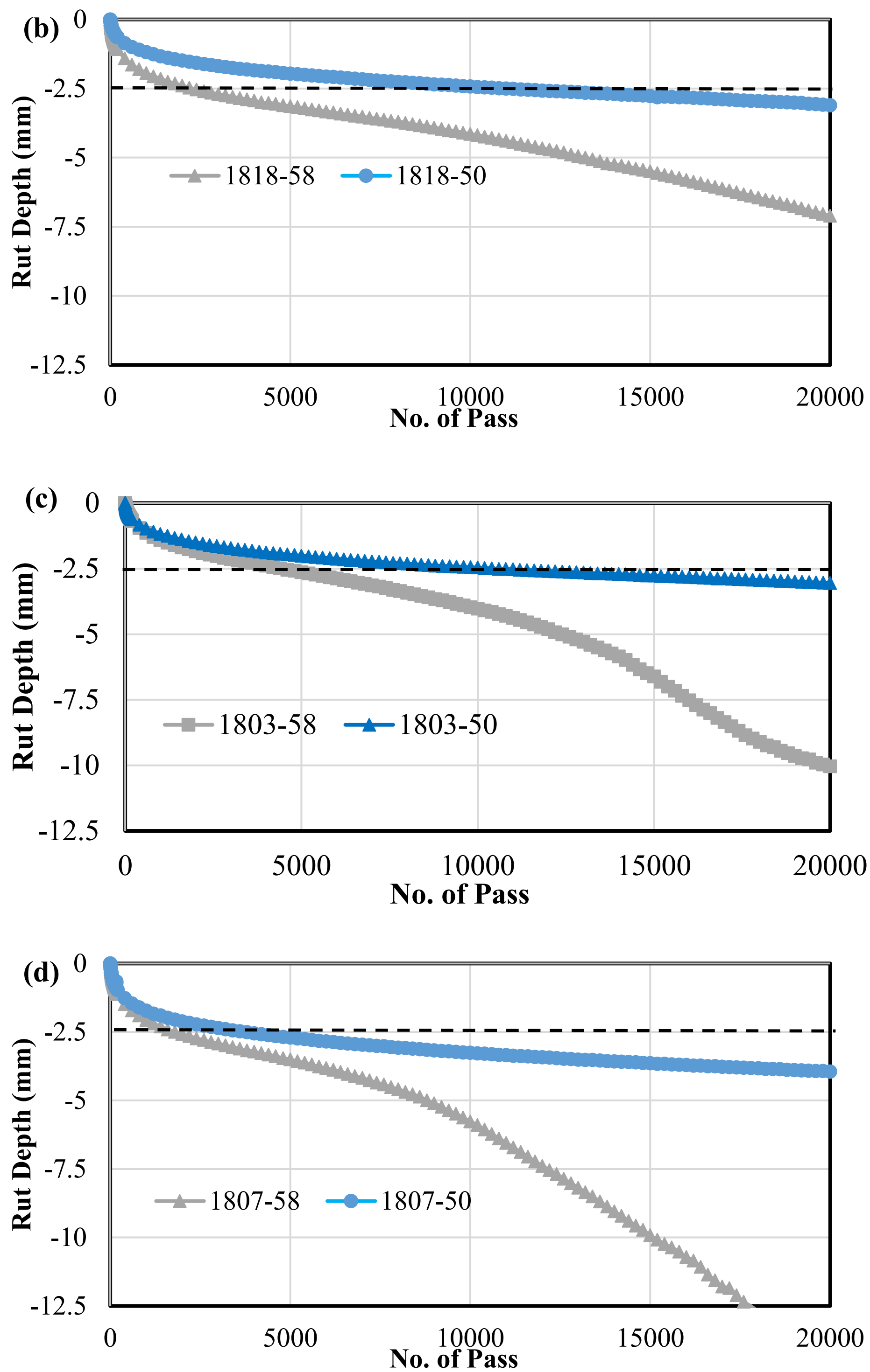


Figure 7-7. Comparing Hamburg testing results at two different temperatures, a) 1829, b) 1818, c) 1803, d) 1807

Table 7-1. Number of passes to reach 2.5 mm rut depth

Mix	58 °C	50 °C	Diff. in No. of Passes
1829	2900	6200	3300
1823	1100	4500	3400
1818	2400	11000	8600
1803	4500	10600	6100
1807	1800	3800	2000
		AVG	**4680**

As already shown in Chapter 3, Figure 3-11, the only mixture which could not meet the existing Hamburg requirements was the 1828 mix. This mix is an IL-4.75 type mixture, which is used below the road surface (at least 2 in. underneath the top of the pavement). Previously, the Tollway called for maximum rut depth of 9 mm under 15,000 passes. As the traffic load in the binder course is not as high as the surface due to the reduction in vertical stress with depth, the required number of wheel passes was reduced from 20,000 to 15,000. Further considering the effect of depth on pavement temperature, another 5,000 pass reduction in the number of required wheel passes is recommended. Figure 7-8 shows the recorded rut depth by this mix at two different temperatures including 50 and 40 °C. As shown in Chapter 3, this mix did not perform well at 50 °C and exceeded the existing rut threshold at 15,000 passes. However, now considering a lower recommended required number of wheel passes of 10,000, the rut depth would decrease to 7.5 mm, which is within the allowable rut depth. Assuming that this mix will be used 100 mm (4 in.) below the surface, the pavement temperature would be around 42.2 °C, using the LTPP bind data presented in Figure 7-6. In this case, the rut depth measured at 40 °C (shown in Figure 7-8) assures that the maximum rut depth even at 20,000 passes will be less than 4 mm, which is quite negligible.

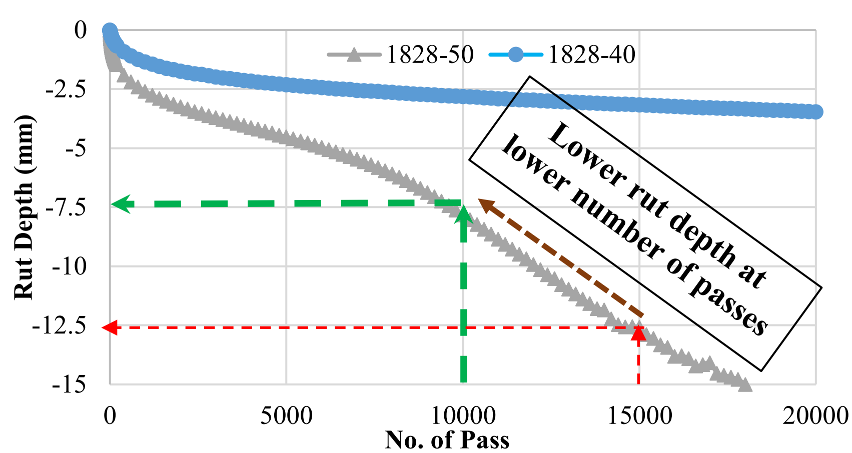


Figure 7-8. Computing a shift in number of passes for 1828 mix

7.4. DC(T) Spec Development

In this section, the flowchart introduced in Figure 7-2 is applied to develop the baseline DC(T) thresholds for different mixture types, which is used in the final consensus/adjustment step. Different borderlines were selected for DC(T) fracture energy based on field observations and stress analysis, such that designing below those limits would very likely result in crack prone mixtures. Reliability was then built in by building on these thresholds. The first upwards adjustment was to take the borderline DC(T) fracture energy levels and raise then to account for the aging that the field samples have experienced. Two levels of aging adjustments (15 and 10%) have been considered for surface and binder mixtures, respectively. Afterwards, assuming that the DC(T) fracture energies obtained from testing the replicates follow a normal distribution, standard deviations according to different reliability levels for differing mixture types were developed. We now review these calculations for each mix type investigated.

7.4.1. SMA Friction Surface Mixtures

Figure 7-9 presents the framework of DC(T) specification development for SMA friction surface mixtures. Considering the SMA mixtures used in sections such as I294, I90-6.6, and I90-6.0, the fracture energy recorded in I90-6.6 (560 J/m^2) was selected as the borderline for the SMA friction surface mixture category. The I90-6.0 section had an SMA fracture surface mix with fracture energy of 830 J/m^2 (see Figure 5-1) and performed very well in-situ. On the other hand, the I294-34 SMA friction surface mix possessed a fracture energy of 451 J/m^2 and experienced significant field cracking. Although the I90-6.6 section used an SMA surface mix, this section is near the I90-6.6 section and therefore had similar environmental and loading conditions. It is also worth mentioning that based on the field performance information, especially "O" cracking data, this section just reached the "O1" severity level meaning that the transverse cracking severity is changing from hairline cracking to infrequent open cracks (see Table 6-1). The 560 J/m^2 borderline set for this mixture type is obtained after testing the sections that are at least eight years old. Based on available literature (Braham et al., 2009) and also preliminary age testing results on laboratory aged samples in this study, a 15% increase was applied to account for aging on the tested field cores and switching the reference aging level for the DC(T) for the short-term aging level used in design. Increasing by an additional 15 % resulted in DC(T) fracture energy of 644 J/m^2.

Like all performance tests, the DC(T) test has an inherent, non-zero COV. SMA friction surface type mixtures have been tested in both plant produced and field core sample types. The standard deviations for both sample types were reported in Chapter 3 and Chapter 4. As a highly simplified and conservative statistical approach, adding the two averaged standard deviations (plant and field core samples) to the previously calculated 644 J/m^2 results in DC(T) fracture energy threshold of 784 J/m^2. After rounding this value to the nearest 25 J/m^2, a DC(T) fracture energy threshold of 775 J/m^2 is completed and recommended for the DC(T) specification for this mix type. It should be mentioned that the two standard deviation level selected for this mix type corresponds to a minimum 95 % reliability (higher due to rounding), which is believed to be an appropriately high level for SMA friction surfaces. This mix is of high criticality, as SMA surface friction mixtures are used on sections with high traffic load (high criticality projects) and/or curves to provide skid resistance. Therefore, this high reliability level helps ensure a high degree of cracking resistance in the mixture, especially in terms of controlling low temperature cracking. This very high fracture energy level will also slow the rate of reflective cracking. Finally, after discussion, consensus was reached to use a threshold of 775 J/m^2 for this mix type, which was supported by the fact that almost all of the plant SMA friction surface mixtures produced in 2018 met this threshold.

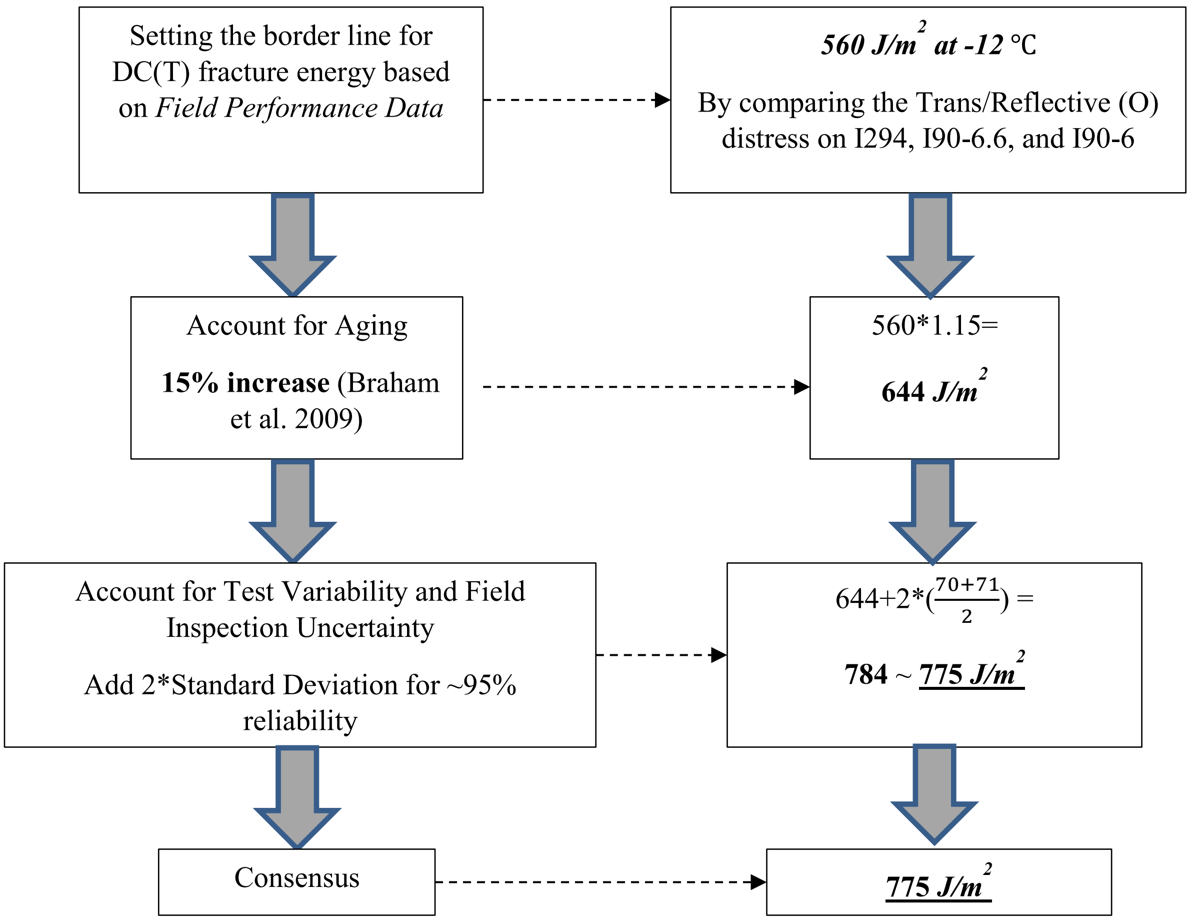


Figure 7-9. Flowchart to Calibrate DC(T) spec for SMA friction surface mixtures

7.4.2. SMA Surface Type Mixtures

The DC(T) specification for SMA surface mixtures is very similar to the one developed for SMA friction surfaces. The only difference in the DC(T) threshold setting is the reliability selected for this type of the mix. For the SMA surface mixtures, 1.5 times the averaged standard deviations was selected, which corresponds to a minimum 87 % reliability in DC(T) fracture energy results based on testing variability. After rounding the calculated threshold to the nearest 25 J/m^2, a fracture energy threshold of 725 J/m^2 was proposed to TRP. However, according to the TRP experience regarding the aggregate types normally used in this mixture type, a consensus to round down to 700 J/m^2 for mix economy was reached. This also provides more spread between the SMA friction surface and SMA surface mixes, which will encourage tailored, unique mix design for the two different categories.

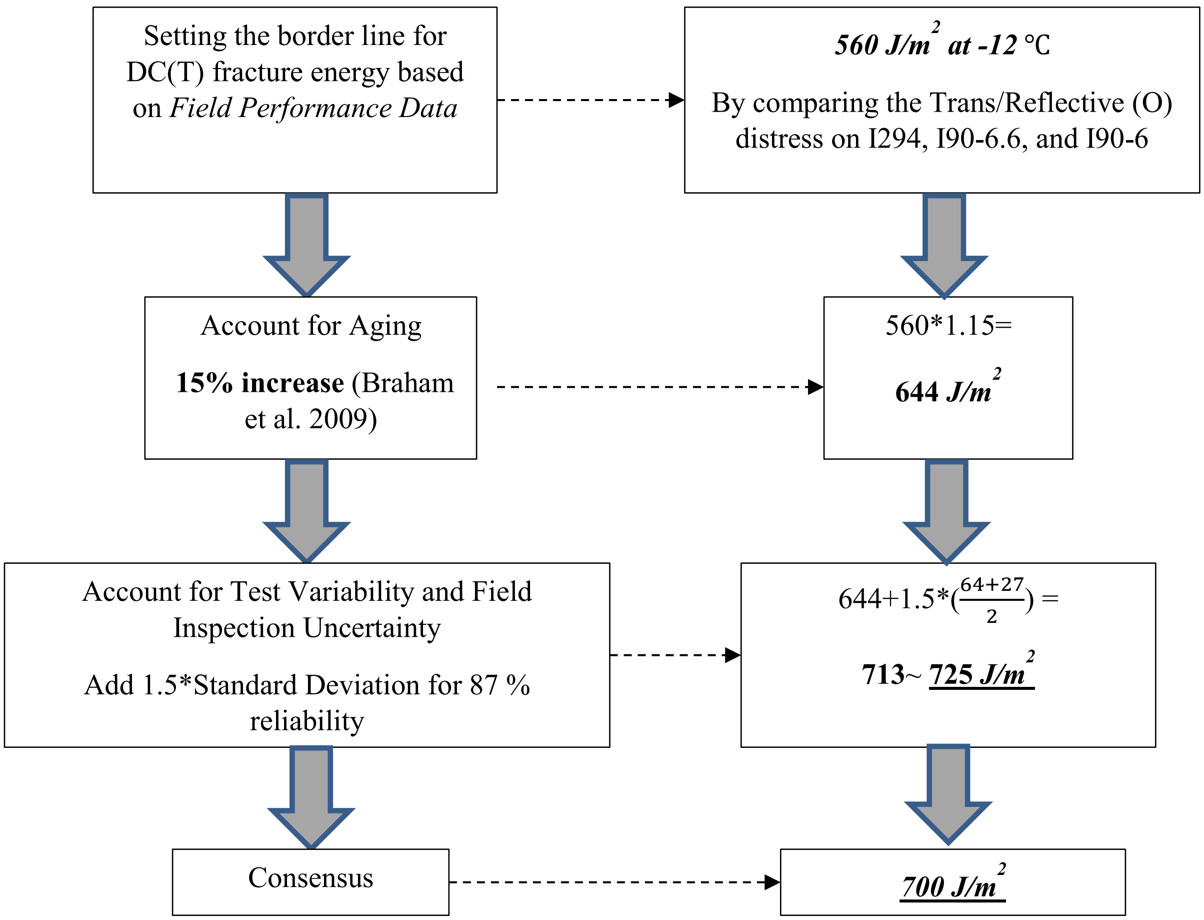


Figure 7-10. Flowchart to Calibrate DC(T) spec for SMA surface mixtures

7.4.3. SMA Binder Type Mixtures

According to the stress analysis discussed in earlier, linear scaling in fracture energy suggests that 70 % of the fracture energy required for the surface mixture should be used to establish a baseline for the SMA binders (560*70/100=392 J/m^2). In order to consider the effect of aging, a 10 % increase in DC(T) fracture energy was assumed. This acknowledges that the aging experienced in pavement sublayers will be lower than that of the surface layer. As the SMA binder mix was only tested in the form of field cores, a reliability of 95 % was achieved by taking two standard deviations based on the field core test results for this high criticality mainline layer. This led to a DC(T) fracture energy threshold of 600 J/m^2 after rounding. During consensus discussions, it was acknowledged that Tollway has already been using a 650 J/m^2 limit for this layer, which encourages higher quality ingredients. Also, since this mix is normally used on jointed concrete pavement, additional fracture energy is thought to help slow down the rate of reflective cracking. Therefore, a consensus was reached that 650 J/m^2 should be retained for the SMA binder type mixture. In the future, this can be revisited, if full-depth asphalt sections gain popularity for major Tollway rehabilitation efforts such as rubblization projects. Less expensive mainline binder courses could be used in these instances where reflective cracking is not of concern. In this case, the 600 J/m^2 fracture energy threshold could be used.

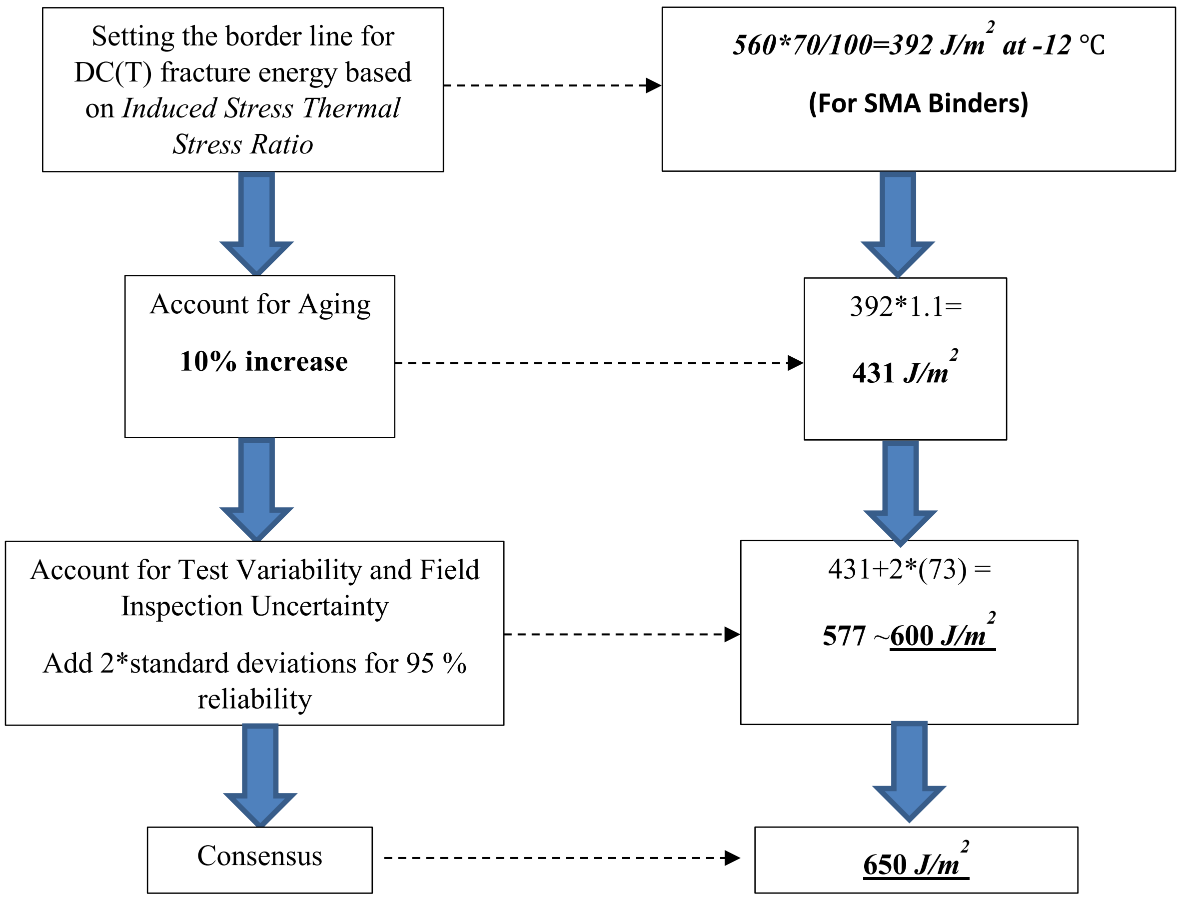


Figure 7-11. Flowchart to Calibrate DC(T) spec for SMA binder mixtures

7.4.4. SMA Shoulder and Dense Shoulder Type Mixtures

The DC(T) specification for two types of shoulder mixtures, including unmodified SMA and dense graded is shown in Figure 7-12. As already shown in Figure 7-1 ("bubble plot"), a DC(T) border line of 400 J/m^2 was set as a limit between highly cracked and low cracked sections in previous studies. Looking at the DC(T) fracture test results shown in Chapter 5 (see Figure 5-1), it can be noted that both good and poor performing sections yielded fracture energy levels around 400 J/m^2. Also, Figure 7-13 indicates that there is a considerable difference between the performance of the shoulder mixtures when the DC(T) fracture energy is above 500 J/m^2. Given all these pieces of evidence, the DC(T) borderline for shoulder mixtures was set to 400 J/m^2 as a starting point in the flow chart. Both mixture types (unmodified SMA and dense graded) are expected to experience the same aging level as SMA surface mixtures. The only difference between unmodified SMA and dense graded shoulders is the criticality of the project such that a higher reliability (87 % or 1.5*SD as opposed to 68 % for dense graded) is set for the unmodified SMA as this mixture type might be exposed to higher traffic during construction. After discussion with the TRP, thresholds of 500 J/m^{2}and 450 J/m^2 were selected for unmodified SMA and dense graded shoulder surface mixtures, respectively.

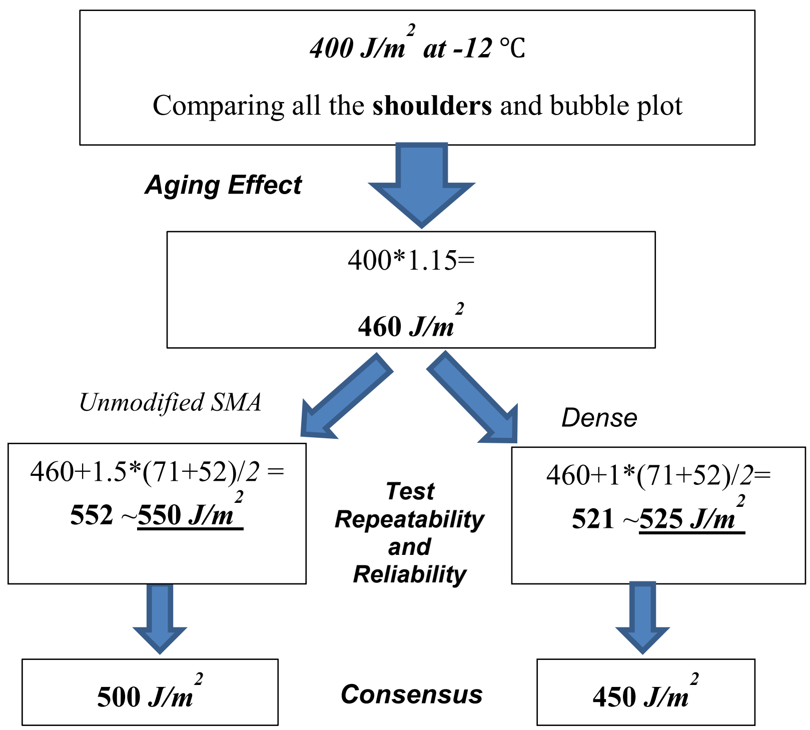


Figure 7-12. Flowchart to Calibrate DC(T) spec for shoulder surface mixtures including “Unmodified SMA”, and “Dense” shoulder surface mixtures

Section	I90-5.12	I90-7.25	I88-57	I-88 (MP:78-Mix:1818)	I355 (Mix1834)
Year	2009	2009	2014	2018	2018
DC(T)	345 J/m^2	370 J/m^2	409 J/m^2	426 J/m^2	512 J/m^2

Figure 7-13. Comparing shoulders with different ages and fracture energies

7.4.5. *Shoulder Binder Mixtures*

The final category for DC(T) specification development is the shoulder binder. This pavement layer can be constructed in different lifts (thicknesses). In this section, two different lifts are studied. The first shoulder binder lift is placed below the shoulder surface and will have a cover of about 2 inches (50 mm) on top; whereas the second binder lift is constructed prior to the top lift binder lift and benefits from a thicker cover (more than 4 inches or 100 mm). Therefore, the required DC(T) fracture energy can be relaxed for the bottom lift of shoulder binder. Considering two different categories (lifts) for the shoulder binder mixture type can result in more economical asphalt mixtures while the environmental and loading conditions have been considered.

Figure 7-14 shows the procedure to calibrate fracture energy for the two different lifts of the shoulder binder mixtures. As shown, a higher reliability is applied for the first lift as compared to the bottom lift. After this step, 450 and 400 J/m^2 were arrived at as thresholds for the top and lower should binder course lifts, respectively. After consultation with the TRP, it was decided that these two thresholds should be consolidated into a single category, using the average value of the two categories, namely 425 J/m^2).

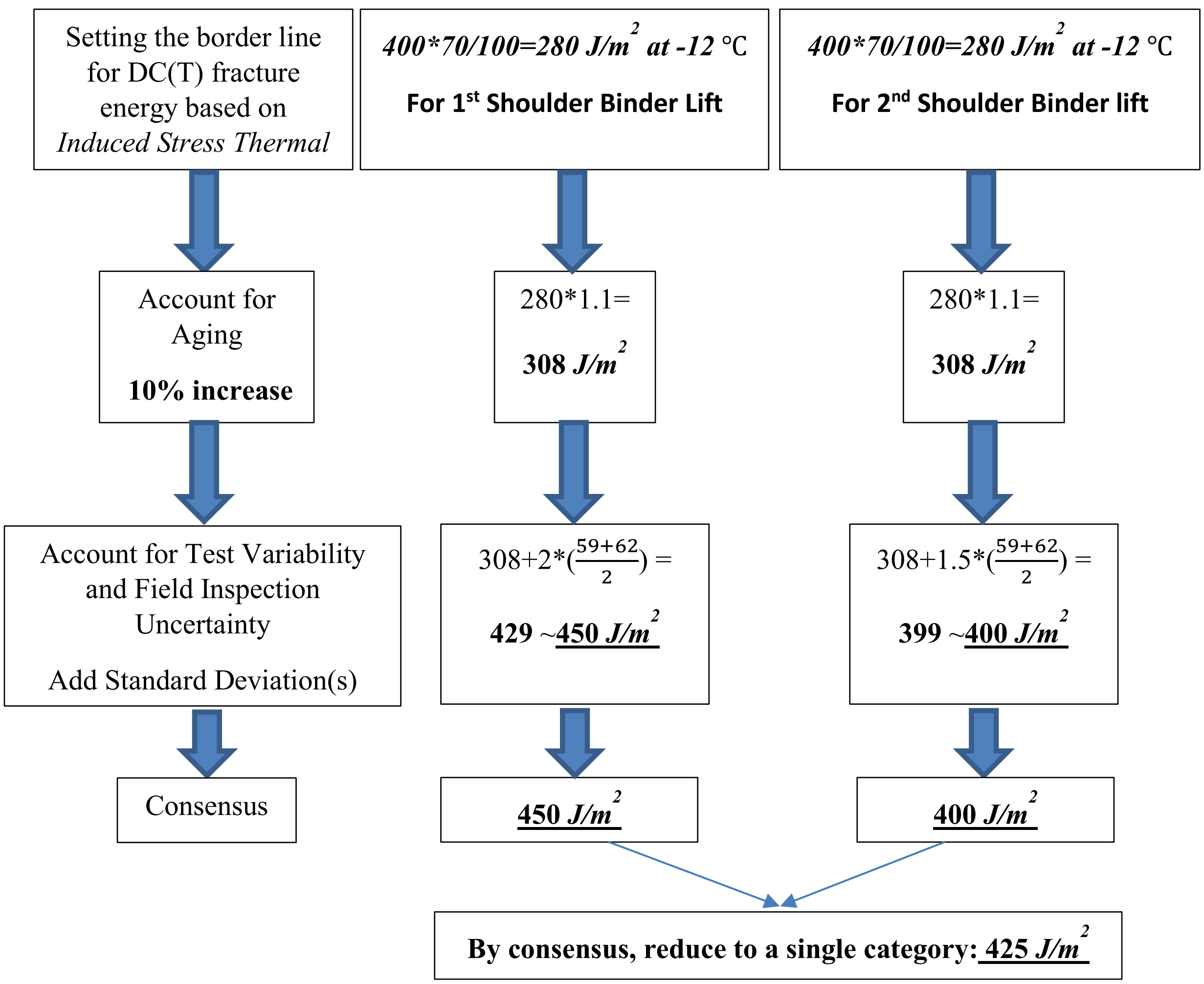


Figure 7-14. Flowchart to calibrate DC(T) thresholds for shoulder binder mixtures

7.5. Performance-Based Specification Levels for DC(T) Fracture Energy

Recommended DC(T) thresholds for different mixture categories are presented in Table 7-2. Experimental results, field performance data, statistical analysis and a final consensus step were used to validate or to adjust the thresholds. The SMA friction surface mixture threshold is recommended to be raised to 775 J/m^2, as shown in Table 7-2. Given the higher aggregate quality used in this mixture and the results from testing of SMAs produced in 2018, it is expected that asphalt producers will be able to meet this threshold with well-designed mixes and high quality materials. In the case of shoulder mixes, it is believed that the elevated DC(T) requirements will lead to lower thermal and block cracking occurrence. Comparing the dense graded shoulder surface mixtures produced in 2018 with the recommended threshold (450 J/m^2), the 1818 mixture would need to be redesigned to meet the new criterion in 2020. This mix used a PG 64-22 binder along with more than 20 % ABR. A softer binder system (e.g. PG 58-28) could help this mixture pass the newly recommended threshold without sacrificing recycled content.

In addition to the mixture types studied herein, the Tollway has two additional mainline binder course mixes in their latest specification: Mainline Binder (N_{design}>50) and Mainline Binder (N_{design}=50). Although not common on the Tollway, these mixes can be used as economical lower layers in full-depth asphalt pavement structures. In setting thresholds for these mixes, it was first acknowledged that shoulder binder course mixture requires a minimum of 425 J/m^2 for durability against environmental cracking. This value was applied to the Mainline Binder (N_{design}>50) mixture, which is near the middle of the pavement structure and has the lowest requirements in terms of cracking resistance. This provides an opportunity to utilize mixtures with higher ABR. On the other hand, the N_{design}=50 mixture will be at-or-near the bottom of the full-depth pavement structure and will therefore carry more bending-related tension. It was decided by consensus to require a slightly higher fracture energy value of 450 J/m^2 for this layer. Some agencies refer to lifts placed at the bottom of full-depth pavement structures as rich-bottom base mixtures, and likewise use specification criteria to promote extra cracking resistance.

Table 7-2. DC(T) thresholds at -12 °C for different mix categories

Mix. Type	Category	Existing	Recommended
SMA	Friction Surface	750 J/m^2	775 J/m^2
	Surface	700 J/m^2	700 J/m^2
	Binder	650 J/m^2	650 J/m^2
	Unmodified	500 J/m^2	500 J/m^2
Dense graded	IL 4.75	450 J/m^2	450 J/m^2
	Mainline Binder (N_{design}>50)	N/A	425 J/m^2
	Mainline Binder (N_{design}=50)	N/A	450 J/m^2
	Shoulder Surface (N_{design}≤70)	N/A	450 J/m^2
	Shoulder Binders	N/A	425 J/m^2

7.6. Performance-Based Specification Thresholds for Hamburg Rut Depth

The recommended thresholds for the Hamburg test to mitigate rutting are presented in Table 7-3. As mentioned before, there were no rutting prone section identified on the Tollway system. This indicates that Tollway has been screening the mixtures in an effective manner in terms of high temperature performance. Therefore, only minor changes have been proposed in the Hamburg requirements. In addition, a procedure was developed to use the Hamburg test as the primary screening tool for mixture stripping, with the classic TSR test used as a second screening step only when failing results are obtained (as explained in section 7.7). This procedure is recommended as a way to avoid the time and testing expense associated with the TSR test in cases where the Hamburg test returns a non-stripping determination.

An SMA binder category has been added to the previous thresholds. The existing thresholds for this mix categoty were not changed, and the same number of passes (20,000) and similar maximum rut depth threshold compared with the SMA friction surface and SMA surface mixtures will be used. For unmodified SMAs, which are used on shoulders, the existing maximum rut depth of 9 mm is suggested to be relaxed to a threshold value of 12.5 mm. This recommendation was made in order to provide additional room for mixture designers to increase the cracking resistance as the shoulders were mainly exhibiting thermal and block cracking, rather than rutting. This threshold is also recommended for the IL4.75 mixtures (9.0 mm limit also recommended to be increased to 12.5 mm).

Table 7-3. Hamburg rut depth thresholds at 50 °C for different mix categories

Mix. Type	Category	Existing		Recommended	
		No. of Passes	Max. Rut Depth	**No. of Passes**	**Max. Rut Depth**
SMA	**Friction Surface**	20,000	6.0 mm	20,000	6.0 mm
	Surface	20,000	6.0 mm	20,000	6.0 mm
	Binder	20,000	6.0 mm	20,000	6.0 mm
	Unmodified	15,000	9.0 mm	15,000	12.5 mm[1]
Dense graded	**IL 4.75**	15,000	9.0 mm	15,000	12.5 mm
	Mainline Binder (N_{design}>50)	15,000	12.5 mm	15,000	12.5 mm
	Mainline Binder (N_{design}=50)	10,000	12.5 mm	10,000	12.5 mm
	Shoulder Surface (N_{design}≤70)	15,000	12.5 mm	10,000	12.5 mm
	Shoulder Binders	10,000	12.5 mm	7,500	12.5 mm

[1]By consensus, the TRP decided to retain a maximum rut depth of 9.0 mm until more field data is available. This is a new mix category for the Tollway.

As mentioned in the previous section, in addition to the mixture types studied herein, the Tollway has two additional mainline binder course mixes in their latest specification: Mainline Binder (N_{design}>50) and Mainline Binder (N_{design}=50). In the Hamburg specification, the effect of pavement depth on reducing temperature has been considered and the required number of passes recommended for the mainline binder course mix (Ndesign=50) and shoulder binder mix were decreased by 5,000 and 2,500 as compared to the existing limits. It is worth mentioning that the requiremenets for mainline binder mix (Ndesign>50) and IL-4.75 mixtures, the recommended rut depth threshold and required number of passes did not change. Based on consultation with the TRP, it was agreed that an intermediate requirement of 12.5 mm maximum rut depth would be appropriate.

7.7. Performance-Based Specification Thresholds for SIP and Use of TSR Test

Table 7-4 presents the recommended SIP thresholds (minimum number of wheel passes at SIP) to control moisture damage. Similar to other Hamburg specifications, the SIP thresholds are either 5,000 or 2,500 cycles less than the required number of passes recommended for a given mix type. The field investigations did not indicate any stripping prone sections, implying that major changes to component material composition or to mix volumetrics is unnecessary for the Tollway. Therefore, the existing thresholds for SIP parameter are recommended to be only modestly changed.

First, recall that the Iowa method for SIP computation involves a pre-screening step. In other words, the first opportunity to specify a mix as non-stripping is in cases where the computed stripping slope-over-creep slope is below the 2.0 threshold. In this pre-screening step, if the criterion is met, there is no need to compute the SIP. There is also no need to check that value against the SIP threshold. <u>In addition, multiple observations led to the recommendation of a second pre-screening step (to be applied only to SMA mixes), to specify a mix as non-stripping</u> (and likewise, eliminating the need to compute SIP and check versus the SIP threshold). The data leading to this observation can be summarized as follows:

- Referring to the SIP results presented in Chapter 3, the slope ratio recorded by SMA mixtures, including sections 1835 and 1845, was above 2.0, while the rut depth accumulated at the end of 20,000 passes was remarkably low (less than 4.0 mm), showing no evidence of moisture susceptibility.

- In addition, the TSR and boiling water test results further evidenced adequate resistance to moisture damage in those SMA mixtures.

- Further investigation revealed that the very low creep slope (close to zero) were the cause of stripping-over-creep slope ratios greater than 2.0. The greater than 2.0 slope ratio triggered the SIP computation, and the flat curves appeared to produce arbitrary SIP numbers. This led to a false-positive stripping detections in the 1835 and 1845 mixes.

- In order to avoid false-positive determinations in highly rut-resistant SMA mixes, a second pre-screen step is recommended: when the rut depth at 20,000 passes is less than or equal to 4 mm, the mix is specified as non-stripping. A subsequent SIP calculation is not required. This can also lead to a drop in average COV of SIP presented in Figure 3-18.

While the use of the Hamburg test as the primary screening tool will certainly save time and testing expense in the mix design stage, it is acknowledged that over-screening of stripping resistant mixes may occur. Thus, until more field data is available, the classic TSR test may be utilized by mix designers as a secondary stripping determination. More specifically, if the asphalt mixture under evaluation does not pass via the Hamburg pre-screening steps or the SIP requirement, the TSR test can be subsequently performed. For mixes meeting or surpassing the 85 % retained tensile strength and 80 psi tensile strength criteria in the TSR, the mix can be specified as non-stripping.

Table 7-4. SIP thresholds at 50 ℃ for different mix categories (applied when the slope ratio is ≥ 2.0)

Mix. Type	Category	Existing	Recommended
SMA*	Friction Surface	15,000	15,000
	Surface	15,000	15,000
	Binder	15,000	15,000
	Unmodified	15,000	10,000
Dense graded	IL 4.75	10,000	10,000
	Mainline Binder (N_{design}>50)	10,000	10,000
	Mainline Binder (N_{design}=50)	5,000	7,500
	Shoulder Surface (N_{design}≤70)	10,000	7,500
	Shoulder Binders	5,000	5,000

* If the measured rut depth for SMA mixes at the required number of passes (determined based on Table 7-3) is lower than 4.0 mm, the mix shall be specified as non-stripping without the need to compute the SIP.

Chapter 8

SUMMARY AND CONCLUSIONS

Pavements are a critical element in our nation's infrastructure, providing personal mobility, economic development and enhancing national security. Asphalt pavements constitute more than 90% of the surfaced roads in the United States. Asphalt mix design methods have steadily evolved over time, but still generally involve the combined use of a detailed volumetric design stage, followed by mechanical testing of the mixture to ensure adequate rutting, cracking, and moisture resistance. The Strategic Highway Research Program (SHRP) culminated in the introduction of the Superior Performing Asphalt Pavements (SUPERPAVE) system, with nationally recognized binder grades and mixture design principles. However, SUPERPAVE failed to deliver simple-yet-effective mixture mechanical tests, commonly referred to as performance tests.

Also driving the need to employ performance tests as part of mixture design is the desire to use high amounts of recycled content from various sources, which are not adequately characterized in Superpave binder, aggregate and mixture volumetric tests and specifications. The FHWA has promoted the concept of Performance-Engineered Mix Design (PEMD), wherein Performance-Related Specifications (PRS) play an important role in the mixture design phase.

Various agencies have begun to research, and to introduce performance-related mixture specifications for their respective road networks. However, a comprehensive evaluation of mixture performance test thresholds, rooted in long-term field performance evaluations, has not been undertaken by many. As a result, the Illinois Tollway commissioned a comprehensive study involving the integration of laboratory performance testing with extensive field performance data for mixtures placed over the past decade, and in some cases before, resulting in the validation and calibration of a performance-related asphalt mix design system. Because the Tollway's asphalt mix design specification was already producing favorable outcomes, e.g., already performance-based and creating high performance/economical mixes with high recycling/sustainability features, it was decided to focus on the mixture performance tests and possibly the relaxation of volumetric requirements. However, changes in volumetric requirements were ruled out through consensus with the TRP during the project, and thus the study was further focused on the selection, validation, and calibration of mixture cracking, rutting, and moisture sensitivity tests in the design specification.

The approaches followed in this study, the key findings, and major recommendations are now summarized.

- Cracking control: Various candidates for a mixture cracking performance test were first selected for evaluation based on the results of the literature review. Among these candidates, the DC(T), IDEAL-CT, I-FIT, and IDT tests were evaluated by testing lab-prepared and field materials. The key findings of this extensive experimental campaign were:
  - The indirect tensile (IDT) strength test -12 °C exhibited the best repeatability (lowest COV); however, it did a poor job of distinguishing between mixtures. This is consistent

with findings in the mid-2000's (see Appendix A), which motivated the development of fracture-based mixture tests and/or tests with little-or-no reliance on peak load carrying capacity (a.k.a., the classic strength-of-materials approach).

- The Illinois flexibility index test (I-FIT) exhibited mixed results with respect to its ability to rank the evaluated mixtures based on performance expectations. Some notable unexpected results were obtained. For instance, sections 1836 and 1828 (dense-graded mixes) yielded FI values significantly higher than those recorded for SMAs. Similar SMA designs have shown to be reliable surface mixes for the Tollway, which have held up to heavy traffic for over a decade without exhibiting materials-related cracking failures. Furthermore, the I-FIT test had the highest COV among the cracking tests investigated. The COV can be lowered by employing a technique developed in Illinois; however, the statistical rigor of eliminating a single outlier is viewed as questionable.
- In addition, I-FIT FI results were found to be highly sensitive to air void content and aging level. A small change in air voids such as +/- 1% could result in a 25% or greater fluctuation in the reported FI value. Counterintuitively, increased air voids leads to increased FI scores (i.e., the mix is characterized as being more crack resistant). This is opposite to performance trends observed in the field, and runs counter to trends observed in the long-established, benchmark 4-pt flexural beam fatigue test.
- For long-term aged samples in the lab, and for field cores, a snap-back behavior is sometimes observed during the post-peak regime of IFIT testing. In the observed cases, the peak load exceeded the 10 kN limit of the device and the post-peak behavior was so brittle that it was not possible to collect enough data points and conduct the load-deflection curve fitting calculation. This posed problems for establishing FI limits based on field coring investigations, rendering the test difficult if not impossible to calibrate to existing field sections, especially when combined with the identified air void and test repeatability issues.
- Three replicates were used for evaluation of the IDEAL-CT test on plant-produced lab compacted and field-cored sections. Compared to the I-FIT, the post-peak slope calculation used in the IDEAL-CT method was found to be more repeatable among the replicates. This resulted in a considerably lower COV for the IDEAL-CT test as compared to the I-FIT.
- A strong correlation was observed between IDEAL-CT and FI values. Unexpectedly, both tests ranked some of the dense-graded mixtures as superior to SMA mixtures.
- The DC(T) test was found to rank mixtures in close accordance with expected, relative field performance trends. Some of the fracture energy values measured for the reheated, plant-produced mixtures were slightly below the specification thresholds used during their design. This was attributed to the effects of sample storage and age effects during reheating, especially in light of the mix of recycled materials and rejuvenators used in some of the mixtures investigated.
- The DC(T) test was found to have excellent repeatability for lab specimens, and good repeatability for field-cored specimens. Limited testing on the samples with different air void content did not show a significant impact on the DC(T) fracture energy. These findings agree with previously reported results.

    - Due to its high repeatability, excellent correlation with field results, and ability to change testing temperature based on environmental conditions of the project location, the DC(T) was recommended to be retained in the Tollway's asphalt mixture design specification. In addition, the DC(T) provides mix designers much more leeway to incorporate higher levels of recycled materials as compared to IFIT and IDEAL. These highly sustainable and economical mixes have been used with success on the Illinois Tollway, but do not always receive passing scores in either the IFIT or IDEAL tests. These mixes include those that contain various combinations of RAP, RAS and GTR. The stiffening effects of these recycled components apparently leads to a significant number of false-negative determinations in the IFIT and IDEAL tests and specifications.
    - A systematic procedure was used to set DC(T) limits, which were then compared to existing DC(T) thresholds, followed by a consensus approval/adjustment step. This procedure involved setting baseline fracture energy thresholds for minimum acceptable cracking resistance based on field results, then adding to this baseline in a conservative fashion according to: differences in fracture energy between short- and long-term aging levels, and the sum of test variability expected in both the lab design stage and field calibration stage.
    - A consensus process was also utilized to allow final rounding of DC(T) thresholds based on practical considerations, such as knowledge of the ability of locally available materials to meet specification thresholds for various mix types, sustainability goals, and economical considerations.
    - It was also recommended to introduce DC(T) specifications for newer Tollway mix types that did not have requirements. This included shoulder mixes, where the primary drivers of deterioration are non-load associated (aging, temperature cycling, moisture, freeze-thaw cycles). The introduction of these new performance requirements are expected to have several positive outcomes for shoulder mixes, including: (1) lessening the occurrence of thermal and block cracking observed on some Tollway shoulders; (2) providing a degree of confidence in shoulder rut resistance in instances where traffic is routed onto shoulders during construction and other lane closure operations, and; (3) allowing increased levels of recycled material to be considered, and allowing new recycling types to be examined, while controlling mixture performance with performance testing.
- <u>Rutting Control</u> - Similar to other agencies, the Illinois Tollway has had a positive experience in using the Hamburg test to conservatively control permanent deformation. The ability of the Hamburg to effectively control stripping in lieu of the TSR test was also of interest to the Tollway.
- Hamburg specification thresholds for the various mix types used by Tollway were evaluated as a function the depth of placement of those layers relative to the surface of the pavement. A summary of key research evaluations applied in support of validating or adjusting Hamburg thresholds include:
    - Based on laser-measured rut depths by ARAN-style pavement conditional assessment vehicles, Hamburg testing at 50 °C at 20,000 passes appeared to be appropriate for the control of rutting in Tollway SMA-type mixtures.

- Considerable efforts were made to evaluate the conservative values currently used for binder course and shoulder mixtures to determine if limits could be relaxed, thereby creating more leeway for mix designers to address cracking resistance, enhanced recycling and mixture economy.
- To this end, computed temperature profiles and plots leading to temperate-wheel pass equivalents were used to arrive at more highly tailored specification thresholds.

- Stripping Control - Degradation of the bond between aggregate and binder in the presence of water leads to stripping distress (or moisture damage). In this project, tests were carried out on loose and compacted asphalt materials to evaluate resistance to moisture damage. The tensile strength ratio (TSR) test and stripping inflection point (SIP) parameter calculated using Hamburg test results were obtained for compacted samples while the boiling water test was performed on loose asphalt mixtures.
  - The SIP parameter was calculated using the Iowa method by means of fitting a 6th degree polynomial curve on rut depth vs. number of passes. This frequently used method of SIP calculation takes into account the stripping-to-creep slope ratio, along with the number of passes at which the creep and stripping lines intersect.
  - The Iowa method triggers the SIP consideration only if the slope ratio is equal to or greater than two. Despite the fact that SIP can detect the stripping potential for mixtures with a considerable amount of rut depth at the end of 20,000 wheel passes, its slope ratio might be misleading for mixtures with very low accumulated rut depths, leading to false-positive stripping determinations based on comparisons to actual field performance.
  - Two SMA mixes (sections 1835 and 1845) possessed low rut depths with very low creep slopes, which created an undesirable artifact where high striping-to-creep slope ratios were computed. Given the fact that the total rut depth on those mixtures at the end of the 20,000 passes at 50 °C was lower than 4 mm, it is very unlikely that they will exhibit significant rutting and shoving due to stripping effects during service life. Thus, it is recommended that the Iowa stripping evaluation be waived for mixtures exhibiting less than-or-equal to 4.0 mm of rutting in the Hamburg test at 20,000 passes. In these cases, the mixture should be reported as non-stripping.
  - The 1835 and 1845 mixtures did not show stripping potential based on the TSR test results. This provided further verification that the SIP was unable to correctly screen these mixtures in terms of moisture damage. In addition, these mixes performed satisfactorily in the boiling water test.
  - Based on the results of this study, it is recommended to use the Hamburg stripping determination in lieu of the TSR test (AASHTO T-283) as the primary testing method for moisture damage control. This could lead to time and testing costs savings. However, as a conservative rollout of this method, for the near future it is recommended that the TSR test be retained as a secondary screening tool for stripping. An issue associated with the SIP parameter is its extremely high variability. The SIP parameter had the highest COV among the conducted tests and various parameters investigated.
  - Thus, if a mixture is identified as having the potential for stripping in the Hamburg procedures outlined above, it is recommended that the TSR test be allowed to be used by designers as the final determination of stripping potential.

- Both TSR and SIP parameters indicated that the 1828 mix has stripping potential. This might be attributed to the very fine aggregate structure (NMAS=4.75 mm) which may not have not provided sufficient room for swelled rubber particles in this mix modified with dry-process GTR. The SIP value for this mix was found to be borderline when compared to the current specification (9,861 < 10,000). The asphalt binder residue observed in boiling water test container provided further evidence that stripping potential may exist in this mix.
- Because there were insufficient poor performing sections encountered at the Tollway with respect to moisture damage, only minor adjustments to existing SIP thresholds were recommended, along with the introduction of new thresholds for new Tollway mix types.

The tested plant produced-lab compacted samples were placed in the field in 2018; therefore, insufficient time had elapsed at the completion of the study in 2020 to enable calibration of the Tollway asphalt mixture PRS based on field performance from these sections alone, especially in terms of cracking resistance. Therefore, the research team collaborated with the project TRP to identify a number of good and poor performing sections placed during 2008 to 2015, which were visually surveyed in May of 2019. These included both mainline and shoulder sections. In some cases, previous mix design and performance test results were available, along with reports, scientific papers, and up-to-date ARAN performance data. This provided an excellent opportunity to link the laboratory performance testing results with actual field performance. The following observations were made after this investigation:

- Sixteen different sections including six mainline and six shoulder mixtures were tested in the lab using most of the candidate cracking and rutting tests evaluated in this study.
  - The COV of the DC(T) test was found to be the lowest among different mix categories, although considerably higher in some cases and more variable overall as compared to the 2018 testing results on plant-produced, lab-compacted specimens.
  - Although general trends were observed between the cracking tests and field cracking trends, several predictions made by the I-FIT and IDEAL-CT indexes greatly underestimated field results.
  - Based on the field observations, a DC(T) fracture energy of 400 J/m^2 was determined to be the threshold, long-term aged fracture energy recommended for shoulder surface mixtures.
  - As an across-the-board anomaly, despite the fact that the I90-17.8 section had exhibited extensive surface block cracking in the field, this mix performed was scored as a good performer in all three cracking tests. This may be attributed to the fact that the particular aggregate used in this section was very hard, and had not been screened by any of the tests in the past. The TRP did not recommend making adjustments to cracking thresholds based on this result, due to the relative obscurity and lack of future supply of this aggregate.
- Performance indices and parameters such as IRI, CRS, and rut depth along with images collected from annual ARAN pavement evaluations were used to compare the performance of different sections and to correlate them to lab and field observations.

  - The I294-34 section is one of the older sections, is subjected to high levels of slow-moving truck traffic, and was paved on existing jointed concrete pavement. A high amount of reflective cracking and a high IRI has been measured on this section, leading to the lowest overall serviceability of the sections investigated as rated by CRS. On the other hand, the two relatively new sections (I88-47 and I88-60.5) recorded the highest CRS.
  - The IRI index had a good correlation with the CRS values reported. The only notable difference was the ability of CRS to detect the effect of the hairline block cracks on the I90-17.8 section, while the IRI did not, since pavement ride does not appear to have been compromised.
  - Somewhat arbitrarily, the measured field rut depth magnitudes on SMA-surfaced mainline sections were in reasonable agreement with rut depths measured in the Hamburg test. This observation also supports the recommendation to retain the 50℃ testing temperature and 20,000 wheel pass level for SMAs as traditionally used in the Tollway's asphalt mix design specification.
- In addition to overall serviceability as characterized by CRS, a detailed analysis of the type, extent, and severity of surface distresses was conducted on the data set provided by ARA. An average severity concept was introduced, and the following observations were then made:
  - Centerline, longitudinal (or 'center *lane*'), and transverse cracking were the most frequent types of cracks recorded on the studied Tollway sections.
  - Because rutting is simply not observed on Tollway asphalt pavements, the frequency and severity of the cracking forms observed suggest that specification changes should be prioritized to address cracking.
  - Centerline and longitudinal cracks are generally believed to be construction and traffic loading related, respectively. Therefore, the transverse cracking data, with extra weight placed on full-depth asphalt sections where reflective cracking does not exist, were used to calibrate the DC(T) specification, along with block cracking observations gathered during field visits.
  - Finally, a framework for the performance specification evaluation and calibration was developed and deployed. This framework took into account the link between lab and field performance results and builds in the effect of aging on desired specification thresholds. In addition, test repeatability and the uncertainty associated with distress detection and measurement were incorporated using a statistically based reliability approach. A subset of the TRP was convened to discuss the rounding and adjustment of specification thresholds based on practical considerations such as material availability and local economics.
  - It is believed that meeting the newly proposed test thresholds will strike an even better balance in mixture performance and mixture economy for the Tollway, and will have particular benefits for the longevity and economics of shoulder mixes. In addition, the specification continues to keep the door open for future innovations. These include the introduction of new, sustainable mixture design approaches and materials as they become available.

APPENDIX A

LITERATURE REVIEW

List of Abbreviations

AASHTO	American Association of State Highway and Transportation Officials
AC	Asphalt Concrete
ALDOT	Alabama Department of Transportation
AHTD	Arkansas State Highway and Transportation Department
AMPT	Asphalt Mixture Performance Tester
APA	Asphalt Pavement Analyzer
ARA	Applied Research Associates
ASTM	American Society for Testing Materials
BBF	Bending Beam Fatigue
BBR	Bending Beam Rheometer
BDWSC	Bridge Deck Water-proofing Surface Course
BF	Flexural Beam-Fatigue
BRBC	Bottom Rich Base Course
BRIC	Bottom Rich Intermediate Course
BSI	British Standards Institution
C	Pseudo Secant Modulus
CMOD	Crack Mouth Opening Displacement
CRM	Crumb Rubber Modifier
CTEC	Coefficient of Thermal Expansion and Contraction
CTOD	Crack Tip Opening Displacement
CTSD	Crack Tip Sliding Displacement
DBN model	Di-Benedetto-Neifar model
DC(T)	Disc-Shaped Compact Tension Test
DM	Dynamic Modulus
DOT	Department of Transportation
EVAC	East Valley Asphalt Committee
FDOT	Florida Department of Transportation
FPBF	Four-Point Beam Fatigue
FI	Flexibility Index
GDOT	Georgia Department of Transportation
GDT	Georgia Development Test
GTR	Ground Tire Rubber
FHWA	Federal Highway Administration
HCT	Hollow Cylinder Tensile Tester
HiPO	High Performance Thin Overlay
HMA	Hot Mix Asphalt
HPTO	High Performance Thin Overlays
HWTT	Hamburg Wheel Tracking Test
IDEAL-CT	Indirect Tensile Asphalt Cracking Test
IDOT	Illinois Department of Transportation
IDT or ITT	Indirect Tensile Test
ITP	Illinois Test Procedure
LDOTD	Louisiana Department of Transportation
LTRC	Louisiana Transportation Research Center
LWT	Loaded Wheel Tester

MEPDG	Mechanistic-Empirical Pavement Design Guide
MDOT	Michigan Department of Transportation or Maryland Department of Transportation
MnDOT	Minnesota Department of Transportation
MoDOT	Missouri Department of Transportation
MOT	Ministry of Transportation
NAPA	National Asphalt Pavement Association
NCAT	National Center for Asphalt Technology
NCDOT	North Carolina Department of Transportation
NCHRP	National Cooperative Highway Research Program
NEPPP	Northeast Pavement Preservation Partnership
NMAS	Nominal Maximum Aggregate Size
ODOT	Oregon Department of Transportation or Ohio Department of Transportation
OT	Texas Overlay Test
PAN	Polyacrylonitrile
PBS	Performance Based Specification
PCC	Portland Cement Concrete
PMS	Pavement Management System
PRS	Performance-Related Specification
PPA	Poly-Phosphoric Acid
QA	Quality Assurance
QC	Quality Control
QRSS	Quality-Related Specification Software
RAP	Recycled Asphalt Pavement
RAR	Reacted and Activated Rubber
RAS	Recycled Asphalt Shingles
RDEC	Ratio of Dissipated Energy Change
RLPD	Repeated Load Permanent Deformation
RSST	Repeated Simple Shear Test
S	Damage Parameter
SBS	Styrene Butadiene Styrene (polymer)
SCB	Semi-Circular Bending Test
SE(B)	Single Edge Notched Beam Test
SGC	Superpave Gyratory Compactor
SHRP2	Strategic Highway Research Program 2
SMA	Stone Matrix Asphalt
SPT	Simple Performance Test
TCAP	Thermal Cracking Analysis Package
TRB	Transportation Research Board
TSP2	Transportation System Preservation Technical Services Program
TSR	Tensile Strength Ratio
TSRST	Thermal Stress Restrained Specimen Test
TxDOT	Texas Department of Transportation
UDOT	Utah Department of Transportation
UIUC	University of Illinois Urbana Champaign
VDOT	Virginia Department of Transportation

VECD	Viscoelastic Continuum Damage
WisDOT	Wisconsin Department of Transportation
WLT	Wheel Load Tracking
WSDOT	Washington Department of Transportation
WMA	Warm Mix Asphalt

A review of the existing literature related to asphalt performance specifications was undertaken. The literature review included the following topics: asphalt performance testing, agency practices regarding Performance Based Specifications (PBS), and studies related to PBS development and test methods and protocols related to PBS. The results of the literature review are described in the following sections.

A.1. Asphalt Performance Testing

A.1.1. Overview of Asphalt Performance Testing

Asphalt roads make up for more than 90% of the USA's pavement infrastructure (NAPA, 2009). Such wide usage warrants a sturdy and robust design of the asphalt mix to make it last long and wear less. Before the Superpave mixture design protocol was in place, a lot of procedures followed for designing mixes to address particular distresses were empirical, for example Hubbard-Field or Marshall Test used to predict permanent deformation (Fujie Zhou, Hu, & Scullion, 2006). Superpave introduced volumetric mix design, which, albeit being a step forward, had considerable gaps in accurately predicting the field performance of the asphalt mixtures. This shortcoming motivated researchers to adopt laboratory mixture testing under simulated loading mimicking the field conditions. These 'performance tests' can provide key insights to the mixture field performance in terms of fatigue cracking, thermal cracking, rutting, moisture resistance, and other key distresses. Table 1 presents few such laboratory mix performance tests.

Table A-1. Assessment of Available Performance Tests for Use in Routine Mixture Design (Federal Highway Administration (FHWA), 2013)

Property	Method	Standardization		Criteria		Complexity		Equipment		
		Method	Precision	Mix Design Pass/Fail	Performance Prediction Model	Method	Analysis	Availability	Reliability	Cost
Modulus	AMPT Dynamic Modulus	Yes	Yes	NA	Yes	Moderate	Moderate	Yes	High	Moderate
	Simple Shear	Yes	Yes		Yes	Moderate	Moderate	Yes	Moderate	High
	Indirect Tension Dynamic Modulus	No	No		Yes	Moderate	Moderate	No	Moderate	High
Permanent Deformation	AMPT Flow Number	Yes	Yes	Yes	Yes	Moderate	Low	Yes	High	Moderate
	Repeated Shear	Yes	Yes	Yes	Yes	Moderate	Moderate	Yes	Moderate	High
	High Temperature Indirect Tensile Strength	Draft	No	Yes	No	Low	Low	Yes	High	Low
	Asphalt Pavement Analyzer	Yes	No	Yes	No	Moderate	Low	Yes	High	Moderate
	Hamburg Wheel Tracking Device	Yes	No	Yes	No	Moderate	Low	Yes	High	Moderate
Load Associated Cracking	Flexural Fatigue	Yes	No	No	Yes	High	Moderate	Yes	High	Moderate
	AMPT Continuum Damage	No	No	No	Yes	High	High	Yes	High	Moderate
	Energy Ratio	No	No	Yes	Yes	High	High	Yes	High	High
	Fracture Energy	No	No	No	Yes	Low	Low	Yes	High	High
Thermal Cracking	IDT Creep and Strength	Yes	No	No	Yes	High	High	Yes	High	High
	Disc Shaped Compact Tension	Yes	Yes	Yes	Yes	High	Low	Yes	High	High
	Semi Circular Bend	Draft	No	Yes	No	Moderate	Low	Yes	High	High
Moisture Sensitivity	Tensile Strength Ratio	Yes	Yes	Yes	No	Moderate	Low	Yes	High	Low
	Hamburg Wheel Tracking Device	Yes	No	Yes	No	Moderate	Low	Yes	High	Moderate

The following sections summarize the current mixture performance tests grouped according to the parameter they measure or the distress they characterize.

A.1.2. Mixture Tests to Mitigate Thermal Cracking

Overview of Tests for Thermal Cracking

Thermal cracking, or low-temperature cracking, is one of the primary distresses of asphalt pavements in cold climates. As the temperature drops, thermal stresses develop due to the differential contraction of the binder and aggregate in the asphalt mastic. When the thermal stresses exceed the tensile strength of the asphalt mixture, the pavement develops cracks. The popular tests for characterizing thermal stresses in asphalt mixtures are Disc-Shaped Compact Tension test (DC(T)), Indirect Tensile creep and strength test (IDT), Semi-Circular Bending test (SCB), and Thermal Stress Restrained Specimen Test (TSRST).

Overview of IDT Test

The IDT test is performed to ascertain the tensile strength and the creep properties of the asphalt mixture specimen, which are critical factors of thermal cracking characteristics. The test is done in accordance with the AASHTO T322 standard (AASHTO, 2010). The specimen is loaded diametrically, inducing horizontal tensile stress in the mid-portion of the specimen. The creep test is done at 0 ºC, -10 ºC, and -20 ºC, and the tensile strength test is done at -10 ºC. The relaxation modulus data, obtained by converting the creep compliance data, is used to estimate the thermal stresses and calculate the critical cracking temperature (Christensen & Bonaquist, 2004; Mandal, 2016).

Applications of IDT Test

Roque et al. modified the Superpave IDT test to obtain fracture parameters of an asphalt specimen by drilling an 8mm hole in the middle of a 150 mm diameter specimen to ensure crack initiation and propagation under loading (Roque, Zhang, & Sankar, 1999). Kim et al. calculated fracture energy of eight asphalt mixtures using the strains at the center of the IDT specimen and calculating the corresponding displacements. The authors used a 50 mm gauge at the center of a 100 mm diameter IDT specimen subjected to a constant rate of ram movement of 50 mm/min at 20 ºC. The authors found a very good correlation of the fracture energy to the fatigue cracking observed in the tracks where the mixtures were laid (Kim & Wen, 2002). Richardson et al. used the IDT test to determine the creep compliance and tensile strength of HMA mixtures used as wearing course in Missouri, using them as inputs in the thermal cracking module of MEPDG (PavementME). The authors tested six different laboratory-prepared wearing courses with varying air voids and RAP content. Increase in air voids led to increased creep compliance and decreased tensile strength, and addition of RAP to the mixtures led to decreased creep compliance and increased tensile strength (Richardson & Lusher, 2008). Although the repeatability of the IDT test is reported to be very high, there are still concerns regarding the correlation of the IDT strength to the field performance and the capability of this parameter to characterize cracking resistance of the mixtures (Walubita et al., 2011). Furthermore, Jahanbakhsh et al.,(2019) observed that the IDT strength parameter could not capture the effect of additives. This test was not also found to be sensitive to the testing temperature (Zegeye et al., 2012). To address these issues, other parameters such as toughness, slope, and fracture energy were recommended to be used instead of the IDT strength (West et al., 2017; Zborowski and Kaloush, 2011; Yin et al., 2018)

Overview of TSRST Test

TSRST (AASHTO TP 10) (AASHTO, 1993) is a simple test wherein a rectangular asphalt mixture specimen is allowed to cool but is restrained on shorter edges, leading to development of thermal stresses within the specimen and ultimately cracking when the thermal stresses exceed the tensile strength of the specimen. It was developed at Oregon State University as a part of SHRP (Aschenbrener, 1995). After fabricating the specimens, they are glued to the plates and conditioned in a cooling chamber at 5 ºC for an hour to impose thermal equilibrium in the specimen. LVDTs are used to measure the deformation of the sample while the temperature of the chamber is reduced. A closed-loop loading frame is used to restrain the shorted edges of the specimen at the original length, inducing thermal stresses. The end result of this test is a thermal stress-temperature plot.

Applications of TSRST

Tapsoba et al. used TSRST to study mixes with different RAP (up to 25%) and RAS content (up to 10%). The authors found that below 15% RAP content and 5% RAS content, the mixes performed very similar to the virgin HMA mixture. Beyond that, the low temperature cracking resistance of the asphalt mixture specimen deteriorated. Further, they tried to simulate the TSRST test using the Di-Benedetto-Neifar model (DBN model) and found a good correlation (Tapsoba, Baaj, Sauzéat, Di Benedetto, & Ech, 2016). Lei et al. used TSRST to see the effects of bio-based and refined waste oil modifiers on low temperature cracking characteristics of asphalt mixtures. The tests showed a much cooler fracture temperature on modification with oil (Lei, Bahia, & Yi-Qiu, 2015). Mohammad et al. used TSRST to characterize and compare the thermal stresses of the sulfur-modified WMA mixture to conventional HMA mixtures. The authors found no statistical difference in the average fracture temperatures of the mixtures (Mohammed, Cooper, & Elseifi, 2010). Jung and Vinson ranked the low-temperature cracking resistance of 14 asphalt mixtures, calculated using TSRST, with variation in air voids %, aggregate types, aging, etc. The authors found a good agreement of the TSRST-based fracture temperature ranking with the ranking by SHRP on basis of asphalt binder properties. The authors further found that the fracture temperature was most sensitive to the asphalt type and the aging levels, whereas the fracture strength was more sensitive to the air voids and aggregate type (Jung & Vinson, 1993).

Correlation between BBR and TSRST

Falchetto et al. tried to develop a correlation between the Bending Beam Rheometer (BBR) creep and strength results with the critical temperature for asphalt mixture obtained by TSRST. The authors conducted TSRST, BBR creep and strength tests on eight asphalt mixtures with varying RAP contents and applied simple size effect theory to extrapolate the BBR strength to compare it to the TSRST results. The results indicated that BBR strength results could not be used with the thermal stress curves of BBR creep tests to determine the critical temperature of the asphalt mixtures without considering the size effect. Even after taking the size effect into account, the strength values from TSRST and BBR were comparable but the difference in the critical temperature values could not be ignored (Falchetto, Moon, & Wistuba, 2017).

Overview of DC(T) Test

The Disk-Shaped Compact Tension (DC(T)) (ASTM D7313-13) (ASTM, 2013) test is used to measure the low-temperature cracking potential of the asphalt mixtures. Wagoner et al. came up with a suitable configuration (shown in Figure A-1) for this test using ASTM E-399 (ASTM, 2012)as a starting point and then modified it for asphalt materials (W. G. Buttlar, Hill, Wang, & Mogawer, 2016). A big advantage of the DC(T) test lies in its ability to test cylindrical cores obtained from field or compacted in Superpave Gyratory Compacter (SGC) and its large fracture surface area (M. P. Wagoner, Buttlar, & Paulino, 2005b). The DC(T) test temperature is generally 10^{o}C higher than the PG low temperature grade of the binder used in the asphalt mixture. The specimen is pulled through the drilled holes, forcing the crack to propagate in perpendicular direction. The notch is made to pre-determine the crack path. The test is conducted at a constant Crack Mouth Opening Displacement (CMOD) rate of 1mm/min (0.017 mm/s). It is stopped when the post-peak loading reaches 0.1kN. A typical load-CMOD curve is shown in Figure A-2. The area under the curve, normalized by the initial fracture area of the specimen, is reported as the fracture energy of the asphalt mixture specimen. The standard method of testing is outlined in ASTM D7313-13 standard (ASTM, 2013).

Sensitivity Studies for DC(T) Tests

Wagoner et al. conducted brief sensitivity studies on specimen thickness, nominal maximum aggregate size (NMAS), and CMOD rate for DC(T) tests. The authors found that the fracture energy increased with thickness and, more importantly, the variation of results was not affected by the specimen thickness. Further, the studies reported higher variation in results for specimen with 19.0 mm NMAS than for 9.5 mm or 12.0 mm NMAS, with a reasoning that the 19.0 mm NMAS resulted in a lower 'representative volume' in the specimen (M. Wagoner, Buttlar, Paulino, & Blankenship, 2005). Wagoner et al, in a separate study, also reported a decrease in fracture energy with an increase in loading rate (M. P. Wagoner et al., 2005b).

Use of DC(T) to Study Effects of RAP and RAS

Behnia et al. used DC(T) to evaluate the effect of addition of RAP in asphalt mixtures on thermal cracking. The authors found that addition of RAP beyond 10% significantly decreased the fracture energy of the specimen (Behnia, Dave, Ahmed, Buttlar, & Reis, 2011). Arnold et al. showed that addition of RAS to asphalt mixture specimen led to an increase in the peak load and a decrease in the overall fracture energies (Arnold et al., 2014). Dave et al. used DC(T) for low temperature fracture characterization of nine mixes with varying RAP content, aging, and air void content. The results showed that DC(T) test was successfully able to capture the effects of temperature, varying content of RAP, binder modifiers, and aging (Dave, Behnia, Ahmed, Buttlar, & Reis, 2013).

Other DC(T) Applications

DC(T) results by Buttlar et al. from a project at UIUC in collaboration with Illinois Tollway and STATE testing laboratory, show that addition of GTR to asphalt mixture specimens led to significantly higher fracture energies (W. G. Buttlar & Wang, 2016). Behnia et al. used the DC(T) test to evaluate the effect of cooling cycles on the asphalt specimens. The fracture results obtained from DC(T) test were reported to be sensitive to the micro-damages induced by cooling cycles and also correlated well with the assessment of low-temperature behavior by non-destructive testing methods, namely the acoustic emission test (Behnia, Buttlar, & Reis, 2014).

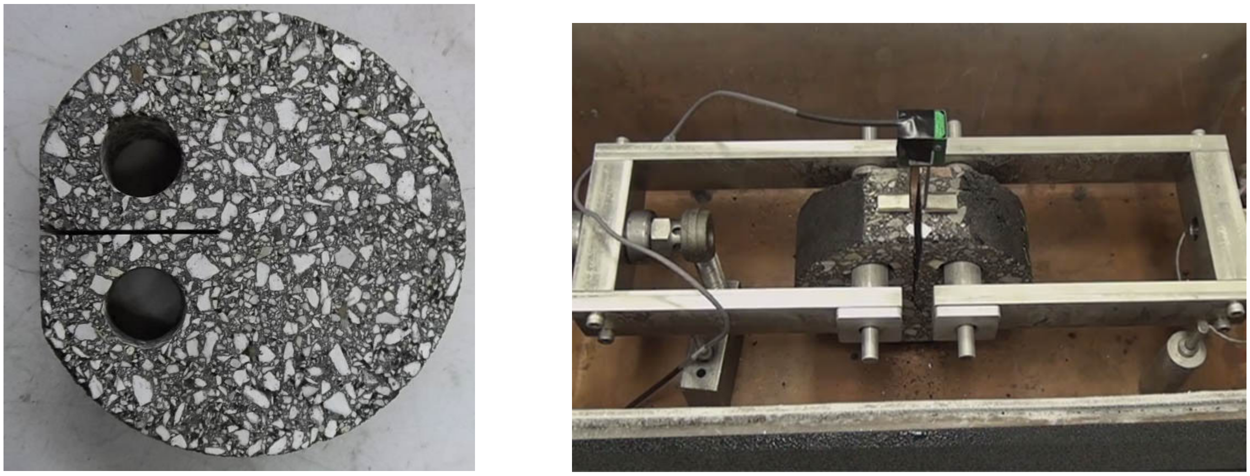


Figure A-1. a) Prepared DC(T) specimen (W. G. Buttlar et al., 2016) b) DC(T) Test set-up

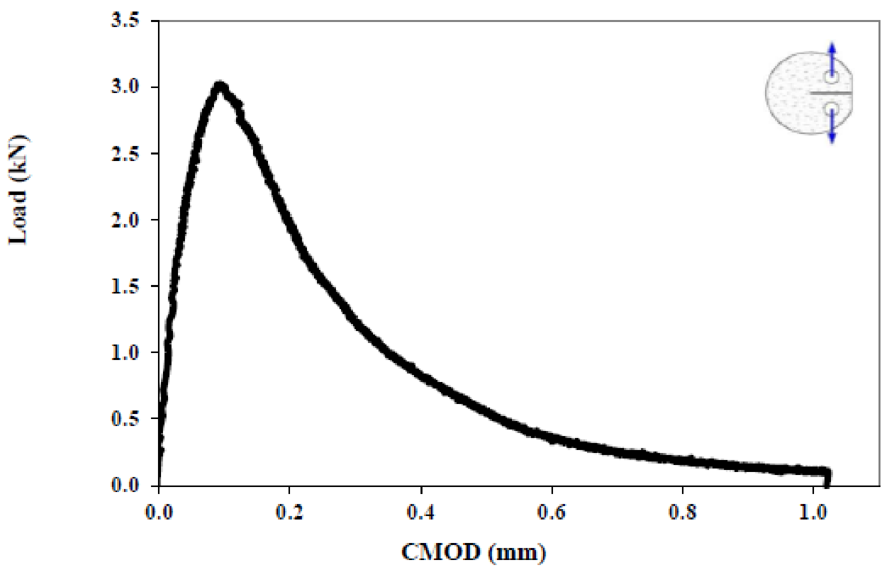


Figure A-2. Typical Load-CMOD curve (W. G. Buttlar & Wang, 2016)

Overview of SCB Test

The Semi-Circular Bending (SCB) test utilizes a simple three-point bending mechanism to determine the cracking resistance of the asphalt mixture specimen. The test uses a semi-circular specimen and load is applied at the center of specimen periphery, as shown in Figure A-3. (Al-Qadi et al., 2015). The test can be conducted in low temperatures as well as intermediate temperatures. The low temperature cracking resistance test standard is outlined in AASHTO

TP105-13 (AASHTO, 2013) and utilizes the same method as the DC(T) test to calculated fracture energy. The intermediate temperature SCB test was developed by Wu et al. in 2005, using 25ºC as the test temperature for the SCB test. The authors used the concept of critical J-integral, which was found to be sensitive to the changes in binder types and nominal maximum aggregate size (NMAS), both affecting the fracture resistance of the mixtures (Ozer, Al-Qadi, et al., 2016; Wu, Mohammad, Wang, & Mull, 2005).

Use of SCB to Study Asphalt Mixture Characteristics

Li et al. used the SCB test to characterize the low-temperature performance characteristics of 90 specimens from six asphalt mixtures with varying mixture components and factors. The authors also investigated the effect of loading rates, temperature, and initial notch length in the specimen on the fracture energy and peak loads obtained from the SCB test. The authors reported a decrease in the fracture energy with colder temperatures and the reverse effect on the peak loads. Further, the increase in the loading rates led to a decrease in the fracture energies of the specimen, with significant effect of test temperature on the results. The authors investigated three different notch lengths in the specimen and reported an increase in the fracture energy with a decrease in notch length at warmer temperatures but no significant effect of notch length for colder temperatures (X. J. Li & Marasteanu, 2010).

Overview and Applications of IL-SCB Method

In 2016, Ozer et al. introduced the IL-SCB method for cracking resistance characterization. Based on previous experiences, the researchers believed that it was difficult to correctly discriminate the asphalt mixtures based only on fracture energies. The researchers observed that the post-peak slope of the load-displacement curve from SCB test was sensitive to the changes in the asphalt mixture specimen and used this to develop the Flexibility Index (FI) - a simple index parameter that is proportional to the crack resistance of the mixture. Ozer et al. used the IL-SCB method to evaluate and discriminate the mixes with increasing high asphalt binder replacement (30%-60%) through addition of RAP/RAS. The authors further validated the index by correlating it successfully to the FHWA's accelerated pavement test sections data (Al-Qadi et al., 2015; Ozer, Al-Qadi, et al., 2016; Ozer, Hasan, et al., 2016).

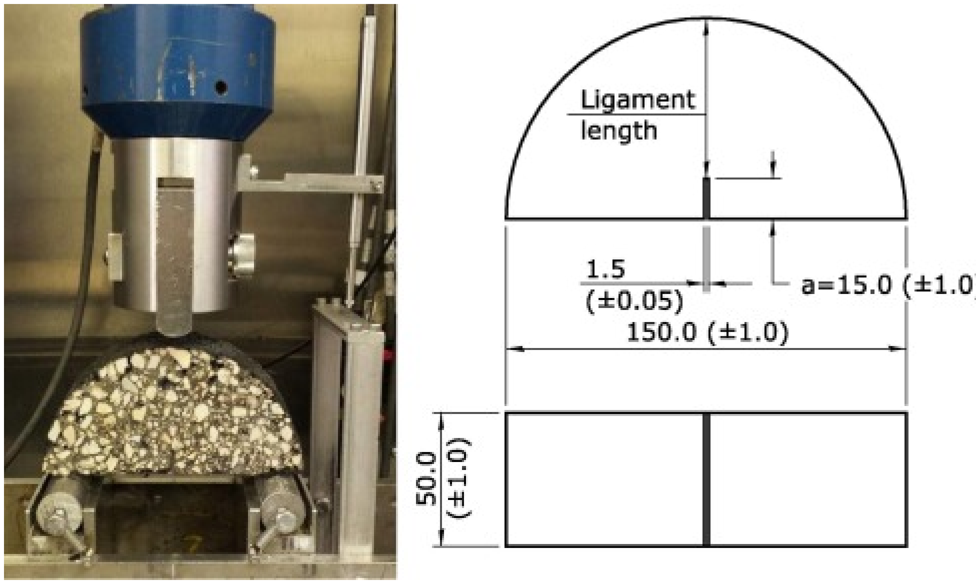


Figure A-3. SCB specimen and test setup (Ozer, Al-Qadi, et al., 2016)

Overview of SE(B) Test

Wagoner et al. developed the Single Edge Notched Beam (SE(B)) test in 2005 to measure the fracture energy of asphalt specimens. It is a simply supported three-point loading beam configuration with a notch to pre-determine the crack path. The beam configuration was chosen due to ease of manufacturing and its K (stress intensity factor)-dominant field under pure mode-I loading. Additionally, the authors reasoned that the ligament length of the specimen could be changed, as required, to encompass the fracture process zone. Further, the SE(B) geometry also allowed researchers to test asphalt specimens under mixed-mode conditions by simply offsetting the mechanical notch from the centerline of the beam (M. P. Wagoner, Buttlar, & Paulino, 2005a).

Use of SE(B) and SCB Tests to Compare SMA and HMA mixtures

Artamendi et al. compared SE(B) and SCB test results for a Stone Matrix Asphalt (SMA) and a dense HMA mixture, computing K and the fracture energy of the specimen. The study used beams with a span of 244 mm and a span to length ratio of 0.8. For Mode-I loading, the authors reported a good agreement between the K_I values, while the fracture energies computed from SCB specimens were twice the values obtained from SE(B) tests. The authors also investigated the mixed-mode loading and reported a higher fracture energy for SMA mixtures than the dense HMA mixtures, indicating an increased resistance to fracture due to introduction of shear component of loading (Artamendi & Khalid, 2006).

Use of SE(B) Test to Investigate Effects of Mixed Mode Loading

Braham et al. used the SE(B) test to investigate the effects of mixed mode loading that a pavement experiences under wheel loads and thermal stresses. The authors induced Mode-II loading by offsetting the notch from centerline. The study tested three asphalt mixtures with varying notch offsets and also measured the crack mouth opening displacement (CMOD), crack tip opening displacement (CTOD), and crack tip sliding displacement (CTSD) to effectively capture the Mode-I and Mode-II characteristics. The authors were successfully able to differentiate between the opening and the sliding fracture work during the loading by using different displacements (CMOD, CTOD, CTSD) to calculate fracture work and the results were as expected. Further, an increase in fracture work with an increase in the notch offset was reported, indicating that it takes more effort to propagate a crack in mixed mode conditions (A. Braham, Buttlar, & Ni, 2010).

Pooled Fund Study of Various Thermal Cracking Tests

Marasteanu et al. conducted various thermal cracking tests in a national pooled fund study funded by FHWA, phase-I in 2007 and phase-II in 2012 (Marasteanu et al., 2007) (Marasteanu, Buttlar, Bahia, & Williams, 2012). The report tested asphalt mixtures that included mixtures modified with Recycled Asphalt Pavement (RAP), Poly-Phosphoric Acid (PPA), and polymers (SBS and Evalvoy), at different aging levels using tests such as IDT, SCB, and DC(T). The IDT strength results showed a poor correlation between the laboratory and the field specimens. The DC(T) and SCB fracture energies fell within similar ranges, with DC(T) fracture values being slightly higher than the SCB fracture values. The report went on to compare the costs associated with DC(T) and SCB tests and more such factors before picking DC(T) test as the preferred method on the basis of an existing ASTM standard for it. The Phase-II of the report suggested development of a simplified method to measure creep compliance of the asphalt mixtures using the fracture tests, namely DC(T) and SCB tests, due to relatively high cost for IDT equipment. The DC(T)-IDT test was proposed to include an extensometer at 10 mm distance from the notch, placed perpendicular to it (Marasteanu et al., 2007, 2012).

DC(T)-IDT Test

Kebede (2012) worked on the DC(T)-IDT idea in his thesis and proposed conducting creep tests at required temperatures and using the same specimens to conduct DC(T) fracture tests after letting the specimens relax for about 24 hours after every test. The author reasoned that since the creep tests were undertaken in the linear viscoelasticity range, the fracture energy of the asphalt specimen should not be affected. Any lingering temperature effects would be avoided by letting the sample relax overnight. Numerical simulations of the test, included in the report by Marasteanu et al. and in Kebede's thesis, showed promising results (Kebede, 2012; Marasteanu et al., 2007, 2012).

DC(T) Creep Test

A modified version of DC(T)-IDT, now called the DC(T) Creep test was used in testing GTR-modified laboratory specimens and field cores for their compliance by Buttlar et al. The DC(T) creep tests measured creep compliance of the GTR-modified asphalt mixtures at 0ºC, -12ºC, and -24ºC. The raw data was fit in a Voigt-Kelvin model and the master curves were plotted with -24ºC as the reference temperature. The master curves obtained were smooth and the varying trends with variation in mixture parameters were as expected (W. G. Buttlar & Rath, 2017).

Application of DC(T) and SCB to Study Factors Related to Fracture Energy

Li et al. used DC(T) and SCB to study the effect of factors expected to affect the fracture energy of asphalt mixtures at low temperatures. The factors investigated were binder type, binder modifier, aggregate type, air voids, asphalt content, and test temperature. The fracture energies obtained from the two tests agreed with each other as the factors were varied in the mixtures except in the case of air voids. The SCB test showed significant variation with change in air voids whereas the DC(T) results were not affected significantly by it. A possible reason suggested was the different loading rates, and specimen geometry used in the respective tests (X. Li, Braham, Marasteanu, Buttlar, & Williams, 2011).

A.1.3. Mixture Tests to Mitigate Fatigue Cracking

Overview of Fatigue Cracking

Fatigue cracking is predominantly associated with repeated traffic loads. The cracks occur in a mesh-shape pattern and propagate throughout the surface resulting in rough riding quality, intrusion of water, and other inconveniences. Researchers have tried to quantify fatigue cracking by incorporating various factors affecting the distresses in empirical equations. In the mechanistic-empirical approach, the response parameters (stresses, strains) are inputted in these empirical equations and pavement fatigue life is predicted. These ME equations need constant calibrations which can be done through laboratory-conducted mix performance tests. PavementME software uses a sigmodal function as the “transfer” function for predicting the fatigue life of the pavement (Wang, Mahboub, & Hancher, 2005).

Use of ITT to Evaluate Fatigue Characteristics of HMA Mixes

Several studies have proposed various laboratory methods of fatigue testing of mixes. Ghulzan and Carpenter used Indirect Tensile Test (ITT) to evaluate the fatigue characteristics of asphalt mixes. The authors used the phenomenological fatigue model or the Stress vs. number of cycles approach (S-N curve) on 480 asphalt specimens under controlled stress and controlled strains. The varying factors in the asphalt specimens were air voids (4% and 7%), binder type (10 sources), aggregate types, and gradation. The authors concluded that the fatigue life of mixtures was highly influenced mode of loading, test temperature, and binder content, while the binder grade, air voids, and aggregate gradation had no significant effect on the fatigue parameters of the asphalt mixtures (Ghuzlan & Carpenter, 2002).

Use of ITT to Study CRM and Fiber Reinforcements

Mashaan et al. studied the effect of adding Crumb Rubber Modifier (CRM) to Stone Matrix Asphalt (SMA) on its fatigue life using ITT. The authors added 6-12% of CRM by weight of bitumen in the SMAs and found out that the rubber particles prevented the growth and propagation of vertical fatigue cracks though the asphalt specimen. Consequently, the fatigue life of the SMAs increase with an increase in the amount of rubber content (Mashaan et al., 2014). Weise et al. used cyclic ITT to evaluate the fatigue characteristics of HMA and SMA modified with various fiber reinforcements (cellulose and Polyacrylonitrile (PAN)). The authors reported a significant increase in fatigue life of the asphalt specimens with fiber reinforcements (Weise & Zeissler, 2016).

General Applications of BF Test

Huang et al. and Saadeh et al. used flexural beam-fatigue (BF) test (AASHTO T-321) (AASHTO, 2017a) (Figure A-4), which is a four point loading test wherein small beams (380x50x63 mm) are subjected to repeated loads, to characterize asphalt mixture specimen in terms of fatigue life. Saadeh et al. used 50% reduction in initial stiffness as the failure criterion on two mixtures with different binder grades, different moisture conditioning and on three replicates. The authors determined that the moisture conditioning of the specimens was one of the critical factors affecting the fatigue characteristics of the mixture, but the binder type had no significant effects (Saadeh & Eljairi, 2011).

Figure A-4. Bending Beam-Fatigue Test (Saadeh & Eljairi, 2011)

Push-Pull Fatigue Test

The push-pull fatigue test, developed by Richard Kim and his coworkers at NCSU, characterizes the fatigue damage in an asphalt mixture specimen using a simple uniaxial test and Viscoelastic Continuum Damage (VECD) principles (Al-Qadi et al., 2015). A cylindrical asphalt mixture specimen is subjected to repeated cyclic tension and compression loading until it fails. A damage characteristic curve, defined by damage parameter (S) and the pseudo secant modulus (C) is made from the test and is used to analyze the fatigue characteristics of the specimen. (Kanaan, 2013; Mbarki, Kutay, Gibson, & Abbas, 2012; Underwood, Baek, & Kim, 2012).

Overview of OT

The Texas Overlay Test (OT) was developed by Robert Lytton in the 1970s to evaluate the asphalt mixture's resistance to reflective cracking (Ma, 2014). Zhou et al. developed and verified the OT test to study fatigue cracking in asphalt mixtures in 2007 (Figure A-5). The test applies a cyclic triangular waveform with a constant maximum displacement of 0.64 in. simulating the opening and closing action of joints. The test is run until failure occurs at a loading rate of 1 cycle per 10 sec at the maximum displacement. This test records the number of cycles to failure, and the data obtained is used to determine the crack initiation and propagation potential of the mixture. It is conducted in accordance to the Tex-248-F standard (DeVol, 2015; F Zhou & Scullion, 2005; Fujie Zhou, Hu, & Scullion, 2007; Fujie Zhou & Scullion, 2003). The OT test has some advantage over the BF test like easier sample fabrication and ability to conduct test in AMPT device. The Asphalt Mixture Performance Tester (AMPT, formerly SPT), is a versatile device which can also employ the principle involved in cyclic push-pull test to determine the fatigue characteristics of asphalt mixtures (Federal Highway Administration (FHWA), 2013, 2016).

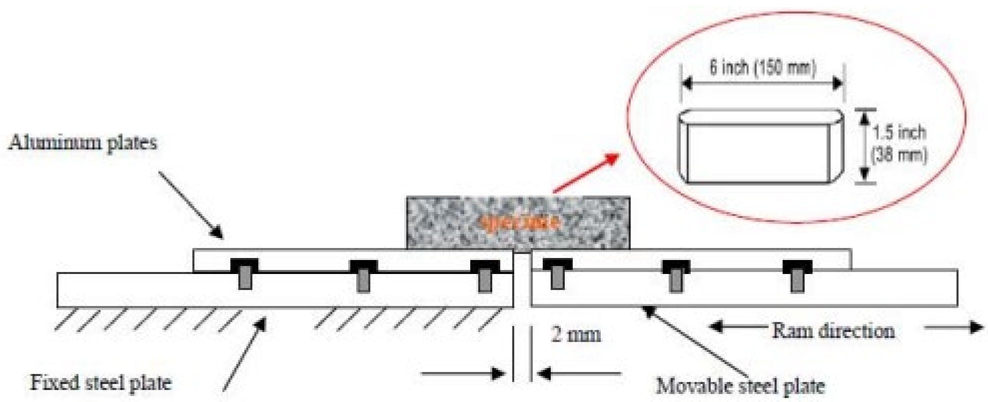


Figure A-5. Concept of Overlay Test (Fujie Zhou, Scullion, & Hu, 2007)

Overview of IDEAL-CT

In a research study by Zhou et al. (2017), the indirect tensile asphalt cracking test (IDEAL-CT) which is similar to the indirect tensile (IDT) test was developed (Fujie Zhou, Im, Sun, & Scullion, 2017). The IDEAL-CT is normally run at room temperature and a loading rate of 50 mm/min. Figure A-6 shows the IDEAL-CT setup with typical results. The IDEAL-CT was compared to the Texas OT and Illinois SCB tests using over 25 laboratory and field plant mixes. All three tests ranked all of these mixes in the same order with respect to crack resistance. The IDEAL-CT showed a strong correlation with the field distresses of fatigue, reflective, and thermal cracking. According to the authors, the IDEAL-CT is straightforward to perform, requires minimal training, and is fast as the test completes within one minute. The IDEAL-CT was found to be rugged with respect to specimen thickness, loading rate, test temperature, and air voids.

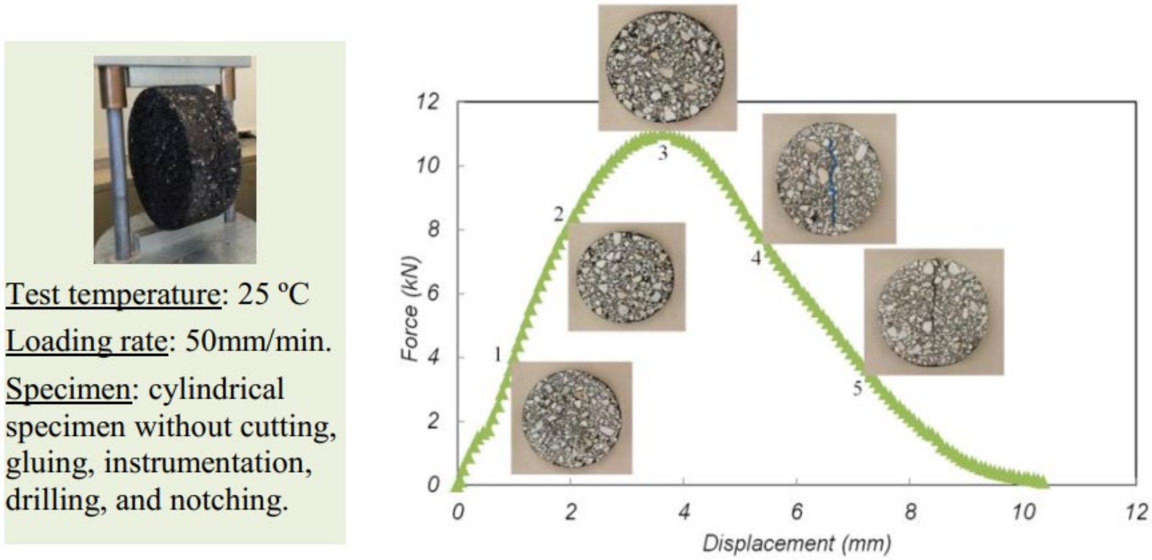


Figure A-6. IDEAL-CT Test Setup and Typical Results (Fujie Zhou et al., 2017)

Study of Ratio of Dissipated Energy Change and Fatigue Characteristics

Shu et al. used criteria proposed by Carpenter et al. of plotting ratio of dissipated energy change (RDEC) and number of cycles on a plot and using the constant plateau value of the curve as a parameter to evaluate the fatigue life characteristics of RAP-modified mixtures. A higher plateau value would mean a higher percentage of input energy was being turned into damage, indicating a lower fatigue life (Shu, Huang, & Vukosavljevic, 2008).

Use of Surface Energy to Assess Fatigue Properties

Cong et al. (2017) used the surface energy concept of asphalt mixture to characterize the fatigue behaviors. They used a uniaxial strain-controlled cyclic tensile test to determine the dissipated pseudostrain energy and determined the final number of cycles to failure for different asphalt mixtures. They also examined the effect of temperature on the fatigue life (Cong, Peng, Guo, & Wang, 2016).

Fatigue Studies of WMA Mixtures

Xiao et al. used BF test to study the fatigue characteristics of rubberized Warm Mix Asphalt (WMA) mixtures. Eight mixtures with variation in WMA additives, aging, and aggregates were tested. The conclusions suggested that addition of rubber in the WMA mixes led to increase in the fatigue life of the mixtures. Further, aggregate source (type) was a major factor affecting the fatigue life (Xiao, Wenbin Zhao, & Amirkhanian, 2009). Fakhri et al. compared the fatigue life of HMA and WMA mixes by computing fatigue characteristics of the mixtures with four-point flexural beam test. The authors found that the fatigue life of WMA mixes was higher than the HMA mixes at lower strain levels. However, at higher strain levels the fatigue lives were comparable. They also studied the effect of polymer modification on fatigue life of WMA mixes and found that the fatigue lives of polymer modified WMA mixes were less than those of non-modified WMA mixes (Fakhri, Ghanizadeh, & Omrani, 2013).

Fatigue Studies of RAP, RAS, and GTR

Huang et al. looked at the fatigue characteristics of HMA mixtures with varying RAP content in them using the similar 50% initial stiffness reduction criteria. The authors found a general increase in the fatigue life of the mixtures with an increase in RAP content with PG64-22 binder, but a decrease in the crack resistance of the mixtures based on the ITT and SCB tests (Huang, Shu, and Vukosavljevic 2011). A similar study by Shu et al. found an increase in the fatigue life of mixtures with RAP content based on the 50% initial stiffness failure criterion. However, the decrease in the dissipated creep strain energy with addition of RAP led the researchers to propose the use of the plateau value failure criterion in the BF test to evaluate RAP-modified HMA mixtures.

Ozer et al. conducted push-pull as well as OT test and found that increase in RAS content in the asphalt mixtures or the use of stiffer binder in the mixtures decreased the fatigue life (Ozer, Al-Qadi, Kanaan, & Lippert, 2013). West et al. found in their research that if the amount of RAP is limited to 30% then the fatigue life of the RAP-modified mixtures is almost the same as the virgin mixtures (Randy West, Michael, Turochy, & Maghsoodloo, 2011). Peralta et al. added GTR in asphalt mixtures and conducted the beam fatigue test. The results revealed an equal or better performance than the non-modified mixtures (Peralta, Williams, Silva, Vera, & Machado, 2013). Sousa et al. introduced a new rubber product called "Reacted and Activated Rubber" (RAR) in 2013, and reported better fatigue characteristics of RAR-modified mixtures than the virgin mixes (Ishai, Amit, Kesler, & Peled, 2015; Sousa, Vorobiev, Rowe, & Ishai, 2012).

Test Track Study to Assess Field Performance

Ma et al. and West et al. studied different sections of a NCAT test track to compare the field performance with the laboratory test results for asphalt mixtures. Both the research groups used OT and beam fatigue test as the laboratory methods. West et al. compared the fatigue properties of test sections with varying amounts of RAP, with some sections further modified using Sasobit and SBS. West et al. found a better correlation of OT with the field performance when the standard displacement of 0.635mm was used in the laboratory tests. Ma et al. first ranked different sections with the aim of building transfer functions to fit the laboratory test results and to determine the number of cycles to failure at a temperature-corrected strain measured in the field. Using the transfer functions, they re-ranked the mixtures and found that OT had a better correlation to the field performance (Ma, 2014; R. C. West, Tran, Taylor, & Willis, 2016).

A.1.4. Mixture Tests to Mitigate Rutting

Overview of Rutting

Although Superpave mixture design protocols have led to a general reduction in the manifestation of rutting on pavements (TRB Superpave Committee, 2005), the ever-increasing traffic and advent of newer materials in asphalt mixture still call for the inclusion of a performance-based test specification. Rutting, or permanent deformation, is an accumulation of unrecoverable strains on the pavement structure due to repeated traffic loads.

WLT Tests

Wheel load tracking (WLT) tests are the most common performance tests for measuring rutting potential of HMA mixes. The WLT methods simulate traffic by passing over standardized wheels simulating real-life traffic loads on HMA specimen at a given temperature. The two most common WLT test devices are Hamburg Wheel Tracking Test (HWTT) (Figure A-7) and the Asphalt Pavement Analyzer (APA) (formerly known as Georgia-loaded wheel tester). The HWTT is performed in accordance to AASHTO T324 standard. A loaded steel wheel, weighing approximately 71.7 kg tracks over the samples placed in a water bath at 50ºC. The vertical deformation of the specimen is noted against the number of wheel passes. The test is stopped when either the specimen deforms by 20mm or the number of passes exceeds 20,000.

Figure A-7. Hamburg Wheel Tracking Device a) Test running b) After test (W. G. Buttlar & Rath, 2017)

APA Test

The APA uses a similar principle with an aluminum wheel and can also simulate the effect of tire pressure, unlike HWTT. It is performed in accordance with the AASHTO TP63 standard (AASHTO, 2007). Both the methods can also be used to determine the moisture sensitivity of the HMA mixtures since the wheel loads are simulated under-water. There are several other test methods to measure permanent deformation such as the Static Creep triaxial test and repeated load triaxial test that use flow time and flow number respectively as parameters to ascertain rutting characteristics of a mixture. Rutting is also related to the dynamic modulus of the mixture (Rushing, Little, & Garg, 2014).

Studies to Compare Rutting Tests

Rushing et al. performed a laboratory study to vet four performance tests characterizing rutting potential of asphalt mixtures to be used for airport asphalt mixture design. The four performance tests were the asphalt pavement analyzer, triaxial creep, triaxial repeated load, and dynamic modulus test. The test results were compared to the proposed threshold values for each test, and the threshold values were further justified using full-scale field testing. The threshold values were justified using data from the previous studies as well as the current study and can be found in the paper. The field trials included two sections, both incorporated with the same HMA mixture, but subjected to different loading and temperature to simulate severe and moderate loading conditions. Furthermore, the test matrix included polymer modified asphalt mixtures to see the effect of polymer modification on rutting resistance of the asphalt mixtures. The polymer modification clearly showed improvement in rutting resistance as measured by the four performance tests. The authors recommended the use of APA for evaluating the rutting resistance of asphalt mixtures to be used in airport pavements due to its ability to test specimens produced by the Superpave Gyratory Compactor (Rushing et al., 2014).

Walubita et al. compared the results and feasibility of a few common methods of characterizing rutting - Hamburg Wheel Track test (HWTT), the dynamic modulus (DM), and the uniaxial repeated load permanent deformation test (RLPD). The study included a variety of laboratory mixtures as well as field cores with different binder grades, aggregate structure, air voids, and traffic levels (for field cores). All three tests showed good correlation to the field results and with one another. However, the HWTT results showed the least variation and was the preferred method by the authors to acquire rutting data on asphalt mixtures (Walubita et al., 2012).

Research Studies to Evaluate Impacts of HMA Mixture Modifications on Rutting

Researchers have found that modification of HMA mixtures with different materials or techniques affects the rutting resistance and the moisture susceptibility of the mixture. Addition of RAP/RAS in HMA mixtures generally imparts a higher rutting resistance as compared to the non-modified mixtures (Hong, Chen, & Mikhail, 2011; Ozer et al., 2013; Vahidi, Mogawer, & Booshehrian, 2014; Randy West et al., 2011). Mejías-Santiago et al. studied the rutting potential of WMA and HMA mixtures used in airport pavements and found them to have similar performance (Mejías-santiago, Doyle, & Rushing, 2014). Rodenzo et al. also compared WMA and HMA mixtures to be used on roadways and found similar results using FN test criteria (repeated load triaxial test) included in AASHTO TP 79-13 (Rodezno, West, & Taylor, 2015). Haghshenas et al. found that addition of rejuvenators to high-RAP mixtures led to their softening and consequently an increase in the rut depth recorded in HWTT for similar number of cycles (Haghshenas & Kim, 2016).

Gillen et al. reported for the Illinois Tollway that addition of GTR led to decrease in rutting potential of dense-graded, OGFC, and SMA asphalt mixtures (Gillen, 2007). Willis used HWTT to compare GTR and SBS modification of HMA mixtures and reported a lower rutting potential for GTR-modified mixtures (Willis, 2013a). Chui et al. used HWTT to study the rutting potential and moisture sensitivity of GTR-modified SMA mixtures and found a decrease in both the rutting potential and moisture sensitivity when compared to non-modified SMA mixtures with same aggregate structure (Chiu & Lu, 2007).

Triaxial Strength Test

In a study by Christensen et al. (2000), the use of triaxial strength testing as a simple performance test for rutting was investigated using ten mixtures from Pennsylvania and New York (Christensen, Bonaquist, & Jack, 2000). In addition, an abbreviated test protocol using the indirect tension (IDT) test was also evaluated. The results found that the abbreviated protocol performed better and more precisely than the standard triaxial procedure. The IDT strength correlated well with the rut resistance, thus showing that the IDT test has the potential to be beneficial for performance testing.

IDT Test

Christensen et al. (2004) evaluated the indirect tension (IDT) test at high temperature as a test for rut resistance (Christensen, Bonaquist, Anderson, & Gokhale, 2004). The results showed that IDT strength performed well at measuring cohesion and was strongly correlated with the observed rut resistance in both the laboratory and actual pavements. Preliminary recommendations for evaluating rut resistance based on IDT results were provided. Overall, the study demonstrated that the high-temperature IDT strength test could be a simple, cost-effective, and accurate test for determining the rut resistance of Superpave mixtures. In the use of the test, attention must be paid to choosing the temperatures and loading rates. The IDT strength test in this study was performed at a loading rate of 3.75 mm/min and at a temperature 20 degrees Celsius below the critical pavement temperature for permanent deformation. Additional research for other loading rates and temperatures was suggested by the authors.

A research study by Christensen and Bonaquist (2007) presented a simplified and faster procedure for using the IDT strength test at high temperature to assess the rut resistance of HMA mixtures (Christensen & Bonaquist, 2007). The previous research (Christensen et al., 2000) (Christensen, Bonaquist, Anderson, & Gokhale, 2004) evaluated the IDT test at a loading rate of 3.75 mm/min and at a temperature 20 degrees Celsius below the critical pavement temperature for permanent deformation. In this study, the researchers applied a simpler procedure by testing at a loading rate of 50 mm/min and at a temperature of 10 degrees Celsius below the critical pavement temperature. Based on the results, the researchers recommended that the IDT strength test be performed at a loading rate of 50 mm/min and at a temperature of 9 degrees Celsius below the critical pavement temperature. The revised protocol simplifies the procedure by allowing the test to be conducted with a standard Marshall press and potentially at room temperature with conditioning due to quick failure of the test specimens. Several case studies were presented. The guidelines for interpreting the test results were also updated. The revised guidelines presented the required IDT strength based on the level of traffic.

A.1.5. Mixture Tests to Characterize Complex Modulus

Overview of Complex Modulus

The Mechanistic-Empirical Pavement Design Guide (MEPDG) predicts the damage accumulation on the pavement based on various key inputs, one of which is complex modulus. Briefly, complex modulus defines the relationship between stress-strain of viscoelastic materials. Complex modulus is calculated in accordance with AASHTO TP 62-03 (AASHTO, 2003). The procedure uses an asphalt concrete specimen of 150 mm height and 100 mm diameter. The test is performed at temperatures -10ºC, 4.4 ºC, 21.1 ºC, 37.8 ºC, and 58.4 ºC, and at frequencies of 0.1 Hz, 0.5 Hz, 1 Hz, 5 Hz, 10 Hz, and 25 Hz. This is a stress-controlled test and the strains are limited within 50 to 150 microstrains. This parameter has been used by various researchers to discriminate among asphalt mixtures based on their composition and aging levels.

Evaluations of Complex Modulus and RAP/RAS

Studies by Sondag et al., Swamy et al., Shahadan et al., and Ozer et al. showed that the complex modulus increases at any particular reference temperature throughout all frequencies for asphalt mixtures with increment in percentage of RAP/RAS (Al-Qadi et al., 2015; Norouzi, Kim, Kim,

& Yang, 2017; Shahadan, Hamzah, Shukri Yahya, & Jamshidi, 2013; Sondag, Bruce A. Chadbourn, & Drescher, 2002; Swamy, Mitchell, Hall, & Daniel, 2011). Additionally, some of the studies reported a decrease in phase angle with increase in RAP/RAS indicating a stiffer or aged mixture. Vavrik et al. found a decrease in the complex modulus value with addition of RAS in Stone Matrix Asphalt mixture (SMA) (Vavrik et al., 2010).

Cooper et al. studied the laboratory performance of five mixtures containing RAP and/or RAS with and without rejuvenator (recycling agents (RAs)) in addition to the control mixture. The authors found a good correlation between the prediction of rutting by complex modulus and by loaded wheel track testing. Further, the low temperature cracking potential, measured by SCB test, showed agreement with the E* results. (S. B. Cooper, Mohammad, & Elseifi, 2016).

Mangiafico et al evaluated the effect of recycling agents (rejuvenators) on asphalt mixture performance by means of two-point bending tests on trapezoidal samples. The asphalt samples containing different binder types and RAP content were tested at 15 ºC, 10 Hz. A requirement of 14,000 MPa for |E*| was set for mixtures according to EN 13108-1:2006 (BSI, 2006). As expected, mixtures with recycling agents were found to have a slightly lower complex modulus compared to corresponding regular mixtures. Contrary to what was observed for mixtures without agent, mixtures produced with recycling agent did not show a remarkable and progressive stiffness increase with increasing RAP content. Therefore, the addition of the recycling agent was observed to generally lower the |E*| (Mangiafico et al., 2016).

Complex Modulus and GTR or Nanoparticles

A report on a GTR asphalt pavement demonstration project by ARA for Illinois Tollway reported an increase in the complex modulus of a dense HMA mix with addition of GTR in it (Gillen, 2007). Willis (2013) compared the complex modulus for GTR- and polymer-modified (SBS) HMA mixes, and reported that the GTR mixes had a higher complex modulus value which was more noticeable in hotter and intermediate temperature ranges (Willis, 2013b). Yao et al. used the complex modulus to see the effects of addition of nanoparticles in asphalt mixture. The authors found an increase in complex modulus with addition of nanoparticles (Yao & You, 2016).

Dynamic Modulus Testing of CRM Modified SMA Mixtures

Using the asphalt mixture performance test (AMPT), Xie and Shen performed dynamic modulus test in load-controlled and axial compression mode to evaluate linear viscoelastic behavior of CRM modified SMA mixtures. Dynamic modulus test results at 45 ºC showed that dry process SMA results in lower dynamic modulus than rubberized SMAs and SBS SMA. This means that dry process SMA might be less resistant to deformation than the wet process, the terminal, and SBS SMA mixes. This finding was in agreement with Hamburg wheel tracking test (Xie & Shen, 2016).

Sensitivity Studies of Dynamic Modulus

Huang et al. studied the temperature, frequency and additive effect on dynamic modulus and phase angle of asphalt mixtures. The two point bend test on trapezoidal specimens showed that dynamic modulus decreases as temperature increases or frequency decreases, while phase angle increases as temperature increases or frequency decreases. Also, while applying different loading levels, it was found that dynamic modulus decreases and phase angle increases as the applied load goes up from 30 to 180 μdef (Y. Huang, Wang, Liu, & Li, 2016). Gedafa et al., studied the correlation between aging and complex modulus and further compared it to some prediction models available for complex modulus (Gedafa, M.Hossain, Romanoschi, & Gisi, 2013). The results indicated that higher aging led to a higher complex modulus of the HMA mixture.

Development of Indirect Tension Dynamic Modulus and Torsion Bar Shear Modulus

Outlining the functions of dynamic modulus in flexible pavements, Yang et al. addressed the challenges associated with testing field cores under uniaxial compression loading. They introduced the indirect tension dynamic modulus (IDT |E*|) and torsion bar shear modulus (torsion bar |G*|) to replace the traditional dynamic modulus geometry. Results of testing 10 different field sections, the IDT |E*|and torsion bar |G*| tests were found to be able to generate consistent master curves which can correlate to each other. Also, these two alternative tests could identify differences between surface course lifts and fairly quantify differences in field performance (Yang, Braham, Underwood, Hanz, & Reinke, 2016).

Development of HCT

In a research study by Buttlar et al. (2004), a hollow cylinder tensile tester (HCT) was developed as a way of determining various asphalt pavement properties including creep compliance, tensile strength, and dynamic modulus (W. Buttlar, Khateeb, & Sherman, 2004). The HCT showed great accuracy in determining creep compliance and dynamic complex modulus when compared with the IDT. However, the HCT currently cannot be used for permanent deformation (rutting).

A.1.6. Mixture Aging for Performance Testing

Asphalt mixtures age by the virtue of interaction between binder and oxygen, leading to a harder and more brittle binder. During the production process of asphalt mixture, the binder is heated at high temperatures before being mixed with the heated aggregates and that leads to aging of the mixture, which continues till it cools down. It is important to simulate this 'short-term' aging in the laboratory specimens, lest we could encounter variation in the mixture properties of plant-compacted and lab-compacted mixtures. Further, asphalt pavements age throughout their service life, albeit at a lower rate than in production and construction. This kind of aging also needs to be simulated in the lab to ascertain the changes in the asphalt mixture properties over time. AASHTO R30 standard, adopted from the SHRP report by Bell et al. in 1994, delineates the procedures for mixture aging at different levels. Bell et al. tried out different aging methods such as forced-draft oven aging, pressure oxidation, extended mixing, and triaxial cell aging for different durations. The authors recommended short-term aging to be achieved by oven-aging the loose mixture for 4 hours at 135 ºC (275 ºF). Further, forced-draft oven aging at 85 ºC (185 ºF) for 5 days (120 hours), applied to the compacted specimens that have undergone short-term aging was recommended for long-term aging (AASHTO R30, 2006; Bell, Abwahab, Cristi, & Sosnovske, 1994). Apart from the AASHTO adopted method, numerous researchers have come up with different aging techniques and many studies have complied those short- and long-term aging methods for asphalt mixtures and binders (Airey, 2003; S. F. Brown & Scholz, 2000; Elwardany, Yousefi Rad, Castorena, & Kim, 2017; Jemere, 2010; Reed, 2010; Yin, Arámbula-Mercado, Epps Martin, Newcomb, & Tran, 2017).

Short-Term Mixture Aging

Yin et al. evaluated the short-term aging protocols for HMA and WMA mixtures by comparing the dynamic modulus and the HWTT rutting results of a wide range of asphalt mixtures. The authors concluded that oven-aging at 135 ºC for two hours in case of HMA mixtures and at 116 ºC for two hours in case of WMA mixtures was an ideal way to simulate short-term aging in the asphalt mixtures. The authors also reported that binder source and production temperature had a significant effect on the performance metrics of the short-term aged asphalt mixtures (Yin, Martin, Arambula, & Newcomb, 2016). R. Bonaquist recommended use of 2 hours of oven aging at the compaction temperature for WMA as well as HMA mixtures in NCHRP Report 691. Further, the author recommended a two-step aging process for WMA mixtures to be tested for rutting using the flow number test. The two-step process included conditioning at compaction temperature for two hours and then at high in-service pavement temperature for a duration of less than 16 hours. Such changes were based on the findings that the current protocol of aging the WMA specimens for 4 hours at 135 ºC (275 ºF) resulted in over-stiffening of binder, misrepresenting the 'as-constructed' condition of asphalt mixture (R. F. Bonaquist, 2011). Liang et al. evaluated the short-term aging effects on rubber modified asphalt mixture and reported an increase in the resilient modulus of the mixtures with short-term aging (Liang & Lee, 1996). Poulikakos et al. conducted Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) spectroscopy on short- and long-term aged asphalt mixtures modified with RAP. Further, the authors conducted performance tests, such as IDT strength test, and flow number test for characterizing rutting. On comparing the modified and non-modified aged asphalt mixtures, the authors reported the absence of 'bee' microstructure in the aged bitumen as compared to virgin binder. Further, the authors concluded that RAP-modified mixtures was at par with the non-modified asphalt mixtures in terms of performance (Poulikakos et al., 2014).

Long-Term Mixture Aging

The AASHTO R30 protocol for long-term aging led to distortion of specimens due to changes in the air voids and softening of asphalt mixture. As a remedy, NCHRP 9-23 recommended using a metal wire mesh to wrap the specimens while aging. However, the wire mesh could not mitigate the distortion problem completely. Further, the aging method was found to induce vertical and radial oxidation gradient in the compacted specimens which was undesirable. This motivated the NCHRP Project 09-54 which aimed at developing a new protocol for long-term aging of compacted asphalt specimens that would be appropriate for fabrication of performance test specimens (Houston, Mirza, Zapata, & Raghavendra, 2006). Kim et al. and Elwardany et al. proposed applying pressure during the aging process of the specimen to reduce the oxidative gradients and the use of smaller specimens to reduce sample distortion. However, the authors reported that application of pressure damaged the asphalt specimens, but aging smaller specimens solved the variable oxidation and the distortion problem. The authors also attempted to determine if long-term aging can be applied to loose mixtures without compromising the compactability and reported that no adjustment in compaction effort was required for the limited mixture cases tested in the study. Long-term aging loose mixture and compacting it to desired specimen geometry was reported as the most promising method. The authors included a series of maps depicting the duration of oven aging required at 95 ºC to simulate different years of field aging at different depths. Table 2 shows the duration required for Illinois climate (Kim et al., 2015; Yousefi Rad, Elwardany, Castorena, & Kim, 2017). Braham et al. investigated the accuracy of the AASHTO R30 aging protocol to simulate field aging of asphalt specimens. The authors conducted IDT creep and tensile strength tests on compacted asphalt specimen aged at 85 ºC for 5 days (AASHTO R30 protocol) and on specimens aged at 135 ºC for 24 hours. The findings suggested that aging at 135 ºC for 24 hours gave a better match of the mixture properties with the field-aged specimens. DC(T) fracture energy calculated for specimens aged at 135 ºC for different time intervals revealed that the 24 hour period was the optimal time for laboratory aging. However, from the results of BBR (Bending Beam Rheometer) and DENT (Double Edged Notch Tension) tests on extracted binder, the authors concluded that the 135 ºC, 24 hours aging might be too conservative for asphalt mixtures (A. F. Braham, Buttlar, Clyne, Marasteanu, & Turos, 2009). Rad et al. compared the effects of long-term aging asphalt mixtures at 95 ºC and 135 ºC. Three types of mixtures were tested in this study with same aggregates but different binders. Aging was also done at 70 ºC and 85 ºC for a better comparison. The authors found that aging at 135 ºC led to decrease in the dynamic modulus and fatigue resistance of the mixtures while altering the oxidation reaction mechanism of the mixture when compared to aging at 95 ºC (Yousefi Rad et al., 2017).

Table A-2. Days of oven aging at 95 ºC required for Illinois climate (Kim et al., 2015)

	Days of Oven Aging at 95 ºC		
Field Aging / Depth	4 yrs	8 yrs	16 yrs
6 mm	3 days	6 days	12 days
20 mm	1 days	3 days	5 days
50 mm	1 days	2 days	4 days

Oxidative aging gradient in asphalt mixtures

Aging of asphalt mixture on field depends on various factors such as solar radiation, air temperature, pavement thickness, and so on. Laboratory-compacted specimen is just a representation of the laid asphalt mixture, but a lot of 'field factors' influence the actual mixture properties of the pavement, e.g. mixture temperature during compaction, lift thickness, variation in rolling compaction, and so on. This leads to variable rate of oxidation with depth. Further, the pavement surface ages much rapidly than the lower layers due to its exposure to the atmosphere (Figure A-8) (Hajj, Alavi, Morian, Kazemi, & Sebaaly, 2014; Quintero, 2007; Zapata & Houston, 2008). Many studies have looked on the effect that the aging gradient has on the mixture properties. Coons et al. tested field cores aged from 1-13 years and found that the viscocity of asphalt did not change with age beyond 1 ½ inches from the pavement surface (Coons & Wright, 1968). Mirza et al. studied 40 field projects across the US and concluded that aging effects are insignificant beyond 38 mm from pavement surface (Mirza & Witczak, 1995). Li et al. extracted binders from field cores part of MnROAD facility at different depths. The authors reported variation in the binder properties at different depth indicating different aging levels (X. Li, Zofka, Marasteanu, & Clyne, 2006). Yin et al. tested field cores from four different projects with varying age and found that the variation in the binder properties at different depths increased with time and cumulative temperature (sum of all high temperature above freezing) (Yin, Epps Martin, Arámbula-Mercado, & Newcomb, 2017). Numerous other studies have also shown similar findings (Farrar, Harnsberger, Thomas, & Wiser, 2006; X. Li et al., 2006; Luo, Gu, & Lytton, 2015; Mirza & Witczak, 1995).

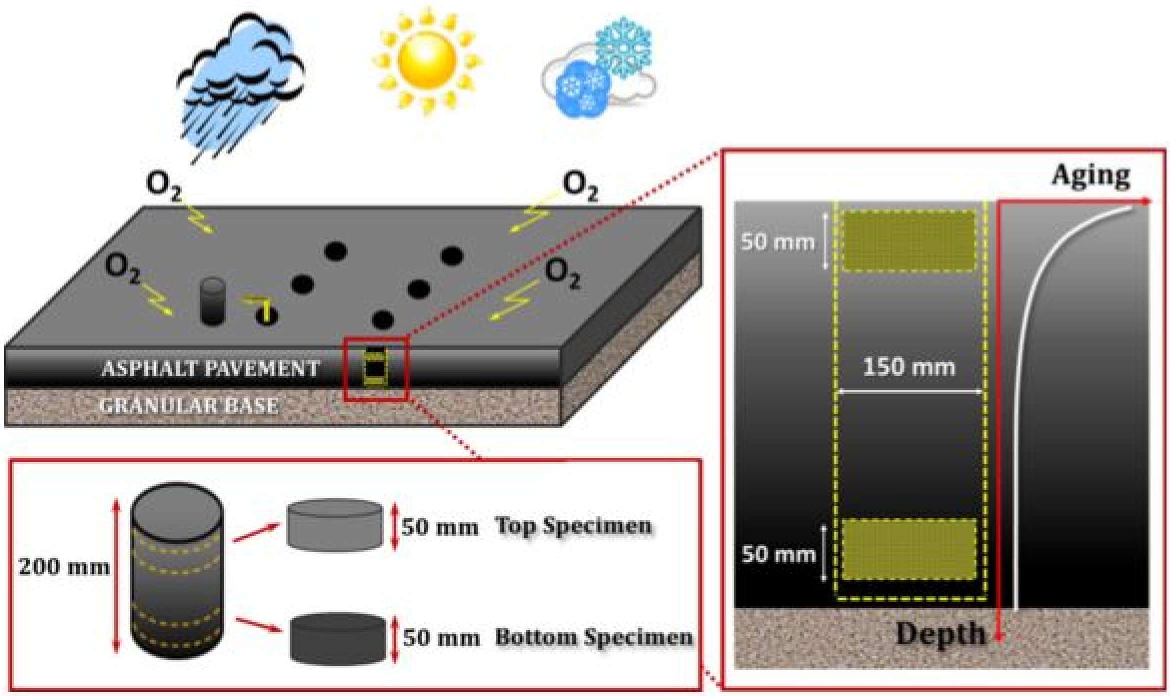


Figure A-8. Schematic representation of pavement aging gradient with respect to depth (Behnia et al., 2014)

A.1.7. Tools Based on Mixture Performance Tests

Pertaining to the scope of this study, it is important to briefly mention a few tools developed and designed based on the results from the mixture performance tools.

Illi-TC

Illi-TC, developed by Dave et al. at the University of Illinois Urbana Champaign (UIUC), primarily uses DC(T) fracture energy, IDT tensile strength, and IDT creep compliance results to first predict the number of critical events that would cause thermal cracks in the pavement and then goes on to perform a viscoelastic FE analysis for predicting the amount of thermal cracking (m/500m). This model addresses the shortcomings of the TCModel, a computer-based thermal cracking model developed in 1992 under a SHRP project. Illi-TC uses a 2-D analysis engine instead of 1-D analysis of TCModel and further, it uses cohesive zone fracture model instead of Paris law (Dave, Buttlar, et al. 2013; Marasteanu et al. 2012). More recently, Buttlar et al. used Illi-TC tool with a compliance input from DC(T) creep tests of GTR-modified asphalt mixtures. The results from the tool related well with the expected values from the mixtures. Further research and a wider sample set should be investigated in this regard (W. G. Buttlar & Rath, 2017).

TCAP

The Thermal Cracking Analysis Package (TCAP) is a newly developed model at University of Nevada, Reno in 2015 by Alavi et al. The model takes into account many more variables of asphalt mixtures, like aging, temperature dependent coefficient of thermal expansion and contraction (CTEC), than its predecessors. At this point, the model only predicts the number of critical events for a particular simulation of asphalt pavement at a location (Alavi, Hajj, & Sebaaly, 2015).

Performance-Space Diagram

Buttlar et al. developed the Performance-Space diagram, a graphical interactive tool suitable to capture the high- and low-temperature mixture performance test results in a single visual. Figure A-9a. shows the performance-Space diagram introduced by Buttlar et al. As shown, the DC(T) fracture energy results are plotted in the X-axis scale versus the Hamburg Wheel Tracking Test results in a reverse Y-axis scale (W. G. Buttlar et al., 2016). Al-Qadi et al. came up with a similar interaction plot by using the results of Flexibility Index, plotted on Y-axis, and Hamburg Wheel Tracking Test, plotted on X-axis. Figure A-9b. shows the interaction plot between FI and rut depth (Al-Qadi et al., 2015).

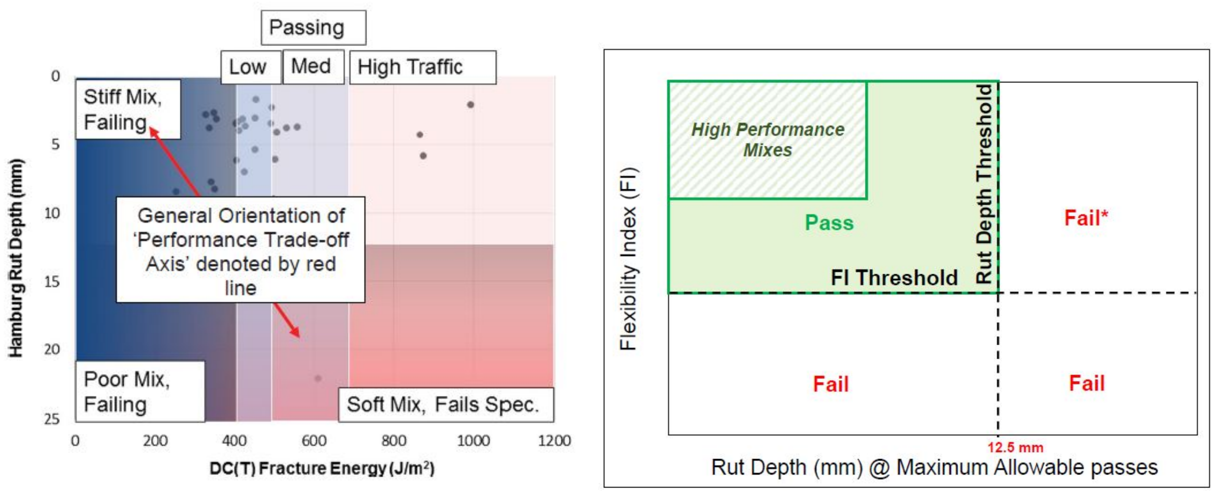


Figure A-9. a) Performance-Space diagram with limits superimposed (W. G. Buttlar et al., 2016) b) Interaction plot for a balanced mix design (Al-Qadi et al., 2015)

A.2. Agency Practices Regarding Performance Based Specifications (PBS)
Many state DOTs currently use or are planning to implement some type of Performance Based Specifications (PBS). The following sections describe current state DOT practices regarding PBS.

A.2.1. NCHRP Synthesis on Asphalt Performance Specifications
A NCHRP Synthesis on current practices regarding the use of performance specifications for asphalt pavements was completed in 2016 (McCarthy, Callans, Quigley, & Scott, 2016). Surveys were distributed to state DOTs, Canadian MOTs, and other agencies. Survey responses were received from ninety percent of DOTs along with Washington, D.C., 11 Canadian MOTs, and a few other agencies. Some of the key survey results include (McCarthy et al., 2016):

- Reasons for the use of PBS include reduced maintenance, greater distress resistance, and increased pavement durability.
- Although a majority of states currently use or are planning to implement some type of PBS, PBS has only been implemented as a standard practice in a limited number of agencies.
- 49% of DOTs currently utilize PBS, but they are frequently based on volumetric properties.
- 98% of DOTs use WMA and RAP for asphalt pavements.
- Agencies use modified tests when working with mixtures that include recycled materials.
- 54% of DOTs agreed that direct measurement of fatigue is an important pavement performance test, while 51% of DOTs agreed that direct measurement of rutting is important.
- The most commonly used performance tests for performance-based designs include the HWTD Test and APA Test.
- 29% of DOTs utilize shadow performance mix design testing for data collection along with volumetric properties for qualification, QC, and acceptance. 21% of DOTs utilize both volumetric properties and performance design specifications for qualification, QC, and acceptance.
- 29% of state DOTs have used data from performance testing for pay factors on pavement projects.
- 36% of DOTs indicated that testing time was a challenge to implementing performance tests, while 36% responded that cost is a significant consideration when deciding whether or not to implement performance testing. Other challenges identified by DOTs include knowledge gaps regarding PBS implementation and a lack of training.

Other key findings from the synthesis include (McCarthy et al., 2016):

- Chicago DOT has implemented a test procedure for DC(T) based on modifications to ASTM D7313.
- A limited number of agencies currently utilize performance tests for mixture acceptance.
- The performance-based properties that are researched the most are measurement of stiffness modulus, thermal cracking and resistance to moisture, and fatigue and durability.
- Volumetric properties are typically used along with performance properties for the design and acceptance of asphalt pavement mixtures.
- A limited number of agencies are in the process of investigating the costs and benefits of PBS. Additional research to help quantify the costs and benefits would be beneficial.
- Current guidance for agencies and contractors regarding the use of PBS is inadequate. There is a need for additional guidance regarding the implementation of PBS for asphalt pavements.

A.2.2. Overview of State Practices

Table 3 below provides a summary of the states that currently include or plan to include PBS for asphalt mixtures. The following sections describe the practices of states and other agencies in greater detail.

Table A-3. DOT's that currently include or plan to include performance-based specifications

Asphalt Mixture Performance Tests	States currently including or planning to include test in mix design
Rutting and moisture sensitivity test (Hamburg Wheel Tracking Test, Asphalt Pavement Analyzer, or TSR)	Arkansas, Alabama, Nevada, Oregon, South Carolina, South Dakota, Georgia, New Jersey, Colorado, Louisiana, Illinois, Wisconsin, North Carolina, Utah, Virginia, Washington
Thermal Cracking tests (DC(T), SCB, TSRST, IDT)	New Jersey, New Hampshire, Vermont, Minnesota, Iowa, Illinois, Wisconsin, Oregon, Washington, Nebraska, Kansas
Fatigue (OT, push-pull, etc.)	New Jersey, New Hampshire, Vermont, Massachusetts
Complex Modulus (and other parameters through APMT) **	Alabama, Colorado, Connecticut, Florida, Georgia, Illinois, Kansas, Kentucky, Maine, Maryland, Nebraska, New Hampshire, New Jersey, New York, North Carolina, Oregon, Pennsylvania, Tennessee, Utah, Virginia, West Virginia, Wisconsin, Wyoming

*The table is not exhaustive, and effort has been made to include as much data as possible
**The DOTs listed are a part of National Pooled Fund study for 'Implementation of Asphalt Mixture performance Tester for Superpave validation', with FHWA as the lead agency (Alabama Department of Transportation, 2012; Arkansas DOT; California Department of Transportation, 2012; Dave, Daniel, Jacques, & Decarlo, 2015; Deusen, 2015; EVAC, 2014; FHWA, 2015; Hanson, 2015; IDOT, 2015; McCarthy et al., 2016; MDOT, 2014; MnDOT, 2014; Mogawer, Austerman, Kluttz, & Mohammad, 2014; NCDOT, 2016; NCHRP, 2016; ODOT, 2011, 2013; TxDOT, 2016a; UDOT, 2012, 2015; VDOT, 2013; WisDOT, 2015; WSDOT, 2017)

A.2.3. Individual Agency Practices

California

Caltrans tested the use of PBS on three northern California Interstate highway rehabilitation projects (Harvey et al., 2014). The specifications included the use of Repeated Simple Shear Test (RSST) for selecting design binder content, flexural fatigue test, HWTT, and Hveem mix design requirements. Several challenges of implementing the PBS were identified, including communication of the meaning of specifications to bidders, procuring laboratory services, writing of PBSs, and schedule constraints for performing the testing.

Florida

The Florida Department of Transportation (FDOT) is currently evaluating several tests: AMPT for flow number and dynamic modulus, HWTD and APA for rutting, IDT and OT for cracking, and interlayer bond-strength test (McCarthy et al., 2016). Challenges to the use of PBS identified by FDOT include practical considerations for the testing and contractor unfamiliarity with performance tests.

Georgia

The Georgia Department of Transportation (GDOT) uses PBS for both acceptance/rejection and pay adjustments (McCarthy et al., 2016). Pay tables have been developed based on AC content, in-place air void content, and gradation. Several tests for Superpave mixtures are used, including tests for volumetric properties, bulk density, short term aging, maximum density and effective gravity, aggregate gravities, moisture susceptibility, rutting susceptibility and permeability. Moisture susceptibility testing is especially important due to possible stripping of aggregates in the mixtures. Abson recovery testing with Georgia Development Test (GDT)-199 is performed on mixtures with more than 20 percent RAP. GDOT emphasizes rutting resistance and resistance to moisture damage. The primary criteria for acceptance of asphalt pavements are in-place air void content and mixture control tolerances, and these attributes are also used to determine the pay factors.

Louisiana

The Louisiana Department of Transportation and Development (LDOTD) accepts asphalt mixtures based on the following characteristics: density, surface tolerance, and dimensional tolerances (McCarthy et al., 2016). LDOTD places emphasis on performance testing for moisture damage and fatigue because it has not experienced much rutting on its asphalt pavements. LDOTD is in the process of creating specifications based on contractor QC tests for acceptance. The LDOTD approach includes shadow testing and mixture qualification based on performance-based mix design properties. One challenge to the implementation of PBS identified by LDOTD is the cost and complexity of the performance testing.

A study by Mohammad et al. (Mohammad, Kim, & Challa, 2010) developed an implementation framework for PBS in Louisiana. In this study, samples were taken from nine asphalt paving projects. Tests considered were the Loaded Wheel Tester (LWT), SCB, and indirect tensile dynamic modulus test. A typical test section layout is shown in Figure A-10. Recommendations included LWT criteria of 6 mm or less (Level 1 pavements) and 10 mm or less (Level 2 pavements) and SCB Jc values of 0.6 kJ/m^2 (Level 1 pavements) and 0.5 kJ/m^2 (Level 2 pavements). A draft plan for PBS testing and sampling was devised. The study recommended a simplified implementation of adding a PBS to current QC/QA specifications. This was preferred to a more complex PBS. Further field and laboratory testing was recommended to further validate the performance criteria and identify and address possible unknown challenges to the PBS approach.

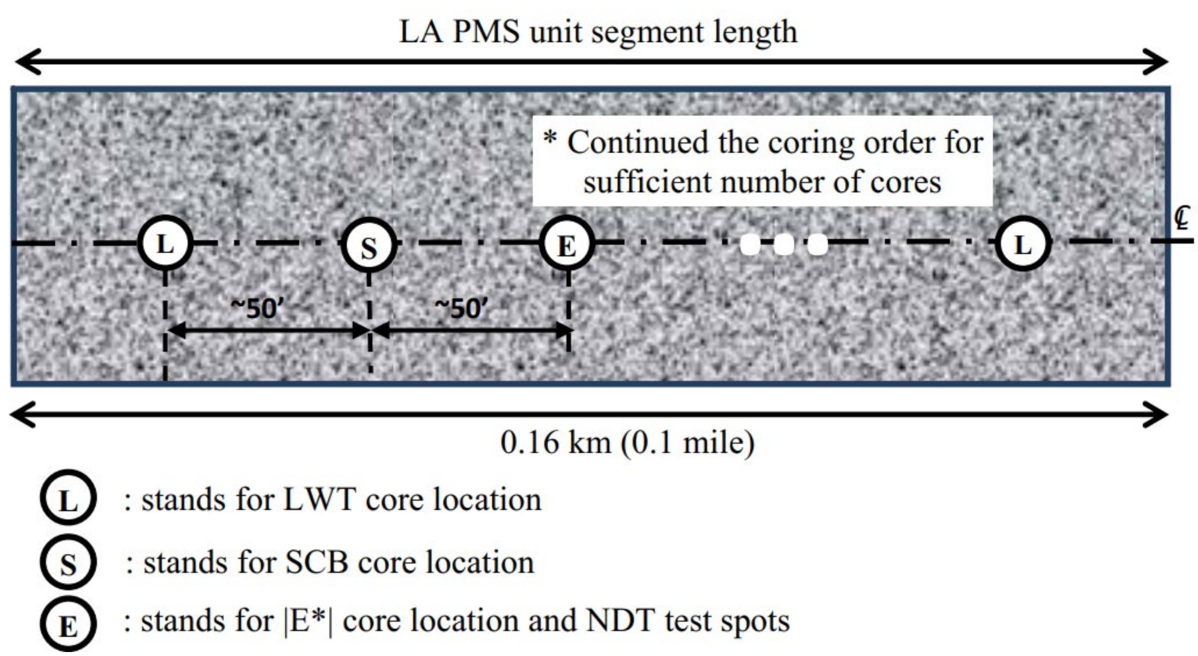


Figure A-10. Typical layout for testing in Louisiana study (Mohammad et al., 2010)

Another previous study (S. Cooper, Mohammad, Kabir, & King, 2014) assessed impacts of modifications to LDOTD specifications aimed at developing a balanced mixture design using Hamburg loaded wheel tester (HLWT) and semicircular bend (SCB) test. Eleven mixtures were sampled from six projects that used the proposed balance mix specification. Results showed that the 11 mixtures performed as well as or better than the mixtures produced from the previous 2006 LDOTD specifications. Fifty percent of the mixtures passed the SCB test (cracking criteria 0.5 kJ/m^2).

Michigan

In a study undertaken for the Michigan Department of Transportation (MDOT) by Williams et al. (Williams, Hill, Hofmann, Zelenock, & Bausano, 2004), laboratory performance test criteria and field specifications for HMA acceptance were developed. APA specifications were established to assess the potential for rutting in mixes, and a model to predict fatigue life based on the four-point beam fatigue (FPBF) apparatus was created. Pay factors for fatigue cracking and rutting were also developed.

Minnesota

Low temperature cracking has been determined to be the most common pavement distress in Minnesota's asphalt pavements (McCarthy et al., 2016). The Minnesota Department of Transportation (MnDOT) developed a specification provision that utilized the DC(T) test to assess the potential for low temperature cracking in asphalt pavement mixtures. MnDOT implementation of DC(T) in PBS included regional validation of PBS, pilot implementation, assessment of the sensitivity of fracture energy to thermal cracking, specification refinement efforts, and round-robin testing (Dave, 2017). For the regional validation studies, 18 sites and 26 sections were used with different binder grades, aggregates, construction types, and traffic. The pilot study included DC(T) tests on five projects. Two of the mixtures met the criteria while the remaining three mixes required adjustments (McCarthy et al., 2016). Distress surveys later found mixed results on these projects as mill and overlay projects experienced more cracking than new construction or full depth reconstruction projects. MnDOT specifications include tables for fracture energy based on traffic levels and PG grade (Tables 4-6) (Dave, 2017). MnDOT is continuing with training and PBS implementation and would like to extend DC(T) specifications to include reflective cracking present in asphalt overlays.

Table A-4. Minimum Average Fracture Energy Mixture Design Requirements for Wearing Course from MnDOT DC(T) Fracture Energy Provisional Performance Specifications (Dave, 2017)

Traffic Level	Fracture Energy
Traffic Level 2-3/PG XX-34	450 J/m^2
Traffic Level 4-5/PGXX-34	500 J/m^2

Table A-5. Allowable Differences between Contractor and Department Test Results from MnDOT DC(T) Fracture Energy Provisional Performance Specifications (Dave, 2017)

Item	Allowable Difference
DC(T) - Fracture Energy (J/m^2)	90
*Test a minimum of six (6) DC(T) test specimens according to ASTM D7313-13 MnDOT Modified revision dated September 1, 2015 to determine the average fracture energy of the submitted mix design (see MnDOT Modified for requirements of when greater than 6 specimens are to be tested).	

Table A-6. Minimum Average Fracture Energy Mixture Production Requirements for Wearing Course from MnDOT DC(T) Fracture Energy Provisional Performance Specifications (Dave, 2017)

Traffic Level/PG Grade	Fracture Energy
Traffic Level 2-3/PG XX-34	400 J/m^2
Traffic Level 4-5/PGXX-34	450 J/m^2

New Jersey

The New Jersey Department of Transportation (NJ DOT) incorporates PBS through both mixture acceptance or rejection and pay adjustments (McCarthy et al., 2016). Reasons NJ DOT is doing performance testing include heavy traffic (increase in truck freight), weather fluctuations, poor pavement performance, and pavement distress issues such as durability and cracking (Blight, 2017). Most of NJ's pavement projects are pavement rehabilitation projects (Bennert, Sheehy, Blight, Gresave, & Fee, 2014).

NJ DOT has developed performance criteria based on fatigue resistance, moisture resistance, and stiffness modulus (McCarthy et al., 2016). Several demonstration projects using APA, Overlay Tester, and flexural beam fatigue test equipment have been undertaken. The NJ DOT performance based process includes verification of volumetrics by NJ DOT regional office, laboratory testing, construction and testing of a test strip (typically the shoulder), and sampling and testing of the material during project construction (Bennert et al., 2014). Performance tests include APA, Overlay Test or Texas Overlay Tester, and Flexural Beam Fatigue.

Instead of using the same asphalt mixture for all applications, NJ DOT implements several performance-based mixtures that are designed based on their purpose (Bennert et al., 2014). NJ DOT's performance-based mixes include High Performance Thin Overlays (HPTO), Bridge Deck Water-proofing Surface Course (BDWSC), Bottom Rich Base Course (BRBC), Bottom Rich Intermediate Course (BRIC), and HMA High RAP (Bennert, n.d.) (Blight, 2017). HPTO

(Figure A-11) is used for maintenance and is rut resistant and durable (Blight, 2017). BDWSC is a waterproof overlay applied on older bridge decks. BRBC is a base course for flexural needs while BRIC is designed to resist reflective cracking on composite pavements. HMA High RAP is used to allow for higher RAP in HMA.

Figure A-11. NJ DOT High Performance Thin Overlay (HPTO) (Blight, 2017)

NJ DOT's performance-based mixes have performed well in the field (Bennert et al., 2014). BDWSC was used on an I-80 bridge with good results as there were no distresses in 1.5 years (Figure A-12) (Bennert). BRBC was applied in summer 2010 on an I-295 rubblization project. NJ DOT believes that using performance based mixes is a good economic investment. Benefit-cost comparisons conducted for different mixes showed benefit-cost ratio ranges between 1.0 and 20.3 (Blight, 2017).

Figure A-12. NJ DOT Bridge Deck Water-proofing Surface Course (BDWSC) on I-80 Bridge (Bennert)

Future work by NJ DOT related to PBS includes adding mixture performance test requirements for all asphalt mixes, including pay adjustments, and QC tests in asphalt plant labs (Blight, 2017). NJ DOT also plans to purchase testing equipment and provide additional training to expand the use of performance testing (McCarthy et al., 2016).

Ohio

The Ohio Department of Transportation (ODOT) uses APA testing for rutting when the design includes at least 15 percent fine aggregates (McCarthy et al., 2016). ODOT also utilizes the Marshall mixture design approach for medium and light mixes, and the HWTD and TSR tests are also used. The Polisher test was developed by ODOT to test friction. ODOT is assessing costs and benefits of performance testing and plans to include a fatigue performance test in the future. Challenges to implementation of PBS identified by ODOT include contractor unfamiliarity with PBS, lack of funding and staff for testing, and knowledge gaps regarding implementation.

Texas

The Texas Department of Transportation (TxDOT) is emphasizing a balanced mix design procedure based on a performance test for cracking (McCarthy et al., 2016). The Texas Overlay Tester (OT) was developed as a potential test for cracking in a study by Zhou et al. (F. Zhou, Scullion, Walubita, & Wilson, 2014). The study recommended using the best three out of five replicate samples due to the sensitivity of the OT to asphalt mix composition and volumetric properties. The research found that due to variations in project conditions such as traffic, climate, and existing conditions, a mix design system based on project-specific conditions should be devised. Figure A-13 shows the concept of the balance mix design that was developed in the study.

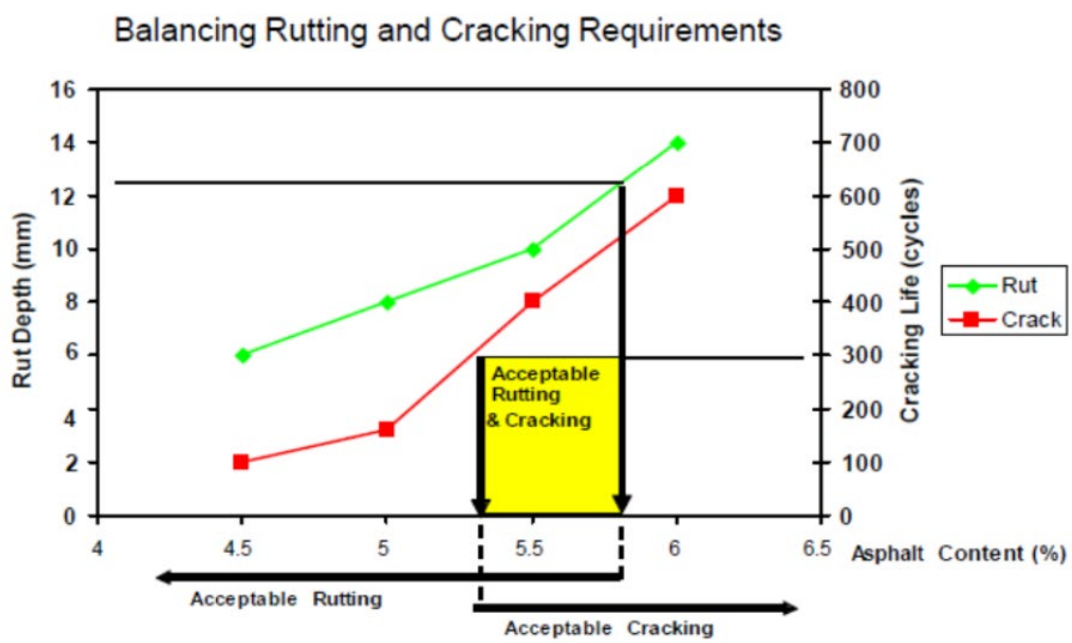


Figure A-13. Balance of Rutting and Cracking Developed in Texas Study (F. Zhou et al., 2014)

TxDOT commonly uses asphalt overlays for highway maintenance (McCarthy et al., 2016). TxDOT hopes to extend pavement service life and improve pavement performance through implementation of PBS. The OT and HWTT are used in (F. Zhou et al., 2014) balance mix design procedure utilized by TxDOT for performance mix design. The HWTT is utilized for possible rutting and moisture sensitivity while the OT tests for cracking resistance (F. Zhou et al., 2014). TxDOT specifies a rutting test according to Tex-242-F, IDT according to Tex-226-F, when required, and an overlay test according to Tex-248-F (TxDOT, 2016b) .

In a previous TxDOT study by Epps et al. (Epps, Glover, & Barcena, 2001), a PBS system for surface treatments was proposed based on PG equipment and grading system. The developed system including recommended limits for high and low surface pavement design temperature based on results of laboratory testing. A field test was proposed to validate the laboratory results.

Another TxDOT study (Elmore, Solaimanian, McGennis, Phromsorn, & Kennedy, 1995) evaluated asphalt binders for seal coats for conformance with the SHRP2 Superpave system since the SHRP2 study did not review seal coats. The study found that asphalts used in Texas can meet the requirements of the PG system based on laboratory testing of samples. The study found that a PG 52-28 could be specified but recommended that selection criteria should be formulated based on traffic, climate, aggregate materials, construction practice, and other factors.

Utah

The Utah Department of Transportation (UDOT) utilizes a standard Superpave HMA mix for asphalt pavements (McCarthy et al., 2016). Acceptance and pay adjustments are determined based on gradation, AC content, longitudinal joint density, and in-place density. UDOT no longer pays for binder as a separate contract item, thus leading to reduced AC content and increased cracking. The HWTD and TSR tests are utilized for rutting and moisture damage resistance. UDOT uses Hamburg tests in its mixture design specs (UDOT, 2015).

Virginia

The Virginia Department of Transportation (VDOT) has incorporated Tensile Strength Ratio (TSR) into its specifications (McCarthy et al., 2016). The Bond test is used to resolve disputes with the contractor regarding mixtures with inadequate performance (McCarthy et al., 2016). VDOT performs Rut testing according to its own standard VTM-110 (VDOT, 2013).

A report by Hughes and Maupin (Hughes III & Maupin, 2000) described plans for implementing PBS in Virginia. Characteristics to be included in the specifications were degree of compaction, thickness, smoothness, segregation, strength, and durability. The report estimated a PBS could be possible in Virginia by 2005. Anticipated challenges to implementation were the ability to devise tests that are linked to performance and the need to assess life-cycle costs that connect the quality characteristic with pavement performance.

Wisconsin

The Wisconsin Department of Transportation (WisDOT) created a specification development team for mixtures with more than 25 percent recycled materials (McCarthy et al., 2016). WisDOT utilizes the HWTD for moisture and rutting and DC(T) test for low temperature cracking for acceptance of mixtures and PG grading. The SCB is used for informational purposes. All the test procedures have been modified by WisDOT (WisDOT, 2015). A test strip is required. The developed specification was used for two pilot projects (one mill and overlay project and one reconstruction project) in 2014. Challenges identified during the pilot study were the effect of additives on performance and impact of aging of DC(T) and SBC on the time needed for project mix design.

Practices of Other States

Other states' practices regarding the use of asphalt PBS are described below:

- Alabama DOT (ALDOT) uses the Asphalt Pavement Analyzer for testing rutting susceptibility of SMA mixes and accepts mixes with rutting less than 4.5mm. The test is done in accordance with ALDOT-401 (Alabama Department of Transportation, 2012).
- The Arkansas State Highway and Transportation Department (AHTD) specifies the Wheel

Tracking Test according to AHTD 480 with rut depth varying from 8 mm to 5 mm depending on the design gyrations (Arkansas DOT, n.d.).

- IDOT has begun implementing the IL-SCB specifications in its 11 newly constructed pilot projects (AASHTO RAC-Sweet 16 High Value Research projects). IDOT has set a specification of FI greater than or equal to 8.0 for acceptance. In addition, IDOT's document for HMA Mixture Design Verification and Production modified for pilot projects only (revised Jan 2016) adopts four performance tests: IL Modified AASHTO 324 (Hamburg Wheel Test), IL Modified AASHTO T283 (TSR), Illinois Test Procedure (ITP) 405 I-FIT (AASHTO TP-124), ASTM D7313 (DC(T)) (ASTM, 2013) (IDOT, 2015).
- The Maryland Department of Transportation (MDOT) is participating in a pooled fund research study aimed at minimizing Asphalt Mixture Performance Tester (AMPT) equipment costs (McCarthy et al., 2016).
- The North Carolina Department of Transportation (NCDOT) performs rutting tests using APA and TSR testing in accordance with NCDOT-T-283 (NCDOT, 2016).
- Oregon DOT (ODOT) includes APA testing (rutting) in their mix designs. ODOT's pavement design guide also includes IDT as a possible laboratory performance test for existing HMAC (ODOT, 2011) (ODOT, 2013).
- Washington DOT (WSDOT) performs both rutting test (AASHTO T324) (AASHTO, 2017b) and IDT test (ASTM D6931) (ASTM, 2017) but they do so with slight variations in the test methods to suit their requirements. They also determine the stripping potential using TSR values (AASHTO T283) (AASHTO, 2014) (WSDOT, 2017).
- The DC(T) spec is being currently implemented by Minnesota DOT (MnDOT) (Minnesota DC(T)), WisDOT, Iowa DOT, Chicago DOT, and Illinois Tollway (Dave et al., 2015).
- Several DOTs have made commitments for "Implementation of the Asphalt Mixture Performance Tester for Superpave Validation" (NCHRP, 2016)

Edmonton, Alberta

In Edmonton, Alberta, each proposed asphalt mixture goes through performance testing before it is implemented in the construction program (McCarthy et al., 2016). Various tests such as the dynamic modulus test, flexural beam fatigue test, APA, and HWTD test are used for qualification, QC, and acceptance. APA testing is required for acceptance of asphalt paving mixtures. Edmonton uses 70 asphalt and concrete test sections in the city for field monitoring. Edmonton meets twice annually with the Alberta Roadbuilders and Heavy Construction Association (ARHCA). Edmonton is evaluating the costs and benefits of its performance testing and the basis for its pay factors. Pay factors are currently determined based on the following characteristics: in situ density, thickness, and binder content.

A.3. Other PBS Studies

Several research studies related to PBS have been undertaken. Some of these studies focused on specification development while other studies concentrated on test methods and protocols for performance tests. These studies are described in the following sections.

A.3.1. PBS Development Studies

Several studies leading to the development of PBS have been performed. Examples of PBS developed in these studies include a PRS using Quality-Related Specification Software (QRSS), a pilot regional specification for high-performance thin overlay (HiPO) mixtures, and AASHTO-formatted performance specification templates for HMA pavement. The following sections describe these studies in greater detail.

Use of QRSS to Develop PRS

In a NCHRP study completed in 2011, a PRS was developed using Quality-Related Specification Software (QRSS) (Fugro & ASU, 2011). The QRSS could be used as a pay adjustment system to calculate penalties or bonuses based on the difference in predicted life between the as-designed and as-built HMA pavement. The QRSS utilizes effective temperature to determine effects of climate on dynamic modulus and simulation of MEPDG predictions of distresses. Methods were established to correlate the level of pavement distress to the predicted pavement life, and the penalty or bonus was then calculated based on the difference in predicted life.

Specification for HiPO Mixtures

A pilot regional specification for high-performance thin overlay (HiPO) mixtures was developed through a partnership between the Northeast Pavement Preservation Partnership (NEPPP), Pennsylvania Asphalt Pavement Association, academia, and industry (Mogawer et al., 2014). Agencies from New Hampshire, Vermont, Massachusetts, Rhode Island, New Jersey, Maryland, and Pennsylvania were involved in this study. The pilot specification was posted on the AASHTO TSP2 website. It includes provision for the use of reclaimed asphalt pavement (RAP) in the mixture and contains the following mixture performance requirements: thermal cracking (TSRST), cracking (OT with and without RAP), fatigue life (AASHTO T 321) (AASHTO, 2017a), and rutting (APA). The specification also includes requirements for surface preparation, material properties, and mixture design requirements. The specification was implemented on two demonstration projects in New Hampshire and Vermont. Two years after construction, the mixtures on both demonstration projects were performing well. Modifications to the pilot specification related to RAP testing, mixture design, surface preparation, and mixture performance testing were suggested.

SHRP2 Study

In a SHRP2 project, AASHTO-formatted performance specification templates were developed for several applications, including HMA and PCC pavement, concrete bridge decks, work zone traffic control, geotechnical applications, and quality management (Scott III et al., 2014). Implementation guidelines were also established in conjunction with the specifications. Performance specifications for pavement foundations and bridge decks were implemented on demonstration projects with the Missouri Department of Transportation (MoDOT) and Virginia Department of Transportation (VDOT), respectively. Recommendations to help further implementation included the use of additional demonstration projects, providing opportunities for outreach and training, developing a performance specification expert technical group, continued development of performance specifications, and the creation of a web-based specification tool.

A.3.2. Studies on Test Methods and Protocols Related to PBS

Some studies to investigate test methods and protocols related to PBS have been undertaken. These studies have assessed pavement distresses, test methods and HMA mixture responses related to pavement distresses, purchase specifications for equipment for performance tests, selection and refinement of cracking tests, characteristics to include in a PRS, and the use of RAP in mix design. These studies are described in the following paragraphs.

Study of Existing Knowledge of Pavement Distresses

A study by Brown et al. (E. R. Brown, Kandhal, & Zhang, 2001) evaluated existing knowledge regarding various pavement distresses including permanent deformation, fatigue cracking, low-temperature cracking, moisture susceptibility, and friction properties and provided some recommendations for performance testing. The results included a list of recommended tests and criteria for permanent deformation. General guidance for minimizing the other distresses was provided.

Assessment of Test Methods

NCHRP Project 9-19, undertaken by Witczak et al. (Witczak, Kaloush, Pellinen, El-Basyouny, & Von Quintus, 2002), sought to choose test method and HMA mixture responses that were closely linked to pavement distresses. Rutting was determined to be the most important distress followed by fracture. Fatigue cracking was also considered in this study. The project recommendations included Simple Performance Test (SPT) method and response parameter combinations for HMA rutting, HMA fatigue cracking, and HMA low-temperature cracking. Preliminary draft protocols for the SPTs were prepared and provided. These SPTs were to be validated in the field in the next phase of this study.

Purchase Specification for Equipment

As part of a NCHRP study by Bonaquist et al. (R. Bonaquist, Christensen, & Stump, 2003), a purchase specification for equipment was developed for three simple performance tests (SPT): dynamic modulus, flow number, and flow time in this project. Two SPT systems were procured and evaluated. The assessment process included mixture testing to investigate material properties, an evaluation of functionality, and testing to determine if the devices met the specifications and were calibrated correctly. The results showed that both devices were user-friendly and satisfied the specification requirements. The flow number test resulted in excessive variability. One of the two devices was not approved while the other device was conditionally approved.

Design of Field Experiments to Validate Laboratory Cracking Tests

The goal of NCHRP Project 09-57 was to facilitate assessment of cracking potential of HMA mixtures through the creation of an experimental procedure for field validation of laboratory cracking tests (F. Zhou et al., 2016). The study process included selection and refinement of cracking tests and development of the field experiment design including time and cost. Laboratory tests for cracking of asphalt mixtures were chosen, and the protocols for the field validation of these tests were developed. A workshop was conducted in which 12 cracking tests were discussed, and seven of these tests were chosen: DC(T), SCB-IL, SCB-TP105, SCB-LTRC, OT, BBF, and IDT-Florida. Field experimental designs including estimated costs and schedules were created for each of four cracking types: thermal, reflection, bottom-up fatigue, and top-down. A follow-up project was suggested to implement these designs in the field.

Identification of HMA Properties for Use in a PRS

NCHRP Project 09-15 sought to identify HMA properties and test methods that could be used for predicting pavement performance and accepting or rejecting mixtures (TRB, 2004). The following five characteristics were identified for possible inclusion in a PRS: segregation, ride quality, in-place density, longitudinal construction joint density, and in-place permeability. The study provided recommendations for test methods and limiting values for these properties.

Development of Guidelines for RAP Mixtures

Another study by West et al. (R. West, Willis, & Marasteanu, 2013) sought to address the need for guidance to manage Recycled Asphalt Pavement (RAP) before mix design. In the study, guidelines for RAP management were created to improve the quality and uniformity of RAP mixes to a level comparable with virgin asphalt mixes. In addition, recommendations to facilitate better mix design standards for mixes with RAP contents between 25 and 55 percent were developed. The report included suggested revisions to AASHTO R 35 and M 323 to improve mix design for high RAP mixes and recommendations for additional tests to assess the mix designs based on their use.

A.4. Conclusion

The review of existing state practices regarding the use of performance specifications for asphalt pavements found that a majority of states currently use or are planning to implement some type of PBS, but PBS has only been implemented as a standard practice in a limited number of agencies (McCarthy et al., 2016). Agencies are using PBS for a variety of reasons, including reduced maintenance, greater distress resistance, heavy traffic (increase in truck freight), weather fluctuations, and increased pavement durability (McCarthy et al., 2016) (Blight, 2017). While the use of PBS faces some implementation challenges such as testing time, cost, training needs, and knowledge gaps regarding PBS, PBS implementation in states such as New Jersey has shown very positive results and economic benefits (McCarthy et al., 2016) (Bennert et al., 2014) (Blight, 2017).

The main aim of agencies involved in construction of asphalt roads is to improve the field performance of the asphalt mixtures. The rising use of recycled and novel materials in asphalt mixture have rendered the previous semi-empirical methods of mixture design partly incapable of accurately predicting the mixture field performance with high precision. Meeting this challenge calls for a shift towards an approach involving mixture performance tests. The preceding sections described a host of mixture performance tests for various parameters or distresses. Combining those tests with binder-related tests and mixture volumetric, the prediction of field performance of mixtures should provide a more robust and reliable design criteria for the current asphalt mixtures leading to better roads.

A.5. References

AASHTO. (1993). *Standard method of test for thermal stress restrained specimen tensile strength (TP 10)*.

AASHTO. (2003). *Standard method of test for determining dynamic modulus of hot mix asphalt (HMA) (TP 62)*.

AASHTO. (2007). *Standard method of test for determining the rutting susceptibility of hot mix asphalt (HMA) using the asphalt pavement analyzer (APA) (TP 63)*.

AASHTO. (2010). *Standard method of test for determining the creep compliance and strength of hot mix asphalt using the indirect tensile test device (T 322). AASHTO T 322*.

AASHTO. (2013). *Standard method of test for determining the fracture energy of asphalt mixtures using the semicircular bend geometry (SCB) (TP 105-13)*.

AASHTO. (2014). *Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage T283-14-UL*.

AASHTO. (2017a). *Standard method of test for determining the fatigue life of compacted asphalt mixtures subjected to repeated flexural bending (T 321)*.

AASHTO. (2017b). *Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Hot Mix Asphalt (HMA) T324-17*.

AASHTO R30. (2006). Standard Practice for Mixture Conditioning of Hot Mix Asphalt, *2*, 1–2.

Airey, G. D. (2003). State of the Art Report on Ageing Test Methods for Bituminous Pavement Materials. *International Journal of Pavement Engineering*, *4*(3), 165–176. https://doi.org/10.1080/1029843042000198568

Al-Qadi, I., Ozer, H., Lambros, J., Lippert, D., El Khatib, A., Khan, T., … Rivera-Perez, J. J. (2015). Testing Protocols to Ensure Mix Performance w/ High RAP and RAS. *Illinois Center for Transportation*, (1), 209.

Alabama Department of Transportation. (2012). Standard Specifications for highway Construction. *Standard Specifications for Highway Construction*, 198,199.

Alavi, M., Hajj, E. Y., & Sebaaly, P. E. (2015). A comprehensive model for predicting thermal cracking events in asphalt pavements. *International Journal of Pavement Engineering*, *8436*(October), 1–15. https://doi.org/10.1080/10298436.2015.1066010

Arkansas DOT. (n.d.). Asphalt Pavements.

Arnold, J. W., Behnia, B., McGovern, M. E., Hill, B., Buttlar, W. G., & Reis, H. (2014). Quantitative evaluation of low-temperature performance of sustainable asphalt pavements containing recycled asphalt shingles (RAS). *Construction and Building Materials*, *58*, 1–8. https://doi.org/10.1016/j.conbuildmat.2014.02.002

Artamendi, I., & Khalid, H. A. (2006). A comparison between beam and semi-circular bending fracture tests for asphalt. *Road Materials and Pavement Design*, *7*(December), 163–180. https://doi.org/10.1080/14680629.2006.9690063

Aschenbrener, T. (1995). Investigation of Low Temperature Thermal Cracking In Hot Mix Asphalt.

ASTM. (2012). ASTM E399-12e3, Standard Test Method for Linear-Elastic Plane-Strain Fracture Toughness KIc of Metallic Material. West Conshohocken, PA: ASTM International.

ASTM. (2013). ASTM D7313-13: Standard Test Method for Determining Fracture Energy of Asphalt-Aggregate Mixtures Using the Disk-Shaped Compact Tension Geometry. *ASTM International Standards*, 2–8. https://doi.org/10.1520/D7313

ASTM. (2017). ASTM D6931-17, Standard Test Method for Indirect Tensile (IDT) Strength of Asphalt Mixtures. West Conshohocken, Pennsylvania: ASTM International.

Behnia, B., Buttlar, W. G., & Reis, H. (2014). Cooling cycle effects on low temperature cracking characteristics of asphalt concrete mixture, 1359–1371. https://doi.org/10.1617/s11527-014-0310-y

Behnia, B., Dave, E. V., Ahmed, S., Buttlar, W. G., & Reis, H. (2011). Effects of Recycled Asphalt Pavement Amounts on Low-Temperature Cracking Performance of Asphalt Mixtures Using Acoustic Emissions. *Transportation Research Record: Journal of the Transportation Research Board*, *2208*(1), 64–71. https://doi.org/10.3141/2208-09

Bell, C. . A., Abwahab, Y., Cristi, M. E., & Sosnovske, D. (1994). Selection of Laboratory Aging Procedures for Asphalt-Aggregate Mixtures, 92.

Bennert, T. (n.d.). Implementation of HMA Mixtures in NJ.

Bennert, T., Sheehy, E., Blight, R., Gresave, S., & Fee, F. (2014). Implementation of Performance-Based Specifications for Asphalt Mix Design and Production Quality Control for New Jersey. *Transportation Research Circular*, *E-C189*(October), 13–31. https://doi.org/10.17226/22248

Blight, R. (2017). NJDOT Experience with Asphalt Mixture Performance.

Bonaquist, R., Christensen, D., & Stump, W. (2003). *Simple Performance Tester for Superpave Mix Design: First-Article Development and Evaluation*. Washington, D.C. https://doi.org/10.17226/21954

Bonaquist, R. F. (2011). *Mix design practices for warm mix asphalt. NCHRP Report 691*. https://doi.org/10.17226/14488

Braham, A., Buttlar, W. G., & Ni, F. (2010). Laboratory Mixed-Mode Cracking of Asphalt Concrete Using the Single-Edge Notch Beam. *Road Materials and Pavement Design*, *11*(3), 681–700. https://doi.org/10.1080/14680629.2010.9690299

Braham, A. F., Buttlar, W. G., Clyne, T. R., Marasteanu, M. O., & Turos, M. I. (2009). The effect of long-term laboratory aging on asphalt concrete fracture energy. *Journal of the Association of Asphalt Paving Technologists*, *78*.

Brown, E. R., Kandhal, P. S., & Zhang, J. (2001). *Performance Testing for Hot Mix Asphalt* (Vol. NCAT Repor). Auburn, Alabama.

Brown, S. F., & Scholz, T. V. (2000). Development of Laboratory Protocols for the Ageing of Asphalt Mixtures. In *2nd Eurasphalt and Eurobitume Congress, Barcelona, Spain.*

BSI. (2006). BS EN 13108-1:2006 Bituminous mixtures. Material specifications. Asphalt Concrete. London, United Kingdom: BSI.

Buttlar, W., Rath, P., Majidifard, H., Dave, E. V., & Wang, H. (2018a). Relating DC (T) Fracture Energy to Field Cracking Observations and Recommended Specification Thresholds for Performance-Engineered Mix Design. *Asphalt Mixtures*, 51.

Buttlar, W. G., Meister, J., Jahangiri, B., Majidifard, H., & Rath, P. (2019). Performance Characteristics of Modern Recycled Asphalt Mixes in Missouri, Including Ground Tire Rubber, Recycled Roofing Shingles, and Rejuvenators (No. cmr 19-002).

Buttlar, W. G., Hill, B. C., Wang, H., & Mogawer, W. (2016). Performance space diagram for the evaluation of high- and low-temperature asphalt mixture performance. *Road Materials and Pavement Design*, *0*(0), 1–23. https://doi.org/10.1080/14680629.2016.1267446

Buttlar, W. G., & Rath, P. (2017). *ILLINOIS TOLLWAY I-88 GROUND TIRE RUBBER TEST SECTIONS: LABORATORY MIX DESIGNS & PERFORMANCE TESTING*.

Buttlar, W. G., & Wang, H. E. (2016). ILLINOIS TOLLWAY PROJECT LABORATORY INVESTIGATION OF ILLINOIS TOLLWAY STONE MATRIX ASPHALT MIXTURES

WITH VARIED LEVELS OF ASPHALT BINDER REPLACEMENT BY.

Buttlar, W., Khateeb, G., & Sherman, D. (2004). Potential Application of the Hollow Cylinder Tensile Tester as a Simple Performance Test. *Transportation Research Circular Number E-C068: New Simple Performance Tests for Asphalt Mixtures*, (September).

California Department of Transportation. (2012). Highway Design Manual. *Media*, (12).

Chiu, C. Te, & Lu, L. C. (2007). A laboratory study on stone matrix asphalt using ground tire rubber. *Construction and Building Materials*, *21*(5), 1027–1033. https://doi.org/10.1016/j.conbuildmat.2006.02.005

Christensen, D. W., & Bonaquist, R. (2007). Using the Indirect Tensile Test to Evaluate Rut Resistance in Developing Hot-Mix Asphalt Designs. *Transportation Research Circular Number E-C124: Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing*, (December).

Christensen, D. W., & Bonaquist, R. F. (2004). *Evaluation of Indirect Tensile Test (IDT) Procedures for Low-Temperature Performance of Hot Mix Asphalt*. https://doi.org/10.17226/13775

Christensen, D. W., Bonaquist, R. F., Anderson, D., & Gokhale, S. (2004). Indirect Tension Strength as a Simple Performance Test. *Transportation Research Circular Number E-C068: New Simple Performance Tests for Asphalt Mixtures*, (September).

Christensen, D. W., Bonaquist, R. F., & Jack, D. P. (2000). *Evaluation of Triaxial Strength as a Simple Test for Asphalt Concrete Rut Resistance*. University Park, Pennsylvania.

Cong, L., Peng, J., Guo, Z., & Wang, Q. (2016). Evaluation of Fatigue Cracking in Asphalt Mixtures Based on Surface Energy. *Journal of Materials in Civil Engineering*, 1–6. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001465.

Coons, R. F., & Wright, P. H. (1968). An Investigation of the Hardening of Asphalt Recovered from Pavements of Various Ages. *Association of Asphalt Paving Technologists*.

Cooper, S. B., Mohammad, L. N., & Elseifi, M. (2016). Multiscale Chemico-Thermo-Hydro-Mechanical Modeling of Early-Stage Hydration and Shrinkage of Cement Compounds. *Journal of Materials in Civil Engineering*, *27*(September), 1239–1247. https://doi.org/10.1061/(ASCE)MT

Cooper, S., Mohammad, L., Kabir, S., & King, W. (2014). Balanced Asphalt Mixture Design Through Specification Modification. *Transportation Research Record: Journal of the Transportation Research Board*, *2447*, 92–100. https://doi.org/10.3141/2447-10

Dave, E., Daniel, J., Jacques, C., & Decarlo, C. (2015). How Does Silo Storage Time Impact Asphalt Pavement Performance and Durability?

Dave, E. V., Behnia, B., Ahmed, S., Buttlar, W. G., & Reis, H. (2013). Low Temperature Fracture Evaluation of Asphalt Mixtures Using Mechanical Testing and Acoustic Emission Techniques (pp. 193–225).

Dave, E. V., Buttlar, W. G., Leon, S. E., Behnia, B., & Paulino, G. H. (2013). IlliTC-Low temperature cracking model for asphalt pavements. *Asphalt Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions*, *82*(January 2014), 91–126. https://doi.org/Doi 10.1080/14680629.2013.812838

Dave, E. V. (2017). Development & Implementation of Low Temperature Cracking Performance Specifications in Minnesota.

Dave, E. V, & Investigator, P. (2011). Synthesis of Performance Testing, (September).

Deusen, D. Van. (2015). DCT Low Temperature Fracture Testing Pilot Project, (May).

DeVol, J. R. (2015). Texas Overlay Tester Project, (August), 3883.

Elwardanya, M. D., Rada, F. Y., Castorenaa, C., & Kima, Y. R. (2018). Climate-, Depth-, and Time-Based Laboratory Aging Procedure for Asphalt Mixtures. *J. Associat. Asphalt Paving Technol.(AAPT)*, *87*.

Elmore, W., Solaimanian, M., McGennis, R., Phromsorn, C., & Kennedy, T. (1995). *Performance-based seal coat asphalt specifications* (Vol. FHWA/TX-96). Austin, Texas.

Elwardany, M. D., Yousefi Rad, F., Castorena, C., & Kim, Y. R. (2017). Evaluation of asphalt mixture laboratory long-term ageing methods for performance testing and prediction. *Road Materials and Pavement Design*, *18*(December), 28–61. https://doi.org/10.1080/14680629.2016.1266740

Epps, A. L., Glover, C. J., & Barcena, R. (2001). A Performance-Graded Binder Specification for Surface Treatments, *FHWA/TX-02*.

EVAC. (2014). Hot asphalt mix criteria. East Valley Asphalt Committee.

Fakhri, M., Ghanizadeh, A. R., & Omrani, H. (2013). Comparison of Fatigue Resistance of HMA and WMA Mixtures Modified by SBS. *Procedia - Social and Behavioral Sciences*, *104*, 168–177. https://doi.org/10.1016/j.sbspro.2013.11.109

Falchetto, A. C., Moon, K. H., & Wistuba, M. P. (2017). Development of a simple correlation between bending beam rheometer and thermal stress restrained specimen test low-temperature properties based on a simplified size effect approach. *Road Materials and Pavement Design*, *18*(sup2), 339–351. https://doi.org/10.1080/14680629.2017.1305147

Farrar, M. J., Harnsberger, P. M., Thomas, K. P., & Wiser, W. (2006). EVALUATION OF OXIDATION IN ASPHALT PAVEMENT TEST SECTIONS \nAFTER FOUR YEARS OF SERVICE. *International Conference on Perpetual Pavement, September 13-15, 2006*, (307), 1–17.

Federal Highway Administration (FHWA). (2013). Asphalt Mixture Performance Tester (AMPT), (February).

Federal Highway Administration (FHWA). (2016). *Testing for Fatigue Cracking in the Asphalt Mixture Performance Tester*.

FHWA. (2015). Asphalt Mixture ETG Meeting Technical Report, (April), 1–58.

Fugro, & ASU. (2011). *A Performance-Related Specification for Hot-Mixed Asphalt*. Washington, D.C. https://doi.org/10.17226/22835

Gedafa, D. ., M.Hossain, Romanoschi, S. A., & Gisi, A. J. (2013). Effect of Aging on Dynamic Modulus and Fatigue Life of Superpave Mixes. *American Society of Civil Engineers (ASCE)*, *Airfield a*(Nchrp 2004), 1040–1050. https://doi.org/10.1061/9780784413005.087

Ghuzlan, K. A., & Carpenter, S. H. (2002). Traditional fatigue analysis of asphalt concrete mixtures. *Urbana*, 25.

Gillen, S. (2007). *Ground Tire Rubber(GTR) Asphalt Pavement Demonstration Project*. Downers Grove, IL.

Haghshenas, H., & Kim, Y. (2016). Research on High-RAP Asphalt Mixtures with Rejuvenators and WMA Additives Nebraska Transportation Research on High-RAP Asphalt Mixtures with Rejuvenators and WMA Additives Hamzeh Haghshenas Department of Civil Engineering.

Hajj, E. Y., Alavi, M. Z., Morian, N., Kazemi, S. F., & Sebaaly, P. (2014). Thermal Cracking Analysis Model and Pavement Temperature Profile Prediction Model. In *FHWA Asphalt Mixture ETG* (pp. 1–47).

Hanson, C. (2015). Minnesota Disk-Shaped Compact Tension Testing (DCT) Why do we need to specify Low Temperature.

Harvey, J., Wu, R., Signore, J., Basheer, I., Holikatti, S., Vacura, P., & Holland, T. (2014).

Performance-Based Specifications: California Experience to Date. *Transportation Research Circular*, *E-C189*(October). https://doi.org/10.17226/22248

Hong, F., Chen, D.-H., & Mikhail, M. M. (2011). Long-Term Performance Evaluation of Recycled Asphalt Pavement Results from Texas. *Transportation Research Record: Journal of the Transportation Research Board*, *2180*, 58–66. https://doi.org/10.3141/2180-07

Houston, W. N., Mirza, M. W., Zapata, C. E., & Raghavendra, S. (2006). *Environmental Effects in Pavement Mix and Structural Design Systems*. https://doi.org/10.17226/23244

Huang, B., Shu, X., & Vukosavljevic, D. (2011). Laboratory Investigation of Cracking Resistance of Hot-Mix Asphalt Field Mixtures Containing Screened Reclaimed Asphalt Pavement, *23*(November), 1535–1543. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000223.

Huang, Y., Wang, X., Liu, Z., & Li, S. (2016). Dynamic modulus test and master curve analysis of asphalt mix with trapezoid beam method. *Road Materials and Pavement Design*, *18*(sup3), 281–291. https://doi.org/10.1080/14680629.2017.1329882

Hughes III, C. S., & Maupin, G. W. (2000). The Journey Toward End Result Specifications for Asphalt Concrete, *VTRC 01-TA*.

IDOT. (2015). Illinois Flexibility Index Test - Pilot Projects (Circular Letter 2015-19). Springfield, Illinois.

Ishai, P. I., Amit, E. M., Kesler, E. T., & Peled, R. (n.d.). New Advancements in Rubberized Asphalt Using an Elastomeric Asphalt Extender – Three Case Studies, *4*, 1–24.

Jahanbakhsh, H., Hosseini, P., Nejad, F. M., & Habibi, M. (2019). Intermediate temperature fracture resistance evaluation of cement emulsified asphalt mortar. *Construction and Building Materials*, 197, 1-11.

Jahangiri, B., Majidifard, H., Meister, J., & Buttlar, W. G. (2019). Performance evaluation of asphalt mixtures with reclaimed asphalt pavement and recycled asphalt shingles in Missouri. *Transportation Research Record*, *2673*(2), 392-403.

Jemere, Y. (2010). *Development of a laboratory ageing method for bitumen in porous asphalt*. Delft University of Technology.

Jung, D., & Vinson, T. S. (1993). Thermal stress restrained specimen test to evaluate low-temperature cracking of asphalt-aggregate mixtures. *Transportation Research Record*, 12.

Kanaan, A. (2013). *Impact of Recycled Asphalt Shingles on Asphalt Concrete Characteristics*.

University of Illinois at Urbana-Champaign. Retrieved from http://hdl.handle.net/2142/45354

Kebede, N. A. (2012). *DEVELOPMENT OF AN ALTERNATIVE TEST TO OBTAIN ASPHALT MIXTURE CREEP COMPLIANCE AT LOW TEMPERATURES*. UIUC.

Kim, Y. R., Castorena, C., Elwardany, M. D., Rad, F. Y., Underwood, S., Akshay, G., ... & Glaser, R. R. (2018). *Long-term aging of asphalt mixtures for performance testing and prediction*. Transportation Research Board.

Kim, Y. R., Castorena, C., Rad, F. Y., Elwardany, M., Underwood, S., Gundha, A., … Glaser, R. R. (2015). *Long-Term Aging of Asphalt Mixtures for Performance Testing and Prediction*. *National Cooperative Highway Research Program (Nchrp)*. https://doi.org/10.17226/24959

Kim, Y. R., & Wen, H. (2002). Fracture energy from indirect tension testing. *Asphalt Paving Technology 2002, March 18, 2002 - March 20, 2002*, *71*(August), 779–793.

Lei, Z., Bahia, H., & Yi-Qiu, T. (2015). Effect of bio-based and refined waste oil modifiers on low temperature performance of asphalt binders. *Construction and Building Materials*, *86*, 95–100. https://doi.org/10.1016/j.conbuildmat.2015.03.106

Li, X., Braham, A. F., Marasteanu, M. O., Buttlar, W. G., & Williams, C. (2011). Design Effect of Factors Affecting Fracture Energy of Asphalt Concrete at Low Temperature. *Road Materials and Pavement Design*, *629*(June 2013), 37–41. https://doi.org/10.3166/RMPD.9HS.397-416

Li, X. J., & Marasteanu, M. O. (2010). Using Semi Circular Bending Test to Evaluate Low Temperature Fracture Resistance for Asphalt Concrete. *Experimental Mechanics*, *50*(7), 867–876. https://doi.org/10.1007/s11340-009-9303-0

Li, X., Zofka, A., Marasteanu, M., & Clyne, T. R. (2006). Evaluation of field aging effects on asphalt binder properties. *Road Materials and Pavement Design*, *7*(December), 57–73. https://doi.org/10.1080/14680629.2006.9690058

Liang, R. Y., & Lee, S. (1996). Short-term and long-term aging behavior of rubber modified asphalt paving mixture. *Transportation Research Record: Journal of the Transportation Research Board*, *1530*(1), 11–17. https://doi.org/10.3141/1530-02

Luo, X., Gu, F., & Lytton, R. L. (2015). Prediction of Field Aging Gradient in Asphalt Pavements. *Transportation Research Record: Journal of the Transportation Research Board*, *2507*, 19–28. https://doi.org/10.3141/2507-03

Ma, W. (2014). *Proposed Improvements to Overlay Test for Determining Cracking Resistance of*

Asphalt Mixtures. Auburn University.

Mandal, T. (2016). ENHANCEMENT OF THE ASPHALT THERMAL CRACKING ANALYSER (ATCA) TEST TO ALLOW MEASURING CRITICAL PROPERTIES AFFECTING CRACKING OF ASPHALT MIXTURES.

Mangiafico, S., Sauzéat, C., Di Benedetto, H., Pouget, S., Olard, F., & Planque, L. (2016). Complex modulus and fatigue performances of bituminous mixtures with reclaimed asphalt pavement and a recycling agent of vegetable origin. *Road Materials and Pavement Design*, *18*(2), 315–330. https://doi.org/10.1080/14680629.2016.1213509

Marasteanu, M. O., Buttlar, W. G., Bahia, H., & Williams, C. R. (2012). *Investigation of Low Temperature Cracking in Asphalt Pavements, National Pooled Fund Study Phase-II*.

Marasteanu, M. O., Buttlar, W. G., Zofka, A., Turos, M., Li, X., Velasquez, R., … McGraw, J. (2007). Investigation of Low Temperature Cracking in Asphalt Pavements. Phase-I, (May), 1–4.

Mashaan, N. S., Karim, M. R., Abdel Aziz, M., Ibrahim, M. R., Katman, H. Y., & Koting, S. (2014). Evaluation of fatigue life of CRM-reinforced SMA and its relationship to dynamic stiffness. *Scientific World Journal*, *2014*. https://doi.org/10.1155/2014/968075

Mbarki, R., Kutay, M. E., Gibson, N., & Abbas, A. R. (2012). Comparison between fatigue performance of horizontal cores from different asphalt pavement depths and laboratory specimens. *Road Materials and Pavement Design*, *13*(3), 422–432. https://doi.org/10.1080/14680629.2012.685843

McCarthy, L. M., Callans, J., Quigley, R., & Scott, S. V. (2016). *Performance Specifications for Asphalt Mixtures*. https://doi.org/10.17226/23564

MDOT. (2014). HMA Production Manual. *OMFIF Bulletin*, *5*(November). Retrieved from http://www.who.int/bulletin/en/

Mejías-santiago, M., Doyle, J. D., & Rushing, J. F. (2014). Warm-Asphalt for Airfield Pavements. Atlantic City, NJ: FAA WORLDWIDE AIRPORT TECHNOLOGY TRANSFER CONFERENCE.

Mirza, M. W., & Witczak, M. W. (1995). Development of a global aging system for short and long term aging of asphalt cements (with discussion). *Journal of the Association of Asphalt Paving Technologists*, *64*.

MnDOT. (2014). *MnDOT Pavement Design Manual*.

Mogawer, W., Austerman, J., Kluttz, R., & Mohammad, L. (2014). Development and Validation of Performance-Based Specifications for High-Performance Thin Overlay Mix. *Transportation Research Circular*, *E-C189*.

Mohammad, L., Kim, M., & Challa, H. (2010). *Development of Performance-based Specifications for Louisiana Asphalt Mixtures* (Vol. 70808).

Mohammed, L. N., Cooper, S. B., & Elseifi, M. (2010). *Mechanistic Evaluation of Asphalt Mixtures Containing Thiopave Additives* (Vol. 70808). Retrieved from http://www.ltrc.lsu.edu/downloads.html

NAPA. (2009). National Asphalt Pavement Association: Fast Facts.

NCDOT. (2016). *2016 Ashpalt Quality Management System*. Raleigh, North Carolina.

NCHRP. (2016). Study Detail View: Implementation of the Asphalt Mixture Performance Tester (AMPT) for Superpave Validation (Transportation Pooled Fund).

Norouzi, A., Kim, Y. R., Kim, S. S., & Yang, J. (2017). Effect of Reclaimed Asphalt Pavement Content and Binder Grade on Fatigue-Resisting Performance of Asphalt Mixtures in Georgia. *Journal of Materials in Civil Engineering*, *29*(9), 4017115. https://doi.org/10.1061/(ASCE)MT.1943-5533.0001960

ODOT. (2011). *ODOT Pavement Design Guide*.

ODOT. (2013). Contractor Mix Design Guidelines for Asphalt Concrete. Salem, Oregon.

Ozer, H., Al-Qadi, I., Kanaan, A., & Lippert, D. (2013). Performance Characterization of Asphalt Mixtures at High Asphalt Binder Replacement with Recycled Asphalt Shingles. *Transportation Research Record: Journal of the Transportation Research Board*, *2371*, 105–112. https://doi.org/10.3141/2371-12

Ozer, H., Al-Qadi, I. L., Lambros, J., El-Khatib, A., Singhvi, P., & Doll, B. (2016). Development of the fracture-based flexibility index for asphalt concrete cracking potential using modified semi-circle bending test parameters. *Construction and Building Materials*, *115*, 390–401. https://doi.org/10.1016/j.conbuildmat.2016.03.144

Ozer, Hasan, Al-Qadi, Imad L, Singhvi, Punit, … Ahmad. (2016). Fracture Characterization of Asphalt Mixtures with High Recycled Content Using Illinois Semicircular Bending Test Method and Flexibility Index. *Transportation Research Record: Journal of the Transportation Research Board*, (2575), 130–137. https://doi.org/10.3141/2575-14

Peralta, J., Williams, R. C., Silva, H. M. R. D., Vera, A., & Machado, A. (2013). Recombination of Asphalt with Bio-Asphalt : Binder Formulation and Asphalt Mixes Application, (i), 1–36.

Poulikakos, L. D., dos Santos, S., Bueno, M., Kuentzel, S., Hugener, M., & Partl, M. N. (2014). Influence of short and long term aging on chemical, microstructural and macro-mechanical properties of recycled asphalt mixtures. *Construction and Building Materials*, *51*, 414–423. https://doi.org/10.1016/j.conbuildmat.2013.11.004

Quintero, N. M. (2007). *Validation of the Enhanced Integrated Climatic Model (EICM) for the Ohio SHRP Test Road at U.S. 23*. Ohio University.

Reed, J. (2010). *Evaluation of the Effects of Aging on Asphalt Rubber Pavements*. Arizona State University.

Richardson, D. N., & Lusher, S. M. (2008). Determination of Creep Compliance and Tensile Strength of Hot-Mix Asphalt for Wearing Courses in Missouri. *Missouri Department of Transportation*, (April).

Rodezno, M. C., West, R., & Taylor, A. (2015). Flow Number Test and Assessment of AASHTO TP 79-13 Rutting Criteria. *Transportation Research Record: Journal of the Transportation Research Board*, *2507*, 100–107. https://doi.org/10.3141/2507-11

Roque, R., Zhang, Z., & Sankar, B. (1999). Determination of Crack Growth Rate Parameters Of Asphalt Mixtures Using the Superpave IDT. *Annual Meeting of the Association of Asphalr Paveing Technologists.*, (352), 404–420.

Rushing, J. F., Little, D. N., & Garg, N. (2014). Selecting a rutting performance test for the airport asphalt mixture design. *Road Materials and Pavement Design*, *15*(July 2014), 1–23. https://doi.org/10.1080/14680629.2014.926626

Saadeh, S., & Eljairi, O. (2011). *Development of a Quality Control Test Procedure for Characterizing Fracture Properties of Asphalt Mixtures*.

Scott III, S., Konrath, L., Ferragut, T., Anderson, S., Damnjanovic, I., Huber, G., … Schaeffer, V. R. (2014). *Performance Specifications for Rapid Highway Renewal*. Washington, D.C.: Transportation Research Board. https://doi.org/10.17226/22560

Shahadan, Z., Hamzah, M. O., Shukri Yahya, A., & Jamshidi, A. (2013). Evaluation of the dynamic modulus of asphalt mixture incorporating reclaimed asphalt pavement. *Indian Journal of Engineering and Materials Sciences*, *20*(5), 376–384.

Shu, X., Huang, B., & Vukosavljevic, D. (2008). Laboratory evaluation of fatigue characteristics of recycled asphalt mixture. *Construction and Building Materials*, *22*(7), 1323–1330. https://doi.org/10.1016/j.conbuildmat.2007.04.019

Sondag, M. S., Bruce A. Chadbourn, A., & Drescher, A. (2002). Investigation Of Recycled Asphalt Pavement (RAP) Mixtures, Report MN/RC-2002-15,.

Sousa, J. B. ., Vorobiev, A., Rowe, G. M., & Ishai, I. (2012). REACTED AND ACTIVATED RUBBER – AN ELASTOMERIC ASPHALT EXTENDER, *2012*.

Swamy, A. K., Mitchell, L. F., Hall, S. J., & Daniel, J. S. (2011). Impact of RAP on Volumetrics, Stiffness, Strength, and Low-Temperature Properties of HMA. *Journal of Materials in Civil Engineering*, *23*, 1490–1497. https://doi.org/10.1061/(ASCE)MT.1943-5533

Tapsoba, N., Baaj, H., Sauzéat, C., Di Benedetto, H., & Ech, M. (2016). 3D Analysis and Modelling of Thermal Stress Restrained Specimen Test (TSRST) on Asphalt Mixes with RAP and Roofing Shingles. *Construction and Building Materials*, *120*, 393–402. https://doi.org/10.1016/j.conbuildmat.2016.05.092

TRB. (2004). Quality Characteristics and Test Methods for Use in Performance-Related Specifications of Hot Mix Asphalt Pavements. *NCHRP Research Results Digest*, *291*. https://doi.org/10.17226/23356

TRB Superpave Committee. (2005). *Superior Performing Asphalt Pavement SUPERPAVE Performance by Design*.

TxDOT. (2016a). Design of Bituminous Mixtures.

TxDOT. (2016b). Test Procedure for Design of Bituminous Mixtures. TxDOT Designation: Tex-204-F.

UDOT. (2012). Standard Specifications for Road and Bridge Construction. Salt Lake City, Utah.

UDOT. (2015). GUIDELINES FOR SUPERPAVE VOLUMETRIC MIX DESIGN AND VERIFICATION, (November), 1–14.

Underwood, S., Baek, C., & Kim, R. (2012). Simplified Viscoelastic Continuum Damage Model as Platform for Asphalt Concrete Fatigue Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, *2296*, 36–45. https://doi.org/10.3141/2296-04

Vahidi, S., Mogawer, W. S., & Booshehrian, A. (2014). Effects of GTR and Treated GTR on Asphalt Binder and High-RAP Mixtures Materials, *26*(April), 721–727. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000858.

Vavrik, W. R., Gillen, S., Haugen, D., Behnke, J., Garrott, F., & Carpenter, S. H. (2010). Evaluation of HMA Modified with Recycled Asphalt Shingles (RAS) Mixtures, (October).

VDOT. (2013). DIVISION II – MATERIALS SPECIAL PROVISION COPIED NOTES (SPCNs), SPECIAL PROVISION (SPs) and SUPPLEMENTAL SPECIFICATIONS (SSs), 9–28.

Wagoner, M., Buttlar, W., Paulino, G., & Blankenship, P. (2005). Investigation of the Fracture Resistance of Hot-Mix Asphalt Concrete Using a Disk-Shaped Compact Tension Test. *Transportation Research Record: Journal of the Transportation Research Board*, *1929*, 183–192. https://doi.org/10.3141/1929-22

Wagoner, M. P., Buttlar, W. G., & Paulino, G. H. (2005a). Development of a Single-Edge Notched Beam Test for Asphalt Concrete Mixtures. *Journal of Testing and Evaluation*, *33*(6), 12579. https://doi.org/10.1520/JTE12579

Wagoner, M. P., Buttlar, W. G., & Paulino, G. H. (2005b). Disk-shaped Compact Tension Test for Asphalt Concrete Fracture. *Experimental Mechanics*, *45*(3), 270–277. https://doi.org/10.1177/0014485105053205

Walubita, L. F., Jamison, B. P., Das, G., Scullion, T., Martin, A. E., Rand, D., & Mikhail, M. (2011). Search for a laboratory test to evaluate crack resistance of hot-mix asphalt. *Transportation research record*, 2210(1), 73-80.

Walubita, L., Zhang, J., Das, G., Hu, X., Mushota, C., Alvarez, A., & Scullion, T. (2012). Hot-Mix Asphalt Permanent Deformation Evaluated by Hamburg Wheel Tracking, Dynamic Modulus, and Repeated Load Tests. *Transportation Research Record: Journal of the Transportation Research Board*, *2296*, 46–56. https://doi.org/10.3141/2296-05

Wang, Y., Mahboub, K. C., & Hancher, D. E. (2005). Survival Analysis of Fatigue Cracking for Flexible Pavements Based on Long-Term Pavement Performance Data. *Journal of Transportation Engineering*, *131*(August), 608–616. https://doi.org/10.1061/(ASCE)0733-947X(2005)131:8(608)

Weise, C., & Zeissler, A. (2016). Effects of fiber reinforcement on the fatigue and rutting performance of asphalt mixes, (June).

West, R. C., Tran, N. H., Taylor, A. J., & Willis, R. J. (2016). Comparison of Laboratory Cracking Test Results with Field Performance of Moderate and High RAP Content Surface Mixtures on the NCAT Test Track. In F. Canestrari & M. N. Partl (Eds.), *8th RILEM International Symposium on Testing and Characterization of Sustainable and Innovative Bituminous Materials* (pp. 979–991). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-7342-3_78

West, R. C., Winkle, C. V., Maghsoodloo, S., & Dixon, S. (2017). Relationships between simple asphalt mixture cracking tests using Ndesign specimens and fatigue cracking at FHWA's accelerated loading facility. *Road Materials and Pavement Design*, 18(sup4), 428-446.

West, R., Michael, J., Turochy, R., & Maghsoodloo, S. (2011). Use of Data from Specific Pavement Studies Experiment 5 in the Long-Term Pavement Performance Program to Compare Virgin and Recycled Asphalt Pavements. *Transportation Research Record: Journal of the Transportation Research Board*, *2208*(2011 Asphalt Materials and Mixtures 2011, Vol. 2), 82–89. https://doi.org/10.3141/2208-11

West, R., Willis, J. R., & Marasteanu, M. (2013). *Improved Mix Design, Evaluation, and Materials Management Practices for Hot Mix Asphalt with High Reclaimed Asphalt Pavement Content*. Washington, D.C. https://doi.org/10.17226/22554

Williams, R. C., Hill, D., Hofmann, K., Zelenock, M., & Bausano, J. (2004). *Development of Laboratory Performance Test Procedures and Trial Specifications for Hot Mix Asphalt : Final Report* (Vol. RC-1410). Lansing, MI.

Willis, R. J. (2013a). Use of Ground Tire Rubber in a Dense-Graded Asphalt Mixture on US 231 in Alabama: A Case Study. In *Airfield and Highway Pavement 2013*. https://doi.org/10.1061/9780784413005.100

Willis, R. J. (2013b). Use of Ground Tire Rubber in a Dense-Graded Asphalt Mixture on US 231 in Alabama: A Case Study. In *Airfield and Highway Pavement 2013*. https://doi.org/10.1061/9780784413005.100

WisDOT. (2015). Notice to All Contractors: Proposal #19. Madison, WI.

Witczak, M., Kaloush, K., Pellinen, T., El-Basyouny, M., & Von Quintus, H. (2002). *Simple Performance Test for Superpave Mix Design*. Washington, D.C.

WSDOT. (2017). *Materials Manual*.

Wu, Z., Mohammad, L., Wang, L., & Mull, M. (2005). Fracture Resistance Characterization of Superpave Mixtures Using the Semi-Circular Bending Test. *Journal of ASTM International*, *2*(3), 1–15. https://doi.org/10.1520/JAI12264

Xiao, F., Wenbin Zhao, P. E., & Amirkhanian, S. N. (2009). Fatigue behavior of rubberized asphalt concrete mixtures containing warm asphalt additives. *Construction and Building Materials*, *23*(10), 3144–3151. https://doi.org/10.1016/j.conbuildmat.2009.06.036

Xie, Z., & Shen, J. (2016). Performance properties of rubberized stone matrix asphalt mixtures produced through different processes. *Construction and Building Materials*, *104*, 230–234. https://doi.org/10.1016/j.conbuildmat.2015.12.063

Yang, S., Braham, A., Underwood, S., Hanz, A., & Reinke, G. (2016). Correlating field performance to laboratory dynamic modulus from indirect tension and torsion bar. *Asphalt*

Paving Technology: Association of Asphalt Paving Technologists-Proceedings of the Technical Sessions, *85*(0), 131–162. https://doi.org/10.1080/14680629.2015.1267438

Yao, H., & You, Z. (2016). Effectiveness of Micro-and Nanomaterials in Asphalt Mixtures through Dynamic Modulus and Rutting Tests. *Journal of Nanomaterials*, *2016*. https://doi.org/10.1155/2016/2645250

Yin, F., Garita, J., Taylor, A., & West, R. (2018). Refining the indirect tensile (IDT) Nflex factor test to evaluate cracking resistance of asphalt mixtures for mix design and quality assurance. *Construction and Building Materials*, 172, 396-405.

Yin, F., Arámbula-Mercado, E., Epps Martin, A., Newcomb, D., & Tran, N. (2017). Long-term ageing of asphalt mixtures. *Road Materials and Pavement Design*, *18*(December), 2–27. https://doi.org/10.1080/14680629.2016.1266739

Yin, F., Epps Martin, A., Arámbula-Mercado, E., & Newcomb, D. (2017). Characterization of non-uniform field aging in asphalt pavements. *Construction and Building Materials*, *153*, 607–615. https://doi.org/10.1016/j.conbuildmat.2017.07.144

Yin, F., Martin, A. E., Arambula, E., & Newcomb, D. E. (2016). Short-term aging of asphalt mixtures. *Asphalt Paving Technology*, *2015*, 333.

Yousefi Rad, F., Elwardany, M. D., Castorena, C., & Kim, Y. R. (2017). Investigation of proper long-term laboratory aging temperature for performance testing of asphalt concrete. *Construction and Building Materials*, *147*, 616–629. https://doi.org/10.1016/j.conbuildmat.2017.04.197

Zegeye, E. T., Moon, K. H., Turos, M., Clyne, T. R., & Marasteanu, M. O. (2012). Low temperature fracture properties of polyphosphoric acid modified asphalt mixtures. *Journal of materials in civil engineering*, 24(8), 1089-1096.

Zborowski, A., & Kaloush, K. E. (2011). A fracture energy approach to model the thermal cracking performance of asphalt rubber mixtures. *Road materials and pavement design*, 12(2), 377-395.

Zapata, C. E., & Houston, W. N. (2008). *NCHRP Report 602: Calibration and Validation of the Enhanced Integrated Climatic Model for Pavement Design*. NCHRP. https://doi.org/10.17226/23098

Zhou, F. (2018). IDEAL Cracking Test for QC / QA and Associated Criteria.

Zhou, F., Hu, S., & Scullion, T. (2006). *Integrated Asphalt (Overlay) Mixture Design, Balancing Rutting and Cracking Requirements*. *FHWA Report* (Vol. 7).

Zhou, F., Hu, S., & Scullion, T. (2007). DEVELOPMENT AND VERIFICATION OF THE OVERLAY TESTER BASED FATIGUE CRACKING PREDICTION APPROACH, *7*(2).

Zhou, F., Im, S., Sun, L., & Scullion, T. (2017). Development of an IDEAL cracking test for asphalt mix design and QC/QA. *Journal of the Association of Asphalt Paving Technologists (AAPT)*, *86*, 1–23. https://doi.org/10.1080/14680629.2017.1389082

Zhou, F., Newcomb, D., Gurganus, C., Banihashemrad, S., Sakhaeifar, M., Park, E. S., & Lytton, R. L. (2016). Field Validation of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures: An Experimental Design. *NCHRP Research Results Digest*, *399*. https://doi.org/10.17226/23608

Zhou, F., & Scullion, T. (2003). UPGRADED OVERLAY TESTER AND ITS APPLICATION TO CHARACTERIZE REFLECTION CRACKING RESISTANCE OF ASPHALT MIXTURES, *7*(2).

Zhou, F., & Scullion, T. (2005). Overlay tester: a rapid performance related crack resistance test, *7*(2), FHWA/TX-04/0-4467-2. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.6226&rep=rep1&type=pdf

Zhou, F., Scullion, T., & Hu, S. (2007). DEVELOPMENT AND VERIFICATION OF THE OVERLAY TESTER BASED FATIGUE CRACKING PREDICTION APPROACH, *7*(2).

Zhou, F., Scullion, T., Walubita, L., & Wilson, B. (2014). Implementation of a Performance-Based Mix Design System in Texas. *Transportion Research Circular*, *E-C189*(October), 32–50.

APPENDIX B

BINDER EXTRACTION AND RECOVERY

The intent of recovered binder testing is to evaluate the properties of the resultant binder which accounts for both virgin and recycled components. Traditionally only the virgin asphalt binder properties are known and assumptions are made regarding the recycled binders. However, the binder properties of recycled products can vary widely depending upon the original application and aging condition. By knowing the resultant binder properties (instead of relying on assumptions), the overall binder properties can be monitored regardless of how much of the binder is virgin or recycled. By monitoring/ controlling the resultant binder properties agencies can continue to push the limits on increasing recycle contents without detrimentally affecting the binder properties. Binder testing is by no means intended to replace mixture testing but rather supplement it. Ultimately mix performance is the end result agencies are concerned with; however, the binder properties play a key role in mixture performance. Thus, monitoring/controlling resultant binder properties is key to ensuring a cost effective mixture that still meets the design intent.

This appendix provides binder testing results obtained after extraction and recovery of binder from plant-produced mixtures sampled during production in 2018, and field cores, investigated later in the project. These results are then correlated to mixture testing results, although it is acknowledged that the resulting comparisons are not completely 'apples-to-apples.' For instance, binder and mixture properties associated with cracking evaluations are at different aging levels. In addition, it was difficult to accurately assess the binder properties stemming from the mixtures containing ground tire rubber, as a significant portion of rubber was lost during the extraction and recovery process. Those shortcomings notwithstanding, the binder results obtained and correlations with mixture properties provide additional data that can be referred to as these experimental sections are monitored during service life. The binder extraction and recovery tests were performed by S.T.A.T.E. Testing, LLC on both plant-produced mixtures (Chapter 3) and on field cores (Chapter 5).

The detrimental effects of hardening in asphalt pavements are well recognized by pavement engineers. This hardening process, known as asphalt aging, is characterized based on the rheological properties of asphalt binders and/or mixtures. As a result of aging, changes occur in chemical composition during pavement construction and service life. Aging stiffens and embrittles the asphalt binder and leads to durability issues such as the high potential for cracking. Generally, three aging levels are considered for asphalt binders, including neat (or tank), short term, and long term. Two laboratory aging procedures are routinely used for Superpave binder testing - the rolling thin film oven (RTFO) and pressure aging vessel (PAV) for short term and long-term aging, respectively.

The Rolling Thin-Film Oven (RTFO) procedure simulates short term aging experienced during mixing and compaction in the asphalt plant and field. The RTFO also provides a quantitative measure of the volatiles lost during the aging process. Neat (unaged) binders are short-term aged in the RTFO apparatus. To this end, they are exposed to a high temperature (163 ℃) and constant exposure to moving air as a carrousel of bottles is rotated at 15 rounds per minute (rpm) for 85 minutes in the RTFO device per AASHTO T 240 and ASTM D 2872. Samples are then stored for use in physical properties tests or the PAV. Per AASHTO R 28 standard, the RTFO aged binders are placed in stainless steel pans and then in a heated vessel pressurized to 2.10 MPa (305 psi) for 20 hours to simulate in-service aging (thought to represent around 7 to 10 years of field aging).

To evaluate the actual binder system present in plant produced, modern, heterogeneous recycled asphalt mixtures, the ASTM D8159 and ASTM D5404 procedures can be used to extract and recover binders from mixture samples. Successful extraction and recovery of polymer modified mixtures has been reported in the literature. Difficulties associated with the extraction and recovery of crumb rubber modified binders has also been reported in the literature (Mturi et al. 2014; Ma et al. 2016). For instance, Heitzman (Heitzman, 1992) reported inaccurate binder and rubber content measurements after extraction and recovery.

Table B-1 introduces the properties of 2018 plant-produced mixtures. The last five columns of Table B-1 indicate the testing results of the recovered binder from the plant-produced mixtures. It is assumed that the recovered binder has gone through short-term aging. Thus, the rutting criterion of $G^*/\sin\delta$ of 2.2 kPa minimum was applied to obtaine the continuous high temperature grade. Testing at multiple high, low and intermediate temperatures was needed to establish the continuous binder grade. For instance, a maximum creep stiffness of 300 MPa and a minimum m-value of 0.300 were used to establish the continuous extracted binder PG at low temperature (PGLT). In addition to the performance grade testing results, the multiple stress creep and recovery (MSCR) test was performed per ASTM D7415 at different temperatures, with results presented in Table B-2. As noted earlier, results from binder testing of recovered material from rubber-modified binders (e.g. sections 1835, 1845, and 1840) should be viewed with an understanding that the rubber particles are not all completely present in the recovered sample.

Referring to Table B-1, the following observations were made:

- In general, the recovered binder testing results were in excellent agreement with the plan PG grade. This suggests that the heterogeneous, high recycling content mixes used by the Tollway are meeting or exceeding traditional plan binder PG requirements (but in a highly sustainable fashion).
- The effect of not capturing all of the rubber particles in the recovered binder samples from ECR (dry-process) and GTR (terminal-blended) mixes containing crumb rubber were two-fold: (1) as expected, the PGHT continuous grades were found to be lower than the plan PG grade; and; (2) the ΔT_c values obtained were quite low in some cases, most likely due to the disturbance of the binder sample.

Table B-2 presents the MSCR testing results obtained at one, or two testing temperatures (when available). In general, it was found that the recovered binders from the highly recycled Tollway mixtures did a good job at meeting or exceeding the high temperature MSCR binder requirements according to the plan PGHT for the mixtures, and expected traffic levels. Figure B-1 shows a reasonably good correlation between the recovered binder properties at high temperatures and the Hamburg wheel tracking test results.

Table B-1. Mixture and binder properties

Mix. ID	Base Binder	Plan Grade	ABR by RAP	ABR by RAS	Total ABR	Extracted PG	Extracted Cont. PG	Failing Temp. Stiff.	Failing Temp. m-value	Delta Tc
1844	SBS 70-28	76-22	10.8	16.0	26.8	88-16	88.2-21.7	-30.4	-21.7	-8.7
1835	46-34 +10%ECR	76-22	25.1	16.1	41.2	70-22	75.3*-22.5	-33.9	-22.5	-11.4
1824	SBS 64-34	76-22	20.4	16.7	37.1	82-28	85.6-29.8	-33	-29.8	-3.2
1845	46-34 +10.5%Lehigh	76-22	23.9	15.4	39.3	70-28	75*-29.1	-33.2	-29.1	-4.1
1836	SBS 64-34	76-22	16.2	16.3	32.5	82-28	85.7-30.5	-33.7	-30.5	-3.2
1840	58-28 +12%GTR	76-22	15.9	9.8	25.7	70-28	72.2*-29	-32.3	-29	-3.3
1829	58-28 +12%GTR	76-22	17.8	9.3	27.0	70-28	74*-29.7	-33.4	-29.7	-3.7
1828	46-34 +10%ECR	76-22	35.3	9.2	44.6	70-22	71.4*-24.6	-37.6	-24.6	-13
1823	SBS 64-34	76-22	24.1	14.2	38.3	82-28	84.6-29	-33	-29	-4
1818	64-22	64-22	20.4	0.0	20.4	70-22	72.9-23.8	-26.4	-23.8	-2.6
1834	58-28	64-22	20.0	0.0	20.0	64-28	67.5-28.8	-31.7	-28.8	-2.9
1826	46-34	64-22	27.6	18.1	45.7	---	---	---	---	---
1807	46-34	64-22	34.4	14.0	48.4	76-16	76.2-16	-31.9	-16	-15.9
1803	58-28	64-22	26.5	16.6	43.1	70-22	71.2-25.6	-29.4	-25.6	-3.8

*Rubber-modified mixtures. It was not possible to recover all rubber particles in the extraction and recovery process.

Table B-2. MSCR testing results. Results from two testing temperatures are provided, where available.

Mix. ID	Base Binder	Extracted Cont. PG	Temp °C	J_{nr}	$J_{nr,diff}$	$J_{nr,slope}$	Temp °C	J_{nr}	$J_{nr,diff}$	$J_{nr,slope}$
1844	SBS 70-28	88.2-21.7	***---***	---	---	---	***---***	---	---	---
1835	46-34 +10%ECR	75.3-22.5	***---***	---	---	---	***---***	---	---	---
1824	SBS 64-34	85.6-29.8	***---***	---	---	---	***82***	1.0	89	15
1845	46-34 +10.5%Lehigh	75-29.1	***---***	---	---	---	***---***	---	---	---
1836	SBS 64-34	85.7-30.5	***76***	0.4	51	4	***82***	0.8	78	12
1840	58-28 +12%GTR	72.2-29	***---***	---	---	---	***70***	2.2	57	25
1829	58-28 +12%GTR	74-29.7	***76***	6.1	65	77	***70***	2.1	66	27
1828	46-34 +10%ECR	71.4-24.6	***76***	8.3	44	82	***70***	3.0	48	31
1823	SBS 64-34	84.6-29	***76***	0.7	63	9	***82***	1.9	89	29
1818	64-22	72.9-23.8	***76***	7.3	22	43	***70***	2.9	14	12
1834	58-28	67.5-28.8	***76***	13.1	22	77	***64***	2.4	19	12
1826	46-34	---	***---***	---	---	---	***---***	---	---	---
1807	46-34	76.2-16	***76***	4.3	33	34	***---***	---	---	---
1803	58-28	71.2-25.6	***76***	8.9	17	42	***70***	3.7	17	17

Figures B-1 to B-3 present the performance test results for the investigated mixtures at different temperatures along with extracted binder properties. As seen in Figure B-1, the DC(T) fracture energy test results do not show a strong correlation with the extracted binder PGLT. The major takeaway from this finding is that in nearly all cases, a suitable blended (resultant) binder system was present in the Tollway mixtures studied where low temperatures are concerned. The wide differences in mixture fracture energy can be explained by the differences in mix types studied, and the corresponding breath in mix volumetrics, aggregate types, and aggregate gradation present across the mixes. The data point that stands out is mixture 1807. In this case, it appears that the extracted binder PGLT (-16.0 ºC) provided a good indication of low-temperature cracking susceptibility, which could not be overcome by the toughnening characteristics imparted by the aggregate and other non-bituminous components. The overall data set highlights the importance of both binder and mixture testing when conducting a comprehensive investigation of new mix types, such as high binder replacement mixes used in high traffic applications. Binder testing presents a good first-order screening tool for new and recycled material combinatinos, while mixture testing provides confidence in the expected performance of the final mixture system.

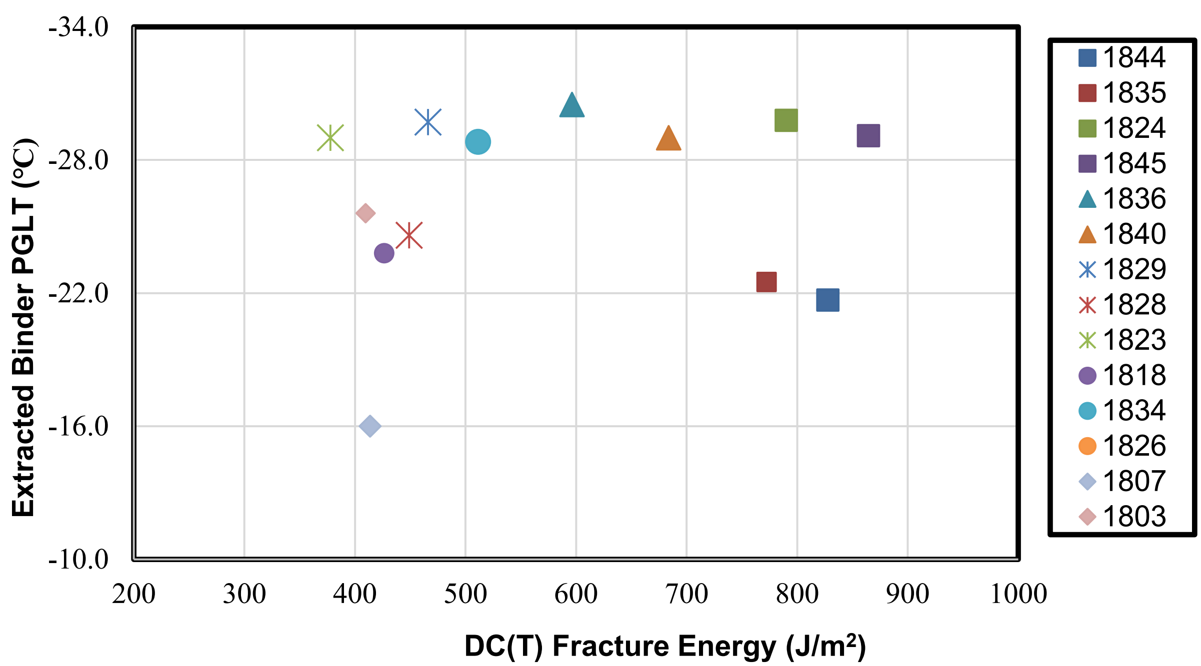


Figure B-1. Comparison between the DC(T) fracture energy and extracted binder PGLT

Figure B-2 compares the DC(T) fracture energy and the ***Δ***Tc parameter obtained from testing of the recovered binders from the investigated mixtures. As shown, similar to the DC(T) and extracted binder PGLT, a strong dependency between the DC(T) fracture energy and the extracted binder ***Δ***Tc was not observed. Hamburg rut depths at the required passes showed a stronger relationship with the extracted PGHT results (Figure B-3). Three general trends were observed: (1) a group of 4 mixtures (1823, 1824, 1836, 1844) with the lowest Hamburg rut depths (under 2.5 mm) had the highest PGLT results (by far, around 85 ºC and higher)); (2) a group of 7 mixes (1835, 1845, 1840, 1829, 1818, 1807, 1803) had slightly higher rut depths (+/-

3 mm) and more intermediate PGHT grades (between 71 and 76 °C), and; (3) two mixes (1834 and 1828) had the highest Hamburg rut depths by far and had among the lowest PGHT grades.

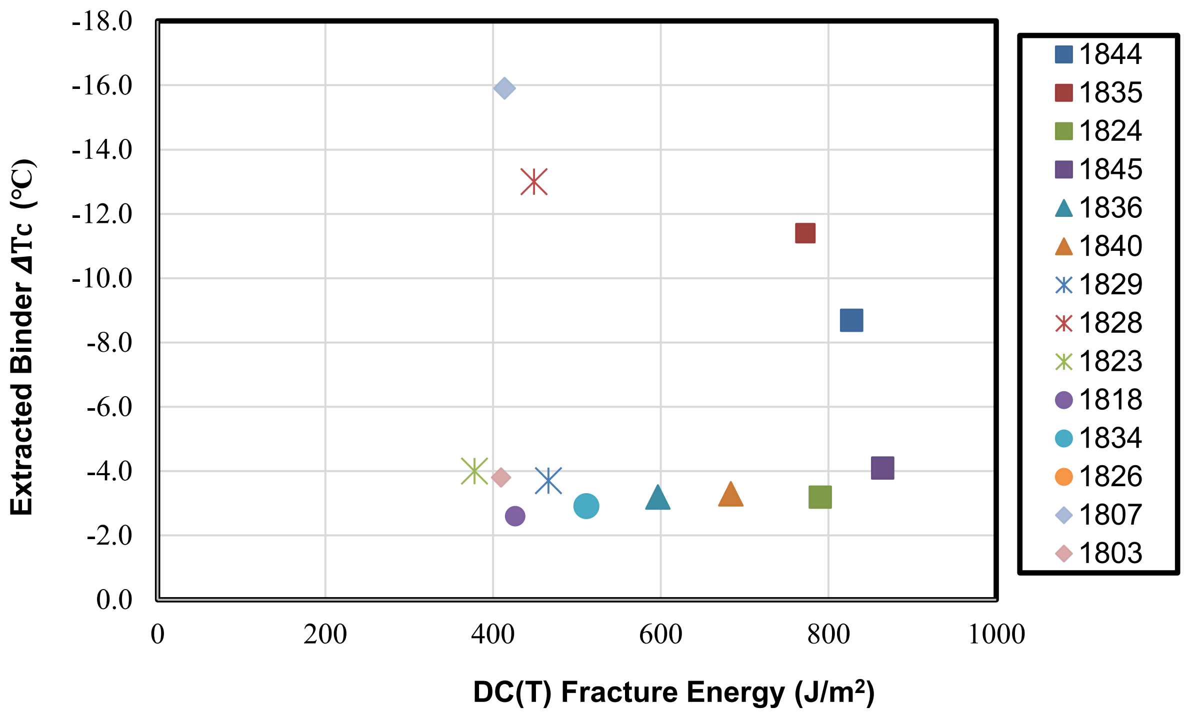


Figure B-2. Comparison between DC(T) fracture energy and extracted binder PGLT

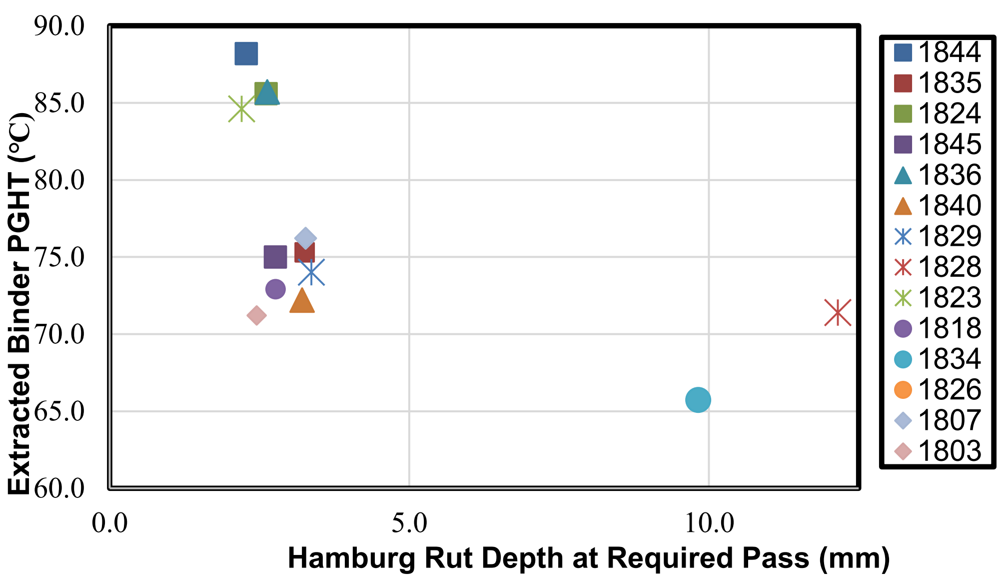


Figure B-3. Comparison between Hamburg rut depth and extracted binder PGHT

The second set of binder extraction and recovery tests were performed on the field cores obtained from the shoulder mixtures. The service life for the last four sections was at least 10 years. Therefore, PAV aging was not performed on the binder recovered from these sections. On the other hand, the first and second sections (I88-52 and I88-57) were resurfaced in 2015 and 2014, respectively. As the service life of these sections was less than 7 years, in addition to testing of the directly recovered binder, PAV-aged samples were also produced and tested. In addition to the continuous PGLT and PGHT, the ***Δ***Tc parameter was calculated and reported in the last column of the Table B-1. It is noted that the only section with a positive ***Δ***Tc parameter is I90-10E. In this case, the stiffness was the controlling parameter in this sample, not the m-value.

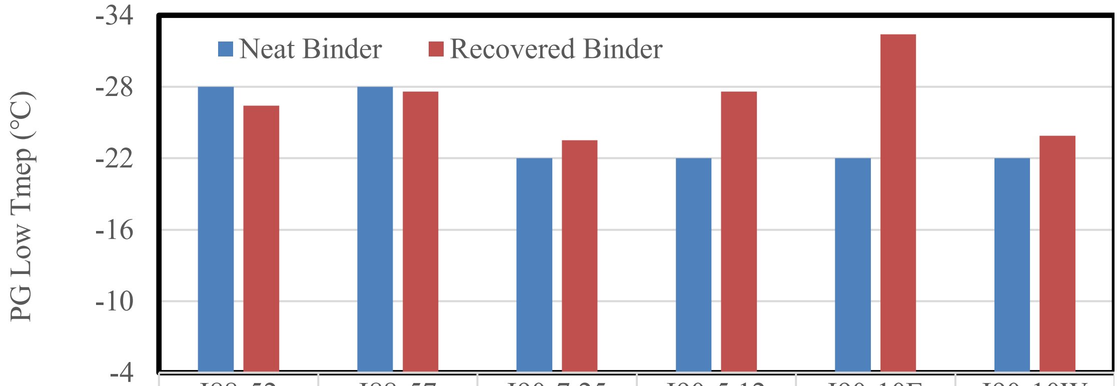
Figure B-4 compares the PGLT of the neat binder with that from the extracted binder without any further application of the aging procedure in the lab. Although the continuous grade of the neat binder was not available, it could be seen that there is generally a difference between the PGLT of the neat binder and the extracted and recovered binder. The PGLT of the recovered binder from section I88-52, which is stiffened by 38.6% ABR, is slightly higher than that of the neat binder (-26.4 vs. -28.0 °C). As shown in Table B-3, this material was produced in 2015 and has a service life of only 4 years prior to extraction and recovery, which falls short of the 7 year threshold typically assumed for long-term aging. Therefore, comparing the PGLT of the PAV-aged recovered binder (now -28 compared to -18.3°C), the effects of aging fall more in line with expectations (although the sample may have been slightly over-aged, beyond the PAV level at this point). Similar results were found for the I88-57 section.

The PGLT of the extracted binder from the I90-10E section was found to be one grade softer than that of the neat binder. The same comparison could be made for the I90-5.12 and I90-10W sections, although the difference between the extracted and neat binder PGLT was not as considerable. In general, the low temperature grades obtained for the recovered binders suggest that the binders did not undergo excessive aging rates in the field. The fact that some of the older sections were found to have thermal and block cracking in the field may be due to the higher, near-surface aging levels present in the pavements, which cannot be easily captured when extracting and recovering binder from field cores having surface layer thicknesses in excess of 37.5 mm (1.5 inches). In addition, unlike the DC(T) mixture test, the binder extraction and recovery process only provides insight toward the rheological properties of the binder, which is only one component of the mixture (does not consider aggregate or other mixture-level, compositional effects).

Figures B-5 and B-6 evaluate low temperature recovered binder properties, such as PGLT and ***Δ***Tc, alongside DC(T) fracture energy. All three data sets point to the potential vulnerability of these older recycled mix designs to thermal and block cracking. A higher degree of thermal and block cracking resistance in the Chicago area generally requires PGLT values of -28 ºC or below, and DC(T) fracture energy values above 500 J/m^2. For block cracking resistance, positive ***Δ***Tc values are preferred. Thus, both the binder and mixture low temperature results are in line with the field observations, which showed that these shoulder sections generally developed thermal and block cracking distress as service time progressed. The new DC(T) thresholds presented in this study should serve to improve the performance of future Tollway asphalt shoulders.

Table B-3. Recovered binder properties from field cores for shoulder mixtures

Sample Type	No.	Location	Year	Mix. Type	Base Binder	ABR by RAP	ABR by RAS	Lab Aging after Binder Recovery by STATE Testing	Resultant High PG	Resultant Low PG	AASHTO M320 Grade	ΔTc
Field Core	**1**	**I88-52**	2015	N70D Surface	58-28	19.1	19.6	**None**	**81.9**	**-26.4**	**76-22**	**-3.5**
								PAV		**-18.3**	**76-16**	**-10.8**
	2	**I88-57**	2014	N70D Surface	58-28	22.8	17.8	**None**	**80.4**	**-27.6**	**76-22**	**-1.5**
								PAV		**-21.0**	**76-16**	**-6.1**
	3	**I90-7.25**	2009	N70D Surface	58-22	16.7	20.1	**None**	**83.6**	**-23.5**	**82-22**	**-2.9**
								PAV		**-**	**-**	
	4	**I90-5.12**	2009	N70D Surface	58-22	24.4	0	**None**	**75.6**	**-27.6**	**70-22**	**-1.5**
								PAV		**-**	**-**	
	5	**I90-10E**	2008	N70D Surface	58-22	24	0	**None**	**67.2**	**-32.4**	**64-28**	**1.1**
								PAV		**-**	**-**	
	6	**I90-10W**	2008	N70D Surface	58-22	16.2	0	**None**	**76.4**	**-23.9**	**76-22**	**-4**
								PAV		**-**	**-**	



	I88-52	I88-57	I90-7.25	I90-5.12	I90-10E	I90-10W
Neat Binder	-28	-28	-22	-22	-22	-22
Recovered Binder	-26.4	-27.6	-23.5	-27.6	-32.4	-23.9
ABR by RAP	19.1	22.8	16.7	24.4	24	16.2
ABR by RAS	19.6	17.8	20.1	0	0	0
Total ABR	38.6	40.7	36.8	24.4	24	16.2

Figure B-4. Comparing PGLT of neat and recovered binders (without further lab aging)

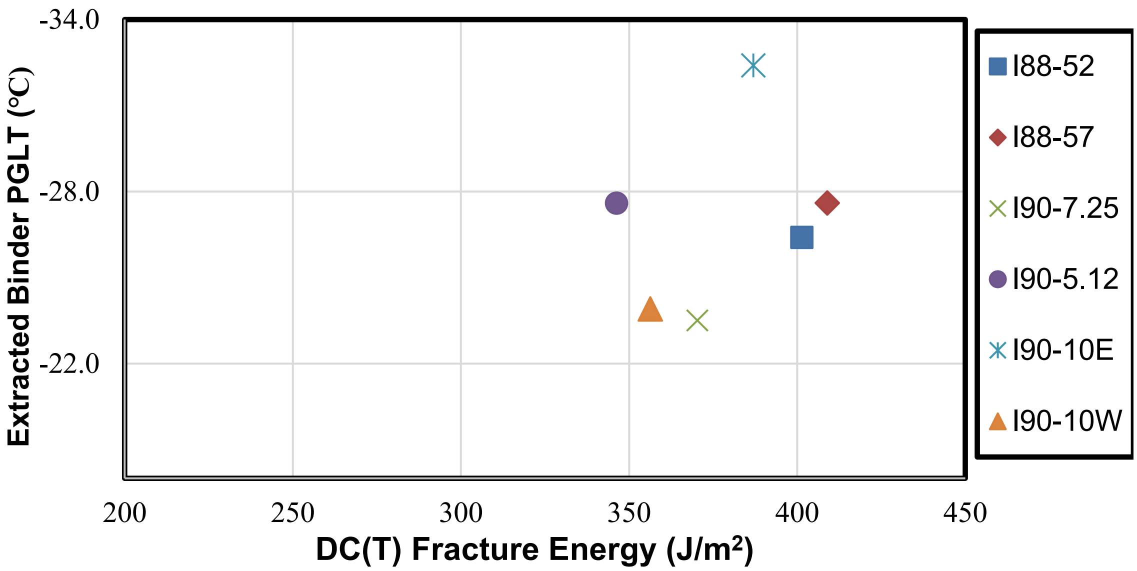


Figure B-5. Comparison between DC(T) fracture energy and extracted binder PGLT for field cores

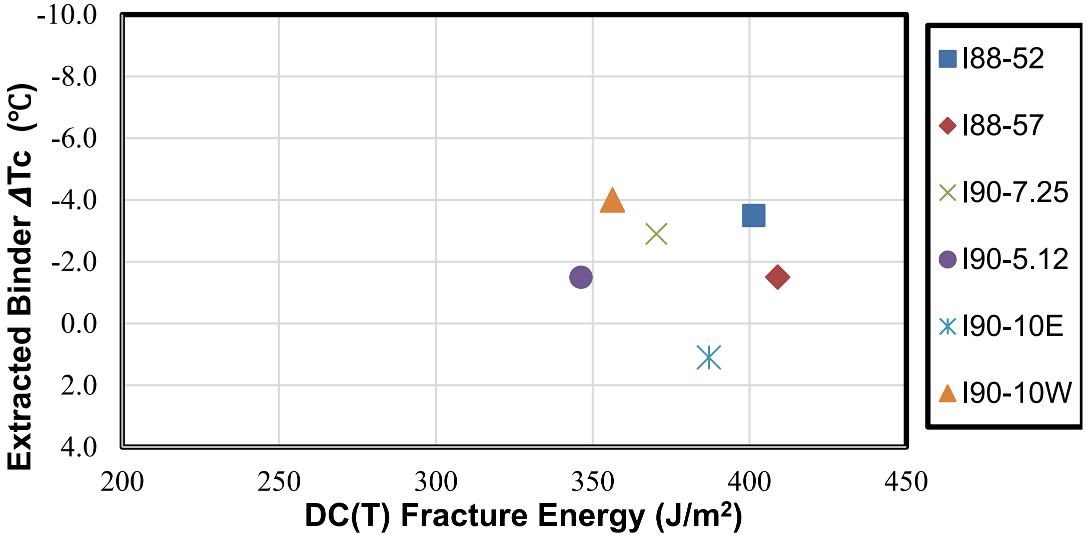


Figure B-6. Comparison between DC(T) fracture energy and extracted binder *Δ*Tc for field cores