DEVELOPMENT OF A BALANCED MIX DESIGN METHOD IN OREGON

Final Report

PROJECT SPR 801

Oregon Department of Transportation

DEVELOPMENT OF A BALANCED MIX DESIGN METHOD IN OREGON

Final Report

SPR 801

By

Erdem Coleri, PhD
Shashwath Sreedhar
Ihsan Ali Obaid

School of Civil and Construction Engineering
Oregon State University
101 Kearney Hall
Corvallis, OR 97331
Phone: 541-737-0944

for

Oregon Department of Transportation
Research Section
555 13th Street NE, Suite 1
Salem OR 97301

and

Federal Highway Administration
1200 New Jersey Avenue SE
Washington, DC 20590

November 2020

Technical Report Documentation Page

1. Report No. FHWA-OR-RD-21-03	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF A BALANCED MIX DESIGN METHOD IN OREGON		5. Report Date November 2020 6. Performing Organization Code
7. Author(s) Erdem Coleri, PhD; 0000-0002-1904-878X Shashwath Sreedhar; 0000-0003-1786-6372 Ihsan Ali Obaid; 0000-0001-9048-8342		8. Performing Organization Report No.
9. Performing Organization Name and Address School of Civil and Construction Engineering Oregon State University 101 Kearney Hall, Corvallis, OR 97331		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Oregon Dept. of Transportation Research Section 555 13th Street NE, Suite 1 Salem, OR 97301	Federal Highway Admin. 1200 New Jersey Avenue SE Washington, DC 20590	13. Type of Report and Period Covered Final Research Report 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract: According to recent surveys conducted with state Department of Transportations (DOTs) and asphalt contractors, commonly used asphalt mixture properties are not directly reflecting the long-term performance of asphalt mixtures. In addition, there are several new additives, polymers, rubbers, and high-quality binder types incorporated into asphalt mixtures today. Volumetric mixture design methods are not capable of capturing the benefits of using all these new technologies on asphalt mixture performance. Furthermore, the interaction of virgin binders with reclaimed asphalt pavement (RAP) mixtures with high binder replacement contents and the level of RAP binder blending into the asphalt mixture are still not well understood. Due to all these complications related to the more complex structure of today's asphalt mixtures, simple volumetric evaluations to determine the optimum binder content may not result in reliable asphalt mixture designs. Two volumetrically identical mixtures may provide completely different rutting and cracking performance according to laboratory tests. For all these reasons, in this study, a procedure to incorporate performance tests for rutting and cracking *(with a new long-term aging protocol developed in this study)* into current asphalt mixture design methods was developed to be able to validate or revise the optimum binder content determined by the volumetric mix design method. Developed balanced mix design method is expected to improve the long-term performance of asphalt-surfaced pavements in Oregon.

17. Key Words Balanced mix design; performance based specifications; asphalt mix design; flow number, Hamburg wheel tracking test; semi-circular bend test; asphalt aging.		18. Distribution Statement Copies available from NTIS, and online at www.oregon.gov/ODOT/TD/TP_RES/	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 117	22. Price

Technical Report Form DOT F 1700.7 (8-72) Reproduction of completed page authorized ♲ Printed on recycled paper

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS					APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	When You Know	Multiply By	To Find	Symbol	Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH					**LENGTH**		
in	inches	25.4	millimeters	mm	mm	millimeters	0.039	inches	in
ft	feet	0.305	meters	m	m	meters	3.28	feet	ft
yd	yards	0.914	meters	m	m	meters	1.09	yards	yd
mi	miles	1.61	kilometers	km	km	kilometers	0.621	miles	mi
		AREA					**AREA**		
in^2	square inches	645.2	millimeters squared	mm^2	mm^2	millimeters squared	0.0016	square inches	in^2
ft^2	square feet	0.093	meters squared	m^2	m^2	meters squared	10.764	square feet	ft^2
yd^2	square yards	0.836	meters squared	m^2	m^2	meters squared	1.196	square yards	yd^2
ac	acres	0.405	hectares	ha	ha	hectares	2.47	acres	ac
mi^2	square miles	2.59	kilometers squared	km^2	km^2	kilometers squared	0.386	square miles	mi^2
		VOLUME					**VOLUME**		
fl oz	fluid ounces	29.57	milliliters	ml	ml	milliliters	0.034	fluid ounces	fl oz
gal	gallons	3.785	liters	L	L	liters	0.264	gallons	gal
ft^3	cubic feet	0.028	meters cubed	m^3	m^3	meters cubed	35.315	cubic feet	ft^3
yd^3	cubic yards	0.765	meters cubed	m^3	m^3	meters cubed	1.308	cubic yards	yd^3
~NOTE: Volumes greater than 1000 L shall be shown in m^3.									
		MASS					**MASS**		
oz	ounces	28.35	grams	g	g	grams	0.035	ounces	oz
lb	pounds	0.454	kilograms	kg	kg	kilograms	2.205	pounds	lb
T	short tons (2000 lb)	0.907	megagrams	Mg	Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)					**TEMPERATURE (exact)**		
°F	Fahrenheit	(F-32)/1.8	Celsius	°C	°C	Celsius	1.8C+32	Fahrenheit	°F

*SI is the symbol for the International System of Measurement

ACKNOWLEDGEMENTS

The authors would like to thank the Oregon Department of Transportation (ODOT) for providing funding for this research. The authors thank the members of the ODOT Project Technical Advisory Committee and ODOT research for their advice and assistance in the preparation of this report. In particular, Norris Shippen, Larry Ilg, Chris Duman, Anthony Boesen, and John Hickey participated on the TAC. The authors would also like to thank Mike Stennett of ODOT who organized material sampling. The authors would also like to thank all graduate and undergraduate research assistants for their help with sieving, batching, and measuring theoretical maximum specific gravity of prepared samples, as well as James Batti for his help in the laboratory.

DISCLAIMER

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

Cracking is a common failure mechanism in asphalt concrete pavement structures. It is one of the main reasons for large road maintenance and rehabilitation expenditures, as well as reduced user comfort and increased fuel consumption due to high road roughness. The resistance of the pavement to this distress mechanism is dependent upon the ductility of the asphalt pavement mixture. The increased use of recycled asphalt materials with high binder replacement rates results in a significant reduction in ductility of the asphalt mixtures used in construction, which causes a significant reduction in the fatigue life of the pavement in many cases. In Oregon, asphalt cracking is the major distress mode, necessitating costly rehabilitation and maintenance at intervals of less than half of the intended design lives in some cases. For this reason, it is necessary to accurately quantify the impact of increasing the recycled asphalt content on the structural cracking and rutting resistance of the pavement through use of low-cost and efficient testing and design procedures that can easily be implemented.

Asphalt mixtures are designed to be used in pavements to withstand vehicular loads under different climatic conditions. The goal of asphalt mix design is to determine an economic blend of aggregates and binder such that the resultant mix provides sufficient stability to resist deformation under traffic loading, and flexibility to withstand cracking. The most commonly known asphalt mix design methods are the Marshall, Hveem, and Superpave methods. Marshall and Hveem mix design procedures were widely used until the early 1990s before Superpave procedure was introduced. Superpave was developed as part of the Strategic Highway Research Program (SHRP) and was implemented in 1993. The original objective was to develop a performance-based mix design process. Although performance tests for asphalt mixtures were a part of the Superpave mix design process and several procedures were developed to predict and evaluate mixture performance, the entire process turned out to be too complex and costly and was never implemented by any state Department of Transportation (DOT).

Superpave mix design had three levels (Level 1, Level 2, and Level 3) with increasing complexity (Cominsky et al. 1994). The performance-based specifications were to be incorporated in Level 2 and Level 3 designs but were never implemented. The current asphalt mix design practice (Level 1) involves proportioning of the aggregates and the asphalt binder based on empirical properties of aggregates and volumetric properties such as densities, air voids, voids in the mineral aggregate (VMA) and voids filled with asphalt (VFA). However, most state DOTs and asphalt contractors do not think that commonly used asphalt mixture properties are directly reflecting the long-term performance of asphalt mixtures. For instance, although there are requirements for VMA set by almost all state DOTs, measurement of VMA relies on the accurate measurement of aggregate bulk specific gravity, while considerable issues were observed in terms of accuracy and variability during the measurement of this parameter (West et al. 2018). In addition, there are several new additives, polymers, rubbers, and high-quality binder types incorporated into asphalt mixtures today. Volumetric mixture design methods are not capable of capturing the benefits of using all these new technologies on asphalt mixture performance. Furthermore, the interaction of virgin binders with reclaimed asphalt pavement (RAP) mixtures with high binder replacement contents and the level of RAP binder blending into the asphalt mixture are still not well understood. Due to all these complications related to the more complex structure of asphalt mixtures, simple volumetric evaluations to determine the optimum binder content may not result in reliable asphalt mixture designs. Two volumetrically identical mixtures may provide completely different rutting and cracking performance according to laboratory tests (Coleri et al. 2017b). For all these reasons, performance tests for rutting and cracking need to be incorporated into current asphalt mixture design methods to be able to validate or revise the optimum binder content determined by the volumetric mix design method. Numerous research studies were recently carried out and are currently being conducted to develop new mix design processes with performance verification (Epps et al. 2002; Zhou et al. 2006; Harvey et al. 2014; Cooper III et al. 2014; Williams et al. 2004; Bennert et al. 2014; Hughes and Maupin 2000; Dave and Koktan 2011; Kim et al. 2011; Zhou et al. 2014). However, this approach is not entirely new and draws upon the existing methods and procedures while the existing methods need to be revised and improved by incorporating findings from recent research studies.

Oregon Department of Transportation (ODOT) Research Projects SPR785 and SPR797 (Coleri et al. 2017b; Coleri et al. 2017a; Sreedhar et al. 2018; Haddadi et al. 2019) constructed the beginnings of a performance-based balanced mix design method for Oregon. It was suggested that semi-circular bend (SCB) test is the most effective and practical cracking test that can effectively be used for balanced mix design. It was determined that the typical flexibility index (FI), an energy parameter calculated using SCB test results, values for production mixtures (plant-produced) with polymer-modified binder range from 9 to 14. However, more experiments need to be conducted to determine an exact threshold for FI that will provide acceptable long-term pavement cracking performance. In these two research projects, flow number (FN) test was used as the experiment for rutting performance evaluation. For highways with high traffic levels (ESALs > 30 million), an FN of 740 was suggested by AASHTO TP79-13 (2013) and used in SPR785 and SPR797 as the threshold value for rutting performance acceptance. However, FI and FN threshold numbers used in these two research projects were not validated using test results from actual asphalt production mixtures sampled from different construction projects. The effectiveness of the FN test and other potential laboratory test options, such as the Hamburg Wheel Tracking Test (HWTT), in predicting in-situ rutting performance was also not evaluated in those two ODOT research projects. In addition, the most effective asphalt mixture long-term aging protocols to achieve reliable semi-circular bend (SCB) test parameters that are correlated with in-situ cracking performance are needed to be developed. The developed aging protocol also needs to be integrated into the balanced mix design procedures that are developed for Oregon in this study.

1.1 ORGANIZATION OF THIS RESEARCH REPORT AND CONNECTION TO THE PREVIOUS SPR785 ODOT RESEARCH PROJECT

The general framework followed in this research study is presented in Figure 1.1 in conjunction with the research project SPR785 completed in 2017 (Coleri et al. 2017b). The research presented in this report facilitates the implementation of performance-based specifications for asphalt mixture design to improve the fatigue cracking performance of pavements in Oregon.

Previous research by Coleri et al. (2017b) achieved the following:

- *Chapter 3 in SPR785-Coleri et al. (2017b)* – This part of the study focused on characterizing the cracking performance of asphalt pavements in Oregon by using four tests commonly used to evaluate fatigue cracking resistance. Based on the results of this part of the study, a fatigue cracking test and a test procedure were proposed by considering cost, practicality, simplicity, and accuracy for agencies and contractors. Related manuscript was also published in the *"Construction and Building Materials"* journal (Sreedhar et al. 2018).

- *Chapter 4 in SPR785-Coleri et al. (2017b)* – This chapter presented the impacts of various mixture properties on cracking and rutting resistance of asphalt mixtures. Also, the impact of dust content and dust-to-binder ratio on cracking and rutting performance of asphalt mixtures were discussed in this chapter. The major goal was to provide a better decision-making structure for the asphalt mixture design stage to address fatigue cracking susceptibility, with the intent of avoiding premature pavement failure and expensive early maintenance and rehabilitation. Related manuscript was also published in the *"Journal of Materials in Civil Engineering"* journal (Sreedhar and Coleri 2018).

This research study builds up on the conclusions (summarized in the previous paragraph) derived by Coleri et al. (2017b). This research report is organized as follows:

- **Chapter 1:** This introductory chapter points out the critical need for this research study and outlines the followed research methodology.

- **Chapter 2:** A comprehensive literature review on long-term asphalt mixture aging protocols and balanced mixture design methods is provided in this chapter.

- **Chapter 3:** The third chapter in this report is titled *"Development of a Long-Term Aging Protocol for Asphalt Mixtures"* and discusses the impact of long-term asphalt aging on mixture cracking performance. In this part of the study, the most effective asphalt mixture long-term aging protocol was determined to achieve reliable semi-circular bend (SCB) test parameters that are correlated with in-situ cracking performance. The developed aging protocol was integrated into the balanced mix design procedure that was developed in Chapter 4.

- **Chapter 4:** The fourth chapter in this report is titled *"Developing Performance-Based Specifications for Asphalt Mixture Design in Oregon."* The main objective of this chapter was to develop performance-based specifications to be used in asphalt mixture design. Based on the findings from all work listed above (SPR 785 and this report), a balanced mix design process was developed and proposed for Oregon in this part of the study.

- **Chapter 5:** A summary of major findings and conclusions of the research completed in this study are provided in this chapter.

- **Finally, Chapter 6:** This chapter includes a comprehensive list of references used in this report.

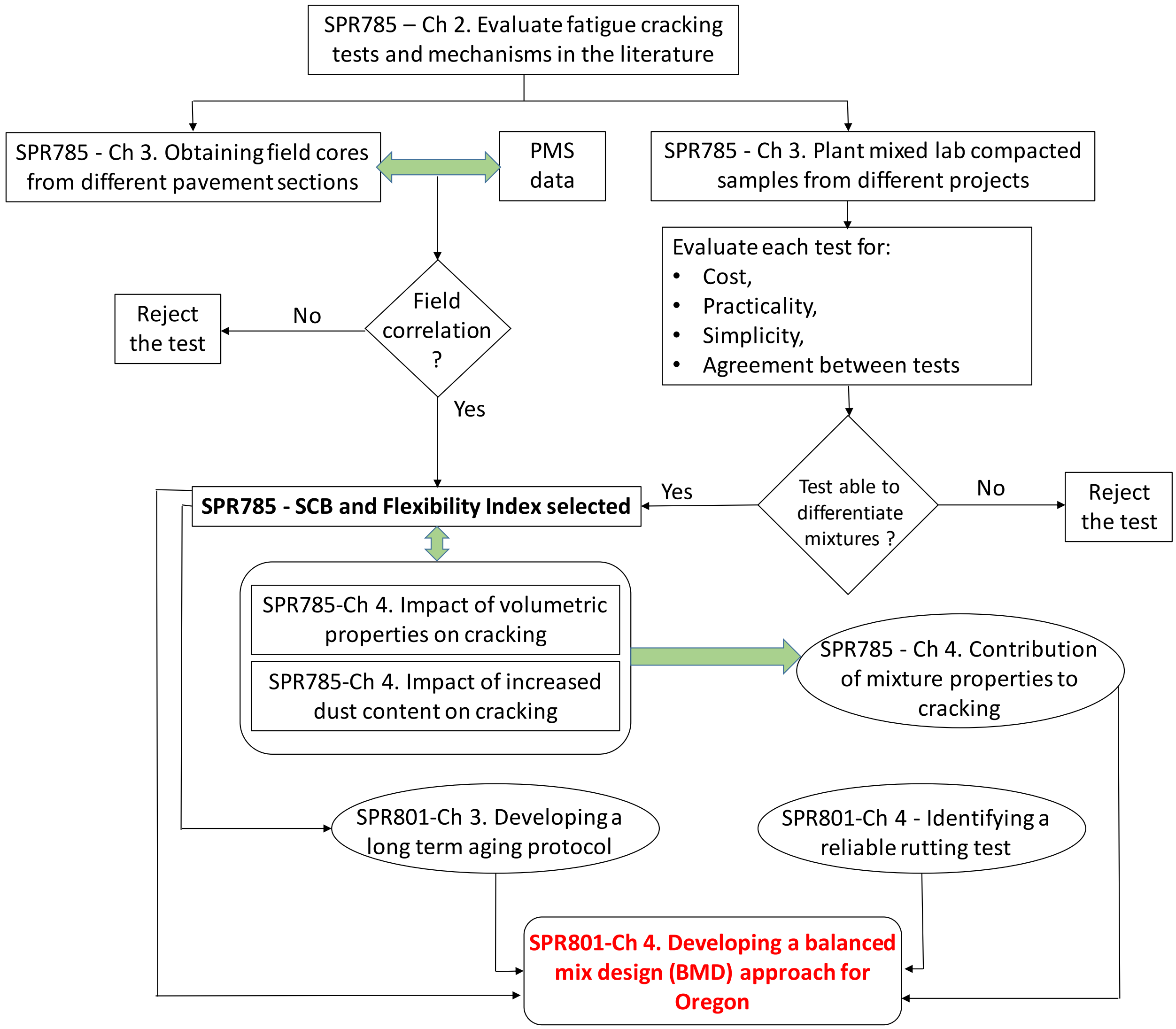


Figure 1.1. Flowchart to achieve the objectives of this research

Note: SPR785 – Coleri et al. (2017b); SPR801 – This research report.

1.2 KEY OBJECTIVES OF THIS STUDY

The main objectives of this study are to:

- develop a long-term aging protocol for Oregon asphalt mixtures for SCB testing;
- determine reliable threshold values for FI, FN, Hamburg Wheel Tracking Test (HWTT) rut depth values (for Level 3 and Level 4 mixtures commonly used for

pavement construction in Oregon) for balanced mix design and performance evaluation;

- determine the most effective laboratory test for rutting performance prediction; and
- develop a balanced asphalt mix design method for Oregon by incorporating performance tests for rutting and cracking into the current volumetric design process.

2.0 LITERATURE REVIEW

In this comprehensive literature review, effectiveness of current volumetric mix design and the potential benefits of a new balanced mixture design method were evaluated by checking the past research studies and surveys conducted with state department of transportation (DOTs) and asphalt contractors.

2.1 PERFORMANCE BASED SPECIFICATIONS AND BALANCED MIX DESIGN PROCESS

Asphalt mixtures are designed to be used in pavements to withstand vehicular loads under different climatic conditions. The goal of asphalt mix design is to determine an economic blend of aggregates and binder such that the resultant mix provides sufficient stability to resist deformation under traffic loading, and flexibility to withstand cracking. The most commonly used asphalt mix design methods are the Marshall method, Hveem method, and the Superpave method. Numerous research studies are currently being conducted to develop a new mix design process with performance verification. However, this approach is not entirely new and draws upon the existing methods and procedures while the existing methods need to be revised and improved by incorporating findings from recent research studies. In order to understand the new approach, it is necessary to know the history of asphalt mix design process.

2.1.1 History of Asphalt Mix Design

In the late 1920s, the Hveem mix design method was developed for asphalt mixtures and was extensively used in some of the Western States. The objective of the entire process is to determine the optimum asphalt content which is assumed to depend on aggregate surface area and absorption. Also, this method assumes that the stability of the mixture is a function of aggregate particle friction and mix cohesion. The stability is measured using a Hveem stabilometer, which applies an increasing load to the compacted asphalt sample at a predetermined rate (Vallerga & Lovering 1984). The mechanical properties as described by stability are used to determine the optimum asphalt content. Air voids were not considered in the design process until the 1990s. However, it is widely observed that the mixtures produced using this method are dry and more susceptible to fatigue cracking (Harvey et al. 2015).

The Marshall method was developed in the early 1940s and was subsequently used by the U.S. Army Corps of Engineers in World War II for designing asphalt mixtures for airports. Similar to the Hveem method, the primary objective of the Marshall method is to determine the optimum asphalt content. The optimum asphalt content is a function of air voids, maximum stability, and maximum density. It is subsequently validated by checking against flow and voids in mineral aggregate (VMA) (Waterways Experiment Station (U.S.) United States 1948). When compared to the Hveem method, the mixtures designed using the Marshall method possess higher asphalt contents (Harvey et al. 2015).

The Marshall and Hveem mix design procedures were widely used until the early 1990s before the Superpave procedure was introduced. Superpave was developed as part of the Strategic Highway Research Program (SHRP) and was implemented in 1993. The original objective was to develop a performance-based mix design process. Although performance tests for asphalt mixtures were a part of the Superpave mix design process and several procedures were developed to predict mixture performance, the entire process turned out to be too complex and was never implemented by any state DOTs. Superpave mix design had three levels (Level 1, Level 2, and Level 3) with increasing complexity (Cominsky et al. 1994). The performance-based specifications were to be incorporated in Level 2 and Level 3 designs but were never implemented.

The current asphalt mix design practice (Level 1) involves proportioning of the aggregates and the asphalt binder based on empirical properties of aggregates and volumetric properties such as densities, air voids, voids in the mineral aggregate (VMA) and voids filled with asphalt (VFA). To arrive at the correct proportioning, determination of volumetric properties, i.e., measurement of the specific gravity of the mix components, is critical. However, over the years, it has been observed that the measurement of these properties are highly variable and measured properties vary from one agency to another. This discrepancy might lead to the faulty determination of optimum binder content selected in the mix design. Asphalt mixes designed with low asphalt content will lead to cracking related distresses while mixes with high binder content are more susceptible to rutting.

Additionally, the increased use of reclaimed asphalt pavement (RAP), recycled asphalt shingles (RAS), and several other additives resulted in serious concerns with using the current simple volumetric mix design methods. Incorporating these recycled materials into the asphalt mix created problems with accurately predicting the performance of the final asphalt mix based on the properties of mix constituents. Besides, it is still not well understood how recycled binders interact with virgin binders (blending of the asphalt binder around the recycled asphalt aggregates and the virgin binder), which ultimately creates more doubt about how these materials affect field performance (West et al. 2018, Coleri et al. 2017a, and Coleri et al. 2017b).

Furthermore, the effects of polymer modification in asphalt, rejuvenators, fibers, and warm-mix asphalt (WMA) additives cannot be assessed in the current volumetric mix design method. Therefore, performance tests need to be included as a part of the mix design procedure in addition to the volumetric properties to help ensure anticipated pavement performance in the field.

2.1.2 Balanced Mix Design Approach

The Federal Highway Administration (FHWA) formed an Expert Task Group to develop a Balanced Mix Design (BMD) process (West et al. 2018). The group defines BMD as "*asphalt mix design using performance tests on appropriately conditioned specimens that address multiple modes of distress taking into consideration mix aging, traffic, climate and location within the pavement structure*". Figure 2.1 illustrates the difference between conventional volumetric mix design and proposed balanced mix design process. In volumetric mix design, an optimum binder content required to achieve 4% air-void content by applying a predetermined compactive effort (number of gyrations in a Superpave Gyratory Compactor) is determined.

However, performance properties of asphalt mixtures are not accounted for in the design process. On the other hand, in a balanced mix design process, performance properties of asphalt mixtures are evaluated in addition to volumetric properties. In the example presented in Figure 2.1, the binder content determined by the volumetric process is 5.7%. This binder percentage satisfies the rutting criteria for asphalt mixtures. However, this binder content does not satisfy the cracking performance requirements (flexibility index of 8 from the IFIT test). On the other hand, the balanced mix design approach yields a binder content ranging between 6.2% and 6.7%. Within this range, both cracking and rutting criteria are met.

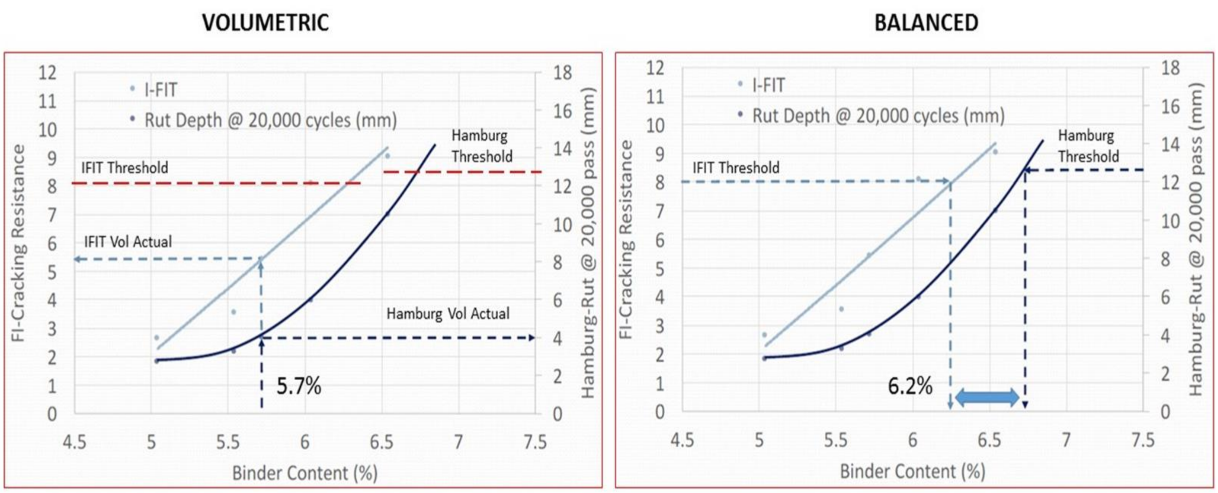


Figure 2.1: Volumetric mix design vs balanced mix design example. (West et al. 2018)

The FHWA group also determined three potential approaches to implement BMD (West et al. 2018), which are briefly described as follows:

Approach 1: Volumetric Design with Performance Verification: This is the most commonly used approach researched and employed by different agencies. In this approach, the mixture is designed based on Superpave specifications. Then, performance tests are conducted to validate whether the mix meets the performance requirements. The mixture should satisfy both volumetric and performance testing criteria. If the mixture does not meet the requirements, the entire mix design process is repeated. The adjustments to the mixture can be made through aggregate source, aggregate gradation, binder source, binder grade, and or additives. This approach is currently being implemented by state department of transportations (DOTs) in Illinois, Texas, Louisiana, New Jersey, and Wisconsin. The process is illustrated in Figure 2.2.

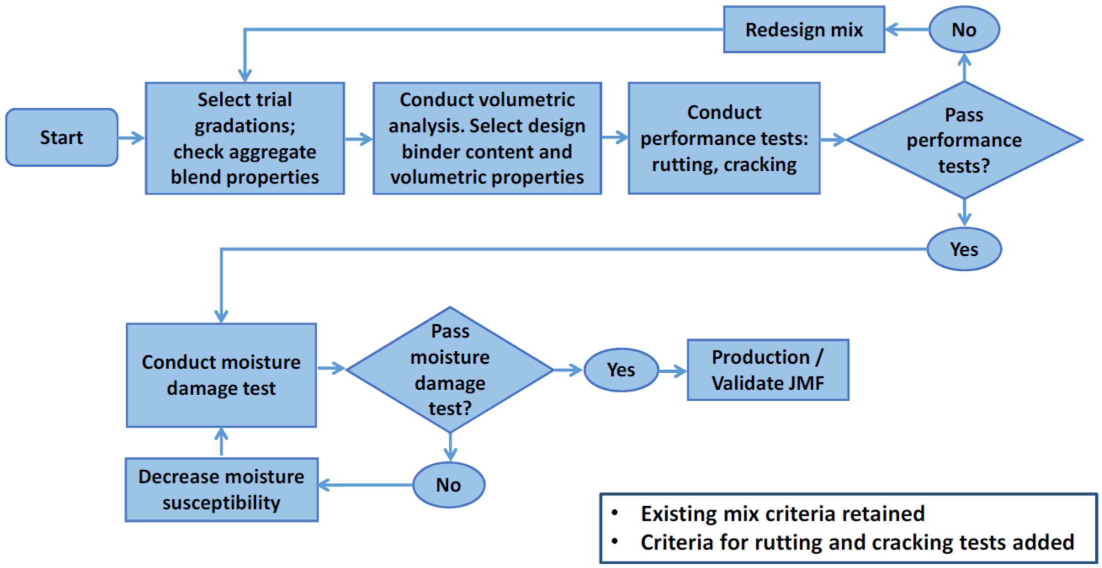


Figure 2.2: Approach 1 - Volumetric design with performance verification. (West et al. 2018)

Approach 2: Performance-Modified Volumetric Mix Design: In this method, the initial aggregate blend and asphalt content are determined using the Superpave mix design process. The mixture proportions are then adjusted to meet the requirements of performance tests. Volumetric mix design requirements are not strictly enforced in this method while performance requirements need to be met. This method is currently being implemented in California. The approach is depicted in Figure 2.3.

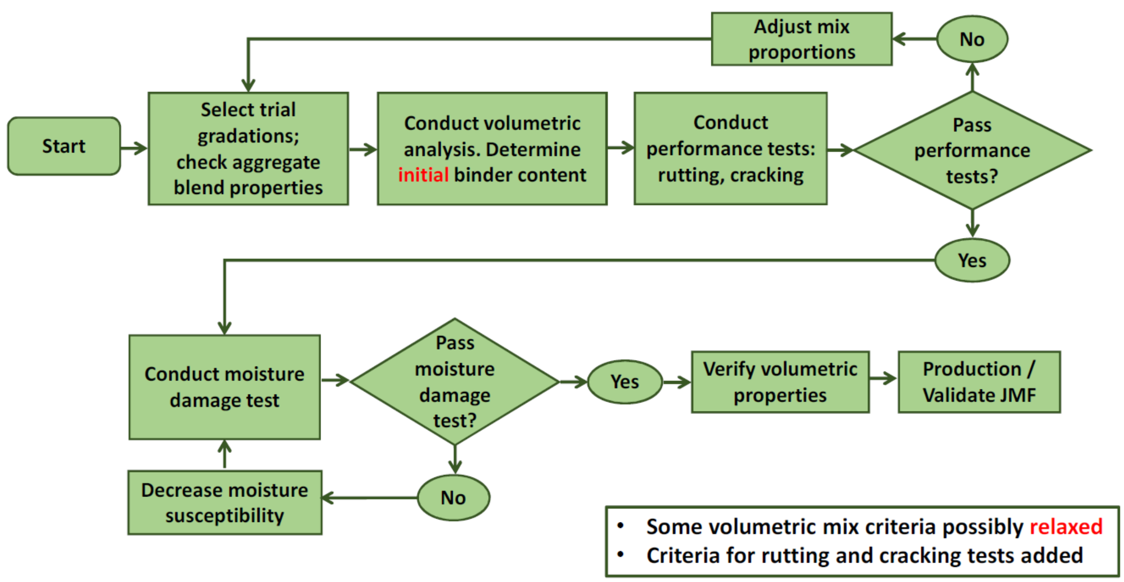


Figure 2.3: Approach 2 – Performance modified volumetric design. (West et al. 2018)

Approach 3: Performance Design: In this approach, the volumetric mix design process is limited or entirely skipped and different trial mixtures are directly evaluated using performance tests as illustrated in Figure 2.4. Here the objective is to use different mixture components in proportion to satisfy the performance test criteria. Therefore, minimum volumetric design criteria may or may not be set for aggregate and binder properties. However, the volumetric criteria such as air voids, minimum asphalt content, aggregate gradation, VMA, and VFA may still be used as a guideline but not as a design criteria. Since there are no field data or knowledge available to validate the effectiveness of this process, this method is not currently being used or implemented by any state DOTs. However, this approach provides a lot of flexibility in design and can be quite rewarding for the contractors and state DOTs. This approach is expected to encourage innovation and direct producers and contractors to evaluate the impact of different additives, gradations, RAP contents, binder types, binder contents, and other variables on asphalt mixture performance. With the objective to reduce mixture costs while meeting the rutting, cracking, and moisture damage failure criteria, different combinations of additives, RAP/RAS contents, gradations, and binder types that will improve pavement longevity can be identified. However, significant level of research combining and evaluating both laboratory and field data is necessary before implementing this high-risk strategy.

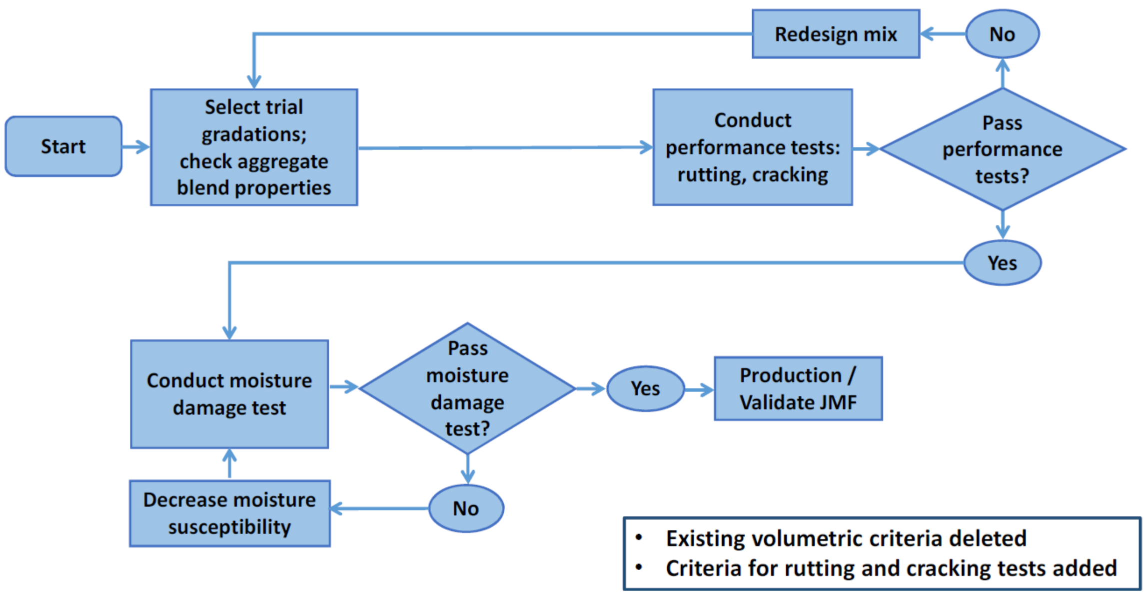


Figure 2.4: Approach 3 – Performance design. (West et al. 2018)

Implementation of a performance-based balanced mix design method is expected to improve the quality and performance of the asphalt mixes used for construction in Oregon. Improved performance will lead to reduced life-cycle costs and increased pavement condition ratings for the Oregon roadway network. Developing a balanced mix design method is also expected to encourage contractors to develop methods and strategies to increase the cracking performance of the pavements while maintaining the required rutting resistance.

2.1.3 The Current Practice of Balanced Mix Design

Several state agencies are investigating the feasibility of integrating performance testing into their mixture design methods. States that are implementing the approaches stated in the previous section are shown in Figure 2.5. The FHWA's task force has identified several DOTs who have begun to implement BMD procedures. Recent research studies conducted to implement BMD methods are summarized in this section.

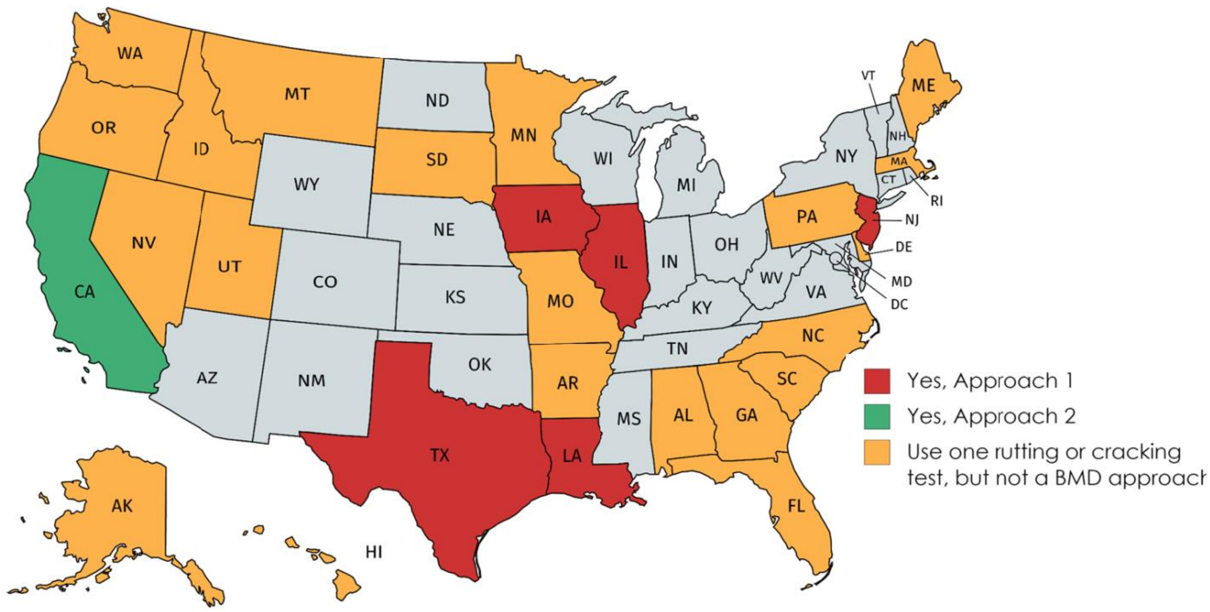


Figure 2.5: U.S. map of current use of BMD approaches (West et al. 2018)

According to the survey conducted by National Center for Asphalt Technology (NCAT) (West et al. 2018), 63% of the states think that VFA requirement in the current volumetric mix design methods should either be eliminated or relaxed (Table 2.1) while this number increases to 69% when the responses from asphalt contractors are evaluated (Table 2.2). These results show that majority of the asphalt specialists do not think that VFA is an effective parameter controlling asphalt mixture performance. A similar conclusion is also valid for VMA when the asphalt contractor responses are considered (Table 2.2). However, 67% of the state DOTs think that VMA is an effective parameter reflecting the long-term performance of the pavements and should not be eliminated or relaxed. It should be noted that VMA is a parameter calculated by using the total aggregate bulk specific gravity. However, problems (high variability) with accurate measurement of aggregate bulk specific results in questionable VMA parameters (West et al. 2018). Thus, the use of VMA as a performance indicator can result in serious issues with projected in-situ mixture performance.

Majority of the state DOTs and asphalt contractors think that TSR is an effective parameter reflecting the moisture sensitivity of asphalt mixtures and should not be changed. About 67% of the asphalt contractors think that dust-to-binder ratio should either be relaxed or eliminated while this number drops to 46% when the DOT responses regarding dust-to-binder ratio are evaluated.

Table 2.1: State DOT Responses on Existing Mix Design Criteria (West et al. 2018)

Mix Design Criteria	No Change	Relaxed	Eliminated
%G_{mm} at N_i	19%	36%	45%
%G_{mm} at N_m	22%	37%	41%
VFA	37%	39%	24%
V_a	53%	42%	5%
D/A Ratio	54%	34%	12%
TSR	63%	15%	23%
VMA	67	24%	10%

Table 2.2: Asphalt Contractor Responses on Existing Mix Design Criteria (West et al. 2018)

Mix Design Criteria	No Change	Relaxed	Eliminated
%G_{mm} at N_i	13%	28%	59%
%G_{mm} at N_m	19%	27%	54%
VFA	31%	43%	26%
V_a	47%	53%	0%
D/A Ratio	33%	49%	18%
TSR	51%	23%	26%
VMA	36%	53%	11%

The ***California Department of Transportation (Caltrans)*** is implementing performance modified volumetric design (Approach 2 outlined in the previous section) (West et al., 2018). For a given aggregate gradation and binder grade, the initial binder content is determined using the existing volumetric approach. Performance tests which include repeated shear, bending beam fatigue, frequency sweep testing, and HWTT are carried out to determine rutting, cracking, and stripping performance of asphalt mixtures. Short-term conditioning (four hours at 135°C) is adopted for repeated shear and HWTT while long-term conditioning in addition to short-term conditioning is used for bending beam fatigue and frequency sweep tests. Based on the results of these performance tests, adjustments to the binder content, binder source, aggregate source, or amount of material passing No. 200 sieve are made. After these adjustments, the mixture is not required to satisfy the volumetric criteria. Performance-based specification developed for California has also been used to evaluate production-mix performance (Tsai et al. 2012). So far, at least seven interstate highways have been constructed using this approach.

The ***Illinois Department of Transportation (IDOT)*** is also employing performance testing in addition to volumetric mixture design (Approach 1) (West et al. 2018). The motivation behind implementing this approach is to address the use of higher contents of RAP/RAS. Binder content is determined using the Superpave volumetric mixture design process after selecting a suitable aggregate gradation and binder grade. I-FIT (Ozer et al. 2016) is used to evaluate the cracking performance after long-term conditioning (a long-term aging protocol is currently being developed) while HWTT is used to evaluate the rutting resistance after short-term conditioning (Two hours of loose mix reheating at 132 ± 3°C). Different thresholds are used for the HWTT for mixes with different performance grades (PG) while a flexibility index of 8 is used as the threshold for the I-FIT cracking test. Different requirements for binder content adjustments, change in binder source, or reduction in quantities of recycled materials are made to achieve the desired mixture performance. However, the final volumetric properties of the mixtures are required to be within the Superpave volumetric mixture design criteria.

The ***Louisiana Department of Transportation and Development (LaDOTD)*** is also using volumetric design plus performance testing approach. This approach has already been implemented in the 2016 LADOTD asphalt specifications and is being used for high and low volume roads on both wearing and binder courses. The optimum binder content is determined using the Superpave specification. Prepared mixtures are then subjected to performance tests which include HWTT after short-term conditioning for rutting resistance and SCB (ASTM D 8044-16) test after long-term conditioning for cracking resistance. Results of performance tests were used to determine the need for changing any mixture properties. After modifying mixture properties to meet volumetric and performance requirements, final set of performance experiments are conducted for the validation of the final mixture design. After the use of balanced mix design procedures, LADOTD decided to reduce the number of gyrations at N_{design} to increase the optimum binder content from volumetric design (Cooper et al., 2014). The major reason for this change was the consistently lower optimum binder contents from the conventional volumetric design when compared to the binder contents suggested by the balanced mix design. It should be noted that LADOTD's balanced mix design has different requirements for two traffic levels. For instance, for lower traffic areas, SCB-J_c (cracking test parameter) should be more than 0.5kJ/m^2 while this parameter was suggested to be more than 0.6kJ/m^2 for highways with high traffic levels.

The ***Minnesota Department of Transportation (MnDOT)*** is also focusing on the implementation of a new mixture design process that involves volumetric design with performance verification (Newcomb and Zhou, 2018). The major distress type considered is low-temperature cracking. To quantify low temperature cracking resistance of asphalt mixtures, Disk-Shaped Compact Tension (DCT) tests are conducted with laboratory compacted and plant mixed (production) specimens. New specifications are currently being developed to determine the correct DCT parameters for different traffic levels. The optimum binder content is determined using the Superpave specification. Performance of asphalt mixtures with poor low-temperature cracking resistance are generally improved by using softer binders or increasing asphalt binder content. Performance of pilot sections constructed in 2013 has been continuously evaluated within the last 6 years to determine the effectiveness of using DCT and performance based specifications in Minnesota.

The ***New Jersey Department of Transportation (NJDOT)*** is also using a volumetric design with performance verification (Approach 1) (West et al., 2018). However, different from other states, they have different performance thresholds for different asphalt mixture types. The five asphalt mixture types designed by using BMD method are high RAP mixtures, bottom-rich base course, bridge deck water proofing surface course, binder-rich intermediate course, and high-performance thin overlay. This approach is being implemented for about 10 percent of the state's total asphalt tonnage on these five mixtures that are subjected to high traffic volumes. For performance verification, Asphalt Pavement Analyzer (APA) (AASHTO T 340) tests on mixtures with short-term conditioning (2 hours at compaction temperature) is conducted to evaluate the rutting resistance and Texas overlay and bending beam fatigue tests (BBF) with again short-term conditioned specimens are conducted to evaluate cracking resistance. Mixture design adjustments generally suggest changes in binder content and inclusion of polymers, rejuvenators, or WMA (warm mix asphalt) additives. Production mixtures are also sampled and tested for performance evaluation.

The ***Texas Department of Transportation (TxDOT)*** is also currently using volumetric design with performance verification (Approach 1) to design specialty mixtures such as stone matrix asphalt and thin overlays. First, conventional volumetric design is conducted to determine the optimum binder content. Then, specimen are prepared at optimum, optimum+0.5%, and optimum+1% binder contents to test for rutting and cracking resistance. HWTT and Texas overlay tests are both conducted on short-term conditioned mixtures (two hours at compaction temperature) for rutting and cracking resistance evaluation, respectively. HWTT results are also used to evaluate moisture susceptibility. If the mixture does not meet the performance criteria, new volumetric mixture design is carried out by adjusting binder content, changing the aggregate source, binder source, or the amount passing No. 200 sieve.

The ***Wisconsin Department of Transportation (WisDOT)*** is also investigating the effectiveness of volumetric design with performance testing verification. HWTT after short-term conditioning of mixtures (four hours at 135°C) is used for rutting assessment, Disk-Shaped Compact Tension (DCT) and SCB after long-term conditioning (twelve hours at 135°C) are used for low-temperature and fatigue cracking performance evaluation, respectively.

2.1.4 Balanced Mix Design Case Studies

This section summarizes the on-going research efforts to implement balanced mix design procedures in Louisiana, New Jersey, and California.

2.1.4.1 Louisiana

LADOTD has been working to improve performance of asphalt mixtures by adapting a balanced mix design procedure. Two comprehensive research studies were conducted by Louisiana Transportation Research Center (Cooper et al., 2014; Mohammad et al., 2016) to determine reliable thresholds for HWTT (rutting) and SCB-J_c (cracking) tests. In these projects, the effects of switching from the 2006 volumetric design specification to 2013 specification were evaluated by conducing these performance tests. 2013 specification required a reduction in the number of gyrations at N_{design} to increase binder content and a slight increase in minimum VMA and VFA requirements. Results showed that asphalt mixtures designed with the 2013 specification have equal or better performance than the mixtures designed according to the 2006 specification. Results also showed that new specification did not have a negative impact on the in-situ performance of evaluated asphalt mixtures.

2.1.4.2 New Jersey

As described in the previous section, performance based mix design method was used to design five special mixtures (about 10 percent of the state's total annual asphalt tonnage) for New Jersey (high RAP mixtures, bottom-rich base course, bridge deck water proofing surface course, binder-rich intermediate course, and high-performance thin overlay) (Bennert 2011). Field performance data collected since 2006 showed that performance of asphalt mixtures designed with the new method has been exceptionally well. Currently, NJDOT is in the process of implementing balanced mix design methods for all asphalt mixtures.

2.1.4.3 California

Since 2000, University of California Pavement Research Center (UCPRC) and Caltrans have been developing, improving, and using the California Mechanistic-Empirical (CalME) design software for pavement design and performance evaluation. CalME was calibrated by using accelerated pavement test and field performance data to improve the predictive capability of existing models. In addition, the software is capable of simulating pavement rehabilitation and maintenance scenarios. Developed software is also able to consider the impact of different variables on the performance variability through a probabilistic model that uses Monte-Carlo simulations (Ullidtz et al. 2010). A comprehensive material database (model coefficients calculated from repeated shear and BBF tests) was developed by UCPRC over the past 15 years to be able to predict in-situ performance for different case studies. The predictions of this software for cracking and rutting performance have been used to evaluate the effectiveness of different asphalt mixtures for different traffic levels and climate regions. In addition to volumetric design, repeated shear tests were generally conducted to determine the rutting performance of the asphalt mixtures at the design binder content. Recommendations to increase the binder content were provided based on the repeated shear test results. This new mechanistic-empirical process combined with laboratory test results was reported to improve the performance of two pilot sections constructed in California (Red Bluff and Weed projects) (Tsai et al. 2012).

2.2 IMPACT OF AGING ON CRACKING PERFORMANCE

Asphalt aging occurs during production, construction, and service life of the mixtures. Aging of the asphalt mixture during production and construction is called as “Short-term aging” while aging during the use phase is called as “Long-term aging”. The aging of asphalt mixtures is mostly affected by the aging of asphalt binder (Bell et al. 1994a). Aging of asphalt binder associated with the oxidation of the binder is a major factor controlling the fatigue performance of asphalt mixtures. As the aromatic compounds in asphalt binders are oxidized, more polar carbonyl compounds are created which results in increased elastic modulus and viscosity, in other words, stiffening of the binder (Glover et al. 2005). Increased viscosity of the binder makes the asphalt mixture less ductile.

Baek et al. (2012) have investigated the effects of aging on the linear viscoelastic response (LVE) and damage characteristics of asphalt mixtures. Four different aging levels were selected: i) short-term aging (STA) – loose mixture conditioned at 135°C for 4 hours; ii) long-term aging level 1 – compacted specimens conditioned for 2 days at 85°C after STA; iii) long-term aging level 2 – compacted specimens conditioned for 4 days at 85°C after STA; and iv) long-term aging level 3 - compacted specimens conditioned for 8 days at 85°C after STA. It was indicated that aging was a significant factor in the damage growth. They also stated that aging influences the distribution of stress and the way damage is accumulated throughout the pavement structure.

A study conducted by Isola et al. (2014) evaluated the effectiveness of laboratory aging methods to simulate the change in asphalt mixture properties in the field. Two aging procedures were used in this study: 1) heat oxidation conditioning (HOC), and 2) cyclic pore pressure conditioning (CPPC) for inducement of moisture-related damage. For short-term and long-term aging simulation, standard short-term oven aging (STOA) and long-term oven aging (LTOA) procedures were used (Bell et al. 1994b). Three asphalt mixtures (lime-treated granite mixture, granite mixture, and limestone mixture) were produced for Superpave IDT testing (indirect tensile test) for four conditioning types (STOA, STOA plus CPPC, LTOA, LTOA plus CPPC). It was concluded that oxidative aging causes the reduction of fracture energy[1] (total energy necessary for fracture inducement) and consequently, stiffening and embrittling mixtures. CPPC created effectively generated additional damage and more reduction in fracture energy (FE) and made the aging process more compatible with the damage observed in the field.

Arega et al. (2013) conducted research on evaluating the fatigue cracking resistance of short-term and long-term aged asphalt mortars with fine aggregate matrix (FAM) and warm mix additives. Two different binders (PG76-28 and PG64-22) with four additives and one aggregate type were tested using dynamic mechanical analyzer (DMA) for this study. Fatigue cracking resistance of specimens were measured before and after long-term aging. For short-term aging, mortars were aged as a loose mix for four hours at 60oC. Then, one batch was compacted with the Superpave Gyratory Compactor (SGC), and another batch was further aged for 30 days in the same environment to simulate long-term aging. Stiffness and fatigue life of FAM is illustrated in Figure 2.6 and Figure 2.7, respectively. It can be observed that short-term aged mixtures have a lower stiffness (G*) with longer fatigue life compared to long-term aged mixtures. However, fatigue resistance rankings of mixtures with and without long-term aging were determined to be the same.

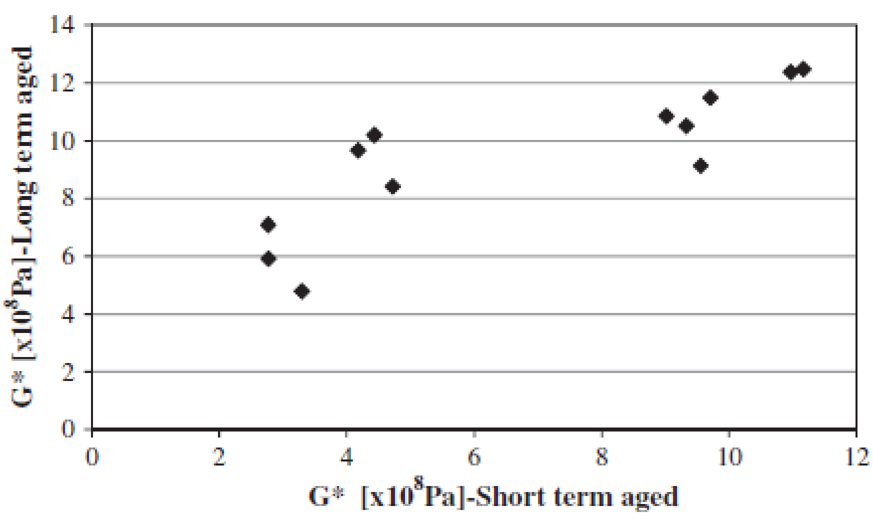


Figure 2.6: Comparison of G* (shear modulus) before aging and after long-term aging for different FAM specimens (Arega et al. 2013)

[1] FE is the total energy necessary for fracture inducement, and it shows the fracture tolerance of the mixture, therefore, represents the cracking performance of the mixture (Roque et al. 2011).

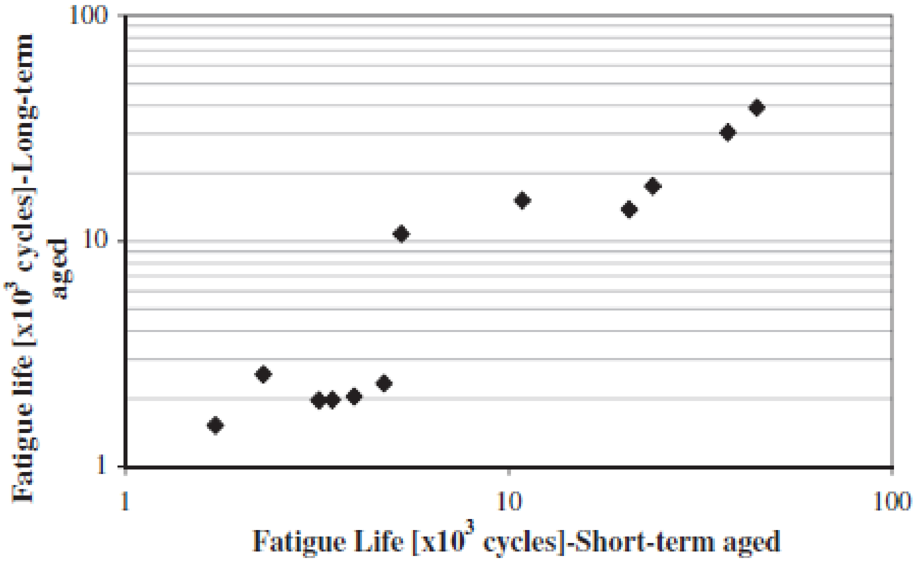


Figure 2.7: Comparison of fatigue life before and after long-term aging for FAM specimens (Arega et al. 2013)

A research study by Yin et al. (2017) had the objectives: i) to develop a correlation between field aging at one to two years after construction and laboratory LTOA protocols; and ii) to identify factors that had significant effects on the long-term aging of asphalt mixtures. Field cores were obtained from seven projects during construction and several months after construction, and also raw materials were procured to produce laboratory specimens that were subjected to selected long-term oven aging protocols. The resilient modulus (M_R) and Hamburg wheel tracking tests were carried out on specimens to evaluate mixture stiffness and rutting resistance of asphalt mixtures with aging. Based on the test results, it was observed that the LTOA protocols (compacted mix) of two weeks at 60°C and five days at 85°C produced mixtures with equivalent in-service field aging of 7-12 months and 12-23 months, respectively. Furthermore, it was also observed that warm mix asphalt (WMA) technology, recycled materials, and aggregate absorption had a significant impact on long-term aging characteristics of asphalt mixtures. Figure 2.8 depicts the M_R ratio results for the post-construction cores and laboratory mixed and compacted specimens with LTOA protocols from two field projects. The M_R ratio values, defined as the ratio of long-term aged specimens' resilient modulus to the short-term aged specimens' resilient modulus, for mixtures with and without recycled materials are plotted against each other. The control mixtures for the Texas (TX) field project were produced using PG 70-22 binder, and the recycled mixtures were produced using PG 64-22 binder with 15% RAP and 3% RAS. The control mixture for New Mexico (NM) field project was produced using a PG 76-28 binder while the recycled mixture was produced using PG 64-28 binder with 35% RAP. The solid line in the plot depicts the line of equality, whereas the dashed line illustrates the shift from the line of equality for the M_R ratio results.

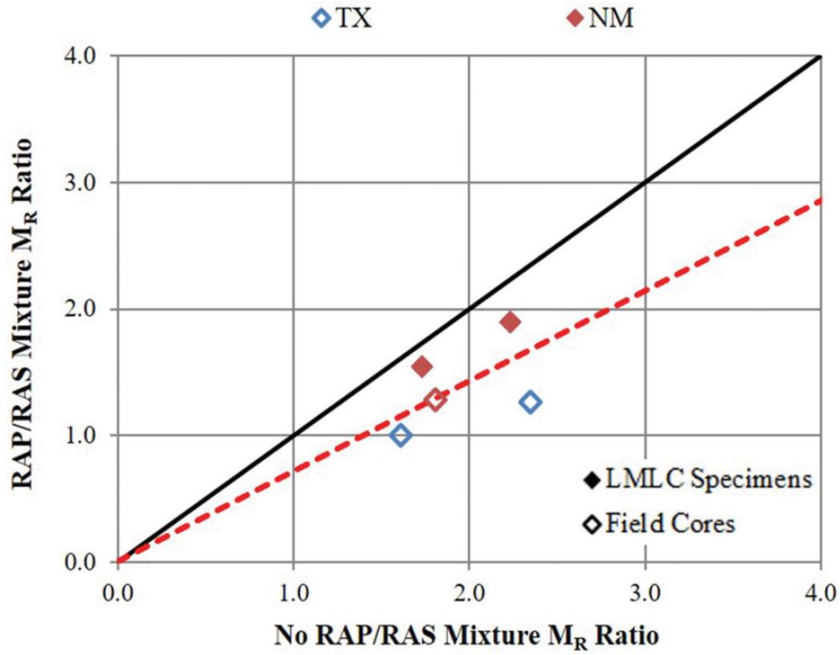


Figure 2.8: M_R ratio comparison for control versus recycled mixtures (Yin et al. 2017)

From the figure, it can be observed that the data points aligned below the line of equality, indicates a significantly higher increase in M_R (stiffness) after long-term aging for the control mixtures compared to the recycled mixtures. Considering the lower initial stiffness but higher M_R ratios for control mixtures versus the recycled mixtures, equivalent mixture stiffness between these two mixture types could be achieved after certain aging periods. Therefore, from this study, it was concluded that recycled materials had a significant impact on the aging of asphalt mixtures and it is possible to have virgin and recycled mixtures present similar stiffness values by excessive aging.

Kim et al. (2018) conducted a comprehensive research study under NCHRP project 09-54 with objectives to develop a long-term aging protocol for asphalt mixtures and to develop an asphalt pavement aging model for mechanistic-empirical (ME) pavement design. In this study, accurate and efficient binder aging index properties (AIPs) were identified to assess aging levels of field cores and laboratory-aged mixtures. The logarithm of binder shear modulus, log G^*, and the total absorbance under the carbonyl and sulfoxide infrared (IR) peaks were selected as the rheological and chemical aging index properties, respectively. For the selection of an aging protocol, three factors were investigated and they were: i) compacted specimen aging versus loose mixture aging; ii) pressure aging versus oven aging; and iii) 95°C aging temperature versus 135°C. It was observed that loose mixture aging led to uniform aging and a significant reduction in aging time compared to compacted specimen aging. However, difficulties were encountered in compaction of aged loose mixtures for laboratory specimen preparation. Pressure aging expedited the process but larger pressure aging vessel (PAV) than the conventional PAV would be required to age a sufficient quantity of loose mixture for test specimen preparation. Comparative tests were carried out between loose mixtures aged at 95°C and 135°C to evaluate the performance repercussions of long-term aging below and above 100°C. Two batches of loose mixtures were first subjected to short-term aging at 135°C for four hours followed by long-term aging at 95°C and 135°C. Small samples of loose mixtures were taken out at periodic intervals for binder extraction and recovery. Chemical and rheological AIPs were used to evaluate the changes in asphalt binder oxidation levels against the aging durations. Figure 2.9 depicts the relationship between the G* value at 64°C and 10Hz frequency and C + S (carbonyl + Sulfoxide) absorbance peaks for binder samples extracted and recovered from mixtures aged at 95°C and 135°C. The carbonyl and sulfoxide are two chemical groups that are formed from the oxidation of asphalt binders. C + S absorbance peaks are obtained from FTIR tests and are an indication of oxidative aging.

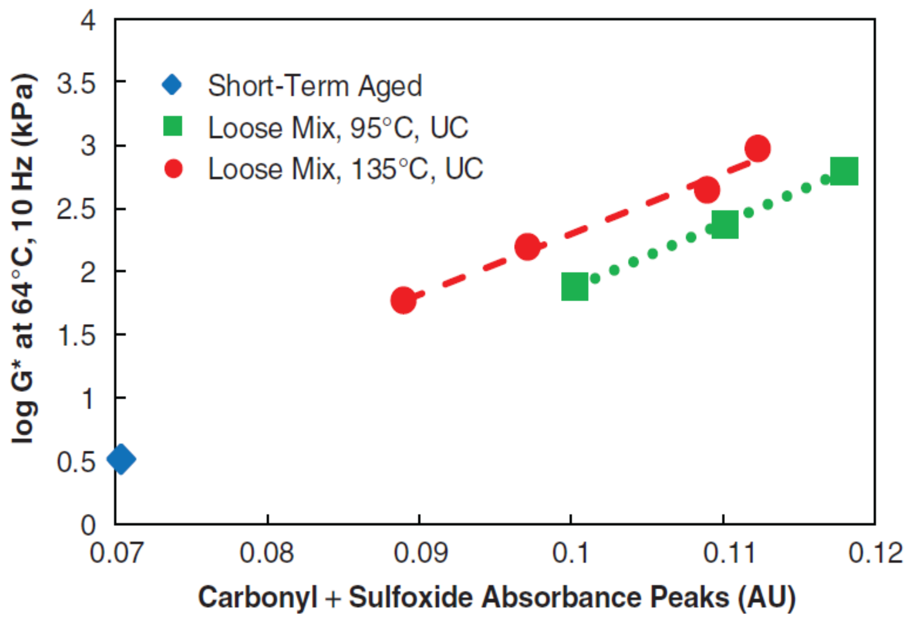


Figure 2.9: Loose mix prepared for long-term aging (Kim et al. 2018)

Figure 2.9 indicates that for the same G* value, binder aged at different aging temperatures have different C + S absorbance peaks thereby suggesting that a change in oxidation reaction occurred when the aging temperature increased from 95°C to 135°C. Figure 2.10 presents the mixture performance results of the specimens fabricated after aging at 95°C and 135°C for 8.9 days and 16.8 hours, respectively. The durations were chosen in such a way that the loose mixture aging at different temperatures yielded similar binder rheology. Figure 2.10 (a) presents the dynamic modulus results and it can be observed that the mixture aged at 135°C for 16.8 hours had a lower stiffness than the mixture aged at 95°C for 8.9 days. Therefore, chemical changes induced by aging at higher temperatures had a significant effect on the performance. Figure 2.10 (b) illustrates the damage characteristic curves (C vs S) for loose mixtures aged at 95°C for 8.9 days and 135°C for 16.8 hours. It can be observed that the damage characteristic curve for mixtures aged at 135°C for 16.8 hours was consistently below the damage curve for the mixture aged at 95°C for 8.9 days indicating that mixture aged at 135°C for 16.8 hours is less stiff than the mixture aged at 95°C for 8.9 days. Also, C value at failure, represented by the end point of C vs S curve, is significantly higher for mixture aged at 135°C than at 95°C suggesting that the mixture aged at 135°C is more brittle than the mixture aged at 95°C. Figure 2.10 (c) depicts the D^R failure criterion results for mixtures aged at 95°C and 135°C. The D^R criterion is an indication of fatigue resistance and uses average reduction in pseudo stiffness (C, a parameter obtained from direct tension cyclic fatigue test) up to failure. Lower the value of D^R, higher is the loss of fatigue resistance (Wang and Kim 2017). From Figure 2.10 (c), it can be observed that loose mixtures aged at 135°C for 16.8 hours had lower D^R failure criterion line as compared to loose mixtures aged at 95°C for 8.9 days suggesting that long-term aging at 135°C for 16.8 hours is more susceptible to reduction of fatigue resistance. Overall, it was concluded long-term aging at 135°C leads to the degradation of fatigue resistance due to the change in the chemistry of the binder. Based on these findings, loose mixture aging in the oven at 95°C was proposed as the long-term aging protocol for asphalt mixtures performance testing.

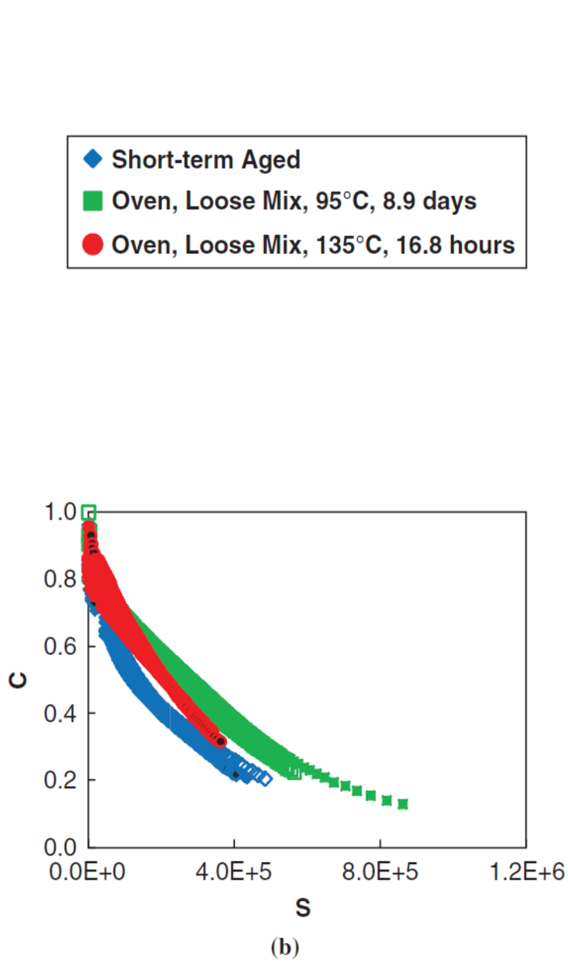


Figure 2.10: Mixture performance test results: (a) dynamic modulus curves, (b) C versus S curves, and (c) D^R failure criterion lines (Kim et al. 2018)

Chen et al. (2018) carried out a research study to select a laboratory loose asphalt mixture aging protocol for the National Center for Asphalt Technology (NCAT) top-down cracking test. In this study, the characterization of asphalt mixtures for field aging was carried out using the cumulative degree days (CDD). CDD was defined as the sum of the daily high temperature above freezing for all the days being considered from the time of construction to the time of coring. This study incorporated materials from five projects in Michigan, Washington, and

Alabama. The loose mixtures were subjected to four different aging protocols: 24 hours at 135ºC, 12 hours at 135ºC, 5 days at 95ºC, and 6 hours at 135ºC. Dynamic shear rheometer (DSR), bending beam rheometer (BBR), and Fourier Transform Infrared Spectroscopy (FT-IR) tests were carried out on the asphalt binders extracted from five mixes subjected to four long-term aging protocols. Figures 2.11, 2.12, and 2.13 show the results of DSR, BBR, and FT-IR tests, respectively. Figure 2.11 presents the G* at 15ºC and 0.005 rad/s of extracted binder plotted against the corresponding δ (phase angle) values. A consistent trend was observed for all mixes with the 24 hours at 135ºC yielded the most significant level of asphalt aging followed by 12 hours at 135ºC, 5 days at 95ºC, and 6 hours at 135ºC protocol, respectively. Figure 2.12 presents the BBR ΔT_c results of extracted and recovered binder from the five mixes. Binders with a more negative ΔT_c are more susceptible to cracking due to reduced relaxation properties. It was observed that the 12 hours and 24 hours at 135ºC protocols produced the lowest ΔT_c values for all mixes. However, no consistent trend was observed for the 5 days at 95ºC and 6 hours at 135ºC protocols. Figure 2.13 presents the FT-IR carbonyl area (CA) results for the extracted and recovered binders for the five mixes subjected to different aging protocols. The CA is defined as the integrated curve (area under the curve) for the wavelength ranging between 1820 and 1650 cm^{-1} (Liu et al. 1998). Higher CA values indicate a greater level of oxidative aging of asphalt binders. The results were in line with the observations from the other two tests. Binders extracted from mixes aged at 135ºC showed a consistent increase in CA with increasing durations suggesting that more polar oxygen-containing functional groups were formed during the process. Unlike the BBR test results, the impact of aging on CA was consistent for aging protocols across all five tested mixtures.

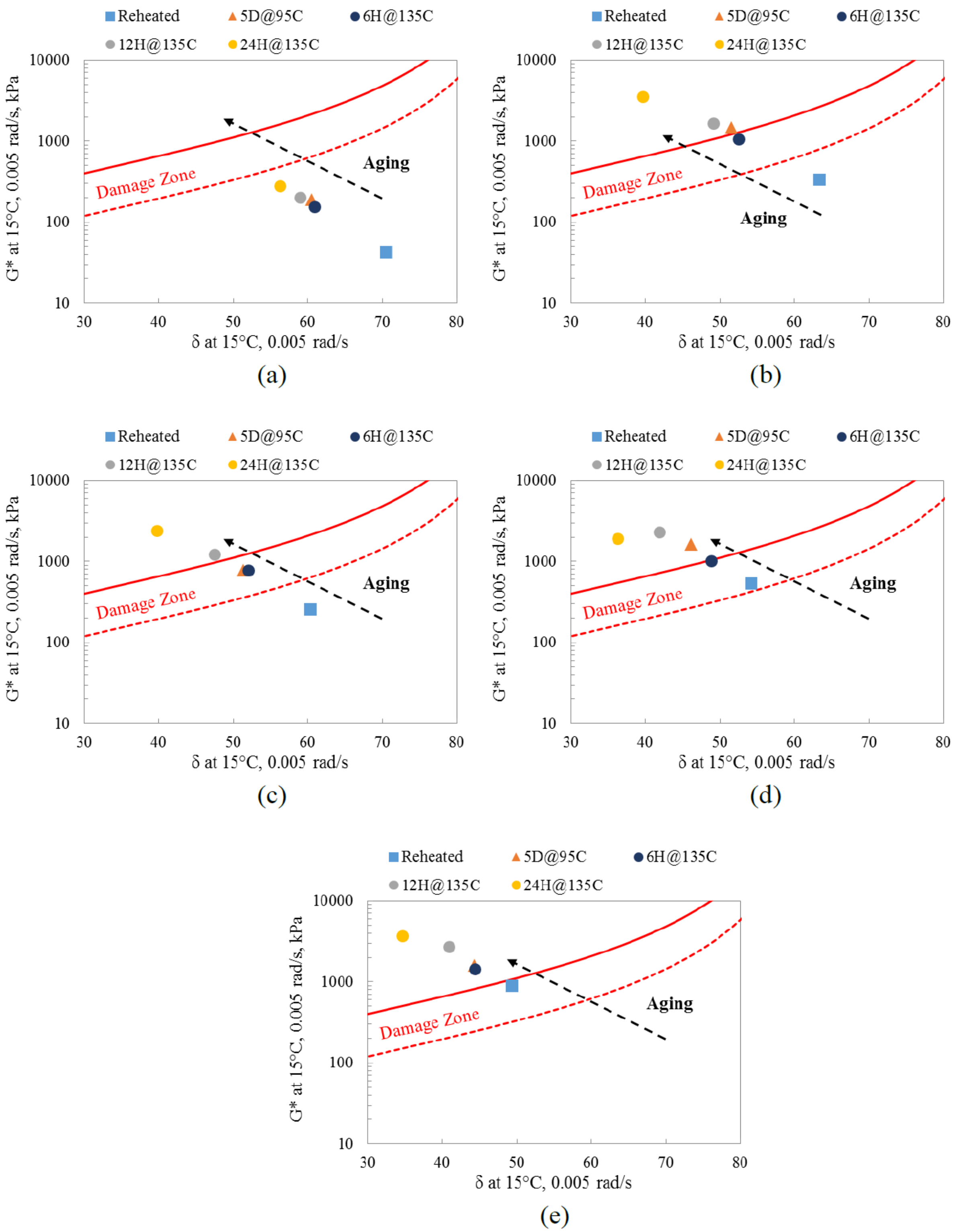


Figure 2.11: DSR test results of extracted asphalt binders with loose mixture aging protocols from five mixes [(a)-(e)] (Chen et al. 2018)

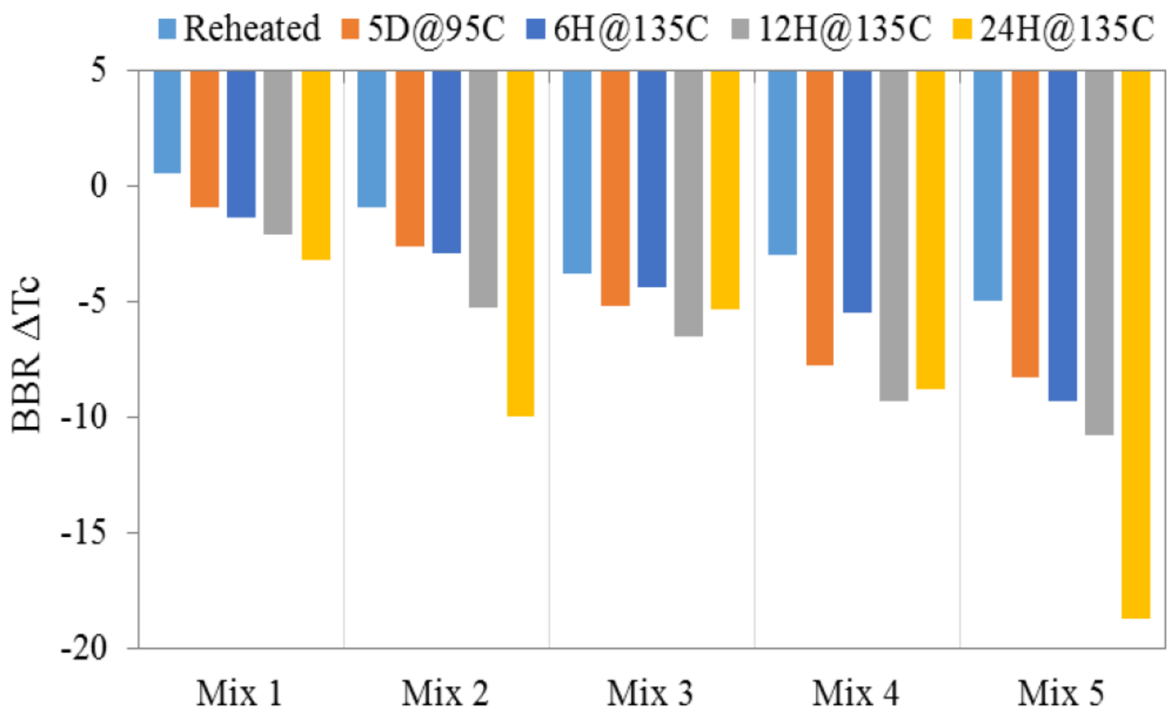


Figure 2.12: BBR test results of extracted binders with loose mixture aging protocols from five mixes (Chen et al. 2018)

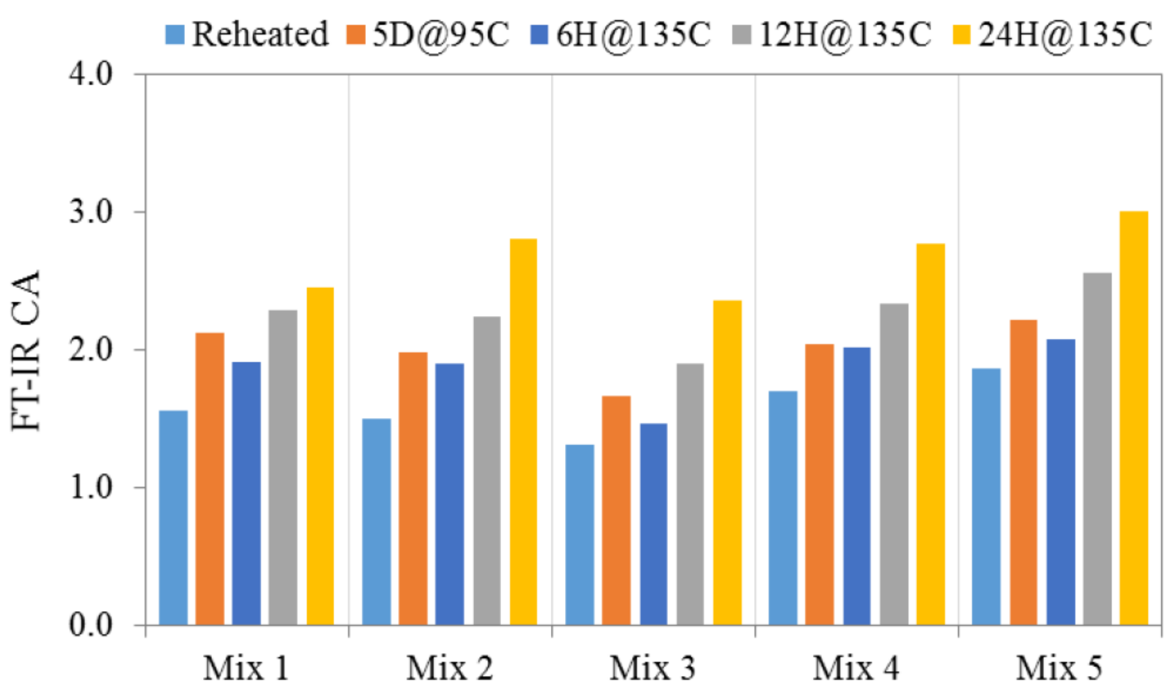


Figure 2.13: FT-IR test results of extracted asphalt binders with loose mixture aging protocols from five mixes (Chen et al. 2018)

After validating the above results against the field data, it was observed that 5 days at 95ºC was the most appropriate protocol to simulate field aging. However, due to the impracticality associated with the implementation due to the longer time span, test results were further analyzed to determine an alternate aging protocol at 135ºC that yielded a similar level of aging as the 5-day 95ºC protocol. The FT-IR CA value obtained from the extracted and recovered binder aged for 5 days at 95ºC was found to be similar to binder aged for about 8 hours at 135ºC (determined by linear interpolation). Figure 2.14 presents FT-IR CA results for an equivalent aging time at 135ºC based on the measured CA of 5 day 95ºC protocol. Finally, this study recommended an aging protocol of 8 hours (rounded 7.4 hours to 8 hours for simplicity) at 135ºC to simulate field aging.

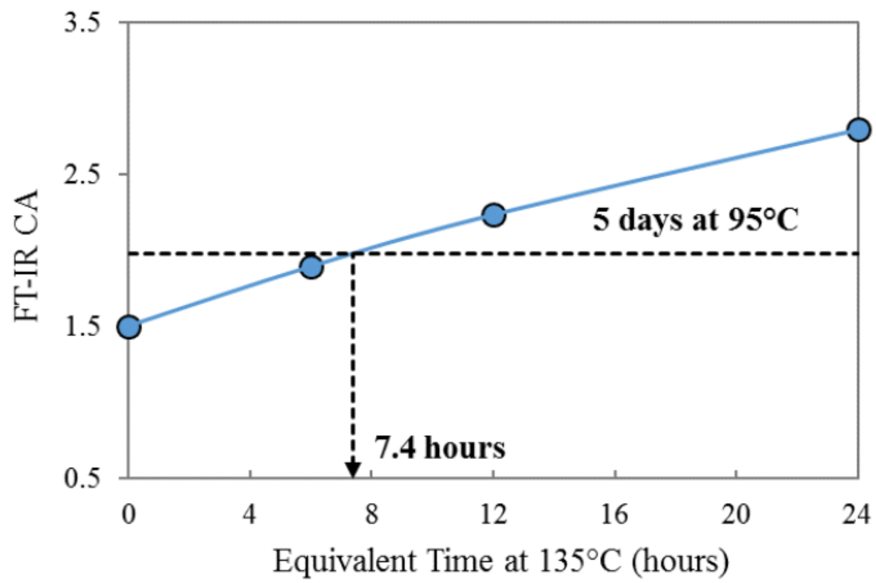


Figure 2.14: Determination of equivalent aging time at 135ºC (Chen et al. 2018)

2.3 LABORATORY TESTS TO EVALUATE PERFORMANCE PROPERTIES OF ASPHALT MIXTURES

A detailed survey with all state DOTs in the U.S. was performed by National Center for Asphalt Technology (NCAT) (West et al. 2018) to determine the most effective performance tests to address several critical pavement distress types. The results are presented in this section.

Table 2.3 shows that fatigue cracking and rutting are the major distress modes in the U.S. followed by thermal cracking. Figure 2.15 illustrate the laboratory tests used by different agencies in the U.S. to determine fatigue cracking, rutting, and thermal cracking resistance of asphalt mixtures. It can be observed that I-FIT (a version of SCB test developed by University of Illinois Urbana Champaign) (Ozer et al. 2016), BBF, and SCB-Jc (a version of SCB developed by Louisiana State University) (Wu et al. 2005) are the most commonly used tests for bottom-up fatigue cracking performance evaluation. I-FIT and direct tension fatigue (S-VECD) are the most commonly used experiments to determine top-down fatigue cracking performance. For rutting performance, HWTT is by far the most commonly used test in the U.S. while APA and FN tests follow HWTT. SCB at low temperatures and disk-shaped compact tension test (DCT) are reported to be the most preferred tests for thermal cracking evaluation. The experiments selected to evaluate balanced mix design procedures in this study are described below in detail.

Answers (DOT)	# (%) Response
Fatigue cracking	40 (88%)
Rutting	33 (70%)
Thermal cracking	30 (64%)
Reflection cracking	29 (62%)
Moisture damage	28 (60%)
Raveling	23 (49%)
Others (block cracking, slippage, etc.)	22 (51%)

Figure 2.15: Pavement distress the state agency wanted to address with mixture performance tests (West et al. 2018)

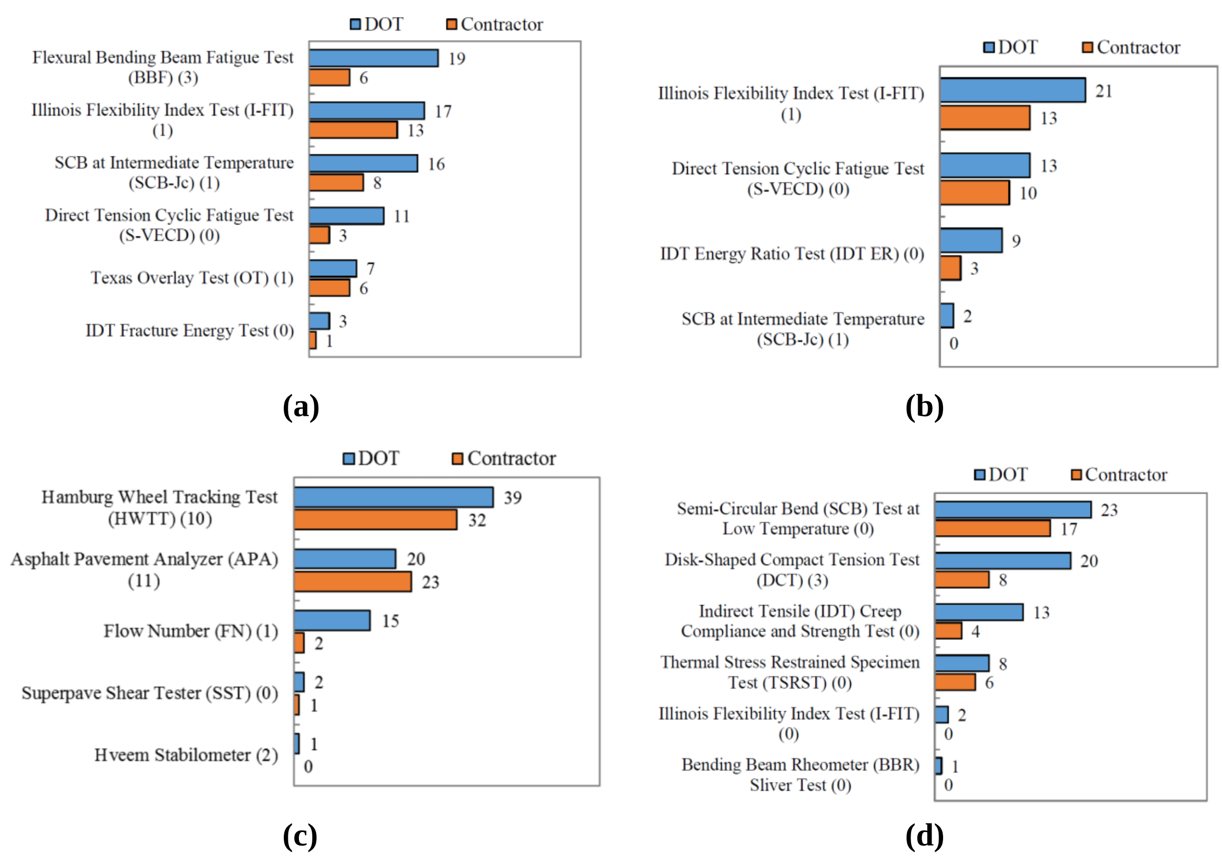


Figure 2.15: Agencies practices for (a) bottom-up fatigue cracking; (b) top-down fatigue cracking (c) rutting; (d) thermal cracking (West et al. 2018)

2.3.1 Semi-Circular Bend (SCB) Test

SCB tests are conducted to determine the cracking performance of asphalt mixtures. 130 mm tall samples were compacted in the laboratory according to AASHTO T 312-12 (2012). Two core samples with the thickness of 57 ± 2 mm were cut from each gyratory compacted sample using a high-accuracy saw. Then, the circular samples (cores) were cut into two identical halves and a notch is introduced in the middle of the test sample.

Wu et al. (2005) suggested performing tests on samples with different notch depths (25.4 mm, 31.8 mm and 38.0 mm). However, Ozer et al. (2016), Nsengiyumva (2015), and Coleri et al. (2017b) showed that reducing the notch depth reduces the variability. For this reason, in this study, a 15 mm notch depth was used for sample preparation. A notch along the axis of symmetry of each half was created with a table saw using another special cutting jig developed at OSU. Notches were 15 ± 0.5 mm in length and 3 mm wide.

Tests were conducted at 25 ^{o}C with a displacement rate of 0.5 mm/min (AASHTO TP 105-13). Samples were kept in the chamber at the testing temperature for conditioning the day before being tested. The flat side of the semi-circular samples was placed on two rollers (Figure 2.16). As a vertical load with constant displacement rate is applied on the samples, the applied load is measured (AASHTO TP 105-13). The test stops when the load drops below 0.5 kN. Fracture energy (G_f), fracture toughness (K_{IC}), secant stiffness (S) and flexibility index (FI) are the testing parameters obtained from this test. Procedures followed to calculate these test parameters are given in the next section.

Figure 2.16: SCB loading set up and test

2.3.1.1 Parameters obtained from SCB test results

This section describes the parameters obtained from SCB test results (displacement vs. load curves) including fracture energy (G_f), fracture toughness (K_{IC}), secant stiffness (S) and flexibility index (FI). Coleri et al. (2017b) suggested the use of FI as a parameter to evaluate fatigue cracking of asphalt mixtures in Oregon.

Fracture Energy (G_f)

Fracture energy (G_f) is obtained by dividing the work of fracture (W_f) by the ligament area (A_{lig}) as shown in Equations (**2-1** to **Error! Reference source not found.**. In Coleri et al. (2017b), a software was developed to calculate FI from test results. As the G_f increases, the work required for crack initiation and propagation increases. Therefore, asphalt mixtures with higher G_f values are expected to show higher resistance to cracking (Ozer et al. 2016). Work of fracture is the area under load versus displacement (P-u) curve (Figure 2.17). The test stops when the load drops below 0.5 kN. The remainder of the curve is extrapolated to estimate the area under the tail of the P-u curve. W_f is the sum of the area under the curve obtained from the test (W) and the extrapolated tail area (W_{tail}) as it is shown in Figure 2.17.

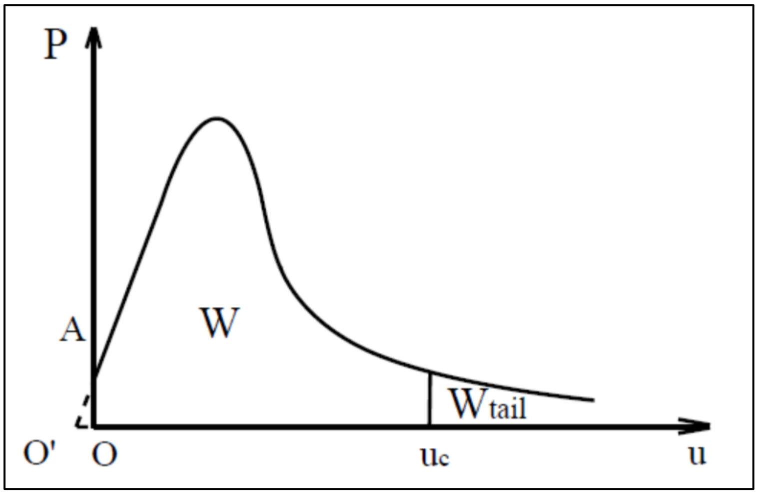


Figure 2.17: Load versus displacement (P-u) curve (AASHTO TP 105-13)

W_f is calculated as follows (AASHTO TP 105-13):

$$\mathbf{G_f = \frac{W_f}{A_{lig}}} \qquad \textbf{(2-1)}$$

$$W_f = \int P \, du$$

(2-2)

$$A_{lig} = (r-a) * t$$

(2-3)

Where:

G_f = fracture energy (kJ/m^2),

W_f = work of fracture (kJ),

P = applied load (kN),

u = load line displacement (m),

A_{lig} = ligament area (m^2),

r = sample radius (m),

a = notch length (m), and

t = sample thickness (m).

The quadrangle rule is used to calculate the area under the curve obtained from the test (W) using Equation **Error! Reference source not found.** (AASHTO TP 105-13):

$$W = \sum_{i=1}^{n} (u_{i+1} - u_i) * (P_i) + \frac{1}{2} * (u_{i+1} - u_i) * (P_{i+1} - P_i)$$

(2-4)

Where:

P_i = applied load (kN) at the i load step application,

P_{i+1} = applied load (kN) at the i+1 load step application,

u_i = load line displacement (m) at the i step, and

u_{i+1} = load line displacement (m) at the i+1 step.

A power function with a coefficient of -2 is used to fit the post-peak part of the P-u curve starting from the point at which the P value is lower than the 60% of the peak load. After fitting the curve, the coefficient c is obtained using Equation **Error! Reference source not found.** (AASHTO TP 105-13). Then, the area under the extrapolated tail portion (W_{tail}) is estimated using Equation **Error! Reference source not found.** (AASHTO TP 105-13).

$$\mathbf{P=\frac{c}{u^2}} \tag{2-5}$$

$$\mathbf{W_{tail}=\int_{u_c}^{\infty} P\, du=\int_{u_c}^{\infty} \frac{c}{u^2}\, du=\frac{c}{u_c}} \tag{2-6}$$

Where:

u = integration variable equal to load line displacement (m), and

u_c = load line displacement value at which the test is stopped (m).

Consequently, the total area under the curve (W_f) is obtained as follows (AASHTO TP 105-13):

$$\mathbf{W_f=W+W_{tail}} \tag{2-7}$$

Fracture Toughness ($\boldsymbol{K_{IC}}$)

Fracture toughness (K_{IC}) is the stress intensity factor at peak load. It shows how much energy is required for crack formation. A higher K_{IC} value indicates higher brittleness of mixtures. The following equations are used to compute K_{IC} (AASHTO TP 105-13):

$$\mathbf{\frac{K_{IC}}{\sigma_0\sqrt{\pi a}}=Y_{I(0.8)}} \tag{2-8}$$

$$\mathbf{\sigma_0=\frac{P_{peak}}{2rt}} \tag{2-9}$$

$$\mathbf{Y_{I(0.8)}=4.782+1.219\left(\frac{a}{r}\right)+0.063*exp(7.045\left(\frac{a}{r}\right))}$$

(2-10)

Where:

P_{peak} = peak load (MN),

r = sample radius (m),

t = sample thickness (m),

a = notch length (m), and

$Y_{I(0.8)}$ = the normalized stress intensity factor (dimensionless).

Secant Stiffness (S)

Secant stiffness (S) is the ratio of the peak load to the vertical deformation required to reach the peak deformation. Higher values for S indicate higher resistance to crack initiation and higher brittleness (Harvey et al. 2015).

$$\mathbf{S\ (KN/mm)=\frac{\Delta y}{\Delta x}=\frac{peak\ load}{vertical\ deformation\ at\ peak\ load}}$$

(2-11)

Flexibility Index (FI)

Flexibility index (FI) is the ratio of the fracture energy (G_f) to the slope of the line at the post-peak inflection point of the load-displacement curve (Figure 2.18). FI correlates with brittleness, and it was developed for asphalt materials by Ozer et al. (2016). Lower FI values show that the asphalt mixtures are more brittle and have a higher crack growth rate (Ozer et al. 2016). Flexibility index is calculated as follows:

$$\mathbf{FI=A*\frac{G_f}{abs(m)}}$$

(2-12)

Where:

G_f = fracture energy (KJ/m^2),

abs(m) = absolute value of the slope at inflection point of post-peak load-displacement curve, A = unit conversion factor and scaling coefficient.

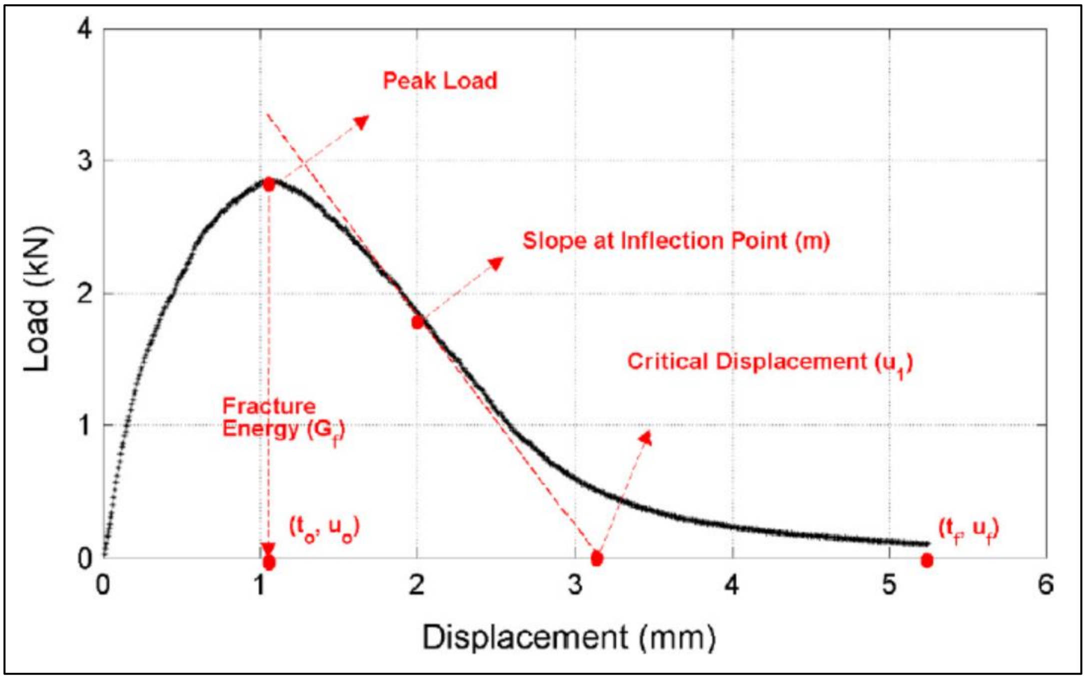


Figure 2.18: Illustration of load-displacement curve and slope at the inflection point (m) (Ozer et al. 2016)

2.3.2 Flow Number (FN) Test

The flow number (FN) test is a performance test for evaluating the rutting resistance of asphalt concrete mixtures (Bonaquist et al. 2003). In this test, while constant deviator stress is applied at each load cycle on the test sample, permanent strain at each cycle is measured (Figure 2.19). Permanent deformation of asphalt pavements has three stages: 1) primary or initial consolidation, 2) secondary, and 3) tertiary or shear deformation (Biligiri et al. 2007). Figure 2.19 shows three stages of permanent deformation. FN is the loading cycle at which the tertiary stage starts after the secondary stage.

In this study, testing conditions and criteria for FN testing described in AASHTO TP 79-13 for unconfined tests were followed. The recommended test temperature, determined by LTPPBind Version 3.1 software, is the average design high pavement temperature at 50% reliability for cities in Oregon with high populations and at a depth of 20 mm (0.79 in) for surface courses (Rodezno et al. 2015). Tests were conducted at a temperature of 54.7°C with average deviator stress of 600 kPa and minimum (contact) axial stress of 30 kPa. For conditioning, samples were kept in a conditioning chamber at the testing temperature for about 4 hours prior to being tested. To calculate FN in this study, the Francken model was used (discussed below).

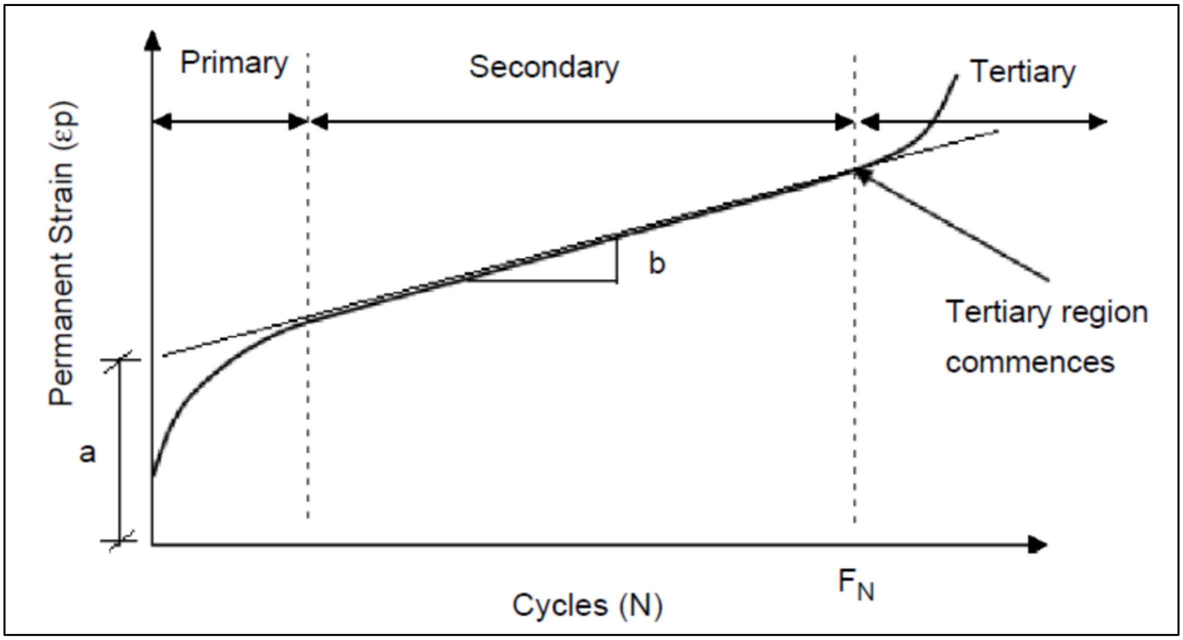


Figure 2.20. Relationship between permanent strain and load cycles in FN test (Biligiri et al. 2007)

Minimum FN values (calculated by using the Francken model) for different traffic levels recommended by AASHTO TP 79-13 are given in Table 2.4 (Rodezno et al. 2015).

Table 2.3: Minimum average FN requirement for different traffic levels (AASHTO TP 79-13)

Traffic (million ESALs)	Minimum Average FN Requirement
<3	NA
3 to <10	50
10 to <30	190
≥30	740

Note: NA= not applicable.

2.3.2.1 Francken model

The Francken Model was developed for triaxial and uniaxial repeated-load tests for different temperatures and stress levels (Francken 1977). A study carried out by Biligiri et al. (2007) showed that this model calculates FN more accurately compared to other mathematical models. This model can also represent all three stages of deformation (1.primary or the initial consolidation of the mix, 2. secondary, and 3. tertiary or shear deformation) more properly. Moreover, Dongre et al. (2009) confirmed the robustness of the Francken model by fitting FN data obtained from field projects.

Details of the Francken model are given as follows:

$$\boldsymbol{\epsilon_p(N)=AN^B+C(e^{DN}-1)}$$

(2.13)

Where:

$\epsilon_p(N)$	= permanent deformation or permanent strain from F_n test,
N	= number of loading cycles, and
A, B, C, D	= regression constants.

The rate of change of the slope of the permanent strain is obtained by taking the second derivative of the Francken model (Equation 2.14). The inflection point, at which the sign of the rate of change of slope changes is considered as the FN and indicates when the tertiary stage begins. FN is the number of cycles at which the second derivative of the Francken model is zero. The second derivative of the model is as follows (Dongre et al. 2009):

$$\boldsymbol{\frac{\partial^2 \epsilon_p}{\partial N^2}=A*B*(B-1)*N^{B-2}+(C*D*e^{D*N})}$$

(2-14)

The model shown in Equation 3.15 is fitted to the permanent strain versus the number of cycles for each sample. After estimating the regression constants (A, B, C, and D), to find the number of load cycles at the inflection point, FN is computed at the point which Equation 2.14 (second derivative of Francken model) is equal to zero. In this study, a code developed by Coleri et al. (2017b) is used to analyze the data and calculate regression constants (A, B, C, and D) of the Francken model to find the FN for each test.

2.3.3 Hamburg Wheel Tracking Test

The Hamburg Wheel-Tracking Test (HWTT) system was developed to measure rutting and moisture damage (stripping) susceptibility of an asphalt concrete sample. Figure 2.20 depicts a typical HWTT system.

Figure 2.19: Hamburg wheel Tracking Device (Instrotek-SmarTracker)

The HWTT follows the AASHTO T 324 standard. According to the specification, either a slab or a cylindrical specimen can be tested. Tests are conducted by immersing the asphalt concrete sample in a hot water bath (40 or 50°C) and rolling a steel wheel across the surface of the sample to simulate vehicular loading. Approximately 20,000 wheel passes are commonly used to evaluate the rutting and stripping resistance of a sample. Figure 2.21 shows an example of typical results obtained from the HWTT. The test provides information related to the total rut depth, post-compaction, creep slope, stripping inflection point and stripping slope of the asphalt concrete sample (Yildirim et al. 2007; Tsai et al. 2016).

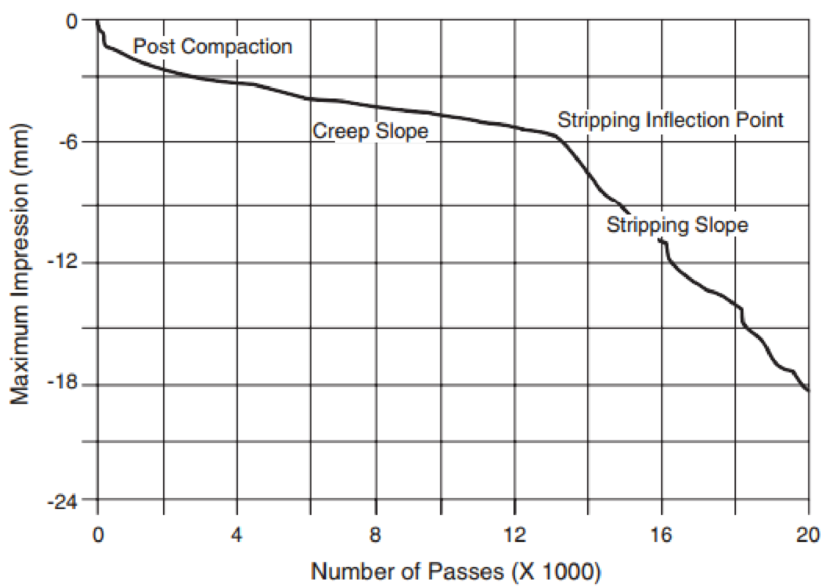


Figure 2.20: Typical HWTD test results (Yildirim et al. 2007)

The deformation after the first 1,000 wheel passes is the post-compaction consolidation and is determined to occur rapidly due to the densification (reduction in existing air-void content) of the asphalt mixture. Creep slope measures the accumulated permanent deformation after post-compaction and prior to the onset of stripping and is an indicator of rutting susceptibility. Creep slope deformation is generally due to a shear flow mechanism other than moisture damage. The number of passes at which the creep slope transitions to the stripping slope is the stripping inflection point. The stripping slope measures the accumulated permanent deformation resulting from moisture and is an indicator of the asphalt mixture's resistance to moisture-induced damage (Yildirim et al. 2007).

Yildirim et al. (2007) conducted an investigation to determine significant factors affecting Hamburg Wheel-Tracking Device test results. The Texas Department of Transportation Hamburg database, consisting of 840 HWTD tests, was utilized for the investigation. The authors explained that the TxDOT uses the HWTD tests on a pass/fail basis. If a sample of a certain mixture produces a rut depth of less than 12.5 mm after 20,000 wheel passes, the mixture is considered passable. A rut depth greater than 12.5 mm after 20,000 wheel passes fails and should not be used for roadway construction. Results of the investigation found that temperature has a large effect on HWTD test results. Of asphalt samples tested at 40°C, 90.2% received a passing rut depth after 20,000 wheel passes. When considering samples tested at 50°C, only 65.2% passed the rut depth criteria. These results suggest that test temperatures should be close to the temperatures that are experienced in the field. The impact of three commonly used binder types in Texas, PG64-22, PG70-22, and PG76-22, on rutting performance were also investigated. As would be expected, stiffer binders were found to improve rutting resistance. For samples tested at 50°C, only 32.7% of samples using PG64-22 binder received passing grades, 56.3% of samples with PG70-22 binder received passing grades and 82.4% of samples using PG76-22 binder passed the rut depth test criteria. The effects of aggregate type and mix type were also investigated but were not found to significantly affect HWTD test results.

However, the current AASHTO T 324 specification possess several disadvantages including: i) two different specimen setups, i.e., either slab or cylindrical specimens are allowed; ii) the definition of rut depth is not clear; and iii) the process involving the visual selection of the first and second portions of an HWTD rutting evolution curve to calculate required test parameters are vague. Tsai et al. (2016) carried out a study in order to address these disadvantages. HWTD tests were simulated using two-dimensional micromechanical finite element (2D-MMFE) models, which are validated with HWTD test data. To determine the consistency of HWTD results obtained from two different specimen setups, the effect of gap/bonding and specimen shape for the cylindrical-core setup were investigated. A dense graded mixture with a nominal maximum aggregate size of 19 mm, a design binder content of 5.1%, and a PG 64-28 polymer modified binder was used for HWTT model development. Based on the simulation results and analyses of the resulting test data, it was observed that the peak maximum principal strain for the model without any bonding between two core samples is about six times more than the sample with bonding. Also, it was observed that rutting in slab specimen accumulates faster than in cylindrical specimens thereby suggesting that the specimen shape effect is an important factor to be considered. This study also recommended the use of the three-stage Weibull approach to fit an HWTD test as it provides a better methodology to interpret the HWTD rutting evolution curve.

2.3.4 Dynamic Shear Rheometer (DSR)

Dynamic Shear Rheometer (DSR) provides the rheological properties of asphalt binder such as complex shear modulus (G^*) and phase angle (δ) for a temperature and frequency range. The testing procedure follows ASTM D7175 2015 specification. In this test, a small sample of asphalt binder is subjected to frequency sweep conducted at temperatures ranging from 20ºC to 80ºC at 20ºC intervals and frequencies ranging from 0.01 Hz to 10 Hz. A typical DSR is illustrated in Figure 2.22.

Figure 2.21: Dynamic Shear Rheometer (DSR)

The master curves are plotted by fitting individual G^* and δ values at each temperature and frequency using Christensen-Anderson-Marasteanu (CAM) model (Santagata et al. 1996). Glover-Rowe (G-R) parameter is then calculated using the following equation (Glover et al. 2005, Anderson, 2011):

$$\boldsymbol{G-R\ Parameter} = \left\{\frac{\boldsymbol{G^*[\cos(\delta)^2]}}{\boldsymbol{\sin(\delta)}}\right\}_{T=15^0C, f=0.005rad/s}$$

(2-15)

The G-R parameter is an indicator of the cracking potential of asphalt binders as it considers both stiffness and embrittlement. High G-R values indicate a greater level of asphalt aging than those with lower G-R parameters.

2.3.5 Fourier Transform Infrared Spectroscopy (FT-IR)

Infrared spectroscopy is the study of the interaction of infrared light with matter. Analysis of infrared spectra can tell what molecules are present in a sample and at what concentrations.

Every FT-IR consists of an interferometer which measures the interference pattern between two light beams. The optical design of an interferometer is shown in Figure 2.23. It consists of four arms. The top arm contains the infrared source and a collimating mirror to collect the light from the source. The bottom arm contains a fixed mirror. The right arm contains a moving mirror which is capable of moving left and right, and the left arm contains the sample and detector. The heart of an interferometer contains a beamsplitter, which transmits some of the light incidents upon it and reflects some of the light incidents upon it. The light transmitted travels towards the fixed mirror and the light reflected travels toward the moving mirror. The reflected beams from these mirrors, travelling back to the beamsplitter, recombine into a single light, interact with the sample, and finally strike the detector. When exposed to infrared radiation, sample molecules absorb radiation of specific wavelengths which causes the change of dipole moment of sample molecules. As a result, the vibrational energy levels of sample molecules transfer from ground state to excited state. The frequency of the absorption peak is determined by the vibrational energy gap. The number of absorption peaks is related to the number of vibrational freedom of the molecule. The intensity of absorption peaks is related to the change of dipole moment and the possibility of the transition of energy levels. Therefore, by analyzing the infrared spectrum, one can readily obtain abundant structure information of a molecule (Smith, 2011).

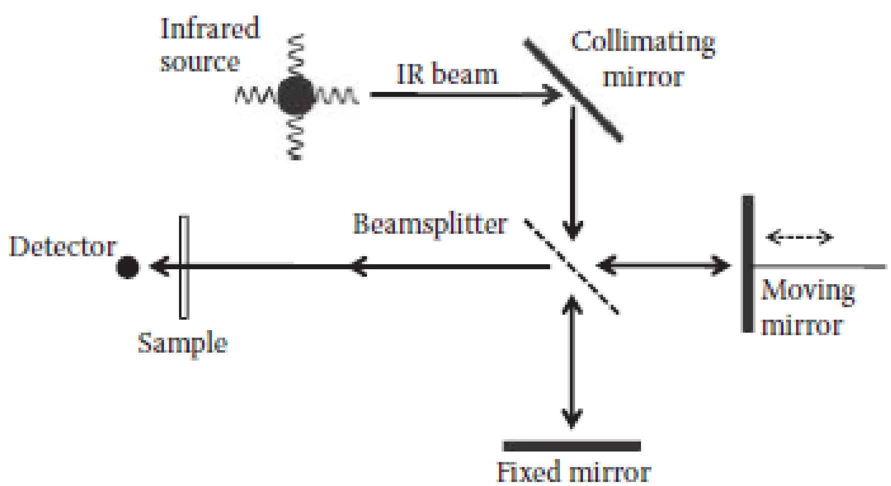


Figure 2.22: The optical diagram of an interferometer (Smith, 2011)

In the current context, infrared spectroscopy measures the infrared spectrum of asphalt binders and evaluate their compositional changes with aging. With oxidative aging, carbonyl groups are formed in asphalt between a wavelength range of 1820 and 1650 cm^{-1}. The carbonyl area (CA) is defined as the integrated peak for the wavelength range between 1820 and 1650 cm^{-1} (Liu et al. 1998). Higher CA values indicate a greater level of oxidative aging of asphalt binders than those with lower CA values.

Ouyang et al. (2006) conducted a research to study the effect of a small amount of zinc dialkyl dithiophosphate (ZDDP), zinc dibutyl dithiocarbamate (ZDBC) and naphthenoid acid on chemical and physical properties of the base asphalt and polymer modified asphalt (PMA) before and after aging. Rotational Thin Film Oven (RTFO) was used to age the asphalt binder. The aging properties were evaluated using FTIR spectroscopy. It was observed that with the addition of additives, the physical properties of asphalt were improved to some extent. Figure 2.24 illustrates the spectra of both base asphalt and PMA as observed through FTIR spectroscopy. Only slight differences were observed in absorbance after the introduction of the additives (ZDDP and ZDBC) before the aging. However, PMA exhibited a higher absorbance after the application of RTFO aging process, indicating that the ZDDP and ZDBC improved the aging resistance of the PMA.

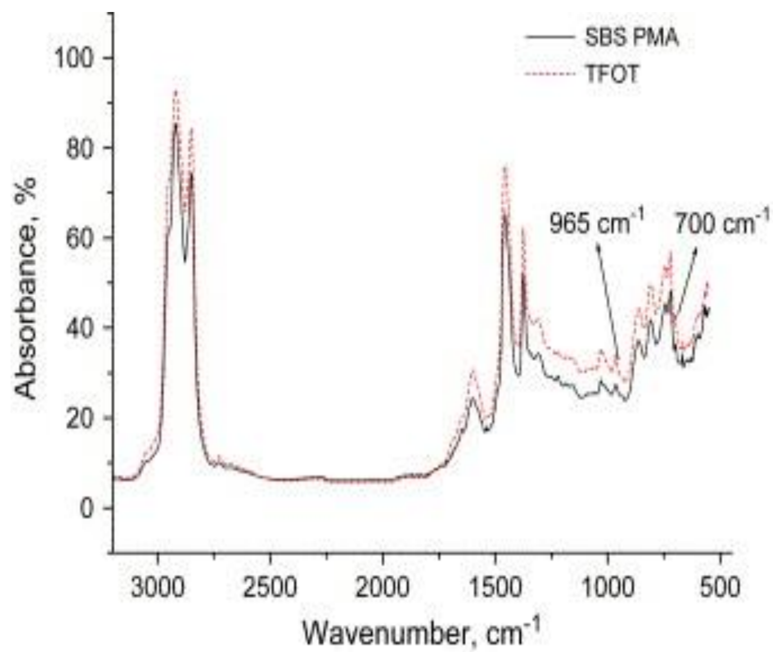


Figure 2.23: FTIR analysis of PMA aged from TFOT (Ouyang et al. 2006)

In a study carried out by Yang et al. (2015), the aging mechanism of asphalt binders blended with bio-binders was studied using FTIR. The aging of the asphalt binder infused with bio-binder was simulated using the RTFO. The binders were conditioned at 120ºC for 20 minutes. Figure 2.25 depicts the FTIR spectra for the control and bio-blended asphalt binders before and after the RTFO aging. A significant mass loss was observed for the bio-blended asphalt during the aging in the form of volatiles. This resulted in faster oxidation and hence faster aging. It was concluded that the chemical components in bio-binders should be stabilized to improve the aging resistance.

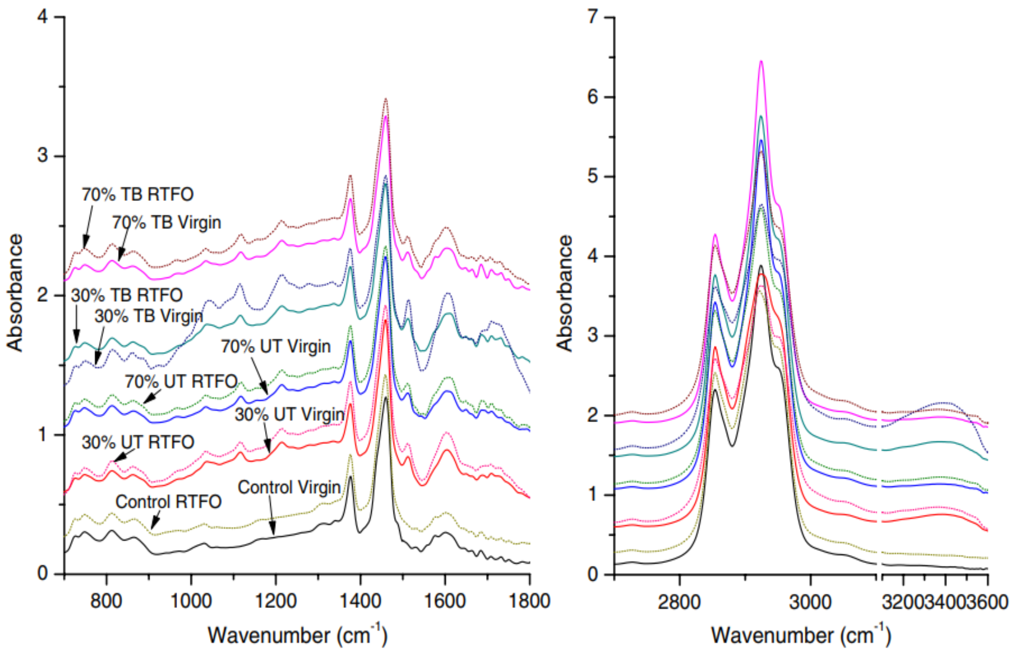


Figure 2.24: Normalized FTIR spectra for control asphalt binder and bio-blended asphalt binders before and after RTFO aging (Yang et al. 2015)

2.4 SUMMARY

In Oregon, fatigue cracking is the major distress mode for asphalt concrete pavement structures. It is one of the main reasons for large road maintenance and rehabilitation expenditures, as well as reduced user comfort and increased fuel consumption due to high road roughness. The resistance of the pavement to this distress mechanism is dependent upon the ductility of the asphalt pavement mixture. According to the literature, aging of asphalt binder associated with the oxidation of the binder is a major factor affecting the fatigue performance of asphalt mixtures. Increasing asphalt binder content, using elastomer-modified binders, and/or using softer binder grades were proved to improve fatigue cracking resistance (Coleri et al. 2017a, Coleri et al. 2017b). Coleri et al. (2017b) showed that binder content of the asphalt mixtures produced with the current volumetric design method can be increased without having rutting failures. The low binder content suggested by the current volumetric design methods results in early fatigue cracking and moisture damage. Increasing density (compactibility) and flexibility by using higher binder contents and/or different types of additives were also recommended to be viable options to improve longevity of Oregon roadway network. However, a robust performance based asphalt mix design method is required to be able to recommend these strategies for performance improvement.

This literature review showed that the existing asphalt mixture design methods do not consider performance criteria and rely on volumetric properties to predict field performance. From an environmental and sustainable standpoint, the use of recycled asphalt materials in asphalt mixtures are becoming increasingly common. A drawback of this practice is a reduction in ductility of the asphalt mixture, which causes a significant reduction in the fatigue life of the pavement in many cases. In Oregon, asphalt pavements are commonly failing prematurely due to cracking-related distresses, necessitating costly rehabilitation and maintenance at intervals of less than half of the intended design lives in some cases. For this reason, it is necessary to accurately quantify the impact of increasing the recycled asphalt content in asphalt mixtures on the structural cracking and rutting resistance of the pavement through use of low-cost and efficient testing procedures that can easily be implemented.

A balanced mix design incorporates performance testing to establish a range of binder contents for a given aggregate gradation. Design method generally provides a maximum binder content for rutting and minimum binder content for cracking considerations. The FHWA formed an expert task group (McCarthy et al. 2016) to develop balanced mix design (BMD) methods. The task force identified three approaches to BMD: i) volumetric design with performance verification; ii) performance modified volumetric design; and iii) performance design. Based on literature, it has been identified that several states such as California, Illinois, Louisiana, Minnesota, New Jersey, Texas, and Wisconsin have adopted or soon implementing some form of BMD procedures. The methodologies followed by these states are summarized in this literature review.

In this comprehensive literature review, effectiveness of current volumetric mix design and the potential benefits of a new balanced mixture design method were evaluated by checking the past research studies and surveys conducted with state department of transportation (DOTs) and asphalt contractors. Most state DOTs and asphalt contractors do not think that commonly used asphalt mixture properties, such as voids in mineral aggregate (VMA), voids filled with asphalt (VFA), and dust-to-binder ratio, reflect the long-term performance of asphalt mixtures. For instance, although there are requirements for VMA set by almost all state DOTs, measurement of VMA relies on the accurate measurement of aggregate bulk specific gravity, while considerable issues were observed in terms of accuracy and variability during the measurement of this parameter (West et al. 2018). In addition, there are several new additives, polymers, rubbers, and high quality binder types incorporated into asphalt mixtures today. Volumetric mixture design methods are not capable of capturing the benefits of using all these new technologies on asphalt mixture performance. In addition, the interaction of virgin binders with reclaimed asphalt pavement (RAP) mixtures with high binder replacement contents and the level of RAP binder blending into the asphalt mixture are still not well understood. Due to all these complications related to the more complex structure of asphalt mixtures, simple volumetric evaluations to determine the optimum binder content may not result in reliable asphalt mixture designs. Two volumetrically identical mixtures may provide completely different rutting and cracking performance according to laboratory test results (Coleri et al. 2017b). For all these reasons, performance tests for rutting and cracking need to be incorporated into current asphalt mixture design methods to be able to validate or revise the optimum binder content determined by the volumetric mix design method.

HWTT is selected as one of the rutting tests as it is fairly well established and used by majority of the state DOTs in the U.S. In this research study, FN was also used as a 2nd rutting tests to evaluate the difference between this test and HWTT. FN was previously determined to be a repeatable experiment to quantify rutting resistance of Oregon asphalt mixtures (Coleri et al. 2017b). However, effectiveness of the FN test for balanced mix design and performance based specifications for Oregon was not investigated in a previous research study. SCB testing, with the analysis protocol developed for Oregon (Coleri et al. 2017b), is chosen as the cracking test. In this study, a long-term asphalt mixture aging protocol, critical for fatigue cracking performance evaluation, was also developed for Oregon asphalt mixtures.

Since asphalt aging reduces the ductility of the asphalt mixture, excessive aging can trigger early fatigue cracking. As the aromatic compounds in asphalt binders are oxidized, more polar carbonyl compounds are created which results in increased elastic modulus and viscosity, in other words, stiffening of the binder. Therefore, with aging, asphalt becomes brittle and more susceptible to cracking. For this reason, a long-term aging protocol for conditioning Oregon laboratory mixtures is needed to be able to achieve realistic cracking performance data from laboratory tests.

Chen et al. (2018) carried out a study to develop a long term aging protocol for loose asphalt mixtures. The loose mixtures were subjected to 6 hours at 135ºC, 12 hours at 135ºC, 24 hours at 135ºC, and 5 days at 95ºC. Based on the results, this study recommended aging loose mixtures for 8 hours at 135ºC. However, another study carried out by Kim et al. (2018) suggested that heating the loose mixtures above 100ºC resulted in serious changes in chemical properties of asphalt binder which are not observed in field. This study proposed conditioning the loose mixtures at 95ºC for durations depending on geographical locations and climatic conditions. Nevertheless, these studies did not consider the impact of using higher RAP contents on aging of asphalt mixtures. Since the asphalt binder present in the RAP is already aged, it is necessary to understand its impact on the virgin mix. Moreover, as these two major studies opine contrary beliefs about aging, it is necessary to evaluate and validate the best long term aging protocol that is suitable for the mixtures used in Oregon.

3.0 DEVELOPMENT OF A LONG-TERM AGING PROTOCOL FOR ASPHALT MIXTURES

3.1 INTRODUCTION

Asphalt aging is recognized as one of the important factors causing distress in asphalt pavements. Asphalt aging occurs during production, construction, and service life of the mixtures. Aging of the asphalt mixture during production and construction is called as "Short-term aging" while aging during the use phase is called as "Long-term aging". The aging of asphalt mixtures is mostly affected by the aging of asphalt binder (Bell et al. 1994a). Aging of asphalt binder associated with the oxidation of the binder is a major factor controlling the fatigue performance of asphalt mixtures. As the aromatic compounds in asphalt binders are oxidized, more polar carbonyl compounds are created which results in increased elastic modulus and viscosity, in other words, stiffening of the binder (Glover et al. 2005; Kumbargeri and Biligiri 2015). Increased viscosity of the binder makes the asphalt mixture less ductile. Baek et al. (2012) have investigated the effects of aging on the linear viscoelastic response (LVE) and damage characteristics of asphalt mixtures. Four different aging levels were selected: i) short-term aging (STA) – loose mixture conditioned at 135°C for 4 hours; ii) long-term aging level 1 – compacted specimens conditioned for 2 days at 85°C after STA; iii) long-term aging level 2 – compacted specimens conditioned for 4 days at 85°C after STA; and iv) long-term aging level 3 – compacted specimens conditioned for 8 days at 85°C after STA. It was indicated that aging was a significant factor in the damage growth. They also stated that aging influences the distribution of stress and the way damage is accumulated throughout the pavement structure.

A study conducted by Isola et al. (2014) evaluated the effectiveness of laboratory aging methods to simulate the change in asphalt mixture properties in the field. Two aging procedures were used in this study: 1) heat oxidation conditioning (HOC), and 2) cyclic pore pressure conditioning (CPPC) for inducement of moisture-related damage. For short-term and long-term aging simulation, standard short-term oven aging (STOA) and long-term oven aging (LTOA) procedures were used (Bell et al. 1994b). Three asphalt mixtures (lime-treated granite mixture, granite mixture, and limestone mixture) were produced for Superpave IDT testing (indirect tensile test) for four conditioning types (STOA, STOA plus CPPC, LTOA, LTOA plus CPPC). It was concluded that oxidative aging causes the reduction of fracture energy[2] (total energy necessary for fracture inducement) and consequently, stiffening and embrittling mixtures. CPPC created effectively generated additional damage and more reduction in fracture energy (FE) and made the aging process more compatible with the damage observed in the field.

Arega et al. (2013) conducted research on evaluating the fatigue cracking resistance of short-term and long-term aged asphalt mortars with fine aggregate matrix (FAM) and warm mix additives. Two different binders (PG76-28 and PG64-22) with four additives and one aggregate type were tested using dynamic mechanical analyzer (DMA) for this study. Fatigue cracking

[2] FE is the total energy necessary for fracture inducement, and it shows the fracture tolerance of the mixture, therefore, represents the cracking performance of the mixture (Roque et al. 2011).

resistance of specimens were measured before and after long-term aging. For short-term aging, mortars were aged as a loose mix for four hours at 60°C. Then, one batch was compacted with the Superpave Gyratory Compactor (SGC), and another batch was further aged for 30 days in the same environment to simulate long-term aging. It was observed that short-term aged mixtures have a lower stiffness (G*) with longer fatigue life compared to long-term aged mixtures. However, fatigue resistance rankings of mixtures with and without long-term aging were determined to be the same.

Research carried out by Yin et al. (2017) had the objectives to develop a correlation between field aging at one to two years after construction and laboratory LTOA protocols, and to identify factors that had significant effects on the long term aging of asphalt mixtures. Field cores were obtained from seven projects during construction and several months after construction, and also raw materials were procured to produce laboratory specimens that were subjected to selected long term oven aged protocols. The resilient modulus (M_R) and Hamburg wheel tracking tests were carried out on specimens to evaluate mixture stiffness and rutting resistance of asphalt mixtures with aging. Based on the test results, it was observed that the LTOA protocols of two weeks at 60°C and five days at 85°C produced mixtures with in-service field aging equivalent to 7-12 months and 12-23 months, respectively. Furthermore, it was also observed that warm mix asphalt (WMA) technology, recycled materials, and aggregate absorption had significant impact on long term aging characteristics of asphalt mixtures.

Kim et al. (2018) conducted a research under NCHRP project 09-54 with objectives to develop a long-term aging protocol for asphalt mixtures and to develop an asphalt pavement aging model for mechanistic-empirical (ME) pavement design and analysis. In this study, an accurate and efficient binder aging index properties (AIPs) were identified to assess aging levels of field cores and laboratory aged mixtures. The logarithm of binder shear modulus, log G^*, and the total absorbance under the carbonyl and sulfoxide infrared (IR) peaks were selected as the rheological and chemical AIP, respectively. For the selection of aging protocol, three factors were investigated: i) compacted specimen aging versus loose mixture aging; ii) pressure aging versus oven aging; and iii) 95°C aging temperature versus 135°C. It was observed that aging loose mixtures led to uniform aging and significant reduction in aging time compared to compacted specimen aging. However, difficulties were encountered in compaction of aged loose mixtures for specimen preparation. Pressure aging expedited the process but larger pressure aging vessel (PAV) than the conventional PAV would be required to age a sufficient quantity of loose mixtures for the specimen preparation. It was also observed that aging at 135°C changed the chemistry of binder to a level that is not observed in the field. Based on these findings, loose mixture aging in the oven at 95°C was proposed as the long-term aging protocol for asphalt mixtures for performance testing.

Chen et al. (2018) carried out a research study to select a laboratory loose asphalt mixture aging protocol for the National Center for Asphalt Technology (NCAT) top-down cracking test. In this study, the characterization of asphalt mixtures for field aging was carried out using the cumulative degree days (CDD). CDD was defined as the cumulative days between time of construction to time of coring for which the daily high temperature was above freezing. This study, incorporated materials from five projects in Michigan, Washington, and Alabama. Loose mixtures were subjected to four different aging protocols: 24 hours at 135ºC, 12 hours at 135ºC, 5 days at 95ºC, and 6 hours at 135ºC. Results from the dynamic shear rheometer (DSR), bending beam rheometer (BBR), and Fourier Transform Infrared Spectroscopy (FT-IR) showed that the 24 hours at 135°C aging protocol yielded the most significant level of asphalt aging. In addition, there was no significant difference in oxidative aging of asphalt binders for mixes aged at 95°C versus 135°C. Finally, this study recommended an aging protocol of 8 hours at 135°C to simulate field aging. However, it should be noted that this study did not evaluate the impact of long-term aging on the cracking resistance of high RAP mixtures.

Zhu et al. (2019) investigated the impacts of asphalt concrete mix design parameters on long-term aging. This study had twofold objectives: (i) to understand the influence of long-term aging on cracking of asphalt concrete using the Illinois Flexibility Index Test (I-FIT), and (ii) to understand the effect of asphalt concrete mix design parameters on aging. Semi-circular samples prepared from projects varying in mixture properties such as mixture type, binder type and content, and amount of total recycled content, were subjected to two aging methods: 5 days at 85°C and 1 day at 95°C. After the designated aging period, I-FIT was carried out on samples and was observed that increase of voids in mineral aggregates (VMA), asphalt binder replacement ratio (ABR), effective asphalt content, and a decrease of low-temperature true grade results in reduction of aging in asphalt concrete. Although the two aging methods considered in this study resulted in similar aging levels, this study concluded that conditioning 1 day at 95°C yields similar performance trend as 5 days at 85°C.

Based on these research studies, it can be observed that there are different schools of thoughts involved in developing an aging method for asphalt concrete, with the major differences being: aging at higher temperatures for shorter durations versus aging at lower temperatures for longer durations, and loose mixture aging versus compacted sample aging. Therefore, this part of the research study was taken up with the following objectives:

1. Comparing field aging levels with different laboratory simulated aging conditions;

2. Understanding the impact of different aging temperatures and durations on cracking resistance and cracking performance rankings of asphalt mixtures with different mix design variables, such as RAP content, binder content, binder type, gradation, additives, etc.; and

3. Developing a long-term aging protocol to be incorporated into the balanced mix design and performance evaluation process (that is presented in Section 4.0).

3.2 EXPERIMENTAL PLAN AND RESEARCH METHODOLOGY

The purpose of this part of the study was to evaluate the impact of laboratory long-term oven aging on stiffness and cracking resistance of asphalt mixtures and thereby, develop a long-term aging protocol for performance based design specifications. To evaluate the impact of aging on plant and laboratory produced asphalt mixtures, this study was divided into three phases:

3.2.1 Phase I - Field aging versus laboratory aging

In the first part, a plant produced asphalt mixture (M0) was subjected to six different aging conditions. Experimental plan for this phase is given in Table 3.1. The gradation curve for M0 is shown in Figure 3.1. Aged asphalt mixtures were used to prepare specimens for semi-circular bend (SCB) testing. Since plant produced mixtures were assumed to have already undergone short-term aging, no short-term aging is simulated. SCB tests were carried out with the samples aged by following different protocols to compare how well they agree with each other. Then, the binders from these samples were extracted and Dynamic Shear Rheometer (DSR) tests were conducted to determine which aging protocol was similar to 1.5 years of field aging. To be able to determine the properties of 1.5 year-aged asphalt mixes, field cores from pavements constructed using M0 mixture were obtained 1.5 years after construction. Asphalt binders from field cores were extracted and recovered for DSR testing. Then, DSR test results from field core binders were compared with the results from lab-produced samples that were subjected to different aging protocols to determine which aging protocol provides DSR test results that are close to 1.5 year-aged asphalt mixes.

Table 3.1: Experimental Plan for Phase I

Test type	Mix Type	Temperature (°C)	Replicates	Aging Protocol (1)	Total Tests
SCB	M0	25	4	6	24
DSR	M0	20, 40, 60 & 70	2	6	72

Note: [1] Aging protocols: 1) No long-term aging (STA); 2) long-term aging for 5 days at 85 °C; 3) long-term aging for 10 days at 95 °C; 4) long-term aging for 6 hours at 135 °C; 5) long-term aging for 12 hours at 135 °C; and 6) long-term aging for 24 hours at 135 °C.

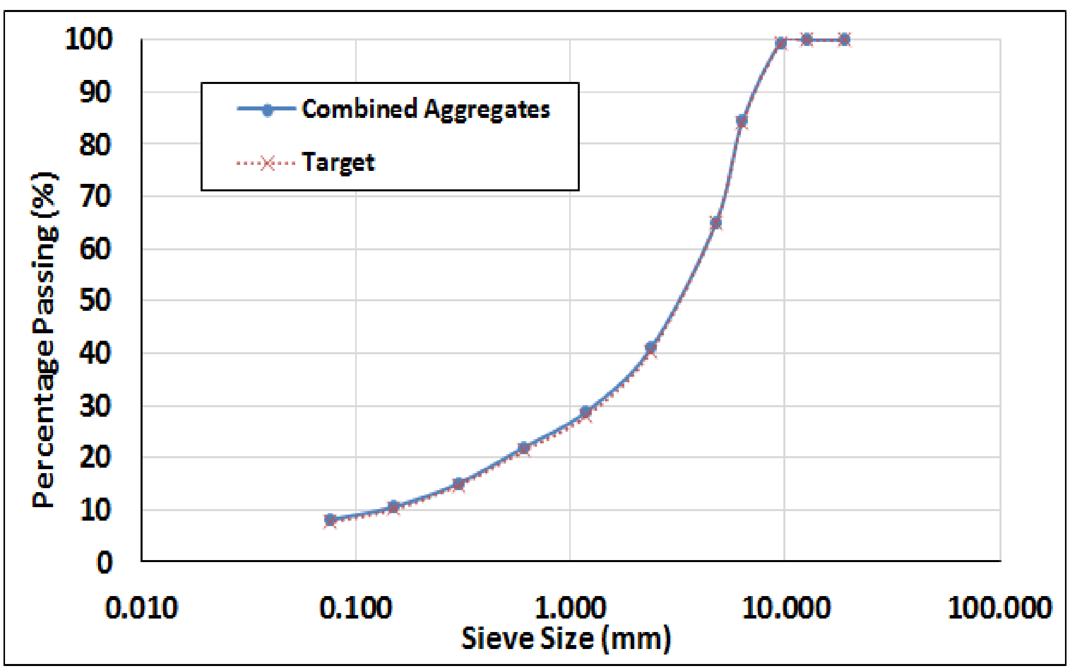


Figure 3.1: Gradation curve for Mix 0 (M0) obtained from the plant

3.2.2 Phase II – The impact of long-term aging on the fatigue cracking resistance of asphalt mixtures with different PG grades and RAP contents

In this phase, laboratory samples were produced using two binder grades (PG64-22 and PG 76-22) and three RAP contents (15%, 30%, and 40%) and the asphalt content was 6% for all mixtures. The purpose of this phase was to determine the impact of RAP content and asphalt binder type (PG grade) on fatigue cracking resistance of asphalt mixtures aged with different protocols in the laboratory. Theoretically, excessive aging is expected to converge the cracking resistance of high RAP asphalt mixtures and make quality evaluation of these mixtures harder. Since excessive mixture aging is expected to get the fatigue cracking resistance of asphalt mixtures closer, it might be possible to have high RAP mixes providing equal or better cracking resistance (due to the inherent variability of test results) than lower RAP mixtures for some of the evaluated cases.

In this phase, aggregate and RAP batches were prepared to reach the target gradation shown in Figure 3.2. Then, aggregates and RAP material were mixed with the asphalt binder to prepare the asphalt mixture. Loose mixtures were then subjected to different long-term aging protocols as shown in Table 3.2 after short-term conditioning for 4 hours at 135°C (AASHTO R 30). It should be noted that samples prepared by using PG76-22 grade binder were subjected to aging conditions 1, 5, 6, and 7 as shown in Table 3.2 while mixtures with PG64-22 binder were subjected to all 7 aging protocols during test specimen preparation.

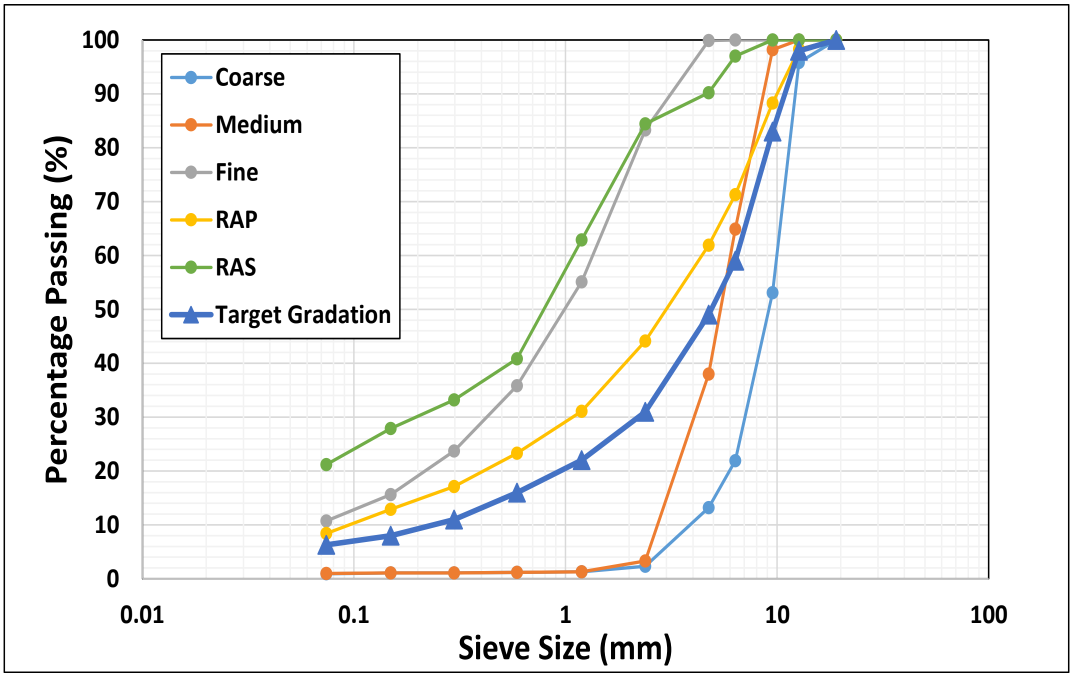


Figure 3.2: Target, extracted RAP, and stockpiled aggregate gradations

Table 3.2: Experimental Plan for Phase II – SCB Test Samples

Binder Grade	Aging Protocol[1]	Temp.	AC (%)	RAP (%)	Replicates	Total Tests
PG 64-22	All 7	25°C	6.0	15 30 40	4	84
PG 76-22	1,5,6,7	25°C	6.0	15 30 40	4	48

Note: [1] Long-term aging protocols: 1) 5 days at 85 °C; 2) 12 hours at 95 °C; 3) 24 hours at 95 °C; 4) 72 hours at 95 °C; 5) 6 hours at 135 °C; 6) 12 hours at 135 °C; and 7) 24 hours at 135 °C.

3.2.3 Phase III – The impact of long-term aging on the fatigue cracking resistance ranking of different plant produced mixtures

In this phase, five production mixtures with different mix designs and properties were sampled from different plants to evaluate the impact of long-term aging on the fatigue cracking resistance of asphalt mixtures with different binder types, additives, and RAP contents. In this phase, actual production mixtures designed with current volumetric procedures were used for the aging evaluation. The purpose of this phase was to determine how long-term aging is changing the fatigue cracking performance rankings between all five mixtures. If rankings were not significantly affected by the aging protocol, it might be possible to use an aging protocol with less duration of aging to be able to complete performance testing and evaluation in a shorter period. The experimental plan and the production mixture properties are given in Table 3.3.

Table 3.3: Mix Design Details and Experimental Plan for Mixtures Used in Phase III

ID[1]	Highway ID	Binder Grade	RAP [2] (%)	AC [3] (%)	RAM AC [4] (%)	P_{be} [5] (%)	P_{200}/P_{be} [6]	Additive	Aging [7]	Total tests
M1	I5	PG 70-22ER	20	5.6	4.63	4.74	1.2	1% Lime	5	20
M2	I84	PG 70-28ER	20	5.1	6.01	4.50	1.5	1% Lime	5	20
M3	US97	PG 70-28ER	20	5.6	5.80	4.72	1.6	1% Lime	5	20
M4	OR11	PG 64-28	30	5.5	5.36	4.71	1.5	0.0375% Latex	5	20
M5	US20	PG 64-28	30	5.6	5.52	4.59	1.5	0.0375% Latex	5	20

Note: [1] Nominal maximum aggregate size was 12.5mm for all five mixes; [2] RAP = Reclaimed asphalt pavement added by mass; [3] AC = Asphalt content added by mass; [4] RAM AC = Asphalt content by mass present in reclaimed asphalt material; [5] P_{be} = Effective asphalt content present by mass in total mix; [6] P_{200}/P_{be} = Dust to binder ratio present in total mix; [7] 1) 6 hours at 135 °C; 2) 12 hours at 135 °C; 3) 12 hours at 95 °C; 4) 24 hours at 95 °C; and 5) 72 hours at 95 °C.

3.3 MATERIALS AND ASPHALT MIXTURE PREPARATION AND CONDITIONING

3.3.1 Phase I

The production mix for this phase (M0) was obtained from a local asphalt plant in Portland, Oregon. Mix 0 (M0) was comprised of 3/8” nominal maximum aggregate size (NMAS) aggregates, 20% RAP and PG 70-22ER (polymer modified binder) grade virgin asphalt binder. The binder content of M0 was 6% by total weight. The gradation curve for M0 is presented in Figure 3.1.

Loose production mixes sampled from the plant were brought to the laboratory. With the help of a mechanical splitter, uniform sampling of the mix was performed. Theoretical maximum specific gravity (G_{mm}) of each mix type was measured to be able to determine the required amount of asphalt mixture to achieve 7% air-void content. Short-term aging was not used for plant-produced mixtures while different long-term aging protocols were followed to simulate long-term aging. After long-term aging, the required amount of asphalt mixtures for different samples were weighed out and again kept in the oven at the compaction temperature for 2 more hours. The mixing and compaction temperatures were obtained from viscosity versus temperature plots for the binder provided by the plants. Cylindrical samples were compacted using a Superpave Gyratory Compactor (SGC) in accordance with the AASHTO T312-12 specification. Then, SCB samples were sawn from those cylindrical SGC samples.

3.3.2 Phase II

Laboratory mixed and laboratory compacted samples were used in this phase of the study to evaluate the impact of RAP content and binder grade on the aging of asphalt mixtures. Table 3.2 shows the mixture properties and the experimental plan followed for this phase of the study. Virgin aggregates were donated by a local source. The virgin aggregates were delivered in three gradations, namely coarse (1/2” to #4), medium (#4 to #8), and fine (#8 to zero). To determine the gradation of each stockpiled aggregate, wet-sieve and dry-sieve analyses were performed on multiple samples of each stockpile following AASHTO T 27-14 (AASHTO 2014). RAP were also provided by the same source. AASHTO T 308-10 (AASHTO 2010) was followed for binder extraction and RAP binder content measurements. The quantity of binder in RAP materials was determined as 6.22%. AASHTO T 30-10 was followed to determine the gradation of extracted RAP and RAS aggregates.

A local producer provided the virgin binders with different binder grades (PG 64-22 and PG76-22) for this study. Temperature curves, mixing temperatures and compaction temperatures were provided by the producer as well. Laboratory mixing and compaction temperatures were estimated by using the viscosity-temperature lines following the procedure described in Asphalt Institute and AASHTO T 316-11 (AASHTO 2011).

Asphalt mixtures were prepared with 6% binder content in this study. This binder content is the percentage of the total binder by the weight of the mix, which includes the recycled binder from RAP and the virgin binder. In this study, it was assumed that all the RAP binder was completely blended with the virgin binder (100 % blending). Target gradation was obtained from an ODOT

Level 4 dense-graded mix design. All mixes were designed to reach the target gradation. Target gradation and the gradations of virgin aggregates and extracted RAP aggregates are presented in Figure 3.2.

For sample preparation, aggregates and RAP were batched to meet the final gradation and 7% target air-void content. Then, batched samples were mixed and compacted using the AASHTO T 312-12 procedure. Before mixing, aggregates were kept in the oven at 10ºC higher than the mixing temperature, RAP materials were kept at 110ºC, and binder was kept at the mixing temperature for 2 hours. After mixing, prepared loose mixtures were kept in the oven for 4 hours at 135ºC (AASHTO R 30) to simulate short-term aging. The goal of short-term aging is to simulate the aging and binder absorption that occurs during mixing and storage phases of the production process. Then, the loose mixtures were further aged at different temperatures and different durations to simulate long-term aging prior to compaction.

3.3.3 Phase III

In this phase, five long-term aging protocols were used to age five production mixtures with different binder grades, RAP contents, binder contents, and additives. Table 3.3 summarizes the mixture properties and the followed aging protocols. The major purpose of this phase was to determine the impact of different aging protocols on the fatigue cracking performance rankings of those five evaluated asphalt mixtures. The sample preparation process used for Phase I production mixtures was also used in this phase although the long-term conditioning protocols were not identical for some cases. Sampled production mixtures were assumed to be already short-term aged since they were obtained from the plant. However, loose mixtures were further subjected to five long-term aging protocols (6 hours and 12 hours at 135°C, and 12 hours, 24 hours, and 72 hours at 95°C) as shown in Table 3.3.

3.4 TEST METHODS

3.4.1 Semi-Circular Bend (SCB) Test

SCB tests were conducted in this study to determine cracking performance of asphalt mixtures after they were subjected to different long-term aging protocols. The test method for evaluating cracking performance of asphalt concrete at intermediate temperatures developed by Ozer et al. (2016) was followed with few modifications. A displacement rate of 0.5 mm/min was used instead of 50 mm/min since this slower loading rate was determined to provide results that are better correlated with actual field performance (Sreedhar et al. 2018).

3.4.1.1 Sample Preparation and Testing

130 mm tall samples were compacted in the laboratory according to AASHTO T 312-12. Two samples with the thickness of 57 ± 2 mm were cut from each gyratory compacted sample using a high-accuracy saw. Then, the cylindrical samples (cores) were cut into two identical halves.

Tests were conducted at 25˚C with a displacement rate of 0.5 mm/min. Samples were kept in the chamber at the testing temperature for conditioning the day before being tested. The flat side of semi-circular samples was placed on two rollers (see Figure 2.16). As vertical load with constant displacement rate is applied on the samples, applied load is measured. The test stops when the load drops below 0.5 kN. Flexibility Index (FI) is the parameter obtained from this test and used for cracking performance evaluation.

3.4.1.2 Flexibility Index (FI)

FI is the ratio of the fracture energy (G_f) to the slope of the line (m) at the post-peak inflection point of the load-displacement curve. FI correlates with ductility, and it was developed for asphalt materials by Ozer et al. (2016). Lower FI values show that the asphalt mixtures are more brittle with higher crack growth rate. Details regarding the calculation of FI from SCB test results is provided in Sreedhar et al. (2018) and in Section 2.3.1.

$$\boldsymbol{FI = A \times \frac{G_f}{abs(m)}} \tag{3-1}$$

Where:

A is a unit conversion and scaling coefficient taken as 0.01.

3.4.2 Binder Extraction and Recovery and DSR Testing

In Phase I, after SCB testing, binders from the tested samples were extracted and recovered by following the AASHTO T 319 (2015) standard. The purpose of extraction and recovery was to test and understand the rheological behavior of the aged binders. In order to determine the rheological properties, DSR testing was conducted with the recovered binders.

DSR tests were conducted with the extracted binder in order to obtain the complex shear modulus (G*) and phase angle (δ) for all extracted binders at various temperatures and loading frequencies (AASHTO T 315, 2012). These parameters describe the expected performance of the binders, such as the ductility and resistance to shear deformation when a load is applied (complex shear modulus), and the time lag between the peak shear stress that is applied to the sample and the resulting peak shear strain that is experienced by the sample (used to calculate phase angle). A higher phase angle indicates more viscous behaviour while a lower phase angle indicates more elastic behaviour. More viscous behaviour will tend to result in slower crack propagation (Harvey et al. 2014). For DSR testing, a frequency sweep test was conducted on all extracted binders at 20°C, 40°C, 60°C, and 70°C temperatures. Two replicate tests were conducted for each binder type and the average results of the two tests were used for the analysis. Selected loading frequencies were 0.0628 Hz, 0.628 Hz, 6.28 Hz, 62.8 Hz, and 94.2 Hz.

3.5 RESULTS AND DISCUSSION

3.5.1 Phase I - Field Aging versus Laboratory Aging

3.5.1.1 SCB Test Results for Production Mixtures Aged According to Different Protocols

SCB test results for plant mixed laboratory compacted mixtures (M0) subjected to six different aging protocols are presented in the Figure 3.3. It can be observed that mixtures aged for 24 hours at 135°C and for 10 days at 95°C yielded similar and lowest flexibility index (FI) values. It was also observed that for 135°C, decreasing the aging duration exponentially increases FI. For instance, the FI of asphalt mixture aged for 6 hours at 135°C was about 5. The same mixture aged for 12 hours at 135°C had a flexibility index of about 2.2 while the one aged for 24 hours had an average FI of 0.76. In addition, the mixture aged for 5 days at 85°C had almost two times higher FI than the mixture aged for 6 hours at 135°C. For this reason, shorter aging durations at higher temperatures will always be ideal to simulate long-term aging in a shorter time period. However, a few studies (Branthaver et al. 1993, Petersen 2009, Kim et al. 2018) have indicated that heating the binders above 100°C will create a significant change in the binder chemistry and may result in unrealistic long-term performance predictions. In addition, damaging the asphalt binder at excessively high aging temperatures (such as 135°C used in this study) during loose asphalt mixture aging can result in compactibility issues during test specimen production. Significantly higher compactive effort (higher number of gyrations in SGC) that might damage the aggregates and mastic phases in the asphalt mixture would be required to compact mixtures that were aged at excessively high temperatures.

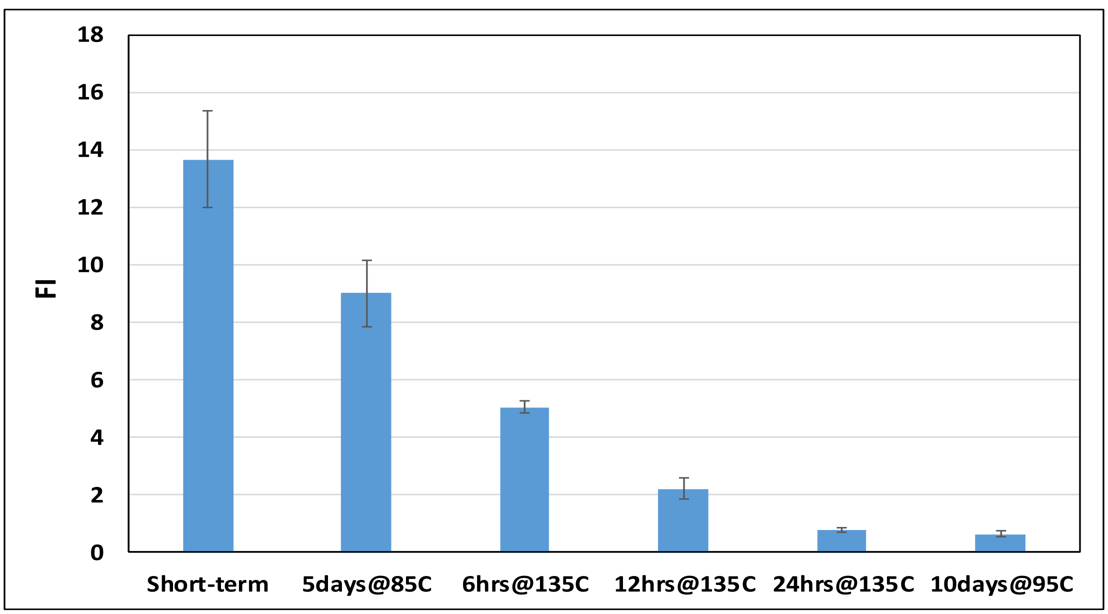
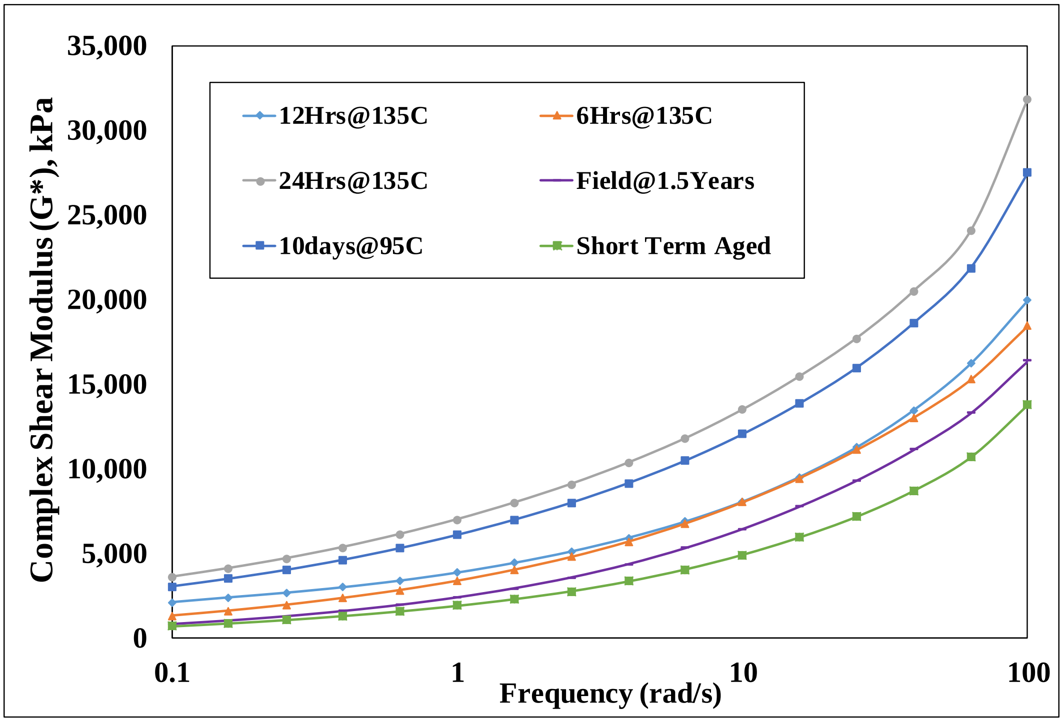


Figure 3.3: SCB test results for FMLC mixtures subjected to different aging protocols (error bar = 1 standard deviation)

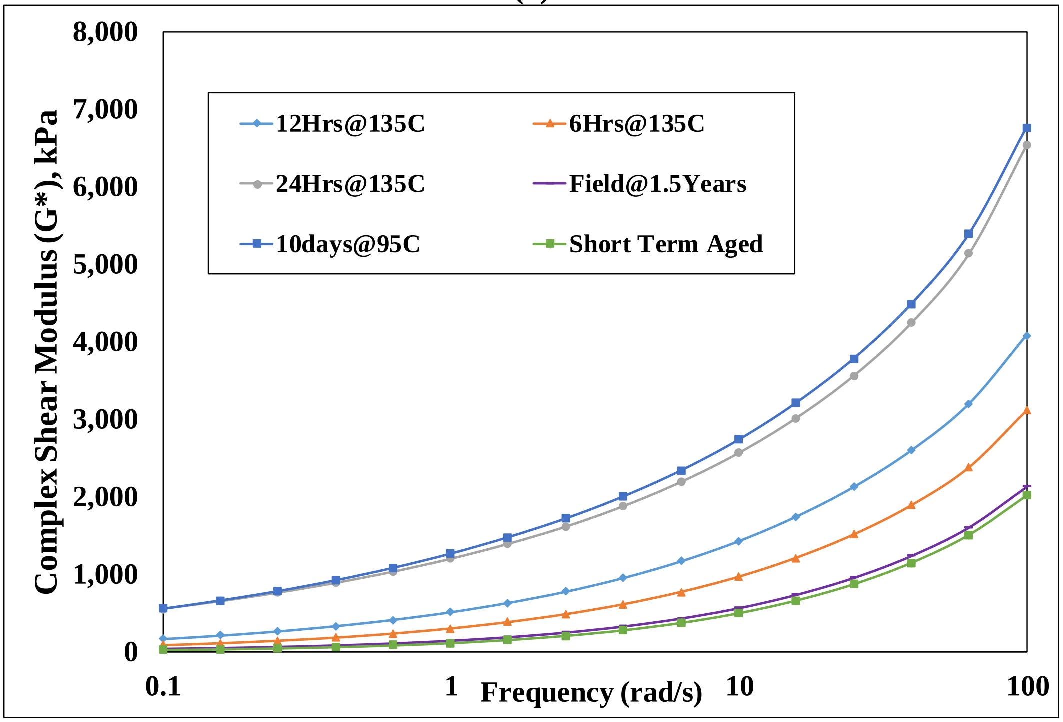
As stated above, there is a significant drop in FI when the aging protocol was changed from 5 days at 85°C to 6 hours at 135°C. This result suggested that increasing the temperature from 85°C to 135°C causes rapid stiffening of the binder and mixture. This stiffening results in excessively low FI values.

3.5.1.2 Direct Comparison of Field and Lab Aging – DSR Test Results for Extracted and Recovered Binders

Binder samples for DSR testing were extracted from the tested SCB samples to understand the rheological behavior of the production mixtures subjected to different aging protocols. Figure 3.4 and Figure 3.5 illustrate the shear modulus (G^*) of binders extracted from asphalt mixtures subjected to different aging protocols. Binders were tested at 20°C, 40°C, 60°C, and 70°C temperatures and with different loading frequencies. As expected, the binder subjected to only short-term aging was the softest. The binder aged for 24 hours at 135°C was the stiffest for all test temperatures except 40°C. However, G^* for the binder samples extracted from mixtures aged for 24 hours at 135°C and 10 days at 95°C were very close for all test temperatures and loading frequencies. This conclusion also agrees with the SCB test results shown in Figure 3.3. In general, the rankings of different aging conditions from DSR tests are in agreement with the FI values obtained from SCB tests.

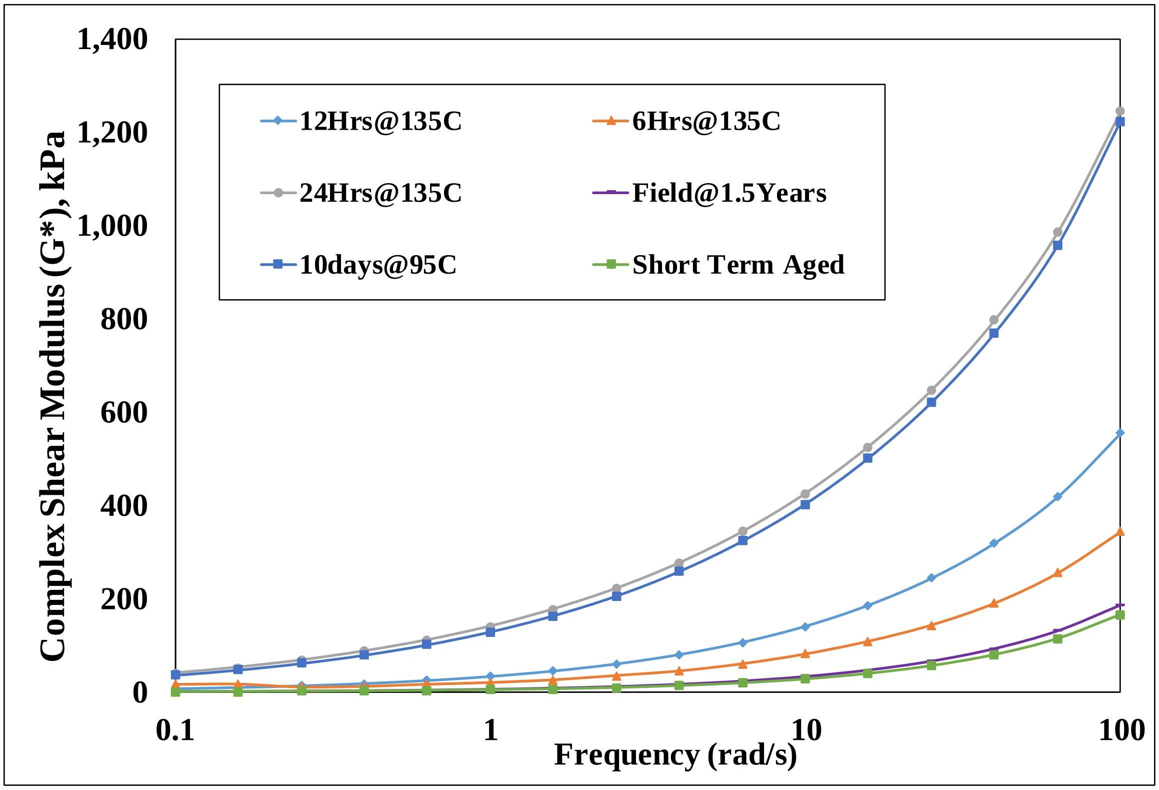


(a)

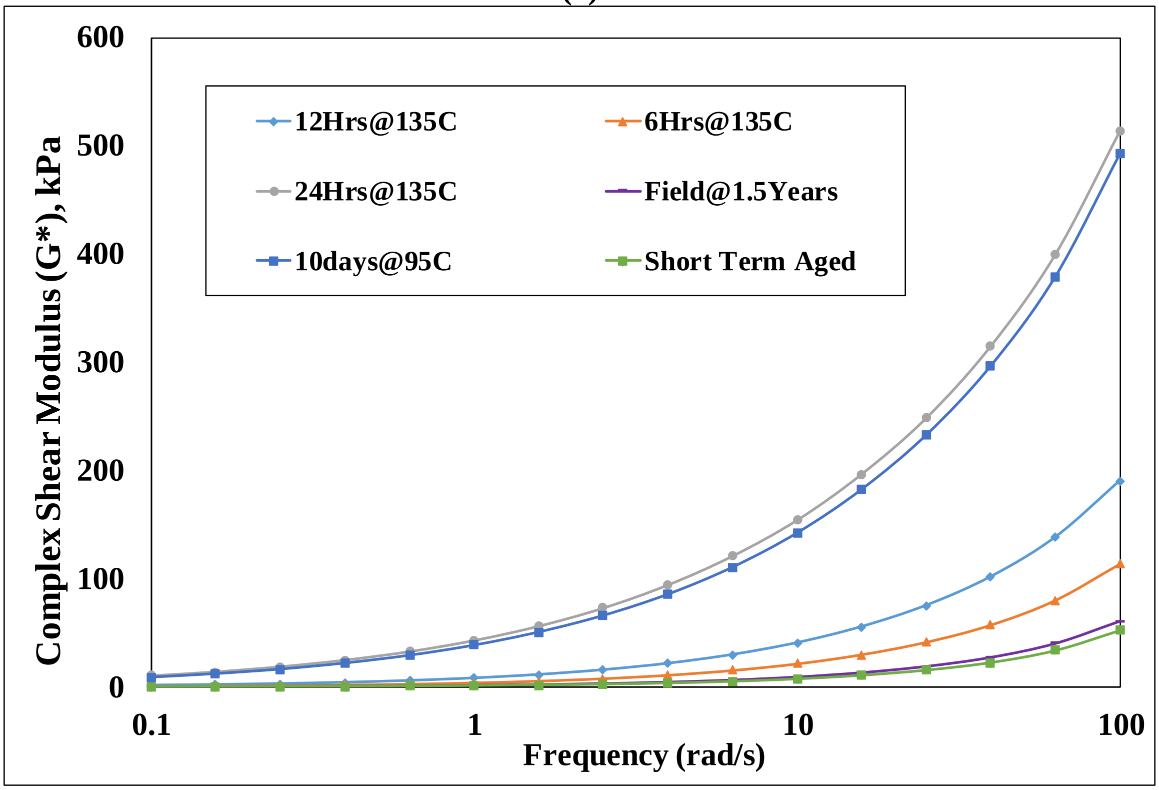


(b)

Figure 3.4: Complex shear modulus of FMLC mixtures subjected to different aging protocols (a) at 20°C, and (b) at 40°C



(a)



(b)

Figure 3.5: Complex shear modulus of FMLC mixtures subjected to different aging protocols (a) at 60°C, and (b) at 70°C

The binder extracted from the field core obtained after 1.5 years of construction had shear modulus values slightly greater than that of short-term aged binder but were significantly lower than the next aging protocol, which is 6 hours at 135°C. Therefore, an aging protocol that is in between short-term conditioning and 6 hours of oven aging at 135°C simulates 1.5 years of field aging. Based on the SCB results from the previous section, it can be concluded that long-term conditioning for 5 days at 85°C is resulting in aging levels close to 2-3 years of field aging. Also, it can be observed from DSR test results from all aging protocols that asphalt binders conditioned at 135°C may lead to higher levels of stiffening which might never be experienced in an actual field condition. A similar result was also observed when the mixture was aged at 95°C but for a significantly longer period (10 days). Therefore, in the subsequent phases of this study, additional lower temperature (95°C) cases were incorporated into our experimental plans to be able to develop an aging protocol that will not excessively age the asphalt mixture and result in unrealistic conclusions in terms of long-term in-situ performance. Some examples of these possible unrealistic conclusions can be: i) higher RAP mixtures providing better cracking resistance; and ii) mixtures with softer binders providing less ductility.

3.5.2 Phase II – The Impact of Long-Term Aging on the Fatigue Cracking Resistance of Asphalt Mixtures with Different Performance Grades and RAP Contents

To evaluate the impact of binder grade and RAP content on aging of asphalt mixtures, SCB test samples were prepared with two binder grades (PG64-22 and PG76-22) and three RAP contents (15%, 30%, and 40%) in the laboratory. Prepared loose mixtures were first subjected to 4 hours at 135°C (AASHTO R 30) for short-term conditioning followed by 5 days at 85°C, and 6 hours, 12 hours, and 24 hours at 135°C as long-term conditioning (four aging protocols). In addition, to understand how aging levels vary at lower temperatures (less than 100°C) as compared to higher temperatures (135°C or higher), additional mixtures prepared by using the PG64-22 binder were subjected to 12 hours, 24 hours, and 72 hours of aging at 95°C (three aging protocols). The samples prepared after the long-term conditioning were subjected to SCB tests and the results are given in Figure 3.6.

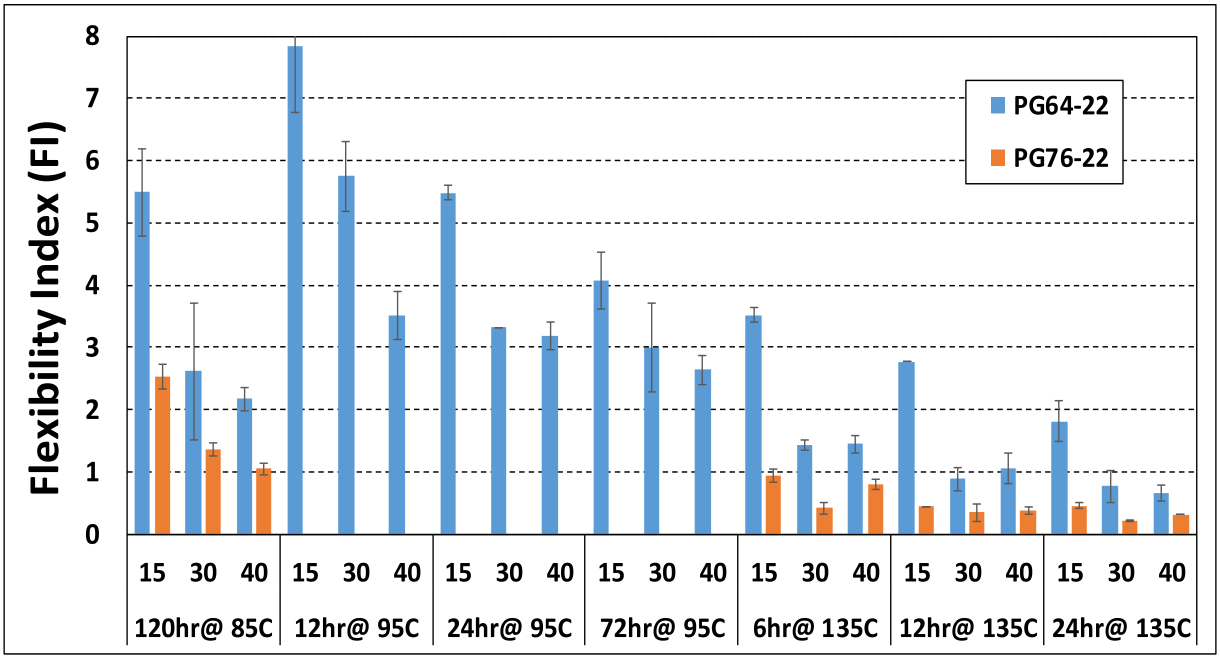


Figure 3.6: SCB test results for LMLC mixtures subjected to different aging protocols (error bar length = 1 standard deviation)

FI values for mixtures aged for 5 days at 85°C and 24 hours at 95°C were very similar indicating similar aging levels. Therefore, it is practical to have 24 hours at 95°C as an alternative for 5 days at 85°C. It was also observed that stiffer binder grade (PG76-22) mixtures were more susceptible to aging when compared to mixtures prepared with a softer binder grade (PG64-22). However, it should be noted that all the mixtures had RAP and that might be significantly controlling the aging rate.

Interestingly, there were no significant differences between the FI values of mixtures with higher RAP contents (30% and 40%) aged at 135°C. For 30% and 40% RAP mixtures with PG64-22 binder aged for 6 hours and 12 hours at 135°C, FI trends did not make sense. 40% RAP mixtures provided higher FI values than 30% RAP mixtures although the FI values very close to each other. A similar result was obtained for all aging protocols with 135°C aging temperature when the FI values for the mixtures with PG76-22 binder was evaluated. This result suggested that aging high RAP mixtures at 135°C creates excessive aging and equalizes fatigue cracking resistance and ductility of asphalt mixtures with 30% and 40% RAP. This result might also be a result of the changing binder chemistry at the 135°C aging temperature (Branthaver et al. 1993, Petersen 2009, Kim et al. 2018).

On the other hand, FI results for the mixtures aged at 85°C and 95°C were all reasonable. Increasing RAP content always expected to result in lower FI values, which is a result of increased stiff binder content (coming from the RAP material) in the mixture and limited blending of the binder around the RAP (Coleri et al. 2017a). In general, field sections with lower RAP mixtures have higher cracking resistance than the sections with higher RAP asphalt mixtures when the structure and traffic conditions are similar (Coleri et al. 2017a). It was also observed in this study that FI values for high RAP mixtures with PG76-22 binder conditioned at

135°C are extremely low (around 0.2-0.4), which is making performance comparisons harder. For these reasons, results of this phase suggested that using 135°C temperature for long-term aging might result in unreasonable results for high RAP and stiff binder mixtures. Cracking resistance of the mixtures conditioned with an aging protocol with 135°C may also not be correlated with the long-term field cracking performance.

3.5.3 Phase III – The Impact of Long-Term Aging on the Fatigue Cracking Resistance Ranking of Different Plant Produced Mixtures

In this phase of the study, the impact of aging on the cracking resistance rankings of different production mixtures was investigated. Theoretically, aging reduces the cracking resistance (measured by flexibility index (FI) from the SCB test in this study) of the asphalt mixture. It was observed in Phase I that increasing aging temperature and duration exponentially decreases measured FI values. The impact of aging duration and temperature on the cracking resistance rankings of production mixtures with different gradations, RAP contents, additives, modifiers, and binder types was investigated in Phase III. The major purpose of this phase was to determine whether aging protocols with shorter duration and lower temperatures provide the same cracking resistance rankings as the more severe aging protocols. Shorter aging durations are expected to make the use of long-term aging in balanced mix design and performance based specifications more practical. In addition, since the temperatures higher than 100°C were determined to change the binder chemistry according to previous research studies (Branthaver et al. 1993, Petersen 2009, Kim et al. 2018) and Phase II of this study also showed inconsistent results for high RAP mixtures aged at 135°C, the effectiveness of an aging protocol with 95°C temperature in providing reasonable cracking performance data was determined in this phase.

Figure 3.7 shows the results of the SCB tests conducted with five production mixtures aged with five different aging protocols. It can be observed that aging the mixtures for 12 hours at 135°C results in the highest level of reduction in FI. In addition, aging at that high temperature almost equalizes the FI values for all mixtures although they were expected to have significantly different mixture properties because of the differences in binder contents, binder types, gradations, additives, and RAP contents (see Table 3.3). For this reason and based on the findings from Phase II, this aging protocol is not selected in this study.

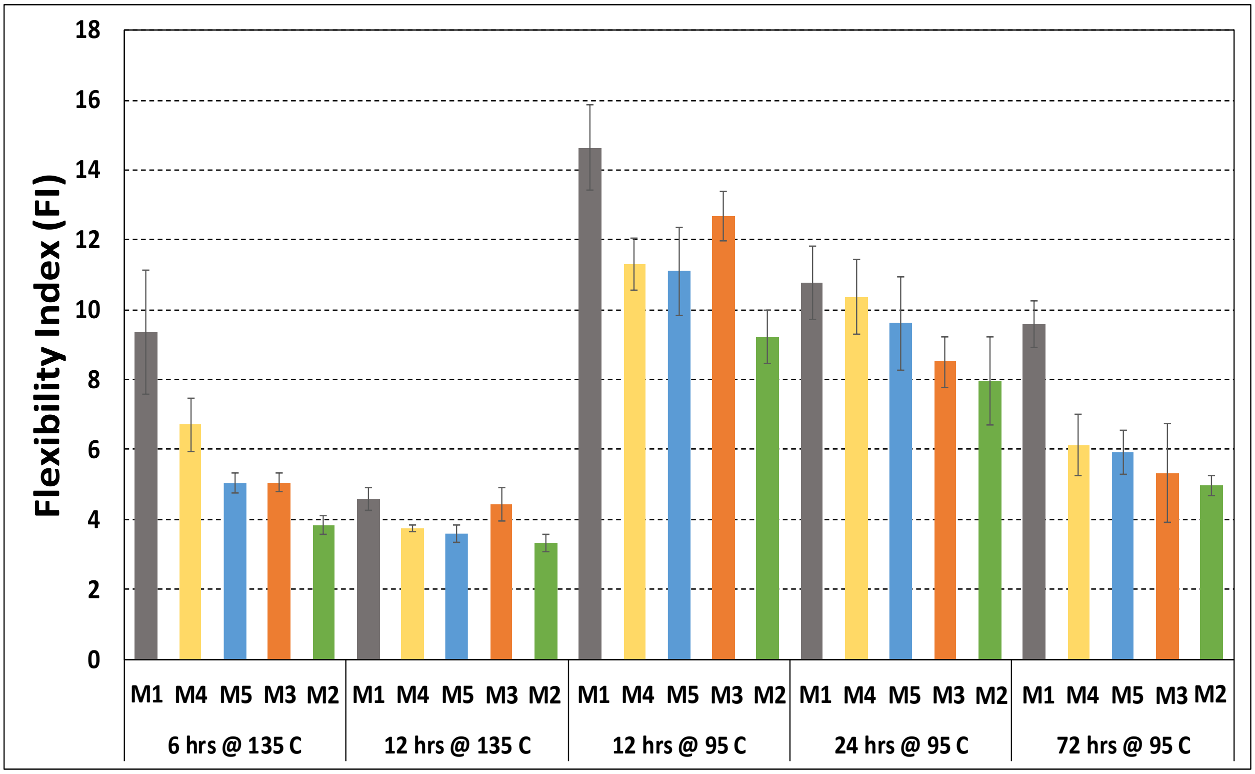


Figure 3.7: SCB test results for PMLC mixtures subjected to different aging protocols (error bar = 1 standard deviation)

The aging protocol with 6 hours of aging at 135°C provided almost equal FI values and performance rankings with the protocol with 72 hours of aging at 95°C. It can also be observed from Figure 3.7 that performance rankings for mixtures aged at 95°C for 24 hours and 72 hours are the same while the FI decreases with increased period of aging (as expected).

Based on all results from all three phases, the protocol with 24 hours of aging at 95°C was selected for the long-term aging protocol for balanced mix design and performance based specifications. Aging protocols with 135°C aging temperature were not selected based on the suggestions from previous studies and the inconsistent test results for high RAP mixtures in Phase II. The protocol with 72 hours of aging at 95°C was also not selected since it requires longer time for mixture preparation for design and performance evaluation testing provided the same performance rankings as the protocol with 24 hours of aging at 95°C. Thus, 24 hours of aging at 95°C was selected as the long-term aging protocol in this study. Although this protocol may not be simulating more than 5 years of aging in the field, it can still provide reasonable levels of aging that is going to provide cracking performance rankings correlated with the performance of mixtures aged for longer periods.

24 hours of aging also allows agencies to easily fit this protocol into their schedule for balanced mix design. Starting mixing in the morning, short-term aging can be completed early afternoon and the mix can be left in the oven at 95°C for 24 hours. Mix can be stirred couple of times during the 24 hour long-term aging period and it can be compacted next day early afternoon. Sample density measurements and cutting can be completed the next day.

3.6 SUMMARY AND CONCLUSIONS

This study focused on developing a long-term laboratory aging protocol for asphalt mixtures. Based on the results of this study, the protocol with 24 hours of aging at 95°C was selected as the long-term aging protocol for balanced mix design and performance based specifications in Oregon. To evaluate the impact of aging on plant and laboratory produced asphalt mixtures, this study was divided into three phases:

- **Phase I - Field aging versus laboratory aging:** In the first part, a plant produced asphalt mixture (M0) was subjected to six different aging conditions and SCB tests were conducted to determine the impact of aging on cracking resistance. Then, DSR test results from field core binders were compared with the DSR results from lab-produced samples that were subjected to different aging **protocols** to determine which aging protocol provides test results that are close to 1.5 year-aged asphalt mixes.

- **Phase II – The impact of long-term aging on the fatigue cracking resistance of asphalt mixtures with different PG grades and RAP contents:** In this phase, laboratory samples were produced using two binder grades (PG64-22 and PG 76-22) and three RAP contents (15%, 30%, and 40%) and the asphalt content was 6% for all mixtures. The purpose of this phase was to determine the impact of RAP content and asphalt binder type (PG grade) on fatigue cracking resistance of asphalt mixtures aged with different protocols in the laboratory.

- **Phase III – The impact of long-term aging on the fatigue cracking resistance ranking of different plant produced mixtures:** In this phase, five production mixtures with different mix designs and properties were sampled from different plants to evaluate the impact of long-term aging on the fatigue cracking resistance of asphalt mixtures with different binder types, additives, and RAP contents. In this phase, actual production mixtures designed with current volumetric procedures were used for the aging evaluation. The purpose of this phase was to determine how long-term aging is changing the fatigue cracking performance rankings between all five mixtures.

Results of the three phases yielded the following conclusions:

Phase I:

- Mixtures aged for 24 hours at 135°C and for 10 days at 95°C yielded similar FI values.

- For 135°C aging temperature, decreasing the aging duration exponentially increases FI.

- Changing the aging protocol from 5 days at 85°C to 6 hours at 135°C results in significantly lower FI values. This result suggested that increasing the temperature from 85°C to 135°C causes rapid stiffening of the binder and mixture.

- An aging protocol that is in between short-term conditioning and 6 hours of oven aging at 135°C simulates 1.5 years of field aging. Based on the SCB results from Phase I, it was concluded that long-term conditioning for 5 days at 85°C is resulting in aging levels close to 2-3 years of field aging.

- It was observed from DSR test results from all aging protocols that asphalt binders conditioned at 135°C may lead to higher levels of stiffening which might never be experienced in an actual field condition. A similar result was also observed when the mixture was aged at 95°C but for a significantly longer period (10 days).

Phase II:

- FI values for mixtures aged for 5 days at 85°C and 24 hours at 95°C were very similar indicating similar aging levels.

- Stiffer binder grade (PG76-22) mixtures were more susceptible to aging when compared to mixtures prepared with a softer binder grade (PG64-22). However, it should be noted that all the mixtures had RAP and that might be significantly controlling the aging rate.

- Aging high RAP mixtures at 135ºC creates excessive aging and equalizes fatigue cracking resistance and ductility of asphalt mixtures with 30% and 40% RAP. Recycled materials had a significant impact on the aging of asphalt mixtures and it is possible to have virgin and recycled mixtures present similar stiffness values by excessive aging. This result might also be a result of the changing binder chemistry at the 135ºC aging temperature (Branthaver et al. 1993, Petersen 2009, Kim et al. 2018). For these reasons, using 135ºC temperature for long-term aging might result in unreasonable results for high RAP and stiff binder mixtures.

- For high RAP mixtures, FI results for the mixtures aged at 85ºC and 95ºC were all reasonable (higher RAP results in lower FI values).

Phase III:

- For the tested production mixtures, aging the mix at 135ºC for 12 hours almost equalizes the FI values for all mixtures although they were expected to have significantly different mixture properties because of the differences in binder contents, binder types, gradations, additives, and RAP contents.

- The aging protocol with 6 hours of aging at 135ºC provided almost equal FI values and performance rankings with the protocol with 72 hours of aging at 95ºC.

- Performance rankings for mixtures aged at 95ºC for 24 hours and 72 hours are the same.

Based on all results from all three phases, the protocol with 24 hours of aging at 95°C was selected for the long-term aging protocol for balanced mix design and performance based specifications. Although this protocol may not be simulating more than 3-5 years of aging in the field, it can still provide reasonable levels of aging that is going to provide cracking performance rankings correlated with the performance of mixtures aged for longer periods.

4.0 DEVELOPING PERFORMANCE-BASED SPECIFICATIONS FOR ASPHALT MIXTURE DESIGN IN OREGON

4.1 INTRODUCTION

Asphalt mixtures are designed to be used in pavements to withstand vehicular loads under different climatic conditions. The goal of asphalt mix design is to determine an economic blend of aggregates and binder such that the resultant mix provides sufficient stability to resist deformation under traffic loading, and flexibility to withstand cracking. The most commonly known asphalt mix design methods are the Marshall, Hveem, and Superpave methods. Marshall and Hveem mix design procedures were widely used until the early 1990s before Superpave procedure was introduced. Superpave was developed as part of the Strategic Highway Research Program (SHRP) and was implemented in 1993. The original objective was to develop a performance-based mix design process. Although performance tests for asphalt mixtures were a part of the Superpave mix design process and several procedures were developed to predict and evaluate mixture performance, the entire process turned out to be too complex and costly and was never implemented by any state Department of Transportation (DOT). Superpave mix design had three levels (Level 1, Level 2, and Level 3) with increasing complexity (Cominsky et al. 1994). The performance-based specifications were to be incorporated in Level 2 and Level 3 designs but were never implemented.

The current asphalt mix design practice (Level 1) involves proportioning of the aggregates and the asphalt binder based on empirical properties of aggregates and volumetric properties such as densities, air voids, voids in the mineral aggregate (VMA) and voids filled with asphalt (VFA). However, most state DOTs and asphalt contractors do not think that commonly used asphalt mixture properties are reflecting the long-term performance of asphalt mixtures. For instance, although there are requirements for VMA set by almost all state DOTs, measurement of VMA relies on the accurate measurement of aggregate bulk specific gravity, while considerable issues were observed in terms of accuracy and variability during the measurement of this parameter (West et al. 2018). In addition, there are several new additives, polymers, rubbers, and high-quality binder types incorporated into asphalt mixtures today. Volumetric mixture design methods are not capable of capturing the benefits of using all these new technologies on asphalt mixture performance. Furthermore, the interaction of virgin binders with reclaimed asphalt pavement (RAP) mixtures with high binder replacement contents and the level of RAP binder blending into the asphalt mixture are still not well understood. Due to all these complications related to the more complex structure of asphalt mixtures today, simple volumetric evaluations to determine the optimum binder content may not result in reliable asphalt mixture designs. Two volumetrically identical mixtures may provide completely different rutting and cracking performance according to laboratory tests (Coleri et al. 2017b). For all these reasons, performance tests for rutting and cracking need to be incorporated into current asphalt mixture design methods to be able to validate or revise the optimum binder content determined by the volumetric mix design method. Numerous research studies were recently carried out and are currently being conducted to develop new mix design processes with performance verification (Epps et al. 2002; Zhou et al. 2006; Harvey et al. 2014; Cooper et al. 2014; Williams et al. 2004; Bennert et al. 2014; Dave et al. 2011; Zhou et al. 2014). However, this approach is not entirely new and draws upon the existing methods and procedures while the existing methods need to be revised and improved by incorporating findings from recent research studies.

Oregon Department of Transportation (ODOT) Research Projects SPR785 and SPR797 (Coleri et al. 2017b; Coleri et al. 2017a; Sreedhar et al. 2018; Haddadi et al. 2019) constructed the beginnings of a performance-based balanced mix design method for Oregon. It was suggested that semi-circular bend (SCB) test is the most effective and practical cracking test that can effectively be used for balanced mix design. It was determined that the typical flexibility index (FI), an energy parameter calculated by using SCB test results, values for production mixtures (plant-produced) range from 9 to 14. However, more experiments need to be conducted to determine an exact threshold for FI that will provide acceptable long-term pavement cracking performance. In these two research projects, flow number (FN) test was used as the experiment for rutting performance evaluation. For highways with high traffic levels (ESALs > 30 million), an FN of 740 was suggested by AASHTO TP79-15 (2015) and used in SPR785 and SPR797 as the threshold value for rutting performance acceptance. However, FI and FN threshold numbers used in these two research projects were not validated using measured field performance. In addition, mix design strategies suggested in SPR797 (high RAP mixtures with different binder contents and binder types) were not validated by plant sampling and laboratory testing. The major objectives of this part of the study are to:

- determine reliable threshold values for FI, FN, and Hamburg Wheel Tracking Test (HWTT) rut depths *[for Level 3 (medium traffic) and Level 4 (highest traffic) mixtures commonly used for pavement construction in Oregon]* for balanced mix design and performance evaluation;
- determine the most effective rutting experiment for balanced mix design by evaluating and comparing results from flow number and HWTT tests; and
- develop a balanced asphalt mix design method for Oregon by incorporating performance tests for rutting and cracking into the current volumetric design process.

4.2 MATERIALS AND SAMPLE FABRICATION

This section provides information about the materials sampled from six different plants in Oregon for this study (including virgin binders, virgin aggregates and RAP materials). The state of Oregon is divided into five different regions by ODOT. At least one construction project was selected from each region to have a representative distribution of aggregate types, plants, and binder sources. All the materials for this study were sampled from asphalt plants located near their respective construction projects. In this study, two types of asphalt samples were used for testing and evaluation:

- *Plant Mixed-Laboratory Compacted (PMLC) samples:* Before construction, loose asphalt mixtures were collected from the asphalt plant of that particular project to prepare PMLC samples in the laboratory.

- *Laboratory Mixed-Laboratory Compacted (LMLC) samples:* Aggregates, virgin binders and RAP material used to produce asphalt mixtures for field construction were sampled from asphalt plants of each project. These materials were used to produce LMLC samples at the Asphalt Materials Performance Laboratory at Oregon State University.

The mix design variables of the eight selected projects considered in this study are summarized in Table 4.1. Location of the construction projects are shown in Figure 4.1. It can be observed that eight projects evaluated in this research study were from several different districts across Oregon. These projects varied in binder grade, binder source and content, aggregate source and gradation, RAP source and gradation, and amount of RAP content. For each mix type, constituents of the production mixtures were identical to the aggregates, RAP, and binder sampled from the plant for laboratory sample preparations. Laboratory produced mixtures replicated field mix design variables except they had four different binder contents: design binder content, -0.5%, +0.5%, and +1%. It should be noted that it is possible to achieve exact binder contents in the laboratory by using high accuracy scales. However, plant produced mixtures are allowed to have ±0.5% variability in production binder content. ODOT is currently in the process of changing this tolerance to ±0.35% to improve accuracy and precision for the binder content of the final product. It should also be noted that Level 3 ODOT mixtures were designed by finding the required optimum binder content to reach 4% air-void content after 80 gyrations of the Superpave Gyratory Compactor (SGC) while Level 4 ODOT mixtures were design for 100 gyrations. Gradation curves for all eight projects are given in Figure 4.2. Level 4 mixtures are generally used for constructing highway sections with high traffic levels while Level 3 mixtures are used for constructing roadways with relatively lower truck traffic levels.

Table 4.1: Mix Design and Production Mixture Variables for Plant Mixed Field Compacted (PMFC) Samples

ID	Highway ID	Mix Design Level	Binder Grade	RAP [b] (%)	AC [c] (%)	RAM AC [d] (%)	P_{be} [e] (%)	P_{200}/P_{be} [f] Ratio	Addi. [g]	VMA[j]-VFA[k] %	TSR[l] %
M1 [a]	I5	Level 4	PG 70-22ER	20	5.6	4.63	4.74	1.2	1% Li[h]	15.5-75	94
M2	I84	Level 4	PG 70-28ER	20	5.1	6.01	4.50	1.5	1% Li	14.9-73	83
M3	US97	Level 4	PG 70-28ER	20	5.6	5.80	4.72	1.6	1% Li	14.9-74	86
M4	OR11	Level 3	PG 64-28	30	5.5	5.36	4.71	1.5	0.0375% La[i]	15.2-74	94
M5	US20	Level 3	PG 64-28	30	5.6	5.52	4.59	1.5	0.0375% La	14.6-73	86
M6	OR569	Level 4	PG 70-22ER	20	5.6	4.8	4.99	1.4	-	15.4-74	83
M7	I205	Level 4	PG 70-22ER	20	5.2	4.9	4.38	1.4	-	14.3-72	89
M8	OR38	Level 3	PG 64-22	30	5.4	5.92	4.94	1.28	-	15.3-74	90

[a] All mixtures had dense gradation and aggregates with a nominal maximum aggregate size of 12.5mm;
[b] RAP = Reclaimed asphalt pavement added by weight;
[c] AC = Asphalt content added by weight;
[d] RAM AC = Asphalt content by mass present in reclaimed asphalt material;
[e] P_{be} = Effective asphalt content present by weight in the total mix;
[f] P_{200}/P_{be} = Dust to binder ratio in the mix;
[g] Addi. = Additive; [h] Li = Lime; [i] La = Latex; [j] VMA = Voids in mineral aggregate; [k] VFA = Voids filled with asphalt; [l] TSR = Tensile strength ratio.

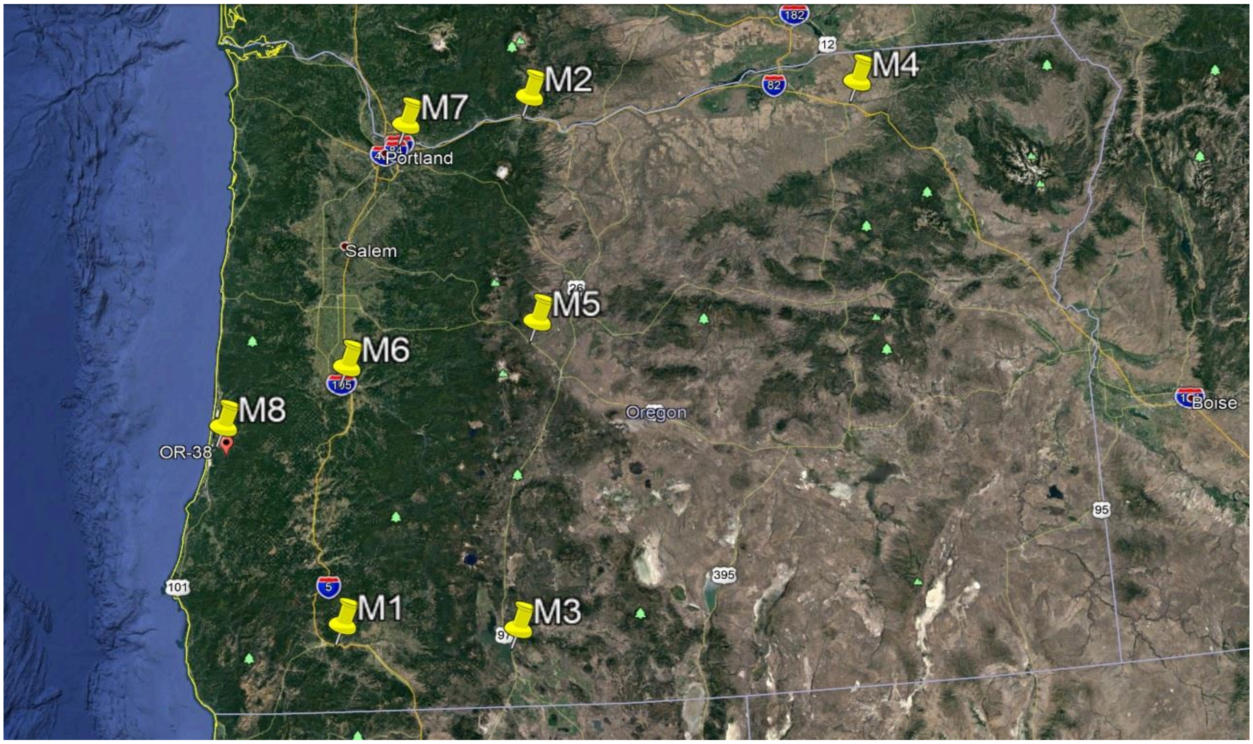


Figure 4.1: Approximate construction project locations across Oregon

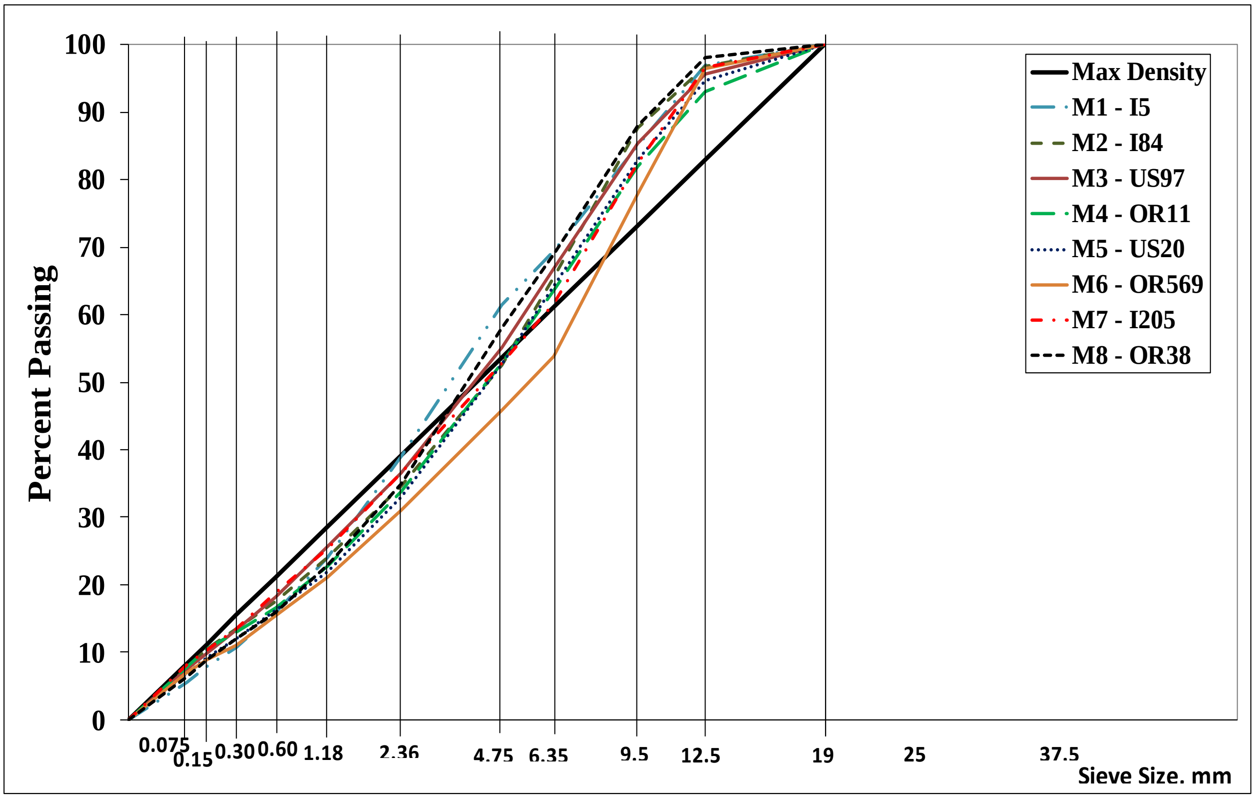


Figure 4.2: Gradation curves for asphalt mixtures from all 8 construction projects on a 0.45 power chart

4.2.1 Preparation of LMLC Specimens

For sample preparation, aggregates and RAP were batched to meet the final gradation and 7% ± 1% air content. Then, batched samples were mixed and compacted by following the AASHTO T 312-12 (2012) specification. Before mixing, aggregates were kept in the oven at 10°C higher than the mixing temperature, RAP materials were kept at 110°C, and binder was kept at the mixing temperature for 2 hours. After mixing, the AASHTO R 30 (2010) recommends conditioning the prepared loose mixtures for 4 hours at 135°C to simulate short-term aging (STA). The goal of short-term aging is to simulate the aging and binder absorption that occurs during the production and silo storage phases. However, based on the suggestions from the NCHRP 815 (Newcomb et al. 2015), a short-term conditioning period of 2 hours at 135°C was adopted. The long-term aging protocol developed for Oregon in Chapter 3.0 was followed for conditioning asphalt mixtures for the SCB cracking tests. Based on the results and recommendations from Chapter 3.0, short-term aged loose mixtures were further aged at 95°C for 24 hours to simulate long-term aging. The conditioning was carried out in a forced draft oven and mixtures were stirred at regular intervals to ensure uniform aging. After LTA conditioning, mixtures were further kept in the oven at compaction temperature for 2 more hours prior to compaction. The mixing and compaction temperatures were obtained from viscosity versus temperature plots for the binder provided by the plants. Cylindrical samples were compacted using a Superpave Gyratory Compactor (SGC) in accordance with the AASHTO T312-12 specification. Asphalt mixtures used for FN and HWTT sample production were only short-term aged (no long-term aging) since rutting generally occurs early in the design life. Asphalt mixtures for only SCB samples were long-term aged to simulate the impact of aging (oxidation and volatilization of different components in the asphalt binder) on long-term cracking resistance.

4.2.2 Preparation of PMLC Specimens

Loose production mixtures sampled from different plants across Oregon were used for PMLC specimen production in the laboratory (See Figure 4.1). With the help of a mechanical splitter, uniform sampling of the mix was carried out in the laboratory. Theoretical maximum specific gravity (G_{mm}) of each mix type was obtained from the plant to be able to determine the required weight of asphalt mixture to achieve 7% air-void content. The required amount of mix to achieve 7% air-void content for different samples was weighed out. Since production mixtures were already exposed to short-term aging during plant mixing and silo storage, short-term conditioning in the laboratory was not performed. For the cracking experiments, loose mixtures were further subjected to long-term aging of 24 hours at 95 ºC. After long-term conditioning, asphalt mixtures for SCB test sample preparation were again kept in the oven at the compaction temperature for 2 more hours before compaction. For the flow number and HWTT (rutting) experiments, after weighing out the amount of mixture required for the samples, asphalt mixtures were kept in the oven at the compaction temperature for 2 hours without performing long-term aging before compaction.

4.3 TEST METHODS

4.3.1 Semi-Circular Bend (SCB) Test

In a previous research study performed at Oregon State University (Coleri et al. 2017b), semi-circular bend (SCB) test was selected as the most effective cracking experiment to characterize asphalt mixtures used in Oregon (Sreedhar et al. 2018). Therefore, SCB tests were conducted in this study to determine the cracking resistance of asphalt mixtures and to determine a suitable threshold for the test's output parameter (flexibility index) to be used as an acceptance criterion in the proposed balanced asphalt mixture design process. Test method for evaluating the cracking performance of asphalt concrete at intermediate temperatures developed by Ozer et al. (2016) was followed with few modifications. A displacement rate of 0.5 mm/min was used instead of 50 mm/min (Sreedhar et al. 2018, Coleri et al. 2017b).

130 mm tall samples were compacted in the laboratory according to AASHTO T 312-12. Two samples with the thickness of 57 ± 2 mm were sawn from each gyratory compacted sample using a high-accuracy saw. Then, cylindrical samples (cores) were cut into two identical halves using a special jig. Tests were conducted at 25ºC with a displacement rate of 0.5 mm/min. Samples were kept in the chamber at the testing temperature for conditioning the day before being tested. Flat side of the semi-circular samples was placed on two rollers (See Figure 2.16). As a vertical load with constant displacement rate is applied to the samples, applied load is measured via a load cell. Test stops when the load drops below 0.5 kN. Flexibility index (FI) is the testing parameter obtained from this test and used for cracking resistance evaluation.

Flexibility Index (FI) is the ratio of the fracture energy (G_f) to the slope of the line (m) at the post-peak inflection point of the load-displacement curve. FI correlates with ductility. Lower FI values show that the asphalt mixtures are more brittle with the higher crack growth rate.

$$\boldsymbol{FI = A \times \frac{G_f}{abs(m)}} \tag{4-1}$$

Where:

A is a unit conversion and scaling coefficient taken as 0.01.

4.3.2 Flow Number (FN) Test

Flow number (FN) test is a performance test for evaluating the rutting resistance of asphalt concrete mixtures (Bonaquist et al. 2003). In this test, while constant deviator stress is applied at each load cycle on the test sample, permanent strain at each cycle is measured. Permanent deformation of asphalt pavements has three stages: 1) primary or initial consolidation, 2) secondary, and 3) tertiary or shear deformation (Biligiri et al. 2007). FN is the loading cycle at which the tertiary stage starts after the secondary stage.

Three 170 mm tall samples were compacted in the laboratory according to AASHTO T 312-12. Cylindrical cores of 100 mm diameter were cored from each gyratory compacted sample using a core drill equipment. In this study, testing conditions and criterion for FN testing described in AASHTO TP 79-15 for unconfined tests were followed. Recommended test temperature of 54.7°C, determined by LTPPBind Version 3.1 software, is the average design high pavement temperature at 50% reliability for cities in Oregon with high populations and at a depth of 20 mm (0.79 in) for surface courses (Rodezno et al. 2015; Coleri et al. 2017b). Tests were conducted at a temperature of 54.7°C with average deviator stress of 600 kPa and minimum (contact) axial stress of 30 kPa. For conditioning, samples were kept in a conditioning chamber at the testing temperature for 6 hours prior to being tested. To calculate FN in this study, the Francken model was used (Francken 1977) (See Section 2.3.2.1 for details).

4.3.3 Hamburg Wheel-Tracking Test (HWTT)

The Hamburg Wheel-Tracking Test (HWTT) system was developed to measure rutting and moisture damage (stripping) susceptibility of an asphalt concrete sample. The HWTT follows the AASHTO T 324 standard. According to the specification, either a slab or a cylindrical specimen can be tested. Tests are conducted by immersing the asphalt concrete sample in a hot water bath (at 40°C or 50°C) and rolling a steel wheel across the surface of the sample to simulate vehicular loading (See Figure 2.19). Approximately 20,000 wheel passes are commonly used to evaluate the rutting and stripping resistance of a sample. The test provides information related to the total rut depth, post-compaction, creep slope, stripping inflection point and stripping slope of the asphalt concrete sample (Yildirim et al. 2007; Tsai et al. 2016). In this study, rut depth after 20,000 wheel passes is used for rutting performance evaluation. Cylindrical specimens were used for testing. In this study, selected test temperature for HWTT was 50°C.

4.4 EXPERIMENTAL DESIGN

This study was performed to determine reliable thresholds for performance parameters (FN and HWTT rut depth for rutting and FI for cracking) and thereby, develop a performance based specification for asphalt mixture design. In order to achieve this objective, virgin materials and production mixtures were obtained from eight different construction projects spread across the state of Oregon (See Figure 4.1). Flow number (FN) and Hamburg Wheel-Tracking Tests (HWTT) were selected as performance tests for rutting. SCB test was used to quantify the cracking performance of the asphalt mixtures. General experimental plan followed in this study is given in Table 4.2. A total of 440 laboratory experiments were conducted for the balanced mix design portion of this study.

Table 4.2: Experimental Plan to Develop a Balanced Mix Design Method

Specimen Type [a]	Mix ID [b]	Test	Temperature (°C)	Asphalt Content (%)	Replicates	Total Tests
LMLC	M1 – M8	SCB	25.0	OBC [c], - 0.5%, + 0.5%, + 1%	4	128
		FN	54.7		3	96
		HWTT	50.0		4	128
PMLC	M1P – M8P	SCB	25.0	OBC	4	32
		FN	54.7		3	24
		HWTT	50.0		4	32

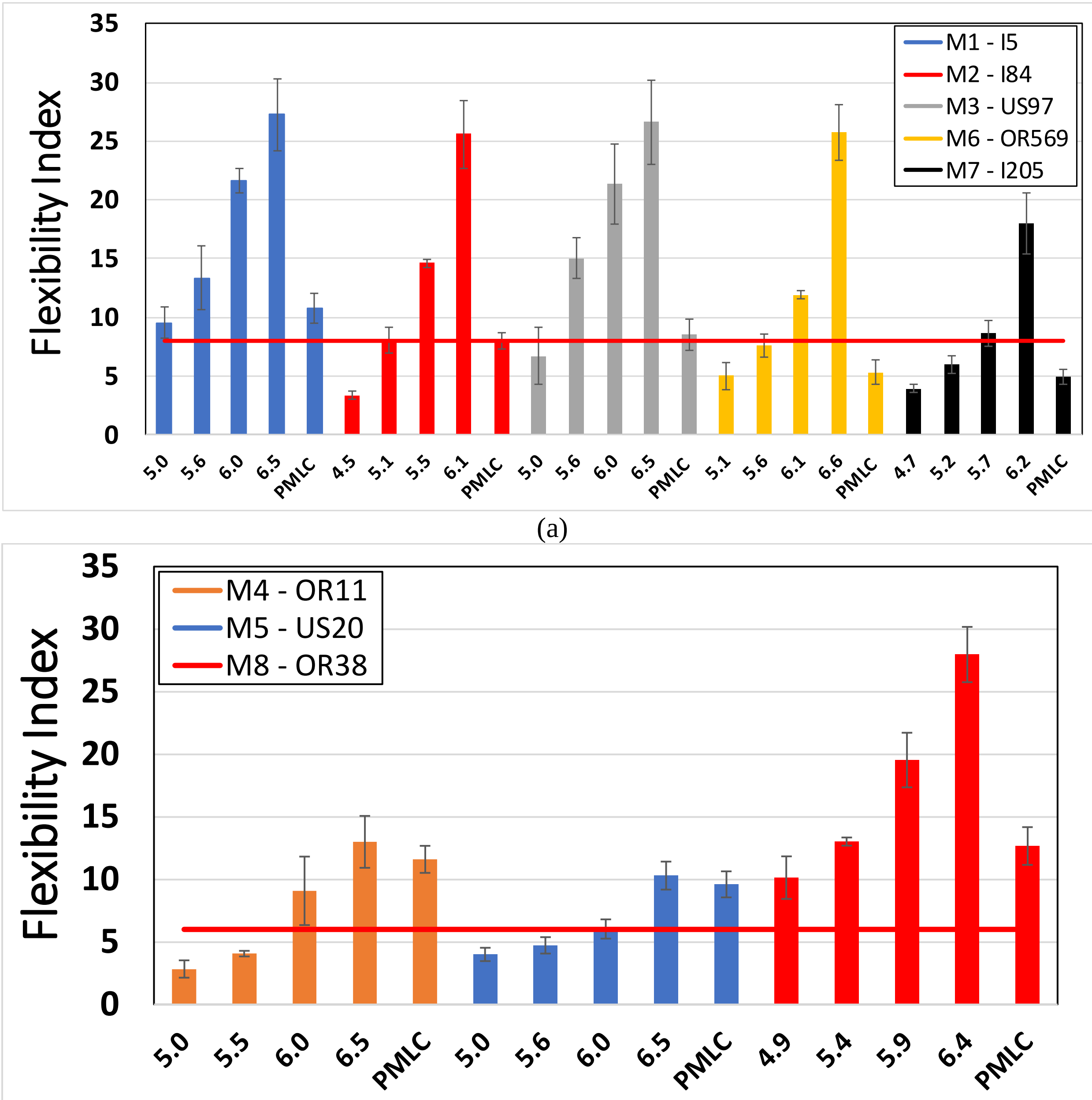
a LMLC = Laboratory mixed, and laboratory compacted; PMLC = Plant mixed, and laboratory compacted.
b M1 – M8/M1P – M8P = LMLC and PMLC samples from eight construction projects as described in Table 4.1.
c OBC = Optimum binder content obtained from volumetric mix design.

4.5 RESULTS AND ANALYSES

Materials for the eight selected projects spread across the State of Oregon (see Table 4.1 and Figure 4.1) were collected, mixed, and compacted to produce test specimens. Four different asphalt contents (AC) were used for each project for balanced mix design: AC_{design} from volumetric mix design, AC_{design}-0.5%, AC_{design}+0.5%, and AC_{design}+1%. Flow number (FN) and Hamburg Wheel-Tracking Tests (HWTT) were used to determine rutting performance of asphalt mixtures. One of these two rutting experiments was recommended as the test for balanced mix design at the end of this research project. SCB test was used to quantify the cracking performance of the asphalt mixtures. Four replicate tests were conducted for SCB tests while three replicate tests were conducted for FN testing and four replicate tests (four core samples with two rut depth measurements) were conducted for HWTT.

4.5.1 SCB Test Results

Figure 4.3 presents the results of tests for cracking (SCB) performance. FI was calculated and used to evaluate the cracking performance of all asphalt mixtures. The process followed to calculate the FI from laboratory SCB test results is described in Section 2.3.1. The horizontal red lines in Figure 4.3 are the FI thresholds selected in this study for Level 3 ($FI_{threshold}$=6) and Level 4 ($FI_{threshold}$=8) mixtures.



(a)

(b)

Figure 4.3: FI test results for LMLC and PMLC specimens (a) Level 4 mixtures (b) Level 3 mixtures (length of the error bar is equal to one standard deviation)

It can be observed from Figure 4.3 that increasing binder content increases Flexibility Index (FI) for all cases, as expected. FI is able to capture the impact of increased binder content on cracking resistance. It should be noted that projects M1, M2, M3, M6 and M7 had Level 4 mixtures (designed with 100 gyrations –SCB test results shown in Figure 4.3a) whereas, M4, M5 and M8 are Level 3 mixtures (designed with 80 gyrations – SCB test results shown in Figure 4.3b).

For Level 4 mixtures, average FI values showed that asphalt mixtures from projects M1, M2, M3 and M6 provided higher FI values than the mixture from M7 project. In Figure 4.3, the second bar for every mixture type shows the FI value for the LMLC samples prepared at the design binder content. It can be observed that M1 and M3 asphalt mixtures had cracking resistances significantly higher than all other Level 4 mixtures. Higher cracking resistance for the mixture from these two projects is likely to be a result of the relatively higher design binder content for these two projects (both at 5.6%). Significantly lower dust-to-binder ratio (see Table 4.1) and high VFA for M1 is probably the major reason for significantly higher cracking resistance. It should also be noted that although M2 had a binder content significantly lower than M1 and M3, it showed cracking performance comparable to M1 and M3. This high cracking resistance for M2 is a result of the high effective binder content and VFA that are close to M1 and M3 mixtures. Test results for the Level 4 PMLC specimens showed that M1, M2, and M3 projects meet the FI threshold of 8 while the PMLC mixtures for M6 and M7 projects fail to meet the requirement with average FI values close to 5. Lower FI value for project M7 can be explained by the relatively lower binder content for this asphalt mixture. The effective binder content and the VFA for the M7 asphalt mixture were lower than all other four Level 4 mixtures evaluated in this study.

Coleri et al. (2017b) concluded that laboratory-mixing process is properly simulating the plant mixing process. SCB test results for Level 4 mixtures in this research study also provided the same conclusion. In Figure 4.3a, it can be observed that average measured FI values for all PMLC specimens from all Level 4 mixtures fall between the test results for LMLC specimens for AC_{design}-0.5% and AC_{design}+0.5%. Since plant produced mixtures are allowed to have ±0.5% variability in production binder content in Oregon, it can be concluded for Level 4 mixtures that laboratory mixing process is not introducing any bias into the measured FI values.

It can be observed from Figure 4.3b that FI values for the Level 3 PMLC mixtures are higher than the AC_{design}+0.5% for projects M4 and M5 while FI falls between AC_{design}-0.5% and AC_{design}+0.5% results for M8. High FI values for M4 and M5 project mixtures might be a result of the production binder content significantly exceeding the design binder content. It should be noted that M4 and M5 mixes were the only two mixtures with latex additives. Less flexible LMLC asphalt concrete specimens may also be a result of the non-uniform coating of the aggregates with latex due to the use of extremely small amount of latex in the mixture (about 0.0375% by weight of mix) during laboratory sample production. The amount of latex in the prepared solution is extremely low for single specimen production in the laboratory and most of the added latex in the water-latex solution may not be uniformly coating the aggregates in the asphalt mixture. On the other hand, significantly larger scale production at the drum plants might result in more uniform coating of the aggregates with latex. For this reason, problems with latex blending for the LMLC specimens might be resulting in lower FI values. In a future study, a special laboratory procedure should be developed to more uniformly coat aggregates with the latex solution to be able to produce latex modified LMLC asphalt mixtures that more closely represent production mixtures.

Correlation matrices and pairs plots were used to determine the correlations between typical asphalt mixture variables [asphalt binder content (AC), effective binder content (P_{be}), dust-to-binder ratio (P_{200}/ P_{be}), voids filled with asphalt (VFA), tensile strength ratio (TSR)] and measured FI values for LMLC specimens at optimum binder content (FI_{OBC}) and for PMLC specimens (FI_{PM}). Pairs plots basically show one-to-one correlations between different variables while correlation matrices show the strength and direction of a linear relationship between variables. Correlation values range from -1 to +1. Minus sign represents an inverse proportion between two variables while plus sign represents the direct proportion. The absolute value of the correlation coefficient closer to 1 represents a strong relationship between the two variables. Pairs plot for SCB test results is given in Figure 4.4. Correlation matrix showing the relationship between asphalt mixture properties and FI values is given in Table 4.3. It can be observed that VFA and effective binder content are highly correlated with both FI_{OBC} and FI_{PM}. With correlation coefficients close 0.57, VFA is the most significant factor controlling the cracking resistance of asphalt mixtures. According to the correlation matrix, there is also a statistically significant relationship between dust-to-binder ratio and FI. Negative correlation between dust-to-binder ratio and FI indicates that higher dust-to-binder ratio results in lower FI and lower cracking resistance, as expected. These statistically significant correlations between asphalt mixture variables and FI also indicate that FI parameter is able to capture the impact of asphalt mixture properties on cracking resistance.

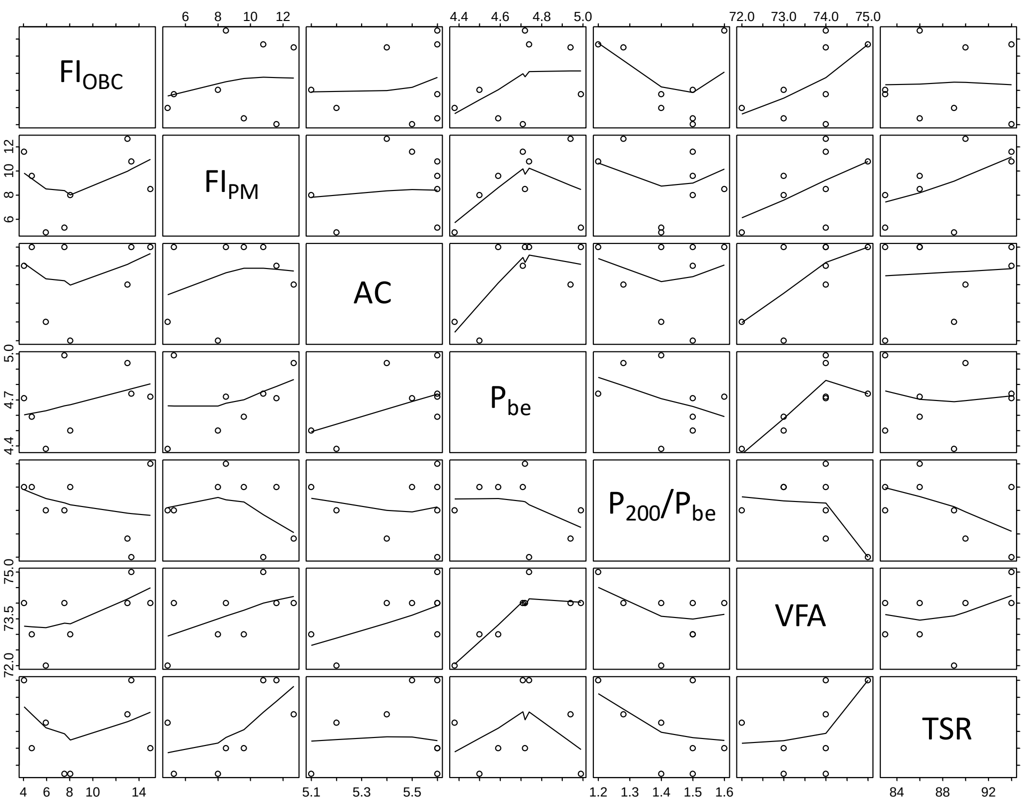


Figure 4.4: Pairs plot to present the relationship between FI values and asphalt mixture variables

Notes: FI_{OBC}=Flexibility index of LMLC specimens at optimum binder content; FI_{PM}=Flexibility index of PMLC specimens; AC= Asphalt content added by weight; P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust-to-binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

Table 4.3: Correlation Matrix Showing the Strength of Correlations between Measured FI Values and Asphalt Mixture Variables

	FI_{OBC}	FI_{PM}	AC	P_{be}	P_{200}/P_{be}	VFA	TSR
FI_{OBC}	**1.00**	0.27	0.23	0.40	-0.30	0.57	0.08
FI_{PM}	0.27	**1.00**	0.27	0.33	-0.26	0.56	0.61
AC	0.23	0.27	**1.00**	0.60	-0.05	0.66	0.20
P_{be}	0.40	0.33	0.60	**1.00**	-0.33	0.74	0.03
P_{200}/P_{be}	-0.30	-0.26	-0.05	-0.33	**1.00**	-0.40	-0.49
VFA	0.57	0.56	0.66	0.74	-0.40	**1.00**	0.40
TSR	0.08	0.61	0.20	0.03	-0.49	0.40	**1.00**

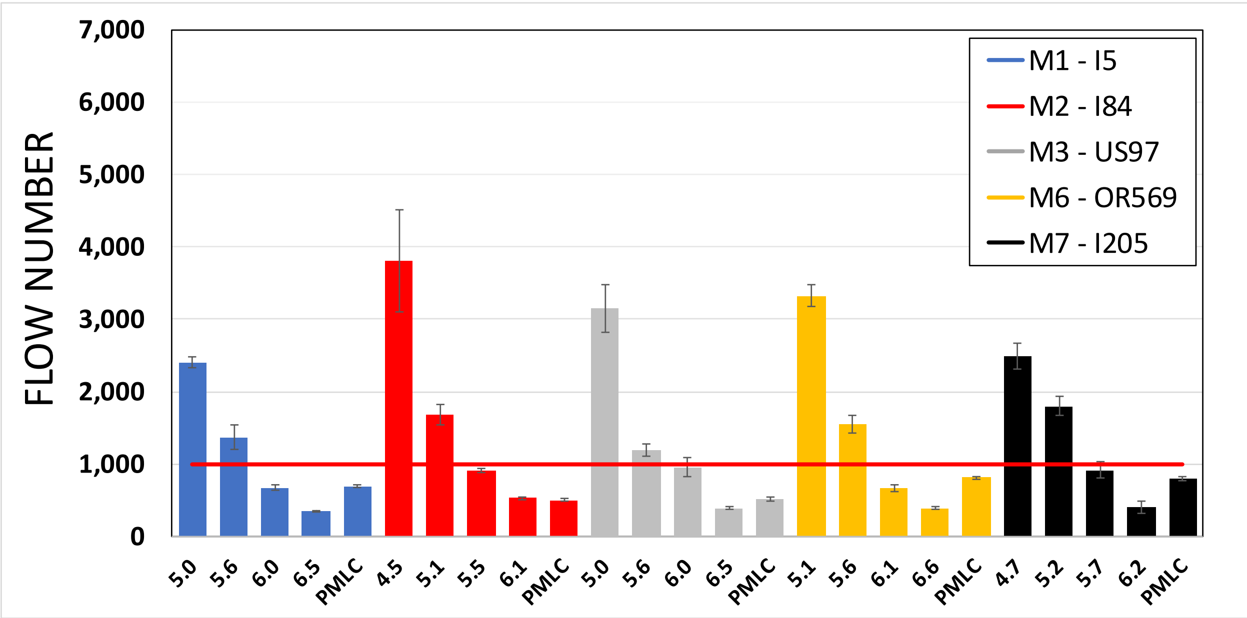
Notes: FI_{OBC}=Flexibility index of LMLC specimens at optimum binder content;
FI_{PM}=Flexibility index of PMLC specimens; AC= Asphalt content added by weight;
P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust-to-binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

4.5.2 FN Test Results

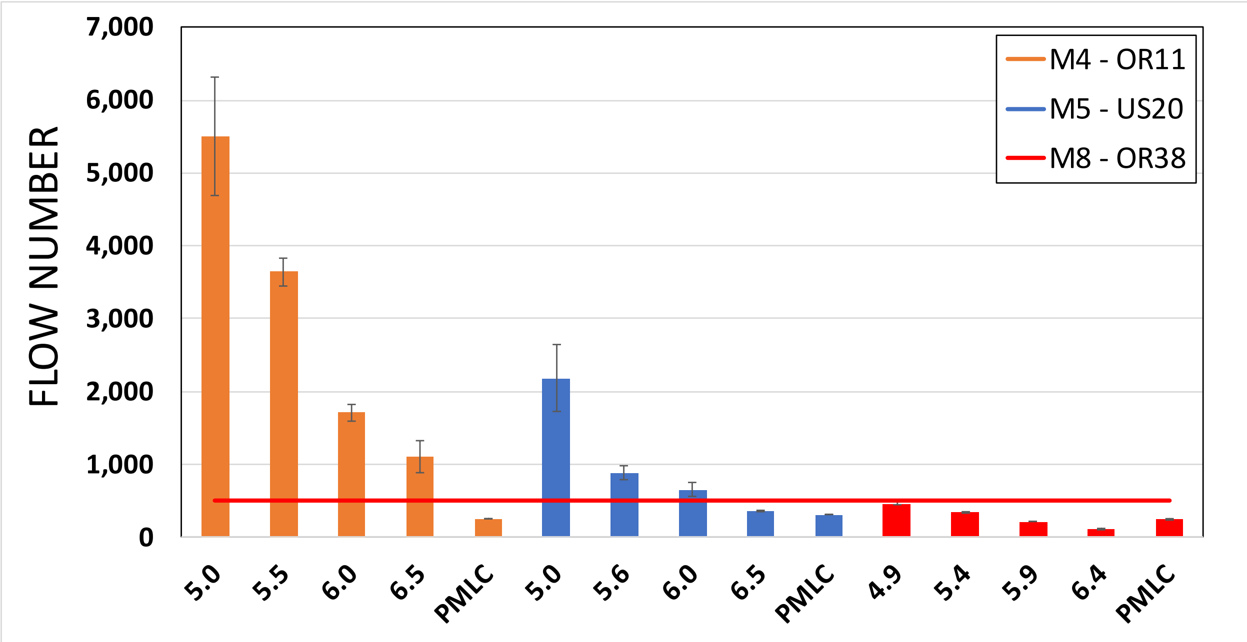
Figure 4.5 presents the results of flow number tests conducted to determine the rutting performance of asphalt mixtures. FN was calculated and used to evaluate the rutting performance of all asphalt mixtures. A mixture with lower flow number is expected to show lower rutting resistance. The process followed to calculate the FN from laboratory FN test results is described in Section 2.3.2.1. The horizontal red lines in Figure 4.5 are the FN thresholds selected in this study for Level 3 ($FN_{threshold}$=500) and Level 4 ($FN_{threshold}$=1,000) mixtures. It should be noted that AASHTO TP 79-15 recommends a minimum FN value of 740 for high traffic (ESALs > 30 million).

It can be observed from Figure 4.5 that increasing binder content decreases FN for all cases, as expected. In other words, FN is able to capture the impact of increased binder content on rutting resistance. It should be noted that projects M1, M2, M3, M6 and M7 had Level 4 mixtures (designed with 100 gyrations –FN test results shown in Figure 4.5a) whereas, M4, M5 and M8 are Level 3 mixtures (designed with 80 gyrations – FN test results shown in Figure 4.5b).

Average FI values for Level 4 PMLC specimens show that all Level 4 mixtures have close FN values while M6 and M7 have FN values slightly higher than M1, M2, and M3 (indicating higher rutting resistance for M6 and M7 mixtures). In Figure 4.5, the second bar for every mixture type shows the FN values for the LMLC samples prepared at the design binder content. It can be observed that all LMLC samples at the design binder content passes the 1,000 FN threshold. However, none of the Level 4 PMLC mixtures fulfills the 1,000 FN requirement while the FN values for M2 and M3 are very low (around 500). This result suggested that differences between the laboratory and plant mixing processes directly control the FN test results. In other words, FN test results are very sensitive to the mixing process. Similar conclusions can also be derived when the FN test results for Level 3 mixtures were evaluated (Figure 4.5b). FN values for the Level 3 PMLC mixtures are significantly lower than the FN values for the laboratory produced specimens. While 500 and 1,000 were suggested as FN thresholds for LMLC Level 3 and Level 4 mixtures for balanced mix design, respectively, those FN values cannot be used as thresholds for PMLC samples. Due to the significant difference in FN values for PMLC and LMLC mixes, the effectiveness of the FN test in terms of characterizing rut resistance becomes questionable.



(a)



(b)

Figure 4.5: FN test results for LMLC and PMLC specimens (a) Level 4 mixtures (b) Level 3 mixtures (length of the error bar is equal to one standard deviation)

Correlation matrices and pairs plots were used to determine the correlations between typical asphalt mixture variables [asphalt binder content (AC), effective binder content (P_{be}), dust-to-binder ratio (P_{200}/ P_{be}), voids filled with asphalt (VFA), tensile strength ratio (TSR)] and measured FN values for LMLC specimens at optimum binder content (FN_{OBC}) and for PMLC specimens (FN_{PM}). Pairs plots basically show one-to-one correlations between different variables while correlation matrices show the strength and direction of a linear relationship between variables. Correlation values range from -1 to +1. Minus sign represents an inverse proportion between two variables while plus sign represents the direct proportion. The absolute value of the correlation coefficient closer to 1 represents a strong relationship between the two variables.

Pairs plot for FN test results is given in Figure 4.6. Correlation matrix showing the relationship between asphalt mixture properties and FN values is given in Table 4.4. It can be observed that the correlations between FN values and asphalt binder content are very weak with values close to 0.09 for both FN_{OBC} and FN_{PM}. The correlations between the FN values and effective binder content are also weak for both LMLC and PMLC specimens. According to the pairs plot and the correlation matrix, changes in VFA do not significantly affect rutting resistance. These weak correlations between asphalt mixture variables and FN indicate that FN parameter is not able to capture the impact of asphalt mixture properties on rutting resistance. For this reason, although FN is a repeatable test that can quantify the impact of 0.5% change in asphalt binder content on rutting resistance, FN test may not be an ideal test for balanced mix design and performance based specifications.

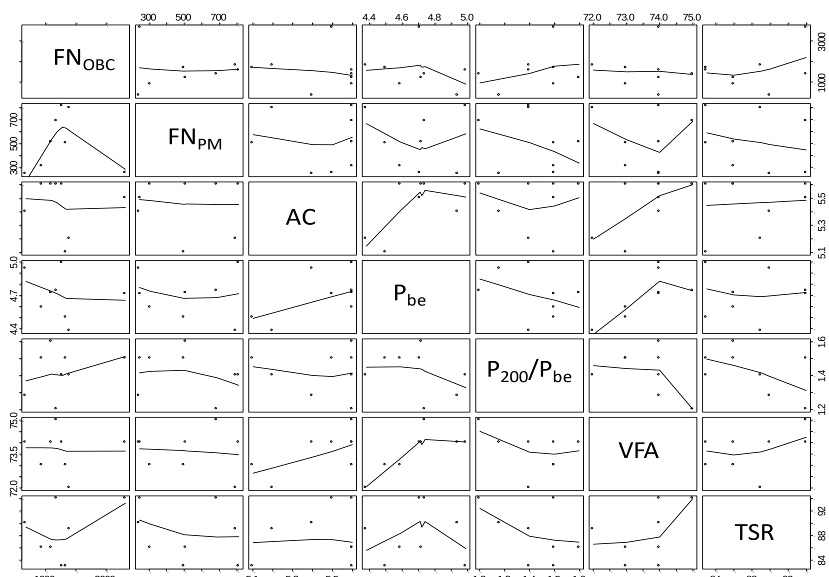


Figure 4.6: Pairs plot to present the relationship between FN values and asphalt mixture variables

Notes: FN_{OBC}=Flow number of LMLC specimens at optimum binder content;
FN_{PM}=Flow number of PMLC specimens; AC= Asphalt content added by weight;
P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust-to-binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

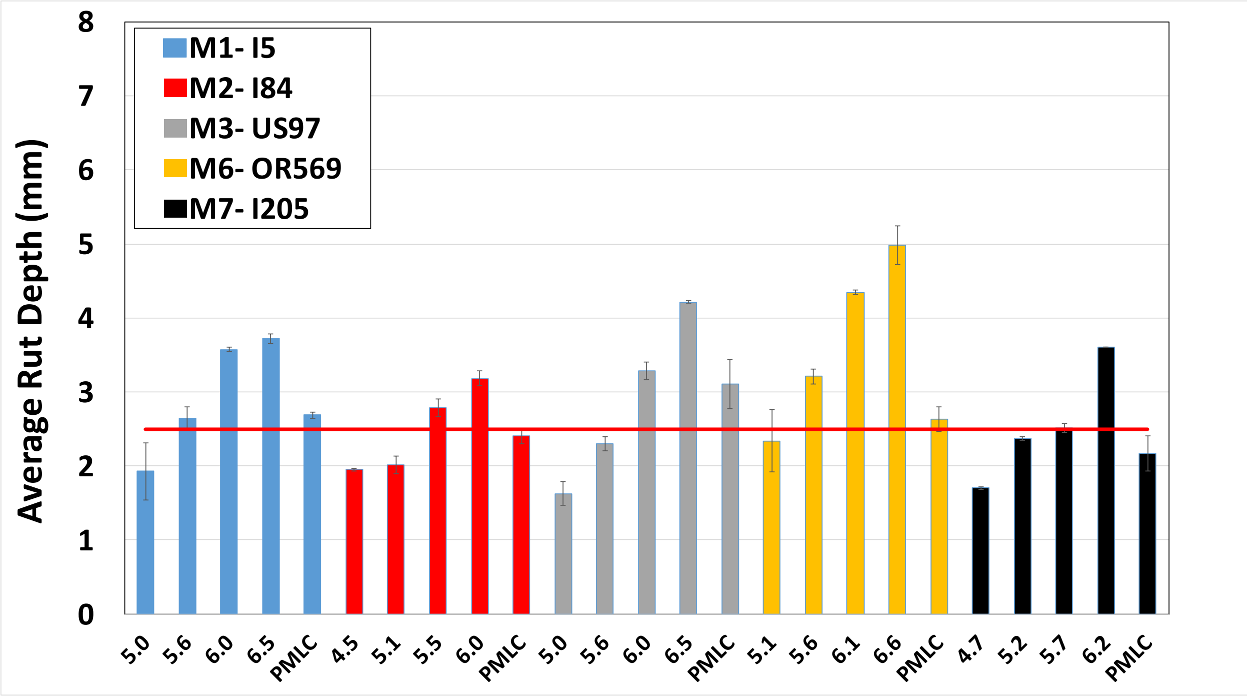
Table 4.4: Correlation Matrix Showing the Strength of Correlations between Measured FN Values and Asphalt Mixture Variables

	FN_{OBC}	FN_{PM}	AC	P_{be}	P_{200}/P_{be}	VFA	TSR
FN_{OBC}	**1.00**	-0.03	-0.09	-0.22	0.30	-0.02	0.36
FN_{PM}	-0.03	**1.00**	-0.08	-0.11	-0.24	-0.11	-0.25
AC	-0.09	-0.08	**1.00**	0.60	-0.05	0.66	0.20
P_{be}	-0.22	-0.11	0.60	**1.00**	-0.33	0.74	0.03
P_{200}/P_{be}	0.30	-0.24	-0.05	-0.33	**1.00**	-0.40	-0.49
VFA	-0.02	-0.11	0.66	0.74	-0.40	**1.00**	0.40
TSR	0.36	-0.25	0.20	0.03	-0.49	0.40	**1.00**

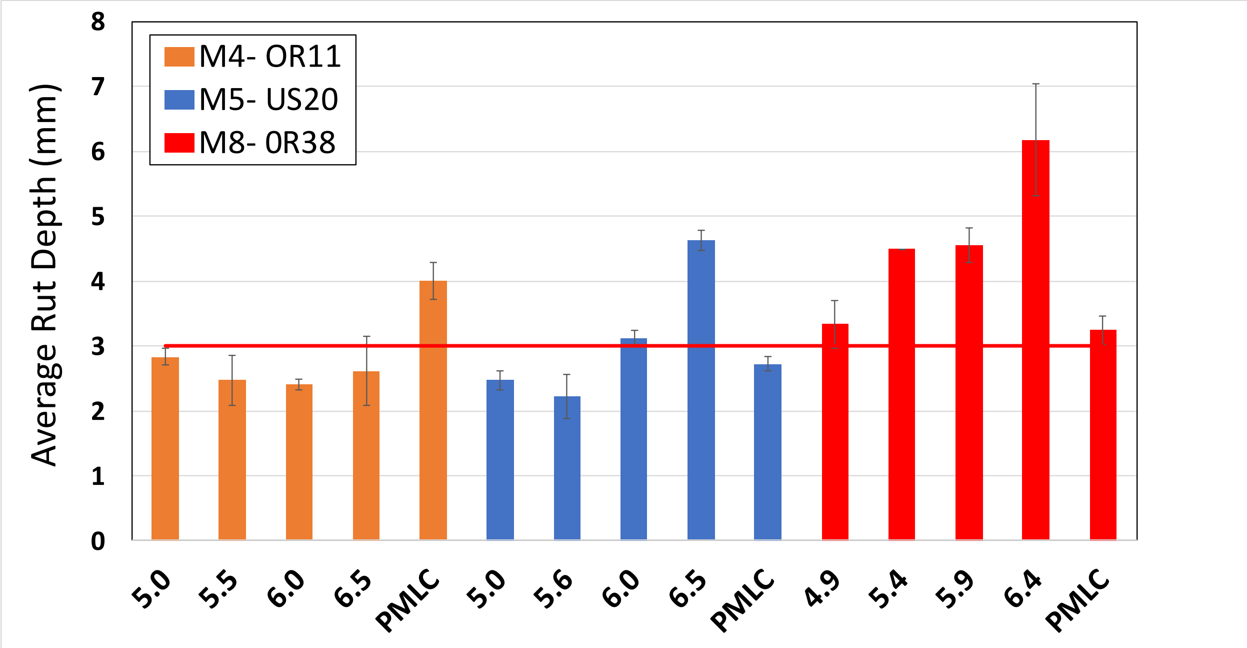
Notes: FN_{OBC}=Flow number of LMLC specimens at optimum binder content;
FN_{PM}=Flow number of PMLC specimens; AC= Asphalt content added by weight;
P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust-to-binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

4.5.3 HWTT Test Results

Figure 4.7 presents the results of HWTT tests conducted to determine the rutting performance of asphalt mixtures. Average surface rut depth after 20,000 wheel passes was used to evaluate the rutting performance of all asphalt mixtures. A mixture with higher rut depth is expected to show lower rutting resistance. General information about the HWTT is given in Section 2.3.3. The horizontal red lines in Figure 4.7 are the HWTT rut depth thresholds selected in this study for Level 3 ($RD_{threshold}$=3 mm) and Level 4 ($RD_{threshold}$=2.5mm) mixtures.



(a)



(b)

Figure 4.7: HWTT test results for LMLC and PMLC specimens (a) Level 4 mixtures (b) Level 3 mixtures (length of the error bar is equal to one standard deviation)

It can be observed from Figure 4.7 that increasing binder content increases rut depth for most of the cases. However, the impact of binder content on measured rutting performance is not as clear as the FN test results. In other words, changes in binder content have less effect on the rut resistance according to the HWTT results. In addition, for M4 and M5 Level 3 mixes, an unexpected trend was observed for lower binder contents. This unexpected trend might be a result of the presence of latex in the mixture. Extremely small amount of latex used for single sample production in the laboratory maybe creating blending issues during laboratory specimen production. Issues with latex blending might be increasing the test results variability. In this study, four replicate asphalt cores were produced for HWTT testing. Since two cores were attached edge-to-edge to run the experiment, a total of two rut depth values were collected from the test system for each case. Increasing replicate test results from two to three is recommended in this study to minimize the impact of high-test results variability on average measured rut depth. In addition, since HWTT experiments were conducted under water, test results are also affected by the moisture susceptibility of the asphalt mixture in addition to rut resistance. Combined effect of moisture and rut resistance reflected in the test results might be increasing the variability of the test.

For Level 4 mixtures, average rut depth values show that all Level 4 mixtures had similar rut resistance while M6 had a slightly higher accumulated rutting than other four Level 4 mixes. This is expected to be a result of the higher effective binder content for the M6 mix. In Figure 4.7, the second bar for every mixture type shows the rut depth value for the LMLC samples prepared at the design binder content. It can be observed that M2 asphalt mixture had rutting resistance higher than all other Level 4 mixtures. Higher rutting resistance for this mixture is likely to be a result of the relatively lower design and effective binder content for this mixture. These results suggested that HWTT is able to capture the impact of changes in the asphalt mixture properties on rut resistance. Test results for the Level 4 PMLC specimens show that M2 and M7 projects meet depth threshold of 2.5mm while the PMLC mixtures for M1, M3, and M6 projects fail to meet the requirement. However, it should be noted that average rut depths for M1 and M6 PMLC mixtures are very close to the 2.5mm rut depth threshold.

In Figure 4.7a, it can be observed that average measured rut depth values for all PMLC specimens from all Level 4 mixtures fall between the test results for LMLC specimens for AC_{design}-0.5% and AC_{design}+0.5%. Since plant produced mixtures are allowed to have ±0.5% variability in production binder content in Oregon, it can be concluded for Level 4 mixtures that LMLC asphalt specimens are representing the specimens produced with mixtures sampled from plants. In other words, laboratory mixing process is not introducing any observable bias into the HWTT results.

It can be observed from Figure 4.7b that rut depth value for the M4 Level 3 PMLC mixture is higher than the AC_{design}+0.5% while measured average rut depth falls between AC_{design}-0.5% and AC_{design}+0.5% results for M5 and M8. High rut depth value for the M4 project mixture might be a result of the production binder content significantly exceeding the design binder content (same result was also observed in the SCB test results for cracking). The latex additive in the M4 mix may also be creating this result. LMLC asphalt concrete specimens with higher rut resistance maybe a result of the process followed to incorporate latex into the asphalt mixtures in the laboratory. In a future study, a special laboratory procedure should be developed to more uniformly coat aggregates with the latex solution to be able to produce latex modified LMLC asphalt mixtures that more closely represent production mixtures.

Correlation matrices and pairs plots were used to determine the correlations between typical asphalt mixture variables [asphalt binder content (AC), effective binder content (P_{be}), dust-to-binder ratio (P_{200}/ P_{be}), voids filled with asphalt (VFA), tensile strength ratio (TSR)] and measured rut depth values for LMLC specimens at optimum binder content (RD_{OBC}) and for PMLC specimens (RD_{PM}). Pairs plot for HWTT test results is given in Figure 4.8. Correlation matrix showing the relationship between asphalt mixture properties and rut depth values is given in Table 4.5. It can be observed that VFA and effective binder content are highly correlated with both RD_{OBC} and RD_{PM}. With a correlation coefficient of 0.74, effective binder content is the most significant factor controlling the rutting resistance of asphalt mixtures. These statistically significant correlations between asphalt mixture variables and HWTT rut depths also indicate that HWTT is able to capture the impact of asphalt mixture properties on rutting resistance. According to the correlation matrix, there is also a statistically significant relationship between TSR and rut depth. Since HWTT is conducted under water, test results are also controlled with the moisture susceptibility of asphalt mixtures in addition to rut resistance. The possibility of using different parameters (such as creep slope shown in Figure 2.20) for rutting performance evaluation should also be investigated in a future study. Results also show that there is a significant negative correlation between TSR and dust-to-binder ratio. Increasing dust-to-binder ratio reduces TSR values. Based on the test results and analyses, HWTT is recommended as the experiment for rutting performance quantification for balanced mix design and performance based specifications**.** Balanced mix design process presented in the next section also used HWTT for rutting performance evaluation.

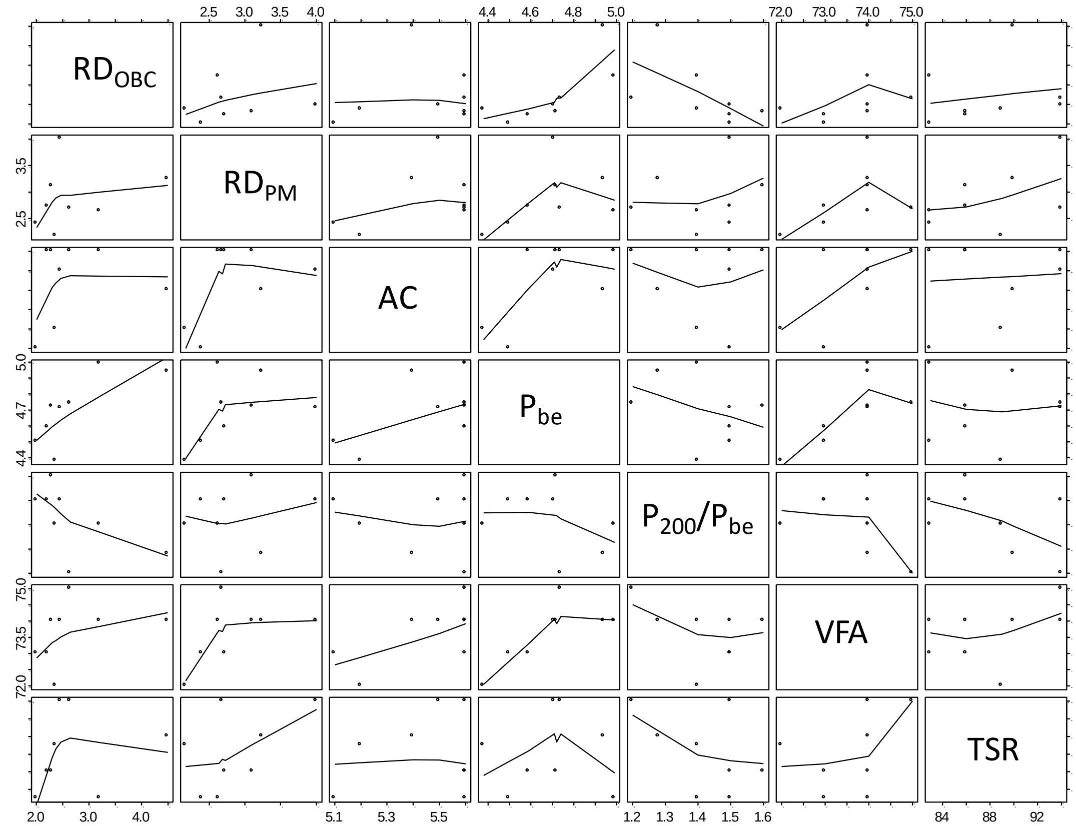


Figure 4.8: Pairs plot to present the relationship between FN values and asphalt mixture variables

Notes: RD_{OBC}=HWTT rut depth of LMLC specimens at optimum binder content;
RD_{PM}= HWTT rut depth of PMLC specimens; AC= Asphalt content added by weight;
P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust to binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

Table 4.5: Correlation Matrix Showing the Strength of Correlations between Measured Rut depth Values and Asphalt Mixture Variables

	RD_{OBC}	RD_{PM}	AC	P_{be}	P_{200}/P_{be}	VFA	TSR
RD_{OBC}	**1.00**	0.26	0.14	0.74	-0.58	0.35	0.17
RD_{PM}	0.26	**1.00**	0.41	0.44	0.17	0.50	0.51
AC	0.14	0.41	**1.00**	0.60	-0.05	0.66	0.20
P_{be}	0.74	0.44	0.60	**1.00**	-0.33	0.74	0.03
P_{200}/P_{be}	-0.58	0.17	-0.05	-0.33	**1.00**	-0.40	-0.49
VFA	0.35	0.50	0.66	0.74	-0.40	**1.00**	0.40
TSR	0.17	0.51	0.20	0.03	-0.49	0.40	**1.00**

Notes: RD_{OBC}=HWTT rut depth of LMLC specimens at optimum binder content;
RD_{PM}= HWTT rut depth of PMLC specimens; AC= Asphalt content added by weight;
P_{be} = Effective asphalt content present by weight in the total mix; P_{200}/P_{be} = Dust-to-binder ratio in the mix;
VFA = Voids filled with asphalt; TSR = Tensile strength ratio.

4.5.4 Balanced Mix Design

Balanced mix design approach helps in determining the binder content range that satisfies both cracking and rutting performance criteria. Minimum binder content is the lowest asphalt binder percentage allowed in the mix to satisfy the FI threshold of 8 for Level 4 mixtures and FI of 6 for Level 3 mixtures. Maximum asphalt content is the highest percentage that satisfies the rutting criteria, rut depth of 2.5mm for Level 4 mixtures and 3mm for Level 3 mixtures. Figure 4.9 depicts an example balanced mix design chart for the M3 project. Based on the volumetric mix design, M3 has an asphalt content of 5.6%. From Figure 4.9, it can be observed that M3 meets the cracking and rutting criteria at the design asphalt content. However, with the balanced mix design approach, the minimum asphalt required is about 5.1% (see Figure 4.9). This reduced binder content is expected to significantly reduce the cost of the M3 asphalt mixture while still keeping it in the acceptable region for rutting and cracking performance. However, to ensure a high long-term cracking performance, 5.7% asphalt binder content (or the design 5.6%) can also be used for production. However, it should be noted that using 5.7% design asphalt content creates a high risk for rutting since plant produced mixtures are allowed to have ±0.5% variability in production binder content in Oregon. ODOT is currently in the process of changing the binder content variability tolerance from ±0.5% to ±0.35%. This change is expected to reduce the risk of rutting or cracking failures due to production binder content variability. For this reason, this study recommends to use the midpoint of the “acceptable asphalt binder content range” (average of minimum and maximum binder content allowed). Based on Figure 4.9, the acceptable range of asphalt content for M3 ranges from 5.1% to 5.7%. A design asphalt binder content of 5.4% (average of 5.1% and 5.7%) should be used for production (slightly lower than the binder content from the volumetric design-5.6%).

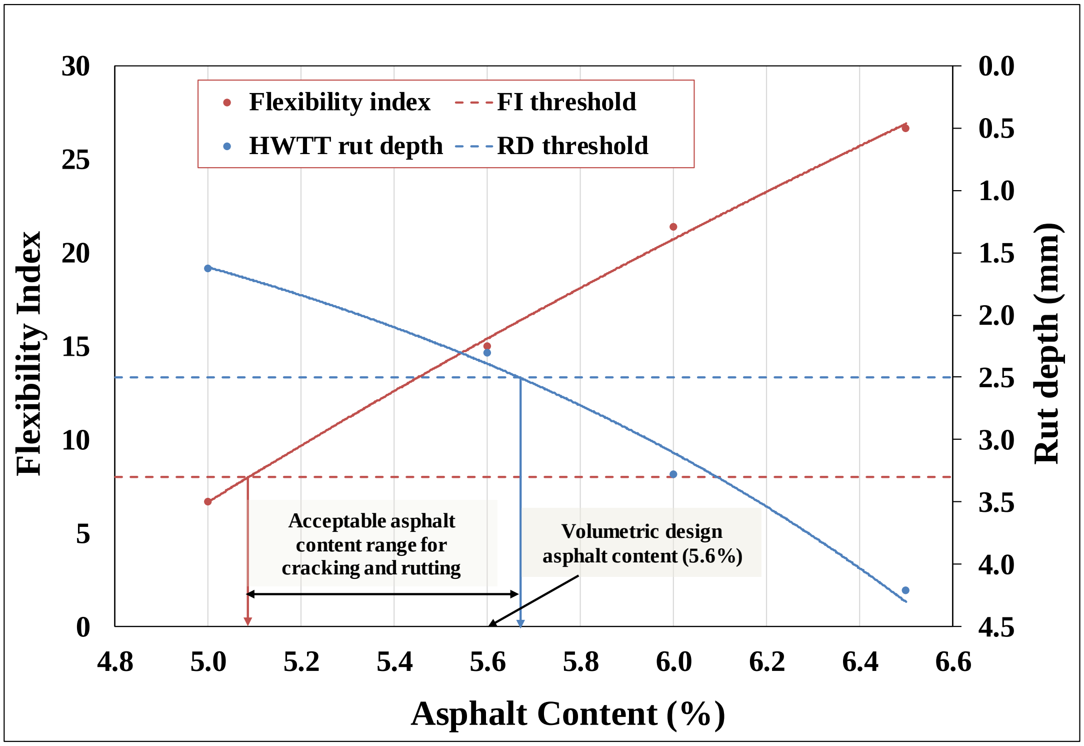


Figure 4.9: Balanced mix design example for M3

By following the balanced mix design process described in Figure 4.9, binder content ranges to meet both rutting and cracking performance criteria for different threshold levels were determined. Table 4.6 presents the ranges of acceptable binder contents obtained for all eight project mixtures evaluated in this study. Binder content intervals for different thresholds were presented for all mixtures to present the impact of selected performance criteria on the binder content range. Although HWTT is recommended as the test for rutting performance evaluation for balanced mix design, thresholds for flow number and the corresponding maximum binder contents were also presented in Table 4.6. The red numbers in Table 4.6 form the asphalt binder content intervals determined by using the selected thresholds for SCB and HWTT experiments (Level 3-$FI_{threshold}$=6; $RD_{threshold}$=3 mm and Level 4-$FI_{threshold}$=8; $RD_{threshold}$=2.5mm).

It is expected to have the binder content from the FI threshold to set the minimum binder content limit while HWTT rut depth should provide the maximum binder content that can be allowed. It can be observed that all mixtures except M5 and M6 have reasonable binder content intervals. The minimum binder content allowed for M5 is 6% according to the cracking criterion while the maximum binder content allowed according to the rutting criterion is 5.9%. Since these two percentages do not create an interval, produced mixture can be redesigned. However, since the difference is minute in this case, a balanced mix design binder content of 5.9% can be recommended to avoid any early rutting failures. However, a decision to use 5.2% as the design binder content for mix M6 can result in cracking related failures since the binder contents suggested by FI and RD thresholds are significantly different. In cases similar to M6, it is required to re-design (starting from volumetric design followed by balanced mix design) the mixture by changing the gradation, RAP content, binder type, and/or binder grade. For all other mixtures, a binder content interval from balanced mix design was achieved. By comparing the volumetric design binder contents to developed balanced mix design binder content intervals, it can be observed that mixes M4, M5, and M7 have volumetric design binder contents lower than the balanced mix design interval (on the dry side) while M1, M6, and M8 have binder contents above the balanced mix design suggested interval (on the wet side). M2 and M3 mixes have volumetric design binder contents in the balanced mix design suggested interval. It should be noted that thresholds for FI and RD are selected to create reasonable binder content intervals. If the balanced mix design process is implemented in Oregon in a shadow specification, these thresholds can be modified within the first year based on the laboratory measured rutting and cracking performance. Potential implementation strategies are summarized in Section 5.2.

Table 4.6: Acceptable Asphalt Binder Content Intervals for Various Thresholds

Test	Thresh-olds	LEVEL 3			LEVEL 4				
		(Mix/ID, AC Design %)			(Mix/ID, AC Design %)				
		M4, 5.5	M5, 5.6	M8, 5.4	M1, 5.6	M2, 5.1	M3, 5.6	M6, 5.6	M7, 5.2
FI	6	**5.7**	**6.0**	**4.2**	-	-	-	-	-
	8	5.9	6.2	4.5	**4.8**	**5.1**	**5.1**	**5.7**	**5.6**
	9	-	-	-	4.9	5.2	5.2	5.8	5.7
	10	-	-	-	5.1	5.2	5.2	5.9	5.8
HWTT	2.5mm	6.2	5.7	4.5	**5.5**	**5.4**	**5.7**	**5.2**	**5.7**
	3.0mm	**7.4**	**5.9**	**4.8**	5.8	5.8	5.9	5.5	5.9
FN	500	**7.0**	**6.3**	**4.8**	-	-	-	-	-
	800	6.8	5.7	3.6	5.9	5.6	6.1	6.0	5.8
	1,000	-	-	-	**5.8**	**5.5**	**5.9**	**5.9**	**5.6**

4.5.4.1 Sensitivity analysis to determine the possibility of reducing trial design binder contents

In this study, four different asphalt contents (AC) were used for each project for balanced mix design: AC_{design} from volumetric mix design, AC_{design}-0.5%, AC_{design}+0.5%, and AC_{design}+1%. Determined acceptable asphalt binder content intervals for different thresholds are given in Table 4.6. The possibility of reducing the four trial asphalt binder contents to three binder contents to improve the practicality of the balanced mix design process is investigated in this section. Table 4.7 and

Table 4.8 show the asphalt binder content intervals when "AC_{design}-0.5%" and "AC_{design}+1.0%" were excluded from the balanced mix design process, respectively. The design binder contents that changed after excluding one of the four binder contents are shown in yellow color. It can be observed that several binder contents for the selected thresholds for RD and FI change when the "AC_{design}-0.5%" is excluded from the analysis. However, none of the selected balanced mix design binder content intervals change when the "AC_{design}+1.0%" is excluded. This result suggested that it might be possible to perform balanced mix design by just using three binder contents (AC_{design} from volumetric mix design, AC_{design}-0.5%, AC_{design}+0.5%). However, this conclusion should be further evaluated in a future study by using more extensive datasets (or during the shadow balanced mix design implementation phase).

Table 4.7: Acceptable Asphalt Binder Content Intervals for Various Thresholds – AC_{design}-0.5% Excluded from the Analysis

Test	Thresh-olds	LEVEL 3			LEVEL 4				
		(Mix/ID, AC Design %)			(Mix/ID, AC Design %)				
		M4, 5.5	**M5, 5.6**	**M8, 5.4**	**M1, 5.6**	**M2, 5.1**	**M3, 5.6**	**M6, 5.6**	**M7, 5.2**
FI	6	**5.7**	**6.0**	**4.9**	-	-	-	-	-
	8	5.9	6.2	5.0	**5.3**	**5.1**	**5.2**	**5.7**	**5.6**
	9	-	-	-	5.4	5.2	5.2	5.8	5.7
	10	-	-	-	5.4	5.2	5.3	5.9	5.8
HWTT	2.5mm	6.2	5.7	4.2	**5.5**	**5.4**	**5.7**	**5.2**	**5.7**
	3.0mm	**7.4**	**5.9**	**4.5**	5.8	5.8	5.9	5.5	5.9
FN	500	**7.0**	**6.3**	**4.8**	-	-	-	-	-
	800	6.8	5.7	3.6	5.9	5.6	6.1	6.0	5.8
	1,000	-	-	-	**5.8**	**5.5**	**5.9**	**5.9**	**5.6**

Table 4.8: Acceptable Asphalt Binder Content Intervals for Various Thresholds - AC_{design}+1.0% Excluded from the Analysis

Test	Thresh-olds	LEVEL 3			LEVEL 4				
		(Mix/ID, AC Design %)			(Mix/ID, AC Design %)				
		M4, 5.5	**M5, 5.6**	**M8, 5.4**	**M1, 5.6**	**M2, 5.1**	**M3, 5.6**	**M6, 5.6**	**M7, 5.2**
FI	6	**5.7**	**6.0**	**4.2**	-	-	-	-	-
	8	5.9	6.6	4.5	**4.8**	**5.1**	**5.1**	**5.7**	**5.6**
	9	-	-	-	4.9	5.2	5.2	5.8	5.8
	10	-	-	-	5.1	5.2	5.2	5.9	6.0
HWTT	2.5mm	6.2	5.7	4.5	**5.5**	**5.4**	**5.7**	**5.2**	**5.7**
	3.0mm	**7.4**	**5.9**	**4.8**	5.8	5.6	5.9	5.5	5.7
FN	500	**6.3**	**6.3**	**4.8**	-	-	-	-	-
	800	6.2	5.7	3.6	5.9	5.6	6.3	6.0	5.7
	1,000	-	-	-	**5.8**	**5.5**	**5.9**	**5.9**	**5.6**

4.6 CONCLUSIONS

The major conclusions derived from this part of the study are as follows:

Conclusions from SCB testing are provided below:

1. For Level 4 mixtures, asphalt mixtures from projects M1, M2, M3 and M6 provided higher FI values than the mixture from M7 project.

2. For LMLC specimen test results at the design binder content, M1 and M3 asphalt mixtures had cracking resistances significantly higher than all other Level 4 mixtures. Higher cracking resistance for the mixture from these two projects is likely to be a result of the relatively higher design binder content for these two projects (both at 5.6%). Significantly lower dust-to-binder ratio (see Table 4.1) and high VFA for M1 are also the major reasons for significantly higher cracking resistance.

3. Although M2 had a binder content significantly lower than M1 and M3, it showed cracking performance comparable to M1 and M3. This high cracking resistance for M2 is a result of the high effective binder content and VFA that are closer to M1 and M3 mixtures.

4. Test results for the Level 4 PMLC specimens show that M1, M2, and M3 projects meet the FI threshold of 8 while the PMLC mixtures for M6 and M7 projects fail to meet the requirement with average FI values close to 5. Lower FI value for project M7 can be explained by the relatively lower binder content for this asphalt mixture. The effective binder content and the VFA for the M7 asphalt mixture were lower than all other four Level 4 mixtures evaluated in this study.

5. Average measured FI values for all PMLC specimens from all Level 4 mixtures fall between the test results for LMLC specimens for AC_{design}-0.5% and AC_{design}+0.5%. Since plant produced mixtures are allowed to have ±0.5% variability in production binder content in Oregon, it can be concluded for Level 4 mixtures that laboratory mixing process is not introducing any bias into the measured FI values.

6. FI values for the Level 3 PMLC mixtures are higher than the AC_{design}+0.5% for projects M4 and M5 while FI falls between AC_{design}-0.5% and AC_{design}+0.5% results for M8. High FI values for M4 and M5 project mixtures might be a result of the production binder content significantly exceeding the design binder content. It should be noted that M4 and M5 mixes were the only two mixtures with latex additives. Less flexible LMLC asphalt concrete specimens maybe a result of the process followed to incorporate latex into the asphalt mixtures in the laboratory.

7. VFA and effective binder content are highly correlated with both FI_{OBC} and FI_{PM}. With correlation coefficients close 0.57, VFA is the most significant factor controlling the cracking resistance of asphalt mixtures. According to the correlation matrix, there is also a statistically significant relationship between dust-to-binder ratio and FI. These statistically significant correlations between asphalt mixture variables

and FI also indicate that FI parameter is able to capture the impact of asphalt mixture properties on cracking resistance.

8. Based on the results of all analyses, an FI threshold of 6 was recommended for Level 3 mixes while the threshold for Level 4 was selected as 8.

Conclusions from FN testing are provided below:

1. The differences between the laboratory and plant mixing processes directly control the FN test results. FN results for PMLC mixtures are always significantly lower than results for LMLC mixes. This difference avoids any direct comparisons between the rut resistance of PMLC and LMLC asphalt mixtures. Due to the significant difference in FN values for PMLC and LMLC mixes, the effectiveness of the FN test in terms of characterizing rut resistance also becomes questionable.

2. Weak correlations between asphalt mixture variables and FN indicate that FN parameter is not able to capture the impact of asphalt mixture properties on rutting resistance.

3. Due to the conclusions 9 and 10 above, FN test may not be an ideal test for balanced mix design and performance based specifications.

Conclusions from HWTT testing are provided below:

1. For the HWTT results, the impact of binder content on measured rutting performance is not as clear as the FN test results. In addition, for M4 and M5 Level 3 mixes, an unexpected trend was observed for lower binder contents. This unexpected trend might be a result of the presence of latex in the mixture. In addition, since HWTT experiments were conducted under water, test results are also affected by the moisture susceptibility of the asphalt mixture in addition to rut resistance. Combined effect of moisture and rut resistance reflected in the test results might be increasing the variability of the test.

2. For Level 4 mixtures, average rut depth values show that all Level 4 mixtures had similar rut resistance while M6 had a slightly higher accumulated rutting than other four Level 4 mixes. This is expected to be a result of the higher effective binder content for the M6 mix.

3. For LMLC specimen test results at the design binder content, M2 asphalt mixture had rutting resistance significantly higher than all other Level 4 mixtures. Higher rutting resistance for this mixture is likely to be a result of the relatively lower design and effective binder content for this mixture. Since M2 also had high cracking resistance (See bullet 3 above) despite its significantly lower binder content than all other Level 4 mixes, it can be considered to be a high performance and cost effective asphalt mixture.

4. Strong correlations between asphalt mixture variables and HWTT rut depths indicate that HWTT test is able to capture the impact of asphalt mixture properties on rutting resistance.

5. For the HWTT experiment, unlike the flow number test, LMLC asphalt specimens are representing the specimens produced with mixtures sampled from plants. In other words, laboratory mixing process is not introducing any bias into the HWTT results.

6. HWTT rut depths are highly correlated with the asphalt mixture properties. In addition, HWTT is a practical experiment that does not require a hydraulic test equipment. Laboratory mixing and specimen production are also not introducing any significant bias into the HWTT results and produce test samples that represent specimens prepared with production mixtures. For these reasons, HWTT is recommended as the experiment for rutting performance quantification for balanced mix design and performance based specifications.

7. Based on the results of all analyses, an RD threshold of 3mm was recommended for Level 3 mixes while the threshold for Level 4 was selected as 2.5mm.

8. Reasonable results were achieved by following the developed balanced mix design process.

9. It might be possible to perform balanced mix design by just using three binder contents (AC_{design} from volumetric mix design, AC_{design}-0.5%, AC_{design}+0.5%). However, this conclusion should be further evaluated in a future study by using more extensive datasets (or during the shadow balanced mix design implementation phase).

Suggestions for potential future research are provided below:

1. Problems with uniformly coating aggregates with latex for the LMLC specimens resulted in FI values that are significantly lower than the FI values for the PMLC specimens. In a future study, a special laboratory procedure should be developed to more uniformly coat aggregates with the latex solution to be able to produce latex modified LMLC asphalt mixtures that more closely represent production mixtures.

2. Since HWTT is conducted under water, test results are also controlled with the moisture susceptibility of asphalt mixtures in addition to rut resistance. The possibility of using different parameters (such as creep slope shown in Figure 2.20) for rutting performance evaluation should also be investigated in a future study.

3. In this study, thresholds for FI and RD are selected to create reasonable binder content intervals. If the balanced mix design process is implemented in Oregon in a shadow specification, these thresholds can be modified within the first year based on the laboratory measured rutting and cracking performance.

5.0 SUMMARY AND CONCLUSIONS

Existing asphalt mixture design methods do not consider performance criteria and rely on volumetric properties to predict field performance. From an environmental and sustainable standpoint, the use of recycled asphalt materials in asphalt mixtures are becoming increasingly common. A drawback of this practice is a possible reduction in ductility of the asphalt mixture, which causes a significant reduction in the fatigue life of the pavement in many cases. In Oregon, premature cracking-related asphalt pavement failures are likely to increase due to the increasing recycled asphalt use in asphalt concrete mixtures. These premature failures are necessitating costly rehabilitation and maintenance at intervals of less than half of the intended design lives in some cases.

Most state DOTs and asphalt contractors do not think that commonly used asphalt mixture properties, such as voids in mineral aggregate (VMA), voids filled with asphalt (VFA), and dust-to-binder ratio, reflect the long-term performance of asphalt mixtures. In addition, there are several new additives, polymers, rubbers, and high quality binder types incorporated into asphalt mixtures today. Volumetric mixture design methods are not capable of capturing the benefits of using all these new technologies on asphalt mixture performance.

For all these reasons, by using low-cost, practical, and efficient performance testing procedures, it is necessary to accurately quantify the impact of using higher RAP content, new additives, and higher quality binders on the cracking and rutting resistance of the pavements. ODOT Research Projects SPR785 and SPR797 (Coleri et al. 2017b; Coleri et al. 2017a; Sreedhar et al. 2018; Haddadi et al. 2019) constructed the beginnings of a performance-based balanced mix design method for Oregon. It was suggested that SCB test is the most effective and practical cracking test that can effectively be used for balanced mix design. It was determined that the typical flexibility index (FI), an energy parameter calculated using SCB test results, values for production mixtures (plant-produced) range from 9 to 14. However, more experiments needed to be conducted to determine an exact threshold for FI that will provide acceptable long-term pavement cracking performance. However, the most effective laboratory test for rutting performance prediction was not determined in those previous ODOT research projects. In this study, a long-term asphalt mixture aging protocol and a balanced mix design process were developed to design asphalt mixtures by evaluating rutting (by conducting HWTT tests) and cracking (by conducting semi-circular bend tests) performance.

5.1 MAJOR CONCLUSIONS

The major conclusions drawn from the results of this study are as follows:

Development of a long-term aging protocol for asphalt mixtures

1. FI values for mixtures aged for 5 days at 85°C and 24 hours at 95°C were very similar indicating similar aging levels.

2. Aging high RAP mixtures at 135ºC creates excessive aging and equalize fatigue cracking resistance and ductility of asphalt mixtures. Recycled materials had a significant impact on the aging of asphalt mixtures and it is possible to have virgin and recycled mixtures present similar stiffness values by excessive aging. This result might also be a result of the changing binder chemistry at the 135ºC aging temperature (Branthaver et al. 1993, Petersen 2009, Kim et al. 2018). For these reasons, using 135ºC temperature for long-term aging might result in unexpected results for high RAP and stiff binder mixtures.

3. For high RAP mixtures, FI results for the mixtures aged at 85ºC and 95ºC were all reasonable (higher RAP results in lower FI values).

4. For the tested production mixtures, aging the mix at 135ºC for 12 hours almost equalizes the FI values for all mixtures although they were expected to have significantly different mixture properties because of the differences in binder contents, binder types, gradations, additives, and RAP contents.

5. Performance rankings for mixtures aged at 95ºC for 24 hours and 72 hours are the same.

6. Based on all results from all three phases, the protocol with 24 hours of aging at 95°C was selected for the long-term aging protocol for balanced mix design and performance based specifications. Although this protocol may not be simulating more than 3-5 years of aging in the field, it can still provide reasonable levels of aging that is going to provide cracking performance rankings correlated with the performance of mixtures aged for longer periods.

Developing performance-based specifications for asphalt mixture design in Oregon

1. Test results for the Level 4 PMLC specimens show that M1, M2, and M3 projects meet the FI threshold of 8 while the PMLC mixtures for M6 and M7 projects fail to meet the requirement with average FI values close to 5. Lower FI value for project M7 can be explained by the relatively lower binder content for this asphalt mixture. The effective binder content and the VFA for the M7 asphalt mixture were lower than all other four Level 4 mixtures evaluated in this study.

2. Average measured FI values for all PMLC specimens from all Level 4 mixtures fall between the test results for LMLC specimens for AC_{design}-0.5% and AC_{design}+0.5%. Since plant produced mixtures are allowed to have ±0.5% variability in production binder content in Oregon, it can be concluded for Level 4 mixtures that laboratory mixing process is not introducing any observable bias into the measured FI values.

3. VFA and effective binder content are highly correlated with both FI_{OBC} and FI_{PM}. With correlation coefficients close 0.57, VFA is the most significant factor controlling the cracking resistance of asphalt mixtures. According to the correlation matrix, there is also a statistically significant relationship between dust-to-binder ratio and FI. These statistically significant correlations between asphalt mixture variables

and FI also indicate that FI parameter is able to capture the impact of asphalt mixture properties on cracking resistance.

4. Based on the results of all analyses, an FI threshold of 6 was recommended for Level 3 mixes while the threshold for Level 4 was selected as 8.

5. The differences between the laboratory and plant mixing processes directly control the FN test results. FN results for PMLC mixtures are always significantly lower than results for LMLC mixes. This difference avoids any direct comparisons between the rut resistance of PMLC and LMLC asphalt mixtures. Due to the significant difference in FN values for PMLC and LMLC mixes, the effectiveness of the FN test in terms of characterizing rut resistance also becomes questionable.

6. Weak correlations between asphalt mixture variables and FN indicate that FN parameter is not able to capture the impact of asphalt mixture properties on rutting resistance.

7. Due to the conclusions 11 and 12 above, FN test may not be an ideal test for balanced mix design and performance based specifications.

8. For the HWTT results, the impact of binder content on measured rutting performance is not as clear as the FN test results. In addition, for M4 and M5 Level 3 mixes, an unexpected trend was observed for lower binder contents. This unexpected trend might be a result of the presence of latex in the mixture. In addition, since HWTT experiments were conducted under water, test results are also affected by the moisture susceptibility of the asphalt mixture in addition to rut resistance. Combined effect of moisture and rut resistance reflected in the test results might be increasing the variability of the test.

9. For LMLC specimen test results at the design binder content, M2 asphalt mixture had rutting resistance significantly higher than all other Level 4 mixtures. Higher rutting resistance for this mixture is likely to be a result of the relatively lower design and effective binder content for this mixture. Since M2 also had high cracking resistance despite its significantly lower binder content than all other Level 4 mixes, it can be considered to be a high performance and cost effective asphalt mixture.

10. HWTT rut depths are highly correlated with the asphalt mixture properties. In addition, HWTT is a practical experiment that does not require a hydraulic test equipment. Laboratory mixing and specimen production are also not introducing any significant bias into the HWTT results and produce test samples that represent specimens prepared with production mixtures. For these reasons, HWTT is recommended as the experiment for rutting performance quantification for balanced mix design and performance based specifications.

11. Based on the results of all analyses, an RD threshold of 3mm was recommended for Level 3 mixes while the threshold for Level 4 was selected as 2.5mm.

12. Reasonable results were achieved by following the developed balanced mix design process.

13. It might be possible to perform balanced mix design by just using three binder contents (AC_{design} from volumetric mix design, AC_{design}-0.5%, AC_{design}+0.5%). However, this conclusion should be further evaluated in a future study by using more extensive datasets (or during the shadow balanced mix design implementation phase).

5.2 RECOMMENDATIONS

Based on the comprehensive literature review and the results of this research study, volumetric design with performance verification (Approach 1) and performance modified volumetric mixture design (Approach 2) were recommended as the most effective BMD approaches for Oregon. Recommended processes for the implementation of these two approaches are described below:

5.2.1 Approach 1 - Volumetric design with performance verification

Based on the cracking and rutting test results and the thresholds selected in this study, volumetric design binder content of some asphalt mixtures need to be increased (M4, M5, and M7), some need to be decreased (M1, M6, and M8), and some are within the binder content range (M2 and M3) determined in this study. Since the asphalt mixture needs to satisfy both volumetric and performance testing criteria according to Approach 1, asphalt mixtures that do not meet the performance criteria need to be redesigned volumetrically. The adjustments to the asphalt mixture can be made through aggregate gradation, binder source, binder grade, additives (using ER binder), and recycled asphalt content. The binder content selected by the second volumetric design should be within the binder content interval suggested by the performance test results. Production mixtures should also be tested for cracking and rutting resistance to ensure that they meet the specified rutting and cracking performance criteria. It should be noted that the thresholds for rutting and cracking performance suggested in this study are based on the laboratory test results of 8 mixes (from 8 construction projects) investigated in this study. These suggested thresholds can be changed based on the data from the shadow specification implementation phase (first year).

5.2.2 Approach 2 - Performance modified volumetric mixture design

In this approach, the original asphalt content will be determined using the volumetric design process. Samples will be prepared at the determined optimum asphalt content, at one 0.5% decrement, and at two 0.5% increments. Rutting and cracking tests will be performed with asphalt specimens at these binder contents. The minimum binder content that satisfies cracking criterion and the maximum binder content that satisfies rutting criterion are determined. Then, the midpoint of the “acceptable asphalt binder content range” (average of minimum and maximum binder content allowed) will be used as the design asphalt binder content. In this method, volumetric mix design requirements are not strictly enforced.

5.3 FUTURE WORK

Suggestions for potential future research are provided below:

1. In this study, thresholds for FI and RD are selected to create reasonable binder content intervals. If the balanced mix design process is implemented in Oregon in a shadow specification, these thresholds can be modified within the first year based on the laboratory measured rutting and cracking performance.

2. Problems with uniformly coating aggregates with latex for the LMLC specimens resulted in FI values that are significantly lower than the FI values for the PMLC specimens. In a future study, a special laboratory procedure should be developed to more uniformly coat aggregates with the latex solution to be able to produce latex modified LMLC asphalt mixtures that more closely represent production mixtures.

3. Since HWTT is conducted under water, test results are also controlled with the moisture susceptibility of asphalt mixtures in addition to rut resistance. The possibility of using different parameters (such as creep slope shown in Figure 2.20) for rutting performance evaluation should also be investigated in a future study. The effectiveness of conducting HWTT without submerging the specimens in water (by controlling the temperature via a heating unit instead) should also be investigated in a future study.

6.0 REFERENCES

American Association of State Highway Transportation Officials. (2010). *Determining the Asphalt Binder Content of Hot Mix Asphalt (HMA) by the Ignition Method* (Standard No. T 308-10). Washington, DC.

American Association of State Highway Transportation Officials. (2010). *Standard Practice for Mixture Conditioning of Hot Mix Asphalt (HMA)* (Standard No. R 30-10). Washington, DC.

American Association of State Highway Transportation Officials. (2011). *Viscosity Determination of Asphalt Binder Using Rotational Viscometer* (Standard No. T 316-11). Washington, DC.

American Association of State Highway Transportation Officials. (2012). *Preparing and Determining the Density of Hot Mix Asphalt (HMA) Specimens by Means of the Superpave Gyratory Compactor* (Standard No. T 315-12). Washington, DC.

American Association of State Highway Transportation Officials. (2012). *Standard Method of Test for Determining the Rheological Properties of Asphalt Binder Using a Dynamic Shear Rheometer (DSR)* (Standard No. T 315-12). Washington, DC.

American Association of State Highway Transportation Officials. (2013). *Standard method for determining the dynamic modulus and flow number for hot mix asphalt (HMA) using the asphalt mixture performance tester (AMPT)* (Standard No. TP 79-13). Washington, DC.

American Association of State Highway Transportation Officials. (2014). *Standard Method of Test for Sieve Analysis of Fine and Coarse Aggregates. American Association of State Highway and Transportation Officials* (Standard No. T 27-14). Washington, DC.

American Association of State Highway Transportation Officials. (2015). *Standard Method of Test for Quantitative Extraction and Recovery of Asphalt Binder from Asphalt Mixtures* (Standard No. T 319-15). Washington, DC.

Anderson, R., King, G., Hanson, D., & Blankenship, P. (2011). Evaluation of the Relationship between Asphalt Binder Properties and Non-Load Related Cracking. In *Asphalt paving technology 2011: Journal of the Association of Asphalt Paving Technologists, Tampa Florida, March 27-30, 2011* (Vol. 80). Lancaster, PA: DEStech Publications.

Arega, Z. A., Bhasin, A., & Kesel, T. D. (2013). Influence of extended aging on the properties of asphalt composites produced using hot and warm mix methods. *Construction and Building Materials, 44*, 168-174. doi:10.1016/j.conbuildmat.2013.02.081

Baek, C., Underwood, B. S., & Kim, Y. R. (2012). Effects of Oxidative Aging on Asphalt Mixture Properties. *Transportation Research Record: Journal of the Transportation Research Board, 2296*(1), 77-85. doi:10.3141/2296-08

Bell, C., & Sosnovske, D. (1994). *Aging: Binder Validation.* (Publication No. SHRP-A-384). Washington D.C.: Strategic Highway Research Program National Research Council.

Bell, C., Abwahab, Y., Cristi, M., & Sosnovske, D. (1994). *Selection of Laboratory Aging Procedures for Asphalt-Aggregate Mixtures* (Publication No. SHRP-A-383). Washington D.C.: Strategic Highway Research Program National Research Council.

Bennert, T., Fee, F., Sheehy, E., Blight, R., & Sauber, R. (2011). Implementation of Performance-Based HMA Specialty Mixtures in New Jersey. In *Asphalt Paving Technology 2011* (Vol. 80, pp. 719-758). Tampa, FL: Association of Asphalt Paving Technologists (AAPT).

Bennert, T., Sheehy, E., Blight, R., Gresavage, S., & Fee, F. (2014). Implementation of Performance-Based Specifications for Asphalt Mix Design and Production Quality Control for New Jersey. In *Transportation Research Board 93rd annual meeting: January 12-16, 2014: Washington, D.C.* (pp. 13-31). Washington, DC: Transportation Research Board.

Biligiri, K. P., Kaloush, K. E., Mamlouk, M. S., & Witczak, M. W. (2007). Rational Modeling of Tertiary Flow for Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board, 2001*(1), 63-72. doi:10.3141/2001-08

Bonaquist, R., Christensen, D., & Stump, W. (2003). *Simple performance tester for superpave mix design: First-article development and evaluation* (Vol. 513). Washington D.C.: Transportation Research Board. doi:10.17226/21954

Branthaver, J., Petersen, J., Roberston, R., Duvall, J., Kim, S., Harnsberger, P., . . . Schabron, J. (1993). *Binder Characterization and Evaluation Volume 2: Chemistry* (Publication No. SHRP-A-368). Washington D.C.: Strategic Highway Research Program National Research Council. Retrieved from http://onlinepubs.trb.org/onlinepubs/shrp/SHRP-A-368.pdf

Chen, C., Yin, F., Turner, P., West, R. C., & Tran, N. (2018). Selecting a Laboratory Loose Mix Aging Protocol for the NCAT Top-Down Cracking Experiment. *Transportation Research Record: Journal of the Transportation Research Board, 2672*(28), 359-371. doi:10.1177/0361198118790639

Coleri, E., Haddadi, S., Sreedhar, S., Lewis, S., Zhang, Y., & Wruck, B. (2018). *Binder-Grade Bumping and High Binder Content to Improve Performance of RAP-RAS Mixtures* (Publication No. FHWA-OR-RD-18-05). Salem, OR: Oregon Department of Transportation.

Coleri, E., S. Sreedhar, S. Haddadi, & B. Wruck. (2018). *Adjusting asphalt mixes for increased durability and implementation of a performance tester to evaluate fatigue cracking of asphalt concrete* (Publication No. FHWA-OR-RD-18-06). Salem, OR: Oregon Department of Transportation.

Cominsky, R., Huber, G., Kennedy, T., & Anderson, M. (1994). *The superpave mix design manual for new construction and overlays* (Publication No. SHRP-A-407). Washington D.C.: Strategic Highway Research Program National Research Council.

Cooper, S. B., Mohammad, L. N., Kabir, S., & King, W. (2014). Balanced Asphalt Mixture Design through Specification Modification. *Transportation Research Record: Journal of the Transportation Research Board, 2447*(1), 92-100. doi:10.3141/2447-10

Dave, E., & Koktan, P. (2011). *Synthesis of Performance Testing of Asphalt Concrete* (Publication No. MN/RC 2011-22). St. Paul, MN: Minnesota Department of Transportation. Retrieved from https://pdfs.semanticscholar.org/ad11/b4766562375906462f1e002ddbe44f1577b2.pdf.

Dongré, R., D'angelo, J., & Copeland, A. (2009). Refinement of Flow Number as Determined by Asphalt Mixture Performance Tester. *Transportation Research Record: Journal of the Transportation Research Board, 2127*(1), 127-136. doi:10.3141/2127-15

Epps, J., Hand, A., Seeds, S., Scholz, T., Alavi, S., Ashmore, C., . . . Leahy, R. (2002). *Recommended performance-related specification for hot-mix asphalt construction: Results of the Wes Track project* (Publication No. 455). Washington D.C.: Transportation Research Board. Retrieved from http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_455a.pdf

Francken, L. (1977). Pavement Deformation Law of Bituminous Road Mixes in Repeated Load Triaxial Compression. In *Fourth International Conference Structural Design of Asphalt Pavements: Proceedings: August 22-26, 1977, The University of Michigan, Ann Arbor, Mich*. Ann Arbor, MI: University of Michigan.

Glover, C., Davison, R., Domke, C., Ruan, Y., Juristyarini, P., Knorr, D., & Jung, S. (2005). *Development of a new method for assessing asphalt binder durability with field validation* (Publication No. FHWA/TX-05/1872-2). Austin, TX: Texas Department of Transportation.

Harvey, J., Liu, A., Zhou, J., Signore, J., Coleri, E., & He, Y. (2014). *Superpave Implementation Phase II: Comparison of Performance-Related Test Results* (Publication No. UCPRC-RR-2015-01). Berkeley, CA: University of California Pavement Research Center. Retrieved from http://www.ucprc.ucdavis.edu/PDF/UCPRC-RR-2015-01.pdf

Harvey, J., Signore, J., Wu, R., & Basheer, I. (2014). Performance-based specifications: California experience to date. In *Transportation Research Board 93rd annual meeting: January 12-16, 2014: Washington, D.C.* Washington, DC: Transportation Research Board.

Haddadi, S. S., Coleri, E., & Sreedhar, S. (2019). Strategies to improve performance of reclaimed asphalt pavement-recycled asphalt shingle mixtures. *International Journal of Pavement Engineering*, 1-12. doi:10.1080/10298436.2019.1593411

Hughes, C., & Maupin, G. (2000). *The journey toward end result specifications for asphalt concrete* (Tech. No. VTRC 0I-TAR7). Charlottesville, VA: Virginia Transportation Research Council. Retrieved from http://www.virginiadot.org/vtrc/main/online_reports/pdf/01-tar7.pdf

Isola, M., Chun, S., Roque, R., Zou, J., Koh, C., & Lopp, G. (2014). Development and Evaluation of Laboratory Conditioning Procedures to Simulate Mixture Property Changes Effectively in the Field. *Transportation Research Record: Journal of the Transportation Research Board, 2447*(1), 74-82. doi:10.3141/2447-08

Kim, Y., Castorena, C., Elwardany, M., Rad, F., Underwood, S., Gundla, A., . . . Glaser, R. (2018). *Long-Term Aging of Asphalt Mixtures for Performance Testing and Prediction* (Publication No. 871). Washington D.C.: NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM. doi:10.17226/24959

Kumbargeri, Y. S., & Biligiri, K. P. (2015). A Novel Approach to Understanding Asphalt Binder Aging Behavior Using Asphaltene Proportion as a Performance Indicator. *Journal of Testing and Evaluation, 44*(1), 439-449. doi:10.1520/jte20140490

Liu, M., Ferry, M. A., Davison, R. R., Glover, C. J., & Bullin, J. A. (1998). Oxygen Uptake As Correlated to Carbonyl Growth in Aged Asphalts and Asphalt Corbett Fractions. *Industrial & Engineering Chemistry Research, 37*(12), 4669-4674. doi:10.1021/ie980450o

McCarthy, L., Callans, J., Quigley, R., & Scott III, S. (2016). *Performance Specifications for Asphalt Mixtures* (Publication No. 492). Washington D.C.: National Academies of Sciences, Engineering, and Medicine. doi:10.17226/23564

Mohammad, L. N., Kim, M., & H. Challa. (2016). *Development of Performance-based Specifications for Louisiana Asphalt Mixtures* (Report No. FHWA/LA.14/558). Louisiana Transportation Research Center, Baton Rouge, LA.

Newcomb, D., Martin, A., Yin, F., Arambula, E., Park, E., Chowdhury, A., . . . Signore, J. (2015). *Short-Term Laboratory Conditioning of Asphalt Mixtures* (Publication No. 815). Washington D.C.: National Academies of Sciences, Engineering, and Medicine. doi:10.17226/22077

Newcomb, D., & Zhou, F. (2018). *Balanced Design of Asphalt Mixtures* (Publication No. MN/RC 2018-22). St. Paul, MN: Minnesota Department of Transportation. Retrieved from http://dot.state.mn.us/research/reports/2018/201822.pdf

Nsengiyumva, G. (2015). *Development of semi-circular bending (SCB) fracture test for bituminous mixtures* (Unpublished master's thesis). University of Nebraska - Lincoln.

Retrieved from https://digitalcommons.unl.edu/cgi/viewcontent.cgi ?article=1090&context=civilengdiss.

Ouyang, C., Wang, S., Zhang, Y., & Zhang, Y. (2006). Improving the aging resistance of styrene–butadiene–styrene tri-block copolymer modified asphalt by addition of antioxidants. *Polymer Degradation and Stability*, *91*(4), 795-804. doi:10.1016/j.polymdegradstab.2005.06.009

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. doi:10.1016/j.conbuildmat.2016.03.144c

Petersen, J. (2009). A Review of the Fundamentals of Asphalt Oxidation: Chemical, Physicochemical, Physical Property, and Durability Relationships. *Transportation Research Circular, E-C140*. doi:10.17226/23002

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*(1), 100-107. doi:10.3141/2507-11

Roque, R., Chun, S., Zou, J., Lopp, G., & Villiers, C. (2011). *Continuation of superpave projects monitoring* (Publication No. 00073127). Tallahassee, FL: Florida Department of Transportation.

Santagata, E., Anderson, D., & Marasteanu, M. (1996). A rheological investigation on the physical hardening of modified asphalt binders. In *Eurasphalt & Eurobitume Congress: 7-10 May 1996, Strasbourg* (Vol. 3). Breukelen, Netherlands: Foundation Eurasphalt.

Smith, B. C. (2011). *Fundamentals of Fourier transform infrared spectroscopy*. Boca Raton, FL: CRC Press.

Sreedhar, S., & Coleri, E. (2018). Effects of Binder Content, Density, Gradation, and Polymer Modification on Cracking and Rutting Resistance of Asphalt Mixtures Used in Oregon. *Journal of Materials in Civil Engineering, 30*(11). doi:10.1061/(asce)mt.1943-5533.0002506

Sreedhar, S., Coleri, E., & Haddadi, S. S. (2018). Selection of a performance test to assess the cracking resistance of asphalt concrete materials. *Construction and Building Materials, 179*, 285-293. doi:10.1016/j.conbuildmat.2018.05.258

Tsai, B.-W., Yu, J., Zhang, X., & Monismith, C. (2012). Development of asphalt pavement fatigue cracking prediction model based on loading mode transfer function. *Road Materials and Pavement Design, 13*(3), 501-517. doi:10.1080/14680629.2012.695240

Tsai, B., Coleri, E., Harvey, J. T., & Monismith, C. L. (2016). Evaluation of AASHTO T 324 Hamburg-Wheel Track Device test. *Construction and Building Materials, 114*, 248-260. doi:10.1016/j.conbuildmat.2016.03.171

Ullidtz, P., Harvey, J., Basheer, I., Jones, D., Wu, R., Lea, J., & Lu, Q. (2010). CalME, a Mechanistic–Empirical Program to Analyze and Design Flexible Pavement Rehabilitation. *Transportation Research Record: Journal of the Transportation Research Board, 2153*(1), 143-152. doi:10.3141/2153-16

Vallerga, B., & Lovering, W. (1984). Evolution of the Hveem stabilometer method of designing asphalt paving mixtures. In *Asphalt paving technology 1984: Proc., Association of Asphalt Paving Technologists, Technical Sessions, Scottsdale, Arizona, April 9-11, 1984* (Vol. 54, pp. 243-265). Ann Arbor, MI: AAPT.

Wang, Y., & Kim, Y. R. (2017). Development of a pseudo strain energy-based fatigue failure criterion for asphalt mixtures. *International Journal of Pavement Engineering, 20*(10), 1182-1192. doi:10.1080/10298436.2017.1394100

Waterways Experiment Station (U.S.) United States. Mississippi River Commission. (1948). *Investigation of the design and control of asphalt paving mixtures* (Vol. 1, Tech. No. 3-254). Vicksburg, MS: Army Corps of Engineers.

West, R., Rodezno, C., Leiva, F., & Yin, F. (2018). *Development of a Framework for Balanced Mix Design* (Publication No. NCHRP 20-07/Task 406). Auburn, AL: National Center for Asphalt Technology at Auburn University.

Williams, R., Hill, D., Hofmann, K., Zelenock, M., & Bausano, J. (2004). *Development of Laboratory Performance Test Procedures and Trial Specifications for Hot Mix Asphalt* (Publication No. RC-1410). Lansing, MI: Michigan Department of Transportation Construction and Technology Division. Retrieved from https://pdfs.semanticscholar.org/99d7/096f19f8b71e57ad0a507e1dbe04ecea1c91.pdf.

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. doi:10.1520/JAI12264

Yang, X., You, Z., & Mills-Beale, J. (2015). Asphalt Binders Blended with a High Percentage of Biobinders: Aging Mechanism Using FTIR and Rheology. *Journal of Materials in Civil Engineering, 27*(4). doi:10.1061/(asce)mt.1943-5533.0001117

Yildirim, Y., Jayawickrama, P., Hossain, M., Alhabshi, A., Yildirim, C., Smit, A., & Little, D. (2007). *Hamburg wheel-tracking database analysis* (Publication No. FHWA/TX-05/0-1707-7). Austin, TX: Texas Department of Transportation Research and Technology Implementation Office. Retrieved from https://static.tti.tamu.edu/tti.tamu.edu/documents/0-1707-7.pdf.

Yin, F., Arámbula-Mercado, E., Martin, A. E., Newcomb, D., & Tran, N. (2017). Long-term ageing of asphalt mixtures. *Road Materials and Pavement Design, 18*(Sup1), 2-27. doi:10.1080/14680629.2016.1266739

Zhou, F., Hu, S., & Scullion, T. (2006). *Integrated asphalt (overlay) mixture design, balancing rutting and cracking requirements* (Tech. No. FHWA/TX-06/0-5123-1). Austin, TX: Texas Department of Transportation Research and Technology Implementation Office.

Zhou, F., Scullion, T., Walubita, L., & Wilson, B. (2014). Implementation of a Performance-Based Mix Design System in Texas. In *Transportation Research Board 93rd annual meeting: January 12-16, 2014: Washington, D.C.* (pp. 32-50). Washington, DC: Transportation Research Board.

Zhu, Z., Singhvi, P., Espinoza-Luque, A. F., Ozer, H., & Al-Qadi, I. L. (2019). Influence of mix design parameters on asphalt concrete aging rate using I-FIT specimens. *Construction and Building Materials, 200,* 181-187. doi:10.1016/j.conbuildmat.2018.12.099