TRC1802

Performance-Based Asphalt Mixture Design (PBD) for Arkansas

Kevin D. Hall, Ph.D., P.E.
Elvis Castillo-Camarena
Andrew F. Braham, Ph.D., P.E.
Nathan Parnell
Allyson Richey

University of Arkansas – Fayetteville
DTS, Inc. – Springdale Arkansas
Kimley-Horn, Inc. – Dallas, Texas

Final Report

December 2022

FINAL REPORT

TRC1802

Performance-Based Asphalt Mixture Design (PBD) for Arkansas

By

Kevin D. Hall, Ph.D., P.E.
Elvis Castillo-Camarena
Andrew F. Braham, Ph.D., P.E.
University of Arkansas

Nathan Parnell
DTS, Inc. – Springdale Arkansas

Allyson Richey
Kimley-Horn, Inc. – Dallas, Texas

Conducted by
Department of Civil Engineering
University of Arkansas

In Cooperation with
Arkansas Department of Transportation (ARDOT)
U.S. Department of Transportation
Federal Highway Administration

University of Arkansas
Fayetteville, AR 72701

December 2022

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. TRC1802	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Performance-Based Asphalt Mixture Design (PBD) for Arkansas		5. Report Date December 2022 6. Performing Organization Code:
7. Author(s) Kevin D. Hall Elvis Castillo-Camarena Andrew F. Braham Nathan Parnell Allyson Richey		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Arkansas – Fayetteville 4190 Bell Engineering Center Fayetteville, AR 72701		10. Work Unit No. 11. Contract or Grant No. Project TRC1802
12. Sponsoring Agency Name and Address Arkansas Department of Transportation (SPR) 10324 Interstate 30 Little Rock, AR 72209		13. Type of Report and Period 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This project sought to develop/adapt and implement a 'cracking test' for asphalt mixture design, to use in conjunction with the current Asphalt Pavement Analyzer rutting test to shift mixture design in Arkansas to a performance, rather than volumetric, basis. In addition, the project attempted to re-evaluate the effectiveness of the current method for estimating resistance to moisture damage - and possibly develop/adapt alternate methods as appropriate. Six sites for material collection were leveraged around the state to ensure a diversity in quarried materials, mixes, and binders. Based on an evaluation to use a cracking-related performance testing in a "go/no-go" decision point process during mix design, the preliminary recommendation for Arkansas is to use the AASHTO R30 short-term aging combined with the NCAT long-term aging protocol. Finally, this study clearly indicates the potential for using the IDEAL-CT as a cracking-related 'index' type test for performance-engineered asphalt mixture design. From a practicality perspective, the IDEAL-CT offers attractive advantages: (1) potentially significant reduction in equipment costs – particularly if a given laboratory possesses a suitable load frame, e.g. a Marshall test setup; (2) significant reduction in test specimen preparation, by eliminating the need for precise sawing. From a mixture design perspective, the IDEAL-CT appears to offer similar/comparable test results to the SCB/IFIT system.

17. Key Words Asphalt mixture, balanced mix design, rutting, cracking, aging, IDEAL-CT, Asphalt Pavement Analyzer		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 82	22. Price

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

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

APPROXIMATE CONVERSIONS FROM SI UNITS

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

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE
INTRODUCTION

The current asphalt mixture design procedure used in Arkansas, commonly referred to as the "Superpave" procedure, is based on AASHTO M323 and AASHTO R35. This design system is based on component specifications (quality and properties of the asphalt binder and aggregates) and volumetric properties. With volumetric based design, engineering properties of asphalt mixtures are controlled only indirectly, influenced by properties of the components and proportions of each component. Originally, Superpave was intended to include performance-based tests – that is, tests that measure engineering properties directly related to performance – but these tests proved to be non-implementable. As a result, the Superpave process described in M323 and R35 is based solely on volumetric properties. Arkansas is one of many states which have supplemented the volumetric design system in Superpave with a laboratory 'performance' test, e.g. rutting potential, as measured by the Asphalt Pavement Analyzer (APA).

Performance-Based Mixture Design (PBD) includes performance tests to evaluate rutting potential, cracking potential and moisture resistance during the mixture design process. Performance tests provide a more direct evaluation of expected performance than volumetric properties. Such tests can better characterize the effect of new materials and processes (e.g. RAP, RAS, Warm-Mix) as well as changes to mix design criteria. Nationally, rutting tests are the most common performance-related test currently performed for mixture design. Comparatively few agencies have adopted tests to characterize cracking resistance.

Recent surveys have indicated concern that current mixture design procedures do not ensure adequate field performance. Of prime concern is early-age cracking in asphalt pavements. In addition, surveys suggest possible differences in the cracking performance of mixtures with different aggregates (e.g. sandstone versus limestone) in Arkansas. TRC1404 – Evaluating Performance of Asphalt Pavement Based on Data Collected During IRP – identified mixture design related issues as a potential cause to early-age pavement distress.

This project sought to develop/adapt and implement a 'cracking test' for asphalt mixture design, to use in conjunction with the current APA rutting test (ARDOT Test Method 480) to shift mixture design in Arkansas to a performance, rather than volumetric, basis. In addition, the project attempted to re-evaluate the effectiveness of the current method for estimating resistance to moisture damage (ARDOT Test Method 455A) - and possibly develop/adapt alternate methods as appropriate.

Specific project objectives included:

1. Document the current state-of-the-practice concerning asphalt cracking and moisture damage tests, to include: testing specifications and protocols; state Departments of Transportation policies, procedures, and specifications related to the implementation and use of these tests; agency experiences with implementation; and any other pertinent information.

2. Identify, develop, and/or adapt laboratory tests related to cracking resistance and resistance to moisture damage, for implementation into current mixture design procedures. Provide testing specifications (in AASHTO format) as necessary.

3. Develop initial mixture acceptance criteria for recommended cracking and moisture damage tests.

4. Suggest a framework for procedures to validate recommended acceptance criteria.

5. Provide recommendations for changes to ARDOT's *Standard Specifications for Highway Construction* and/or *Roadway Design Plan Development Guidelines* necessitated by the implementation of cracking and moisture damage tests.

This report is organized around these objectives. **Chapter 2** summarizes the state-of-the-practice concerning asphalt cracking tests and the balanced mixture design concept. **Chapter 3** discusses the selection of specific paving projects, for which field cracking performance was readily available, to include in the study. Mixtures from these projects were re-created in the laboratory and tested; laboratory test results were validated against observed field performance. **Chapter 4** presents and discusses the data generated during the study. **Chapter 5** includes the determination of initial/trial acceptance criteria for the recommended cracking test. **Chapter 6** lists specific recommendations regarding the adoption and initial implementation of a cracking test for Arkansas, and outlines recommendations related to further field studies needed to fully validate the recommended mixture acceptance criteria.

CHAPTER TWO
STATE OF THE PRACTICE

Nationally, there is significant concern that asphalt pavements are experiencing premature distress and failure. A survey conducted for NCHRP Synthesis 492 "Performance Specifications for Asphalt Mixtures" (McCarthy et al., 2016) indicated there is concern among state highway agencies (SHAs) that the current asphalt mixture design procedures do not ensure good field performance. As discussed in Chapter 1, current mix design procedures (Superpave) are based on the volumetric properties of a given mix; no performance-related tests are included in the procedure. To ensure mixes have a higher potential for performance many states have added one or more performance-related tests to the mixture design process. Mohammad (2016) reported survey results which indicated that many SHAs – 21 out of 27 – include laboratory mechanical tests in their mixture design specifications. The most common test was for moisture damage; in addition, a majority of those states (14 out of 21) are using a loaded wheel tracking test to indicate rutting potential (either the Asphalt Pavement Analyzer or Hamburg Wheel Tracking Test). However, most SHAs do not yet require a test to assess the cracking resistance of a mix.

The Arkansas Department of Transportation (ARDOT) noted that several projects constructed during its first interstate rehabilitation program (IRP) showed signs of premature deterioration after only 9-12 years of service. Ten projects were selected for study under TRC-1404. After evaluating the ten projects, it was determined that the bond strength, in-place air voids, and moisture damage contributed to the premature deterioration. Researchers recommended that ARDOT consider updating its moisture susceptibility testing procedures. In addition, it was recommended that mixture specifications be reviewed – including items such as air void, VMA, and 'N-design' (number of gyrations for compaction of laboratory mixes). Implementation of a cracking resistance assessment in the mix design process could also address these concerns.

2.1 Mixture Design

Using performance-related tests to augment volumetric analyses gives rise to the concept of performance-based mix design (also known as "balanced" mix design, or BMD). In September 2015, the FHWA Expert Task Group (ETG) on Asphalt Mixture and Construction formed a Task

Force on Balanced Mixture Design (BMD) to advance changes in the formulation of asphalt mixtures. The task group has defined BMD as ***"Asphalt mixture design using performance tests on appropriately conditioned specimens that address multiple modes of distress taking into consideration mixture aging, traffic, climate and location within the pavement structure."*** The objective of BMD is to design asphalt mixtures for performance using a rational approach instead of relying on strictly volumetric guidelines. The concept is straightforward; Figure 1 provides an illustration. Generally, potential of rutting will increase as the binder content of the mix increases – thus, rutting potential "sets" the upper limit of binder content. Conversely, resistance to cracking will decrease as the binder content decreases – thus, cracking resistance "sets" the lower limit of binder content. There will be a range of 'acceptable' binder contents between these two limits, which can be refined using volumetric specifications.

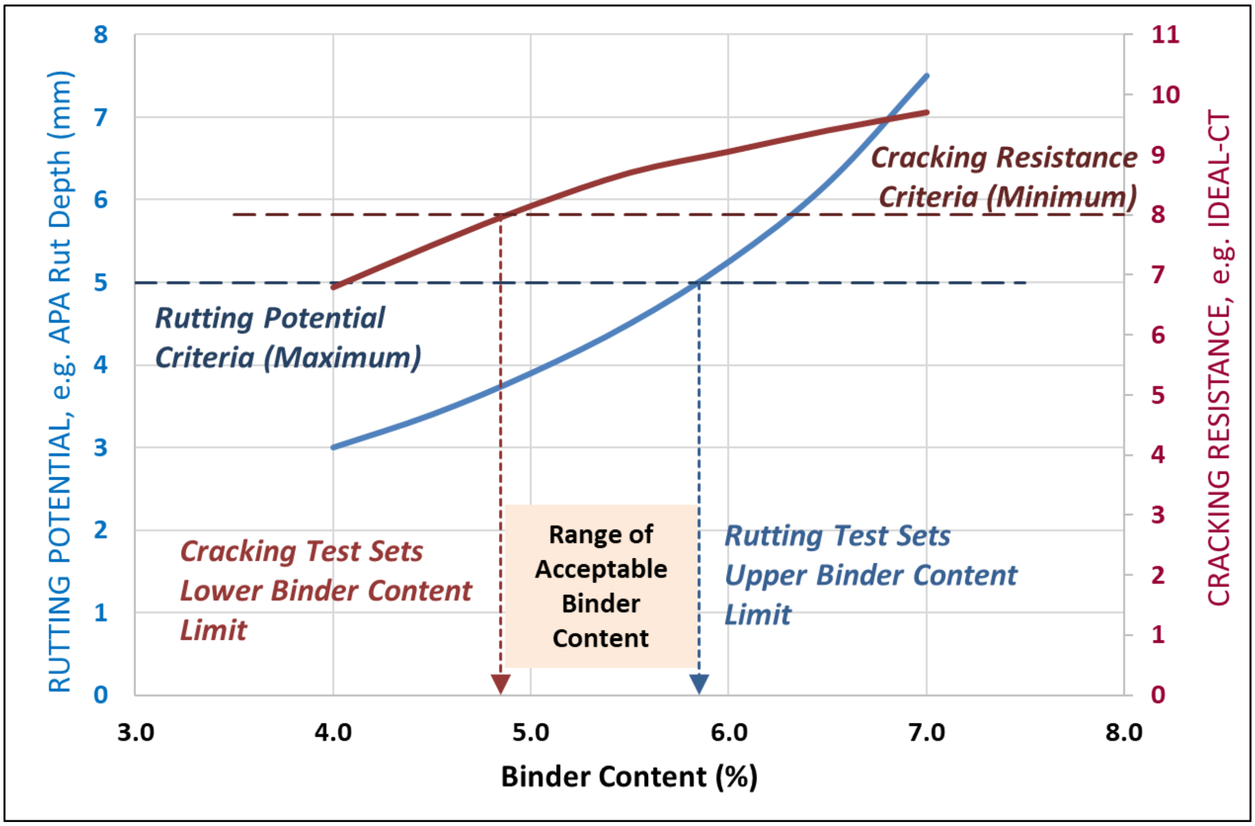


Figure 1. Performance Mix Design Concept

The ETG task force identified three potential approaches for implementing performance mix design principles into routine practice. Figure 2 provides a flow diagram to illustrate the approaches, which are described in the listing which follows.

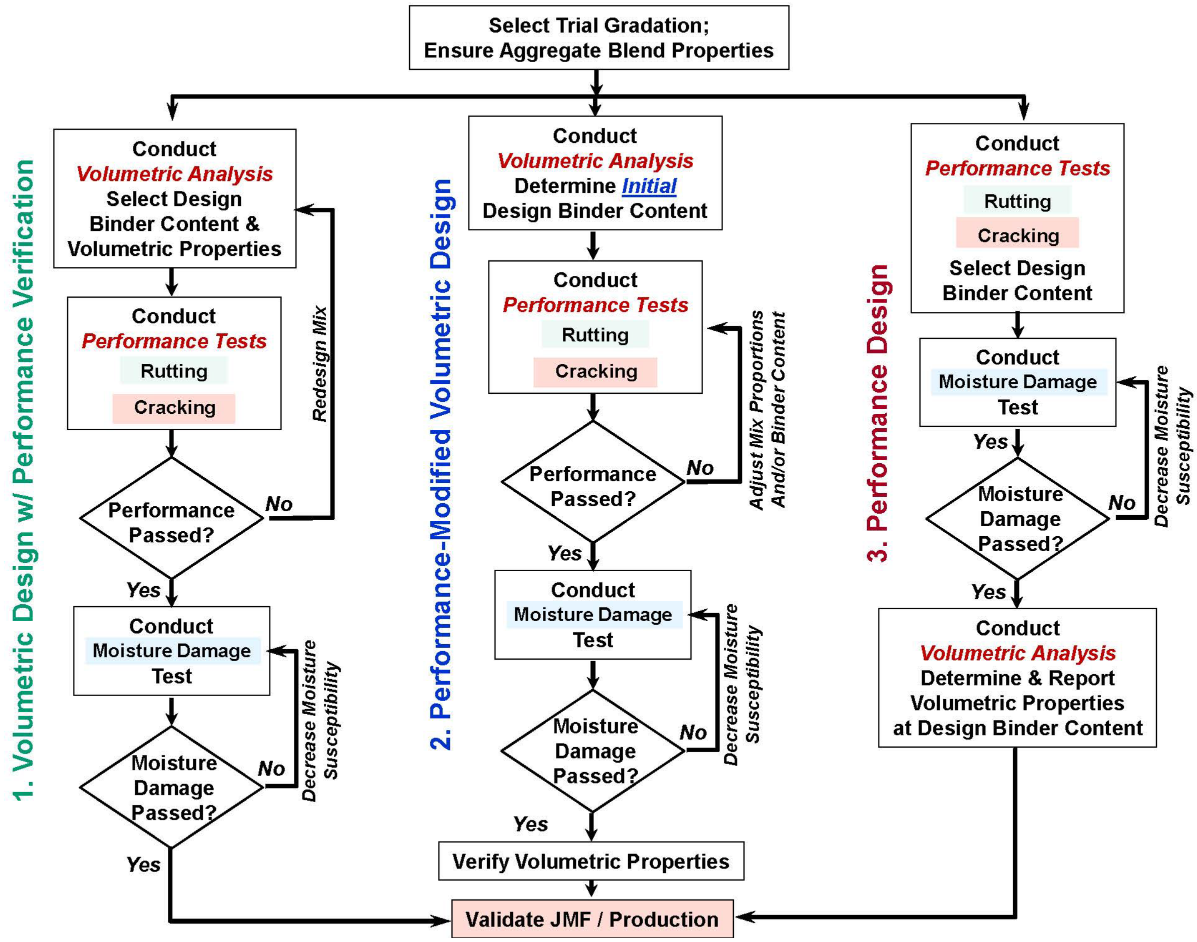


Figure 2. Potential Approaches to Performance Mix Design (after FHWA-ETG).

1. **Volumetric Design with Performance Verification.** Mixes are designed using volumetric properties; performance properties are subsequently confirmed. Both volumetric and performance test results must meet established criteria. If performance properties require adjustment, a new design is executed, i.e. aggregate and/or asphalt binder properties or proportions would be changed. This is the most common approach currently in use by state highway agencies.
2. **Performance-modified Volumetric Design.** Volumetric design is only a preliminary step; mixture adjustments are made to either component or volumetric properties on the basis of performance test results. The final design is primarily driven by performance test criteria. Mixes may be required to meet only a subset of volumetric criteria, or none at all.

3. **Performance Design.** As the name implies, mixture design is based on performance properties, with limited or no requirements related to volumetric properties (outside, perhaps, of minimum requirements for asphalt binder and/or aggregate properties). Component properties are selected to meet the performance test criteria; air voids and VMA might be non-mandatory suggestions.

As part of its work, the ETG task force surveyed states to assess current efforts regarding performance (balanced) mix design. Figure 3 shows the results of the survey; Table 1 provides details on states' approaches to implementation.

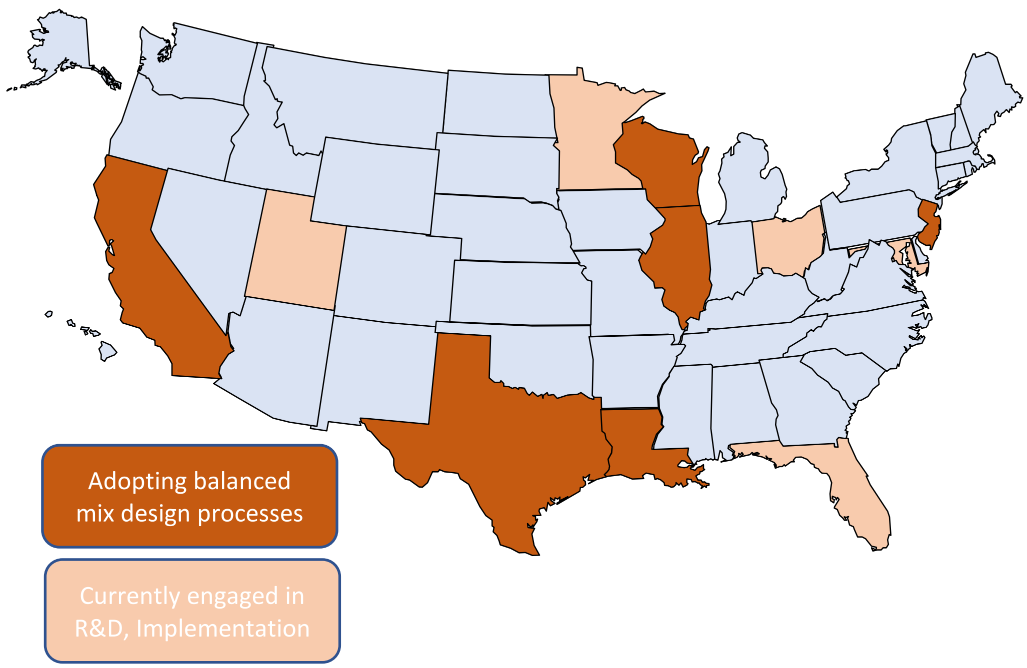


Figure 3. Balanced (Performance) Mix Design Efforts in the U.S.

Arkansas has achieved success in characterizing the rutting potential of asphalt mixtures through the use of the Asphalt Pavement Analyzer (APA). To fully implement the concept of a performance mixture design, a test for characterizing the cracking resistance of mixtures is needed. The subsection which follows describes national efforts related to the identification of tests suitable for characterizing the cracking resistance of asphalt mixtures.

Table 1. FHWA-ETG Survey Results of Balance Mix Design (January 2016)

State	Aggregate Properties	Aggregate Gradation	Binder Grade	Binder Quantity	Observed Mix Design Adjustments
Model A: Superpave plus Performance					
Illinois Building 8 projects this year	Same FAA education	Same	Same	Same Superpave	RAP and RAS quantities Binder source change Construction: silo time, aggregate moisture, plant temperatures
Texas All specialty mixes for 2-3 years	Same	Same	Same	Same Superpave	Asphalt content Binder source change Gradation adjustment for fines (P200) Aggregate source changes
Wisconsin 4 projects last year	Same	Same	Same	Waive VFA Superpave	Binder source and additives Aggregate gradation and fines Rubber
Louisiana	Same	Same	Same	Same	
New Jersey All specialty mixes - 5-10% of statewide tonnage	Same	Same	Open	Same	WMA Rejuvenators Polymers Changing effective asphalt content
Model B: Superpave ± plus Perf.					
California 7 Interstate projects to date.	Same - Min. is starting point; usually have to exceed these	Same	Same	Same - May go outside tolerances pending perf. test results Hveem and Superpave	Binder source / Aggregate source Binder content Dust : Asphalt ratio Currently developing mix guidance steps (easy and least costly to more difficult and costly) – Report will be available in April.
Model C: Performance					
New Jersey Proposed	Same	Same	Open	Optimum AC determined between lowest and highest asphalt contents from performance tests. A field production tolerance is set at ±0.3% on the optimum.	To be determined

2.2 Cracking Resistance

Characterizing the cracking resistance of asphalt mixtures as part of the mixture design process is receiving significant national attention. Work is ongoing to identify test procedures suitable for implementation into mix design. Zhou, et. al. (2016), working under NCHRP 9-57, identified candidate cracking-related test procedures and proposed a field validation plan to assess the efficacy of the various tests. The NCHRP 9-57 project team conducted a workshop with selected individuals (28 participants total) to select candidate tests for possible future field validation efforts. The results of these efforts are summarized in Tables 2 through 4.

Seven attributes of each cracking test were weighted for the selection process:

(a) availability of the test method

(b) test simplicity

(c) test variability

(d) sensitivity to mix parameters

(e) complexity of the data analysis

(f) availability and cost of test equipment

(g) laboratory-to-field correlation

Table 2. Cracking Tests Considered for the NCHRP 9-57 Workshop *(Zhou, et. al. [2016])*

Thermal Cracking Tests	**Reflection Cracking Tests**	**Bottom-Up Fatigue Cracking Tests**	**Top-Down Cracking Tests**
DCT	OT	Beam Fatigue	IDT-Florida
SCB (AASHTO TP105)	BBF	S-VECD	SCB-LTRC
SCB-IL	SCB-LTRC	Repeated Tension	S-VECD
IDT (AASHTO T322)	DCT	OT	Repeated Tension
TSRST/UTSST		SCB-LTRC	Modified OT

DCT: Disc-shaped Compact Tension; OT: Texas Overlay Test
SCB: Semi-Circular Bend Test BBF: Bending Beam Fatigue Test
IDT: Indirect Tension Test S-VECD: Simplified Visco-Elastic Continuum Damage
LTRC: Louisiana Transportation Research Center TSRST: Thermal Stress Restrained Specimen Test
UTSST: Uniaxial Thermal Stress and Strain Test

Table 3. Top Cracking Tests Identified at the NCHRP 9-57 Workshop *(Zhou, et. al. [2016])*

Thermal Cracking Tests	**Reflection Cracking Tests**	**Bottom-Up Fatigue Cracking Tests**	**Top-Down Cracking Tests**
DCT	OT	Beam Fatigue	IDT-Florida
SCB-IL	SCB-LTRC	SCB-LTRC	SCB-LTRC
SCB (AASHTO TP105)	BBF	OT *(added by request of the NCHRP 9-57 panel)*	

DCT: Disc-shaped Compact Tension; OT: Texas Overlay Test
SCB: Semi-Circular Bend Test BBF: Bending Beam Fatigue Test
IDT: Indirect Tension Test LTRC: Louisiana Transportation Research Center

Table 4. Final Selection of Cracking Tests at the NCHRP 9-57 Workshop *(Zhou, et. al. [2016])*

DCT	SCB-IL	SCB-TP105	SCB-LTRC
OT	BBF	IDT-Florida	

DCT: Disc-shaped Compact Tension; OT: Texas Overlay Test
SCB: Semi-Circular Bend Test BBF: Bending Beam Fatigue Test
IDT: Indirect Tension Test LTRC: Louisiana Transportation Research Center

The focus of TRC1802 is load-related cracking; that is, reflection cracking, bottom-up fatigue, and top-down cracking. The combination of Arkansas' climate, the Superpave binder specification, and ARDOT policy regarding binder selection has rendered thermal cracking extremely rare. When considering the various cracking tests identified by NCHRP 9-57, it is instructive to examine the cracking parameter (or property) which is used to assess the cracking potential of an asphalt mixture, and the degree to which the test has been correlated to field performance (to date). Table 5 summarizes findings from Zhou, et. al. (2016) for load-related cracking only.

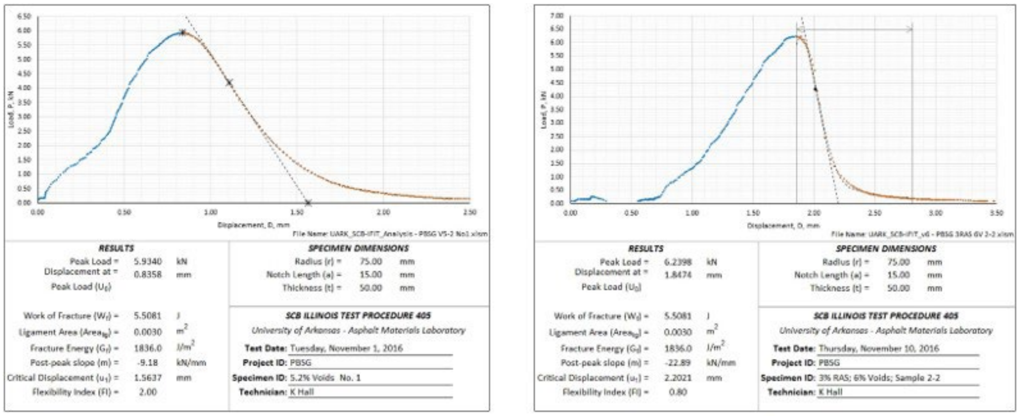
It is anticipated that the cracking tests identified and evaluated in NCHRP 9-47 will be examined for use in TRC1802. For example, the University of Arkansas - Fayetteville tested asphalt samples from a hot-mix asphalt plant located in south-central Arkansas using the semi-circular bend test (SCB) at intermediate (room) temperatures. The data was analyzed using the Illinois Flexibility Index test (iFIT). Figure 4 presents the results of two of the tests; differences in 'ductile' versus 'brittle' behavior is apparent – suggesting that the test is able to discriminate among mixture performance.

Table 5. Cracking Parameters and Field Performance Correlations for Cracking Tests

(after Zhou, et. al. [2016])

Test Name	Cracking Type	Cracking Parameter	Correlation to Field Performance
SCB-LTRC	Bottom-up fatigue; Top-down	Energy Release Rate	Fair correlation from Louisiana pavement management system
SCB-IL	Bottom-up fatigue; Top-down	Flexibility Index (FI) *(related to fracture energy)*	Ongoing validation work in Illinois
OT	Reflection; Bottom-up fatigue	No. of cycles, or fracture parameters A, n	Good correlation with reflection cracking validated in TX, CA, NJ; Promising correlation with fatigue validated with FHWA-ALF and NCAT test track
BBF	Bottom-up fatigue	No. of cycles, or fatigue equation	Correlation with bottom-up fatigue historically validated
IDT-Florida	Top-down	Energy Ratio	Validated with field cores in Florida; confirmed at NCAT test track.

OT: Texas Overlay Test
SCB: Semi-Circular Bend Test
IDT: Indirect Tension Test
FHWA: Federal Highway Administration
ALF: Accelerated Load Facility
BBF: Bending Beam Fatigue Test
LTRC: Louisiana Transportation Research Center
IL: Illinois
NCAT: National Center for Asphalt Technology



(a) (b)

Figure 4. Example of Semi-Circular Bend (SCB) Test Using Illinois Flexibility Index (iFIT) Protocols – for (a) Ductile and (b) Brittle Mixture Behavior

2.3 Resistance to Moisture Damage

As discussed previously, one finding from TRC1404 relates to the presence of moisture damage in poorer-performing pavements. Moisture damage can be defined as the loss of stiffness or strength of asphalt concrete due to the intrusion of water into the binder-aggregate structure. Little, et. al. (2003) identified five mechanisms that contribute to cohesive and/or adhesive failure in asphalt mixtures:

- Detachment: a thin film of water separates aggregate from binder.
- Displacement: a combination of detachment and a break in the binder film.
- Spontaneous Emulsification: formation of water droplets within the binder.
- Pore Pressure: trapped water develops pore pressure due to loads.
- Hydraulic Scour: tire loads push water into the pavement.

Hand (2012) reported that the moisture sensitivity of a given asphalt mixture is mainly determined by the unique combination of aggregate, binder and additives used to produce it, and their individual properties. Examples of these properties are aggregate chemistry, aggregate storage conditions, presence of clay fines, aggregate gradation and angularity, aggregate surface texture, binder and additive chemistry, binder stiffness, and whether or not the mixture contains some antistrip agent.

States have long recognized the importance of resistance to moisture damage, and the effect of moisture-sensitive mixtures on the integrity and performance of their roads. Tests for characterizing the moisture sensitivity of mixtures may be divided by those which are performed on loose mixes to determine coating resistance, and those performed on compacted mixes to evaluate retained strength or stiffness (Solaimanian et al, 2003; Solaimanian et al, 2007). According to Solaimanian et al (2003), tests performed on loose mixes are relatively inexpensive and generally simpler to perform than test on compacted mixtures. However, due to the nature of the test configuration, loose specimens are incapable to develop pore pressure on trapped water (if any). In tests conducted on compacted specimens, on the other hand, phenomena such as pore pressure effects, water and traffic (loads) interaction, among others, can be considered (Solaimanian et al, 2003). This is the reason why laboratory tests conducted on compacted mixes are believed to produce more reliable results. (Solaimanian et al, 2007)

Solaimanian et. al. (2003) presents a list of some well-known tests performed on compacted mixtures. A listing of these tests follows.

- Immersion-compression (ASTM D1075, AASHTO T 165)
- Marshal Immersion (e.g. ARDOT 455A)
- Modified Lottman indirect tension (AASHTO T 283)
- Hamburg Wheel Tracking (AASHTO T 324)
- Asphalt Pavement Analyzer (APA)
- Environmental Conditioning System (ECS) and Asphalt Materials Performance Test Procedure (AMPT)

Tests such as the Immersion-compression, Marshal Immersion, Modified Lottman, and the AMPT use a ratio of mixture 'strength' – of samples subjected to moisture 'conditioning' versus mixtures no so subjected – as the measure of resistance to moisture damage. The tests may feature freeze-thaw cycles in the conditioning process. The modified Lottman test, specified in AASHTO T 283, is the most common test to measure moisture sensitivity in the U.S. (Solaimanian et al, 2007). The Hamburg Wheel tracking (HWT) test and Asphalt Pavement Analyzer (APA) feature repeated applications of a loaded wheel on asphalt specimens submerged in heated water. The HWT, originally developed in the 1970s, has gained increased popularity recently (Varveri et al,2015).

Moisture damage (stripping) is determined by examination of rut depth versus number of passes of the wheel.

The conditioning protocols for moisture damage testing greatly affect the results of a moisture sensitivity analysis. Time, temperature, and pressure are key factors. Unfortunately, similar conditioning is not always possible using the same procedure for all mixtures. For example, Varveri et al (2015) report that a different number of vacuum cycles are needed to obtain the same level of saturation depending on whether the mixture has fine or coarse aggregates. In addition, binder stiffness affects conditioning protocol and time and to produce similar effects on mixtures. A challenge in TRC-1802 relates to adopting/refining a conditioning protocol that accurately represents the field conditions to which a pavement structure is subjected.

CHAPTER THREE
SITE SELECTION

Ideally, the selection of a cracking test and the establishment of associated acceptance criteria would be based on observed performance of asphalt mixtures in the field. This project sought to identify field sites that collectively exhibited a range of cracking-related performance. The asphalt surface mixtures used on those jobs would then be re-created in the laboratory to investigate the efficacy and applicability of cracking tests. This chapter summarizes the site selection procedure used to identify mixtures for subsequent laboratory studies. It is noted that the site selection procedure developed in this project can serve as a template for future studies which propose to use field performance data to identify candidate materials/mixtures for additional study.

3.1 Initial Site Selection

Data Compilation and Organization

The site selection piece of TRC 1802 is the first step in correlating volumetric and performance-based metrics. The Arkansas Department of Transportation (ARDOT) provided information on a subset of thirty projects composed of hot-mix asphalt pavements, in the form of a small database that was initially populated with identifying information useful to the state. The database was also accompanied by hard-copy scans of field data surveys taken at various points in time over the pavement's life. The thirty projects were all relatively new projects, the oldest being constructed no more than 15 years before this study. The pavements, per the field distress surveys, were usually evaluated for distresses between the months of December to March; the uniformity of this repetition varied from district to district within the state. ARDOT also provided construction plan sets for each job; these would be used later to help correlate distresses with the pavement structure for each site.

Field Distress Surveys

The most pertinent piece of data for the candidate projects was the field data surveys. The field data survey was a series of ten grids, each grid containing a total area of 850 square feet. The total area covered by a survey was 8,500 square feet. This same pavement section was re-evaluated each time the pavement was surveyed. Figure 5 is an example of a typical field data survey received for this project.

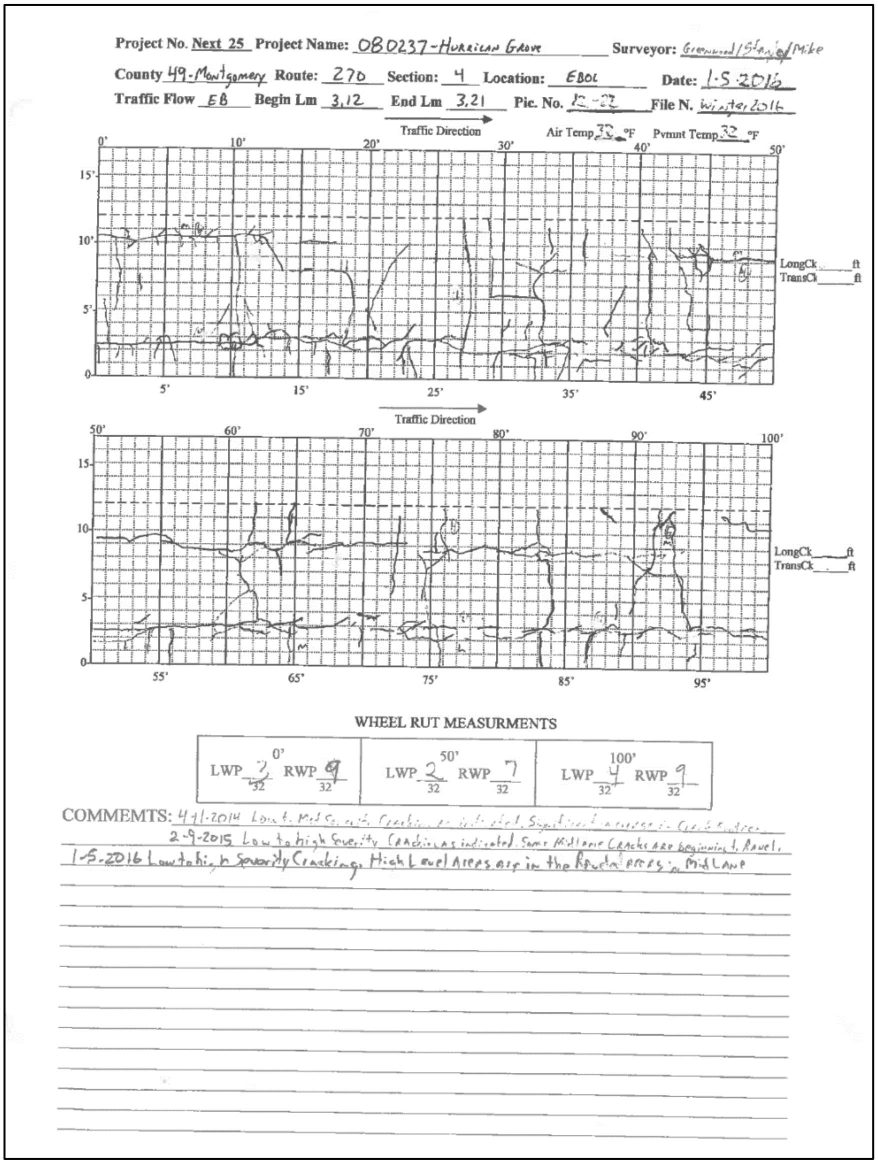


Figure 5. Example Field Distress Survey

Often the field data surveys were filled out by various surveyors from year to year. This introduces an inconsistency in which the distresses were recorded—given the same pavement area and distresses, no two surveyors will record the exact same quantity and severity of pavement distresses. In addition to the visual representation of the cracking as well as the quantity, comments were often left to summarize the state of the pavement. However, these comments were often lost from year to year so the physical representation of the cracking on the grids was almost exclusively considered. If the quantity was not recorded by the surveyor, the quantity recorded was based on counting (to scale) the cracks on the grid itself. A database was developed to record information from the field surveys. Key elements of the database included:

- Site Location
- Cracking
    - Time Stamp – to allow an assessment of cracking propagation over time
    - Type (longitudinal, transverse, fatigue/alligator)
    - Extent (typically linear feet for longitudinal and transverse cracks, and area for fatigue/alligator cracking)
    - Severity

It was noted that the quality of the survey was in large part a function of the surveyor and their attention to detail and continuity. In total, approximately 1,210 fifty-feet-by-seventeen-feet grids were evaluated for the site selection, or 121 field data surveys. Often the surveys were clean enough to be quantified with minimal effort, or there were no recorded distresses at all. In these cases, to populate amount of cracking per job was quick and was as simple as counting boxes with the mouse cursor. However, this was not true for all surveys. For some sites, there were simply too many cracks to track accurately. For others, the presence of alligator cracking muddled the ability to quantify transverse and longitudinal cracking, necessitating another approach.

For sites with large amounts of cracking, the surveys were annotated, first to categorize the cracks, and later counted to quantify them. Figure 6 demonstrates this process. The magenta lines on the Figure represent areas of alligator cracking, the yellow lines represent longitudinal cracking, and the blue lines represent transverse cracking. A very small portion of the field data surveys (approximately 12 percent) required this level of detail for their analysis.

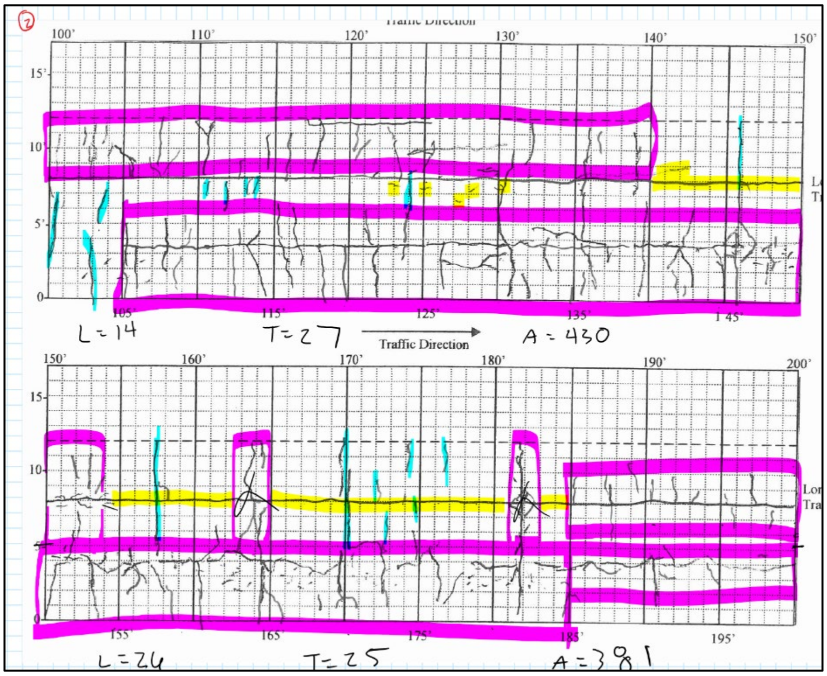


Figure 6. Example Data Reduction on a Field Distress Survey

Structural Cross-Sections

In addition to field distress (cracking) data, elements of the pavement structural cross-section were included in the project database. The construction plan sets that accompanied the original dataset from the state served to detail the pavement structure; no “as built” data was available to the project. Figure 7 provides an example of a pavement cross section and the data extracted for use in the study.

Database Creation

Distress (cracking) data and structural cross-section information were the critical components of the first database iteration. Additional parameters would eventually be added as the project evolved and the sites were evaluated, however this was all the information collected for all sites. Table 6 illustrates this initial database iteration and represents data that was collected for all thirty sites.

Pavement Structure	4" ACHM Surface Course
	3 " ACHM Binder Course
	4.5" ACHM Base Course
	6" Aggregate Base Course
	6" Lime Treated Subgrade
Total:	**23.5 Inches**

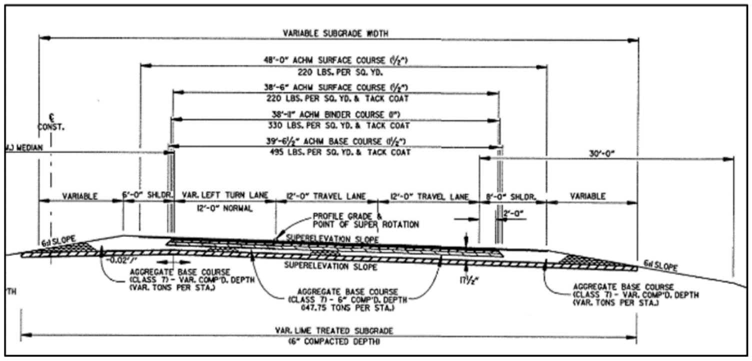


Figure 7. Example Pavement Structural Cross-Section Data

Table 6. Example Initial Database Record for One Field Site

ArDOT District	Date Constructed (year)	Full Depth Thickness (inches)	Total Pvmt Length (feet)	Pvmt Sample Length (feet)	Pvmt Sample Width (feet)	Pvmt Sample Area (sq. feet)	Date of Survey(s)
2	2004	26	12457	500	17	8500	12/7/2012
2	2004	26	12457	500	17	8500	4/7/2014
2	2004	26	12457	500	17	8500	1/12/2015
2	2004	26	12457	500	17	8500	3/2/2017
		Total Longitudinal Crack Length (feet)	**Total Tranverse Crack Length (feet)**	**Fatigue/Alligator Cracking Area (sq. feet)**	**Percent Longitudinal Cracked**	**Percent Transverse Cracked**	**Percent Fatigue/Alligator Cracked**
		32	0	0	6.40%	0.00%	0.00%
		747	5	0	149.40%	1.00%	0.00%
		297	11	334	59.40%	2.20%	3.93%
		308	13	601	61.60%	2.60%	7.07%
	Pavement Structure	2" ACHM Surface Course					
		3 " ACHM Binder Course					
		5" ACHM Base Course					
		10" Aggregate Base Course					
		6" Lime Treated Subgrade					
	Total:	**26 Inches**					

Geospatial Representation

After this initial database was created, it became obvious that the data needed to be represented in a different manner. Although the database located each job site by county and log mile, it was hard to assess how the sites were geographically distributed. It was ultimately proposed that these sites be mapped using geospatial software readily available to most people, Google Earth. Based on the site maps from the construction sets, each site was approximately mapped to its real-world location. The word approximately is used as the pavement area is delineated by log mile, and the actual locations of these log miles was not made clear to those not working with each district in the state. So, the site location was approximated by consistently using the beginning of the job site as noted in the site map from the construction plans. Figure 8 shows the resulting geographic distribution of all thirty sites throughout the state.

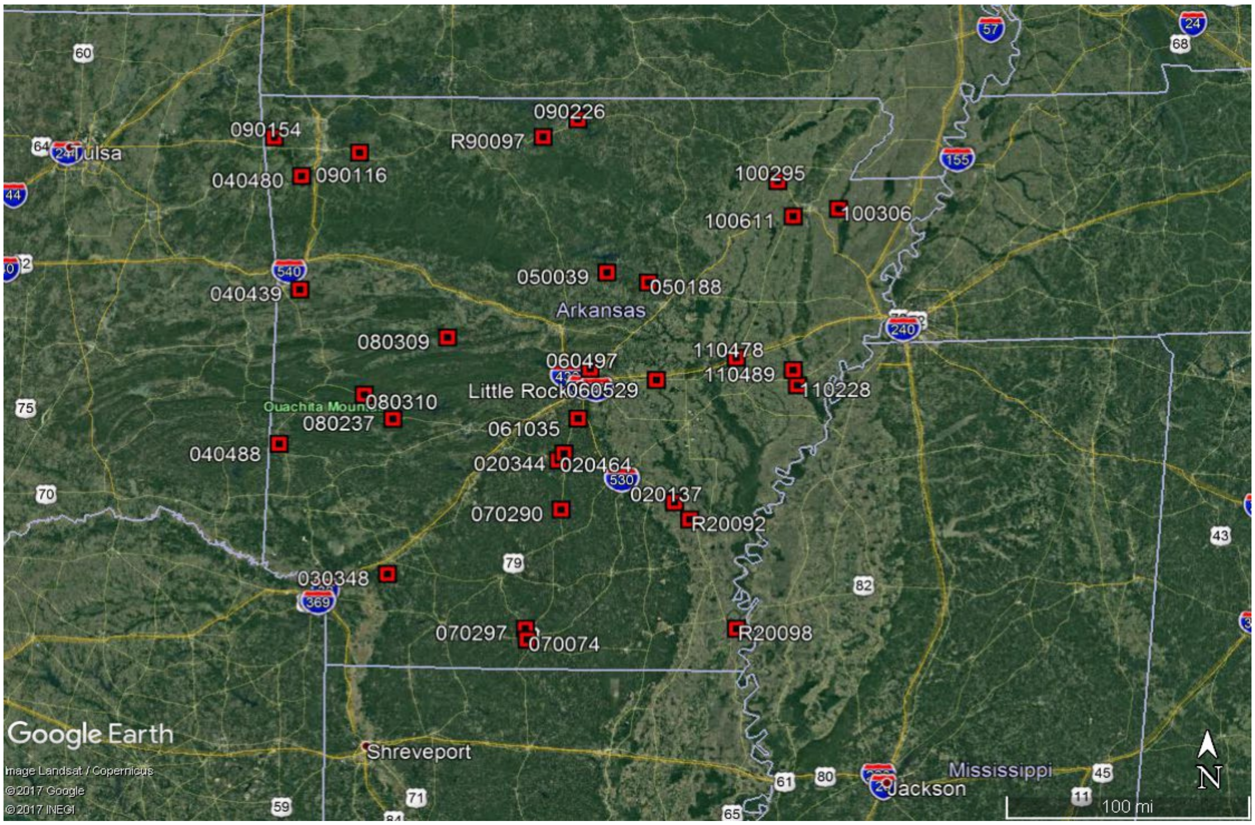


Figure 8. Geographic Distribution of all Thirty Sites

Cracking data quantification, database creation, and the creation of the geographic representation in Google Earth collectively took almost one month. Most of the analysis up to this point had not necessarily been on any one site, but rather generally in terms of the data set as a whole. Once this initial data collection and organization was complete, the database could be refined, shortlists of candidate sites could be made, and additional parameters could be discussed and considered.

3.2 Data Collection and Refinement

Quantification and Classification of Cracking Data

The first level of refinement that was made was to categorize the amount of cracking by type, to enable the categorization of the sites from best to worst, as well as which general time period the cracking data was to represent.

The quantity of cracking greatly varied by the cracking type being considered. This somewhat goes back to the mechanisms that caused them, but ultimately became relevant due to the wide range in which these values existed. It did not appear reasonable to apply the same thresholds to transverse cracking as those in longitudinal cracking if the worst case in longitudinal cracking was 900 feet while the worst case in transverse cracking was 40 feet. The preliminary categories for classifying was a color gradient: green meant little to no cracking, yellow more so than green, orange more so than yellow, and red more so than orange. Different thresholds were applied to each type of cracking based on the natural breaks in the data. The natural breaks discussed simply refers to noting, by observation, bounds in which crack severity increased in a significant way such that ranges could be identified. Figure 9 shows how these specific thresholds were applied to longitudinal cracking.

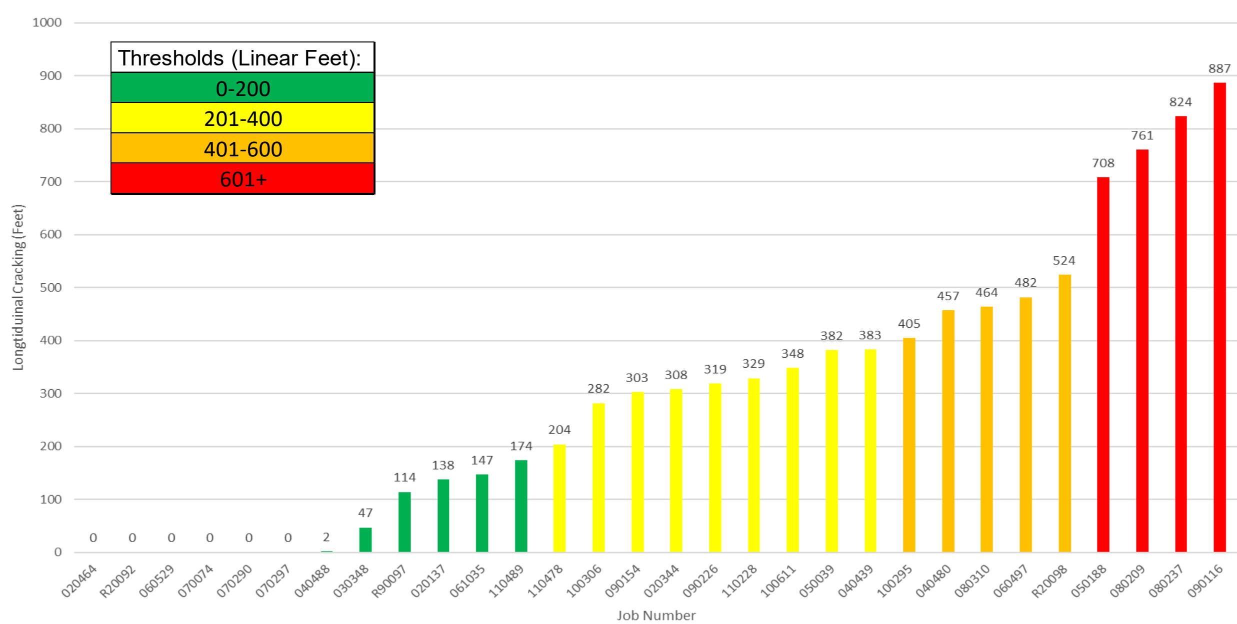


Figure 9. Graphical Representation of Categorical Thresholds by Longitudinal Cracking

While categorizing each type of cracking into these four subsets, it became apparent that the time period over which the quantity was being evaluated could greatly affect which sites were considered good and poor. Two approaches were identified: (1) an average could be taken across all existing surveys and aggregated into a single number; or (2) only the most recent survey (which was 2017 for all sites) would be considered. Each approach had advantages and disadvantages.

If an average was taken, it could eliminate surveyor bias by putting each individual field data survey on an 'even playing field' as no singular survey would be more prevalent then another. However, by taking an average the progression of crack propagation would be lost. In addition, a crack that was initially considered longitudinal could propagate and connect with surrounding cracks. By doing so, in the next field data survey, that same crack that was longitudinal is now considered alligator cracking. An average would not consider this natural progression of crack formation because it would essentially be counting that same count twice in two different categories of cracking. Additionally, taking the average greatly skewed values. If a singular pavement showed little to no distresses and in the most recent survey was recorded as greatly distressed with severe cracking, the magnitude of how severe the cracking developed was lost as it became a lesser value in the average.

If the latest survey was used to represent the data, crack progression would in turn be considered as a singular crack formation and would only be considered in one category. However, this methodology ignores prior field data surveys taken as they have no prevalence as only the latest is relevant. Ultimately, the integrity of crack formation and the underlying explanation that pointed to the mechanisms that caused them was too important to lose within an average. In addition to crack integrity, the most recent survey method was picked over the average method as the amount of field data surveys varied from project to project and there was not a clear way to fairly average a site with five surveys versus a site with two surveys. For this reason, the sites were categorized based on the most recent field data survey. A tacit assumption is included here: field surveys are designed to be 'cumulative' – that is, each subsequent survey should verify the results of the previous survey, and add to the data as new cracks (or expanding cracks) are identified. While this is generally the case in the field surveys supplied for this project, it could not be independently verified.

Identification of Cracking Mechanisms

Now that the time-related methodology was selected, the type of cracking to be considered needed to be chosen. Three types of cracking were quantified for every field data survey: longitudinal cracking, transverse cracking, and alligator (fatigue) cracking. These three cracking distresses each have different mechanisms that cause them, so this distinction was critical to make. The overall objective of this study focuses on material-based characterization related to the susceptibility to pavement distresses and how to abate issues related to such. This clearly differentiates from structurally-based pavement distress mechanisms, as the issue is not with the pavement itself but rather with the underlying base and subgrade. For example, a pavement rehabilitation or repair designed to solve material-based pavement distresses like raveling would do no good against a structural-based deficiency such as a pothole. So, the mechanisms that cause these three types of cracks were critical to understand.

Longitudinal cracking is often first found in the wheel path of a pavement as a symptom of fatigue cracking, or in seams in the pavement along joints. Longitudinal cracks always run parallel to the laydown direction or centerline of a pavement. Typically, they are the first types of distresses to appear in accordance to fatigue cracking (although they are not always indicative of structural issues). Longitudinal cracking is typically referred to as a top-down crack, meaning the crack begins at the surface of the pavement and propagates downward. Traditionally, fatigue cracking is typically considered to initiate at the bottom of the asphalt layer, as the bottom of the asphalt layer experiences the most tensile strain. Therefore, traffic (dynamic loading) would cause this bottom layer to shear and crack starting at the bottom and moving upward. Figure 10 illustrates how these forces work in tandem to create this type of cracking. However, top-down cracking would be the opposite of this; instead of the crack propagating from the bottom of the layer, it begins at the top and works downward through the layer. This is due to the materials in the surface interacting with both shear and tensile forces caused by tires. The reaction of the pavement surface is worsened when binder becomes brittle during the aging process, or by low stiffness in the upper layer of the pavement.

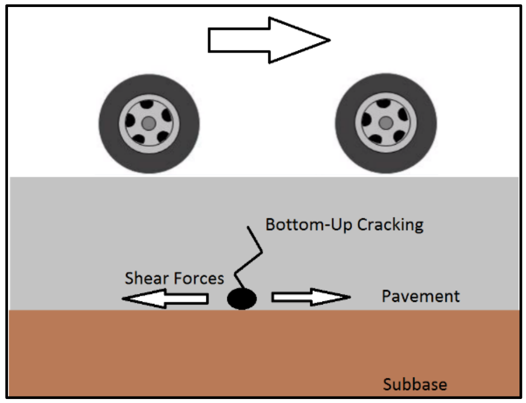


Figure 10. Bottom-Up Cracking

Transverse cracking, by contrast, it almost always related to material properties of the pavement and not typically considered indicative of any structural deficiency (barring reflection cracking). Most frequently, Transverse cracking is a result of shrinkage in the asphalt and typically runs perpendicular to the centerline. This shrinkage happens mostly within the binder, as stated previously, the binder is recommended almost exclusively based on predicted pavement temperatures and latitude of the project site.

Fatigue cracking, also known as alligator cracking, is when these longitudinal and transverse cracks connect and create a gridded appearance on the pavement surface. This gridded appearance is like the appearance of a back of an alligator, explaining why it is also called alligator cracking. Often this surface will deteriorate below the elevation of the surrounding pavement, doing so until it becomes a pothole. If fatigue cracking appears early in a pavement's life, it's typically indicative of a structural deficiency in the pavement structure. However, often over the course of a pavement's life wear-and-tear will cause fatigue cracking that eventually can turn into alligator cracking.

Understanding these mechanisms in light of the goals of TRC 1802 led to the decision to focus on choosing sites based on longitudinal cracking. Although the other types of cracking are not ignored, the focus was on longitudinal cracking. The intent of subsequent laboratory testing is to correlate sites with early-age fatigue cracking as related to pavement material deficiencies, not structural deficiencies. Sites with high alligator cracking could be more indicative of structural issues, which was not in the scope of the project to be considered. Transverse cracking quantities generally trended with the amount of longitudinal cracking, so recommendations between the two usually coincided. Therefore, site categorization and selection was based on longitudinal cracking the most recently available field data survey.

Additional Data

The research team sought additional data to create a more robust profile of each job site. A few important parameters typically considered in roadway and traffic engineering had not yet been considered; namely, what type of construction were these job sites? How similar was the pavement structure of these roadways? What was the traffic on these roadways? Some of this data was relatively easy to obtain, while others were not as readily available. The type of construction was available in the construction plans previously referenced to find the pavement structure. A large majority of the original thirty projects were either applying overlays, adding a travel lane, or both. Due to the difficulty of being able to specifically identify the sample area of the field data surveys, this factor was noted and projects of this type (either overlay or overlay and lane addition) were almost exclusively selected in order to keep the comparisons as similar as possible.

A very common and convenient method for comparing pavement structure is the Structural Number (SN). Structural Number is calculated in accordance with the 1993 AASHTO Pavement Design Guide *(AASHTO 1993)*, using the following equation:

$$SN = a_1 D_1 + a_2 D_2 M_2 + a_3 D_3 M_3$$

Where a_n is the structural layer coefficient, D_n is the layer thickness, and M_n is the drainage coefficient. Structural layer coefficients (a_n) were taken from the ARDOT Roadway Plan Development Guide *(ARDOT, 2017)*. In Arkansas, a value of 1.0 is assumed for the drainage coefficient (M_n).

The last parameter added was to consider the amount of traffic on these roadways. The most readily available source of such information was through traffic maps providing AADT, or average annual daily traffic. Not only was AADT collected, but percent truck traffic was also collected. This was a crucial factor as roadway degradation is expedited in the presence of increased truck traffic.

Traffic loading done by various types of vehicles is compared by ESALs, or Equivalent Single Axle Loads. One ESAL is equivalent to 18,000 pounds on a single axle while a standard passenger car is about 2,000 pounds *(TxDOT, 2005)*. To put the damage to pavement caused by a semi-truck in perspective, a typical ESAL value for a semi-truck is 2.5; for a passenger car, the ESAL value is generally accepted as 0.0004. Thus, compared to passenger cars, a semi-truck can cause up to 10,000 times more damage to a pavement. So, a pavement with increased cracking could actually be performing better than a pavement with less cracking if that pavement has a higher percentage of truck traffic than the other pavement does. Having AADT and percent truck allows another perspective in the data analysis by being able to further explain and justify the type and quantity of pavement distress when combined with the field data survey information. Figure 11 is an example of an AADT map from ARDOT from which the traffic data was sourced. In some cases the exact pavement length did not explicitly have an available truck traffic percentage associated with it. In this case, the road closest in proximity was identified and that truck traffic percentage was extrapolated to the pavement section of interest.

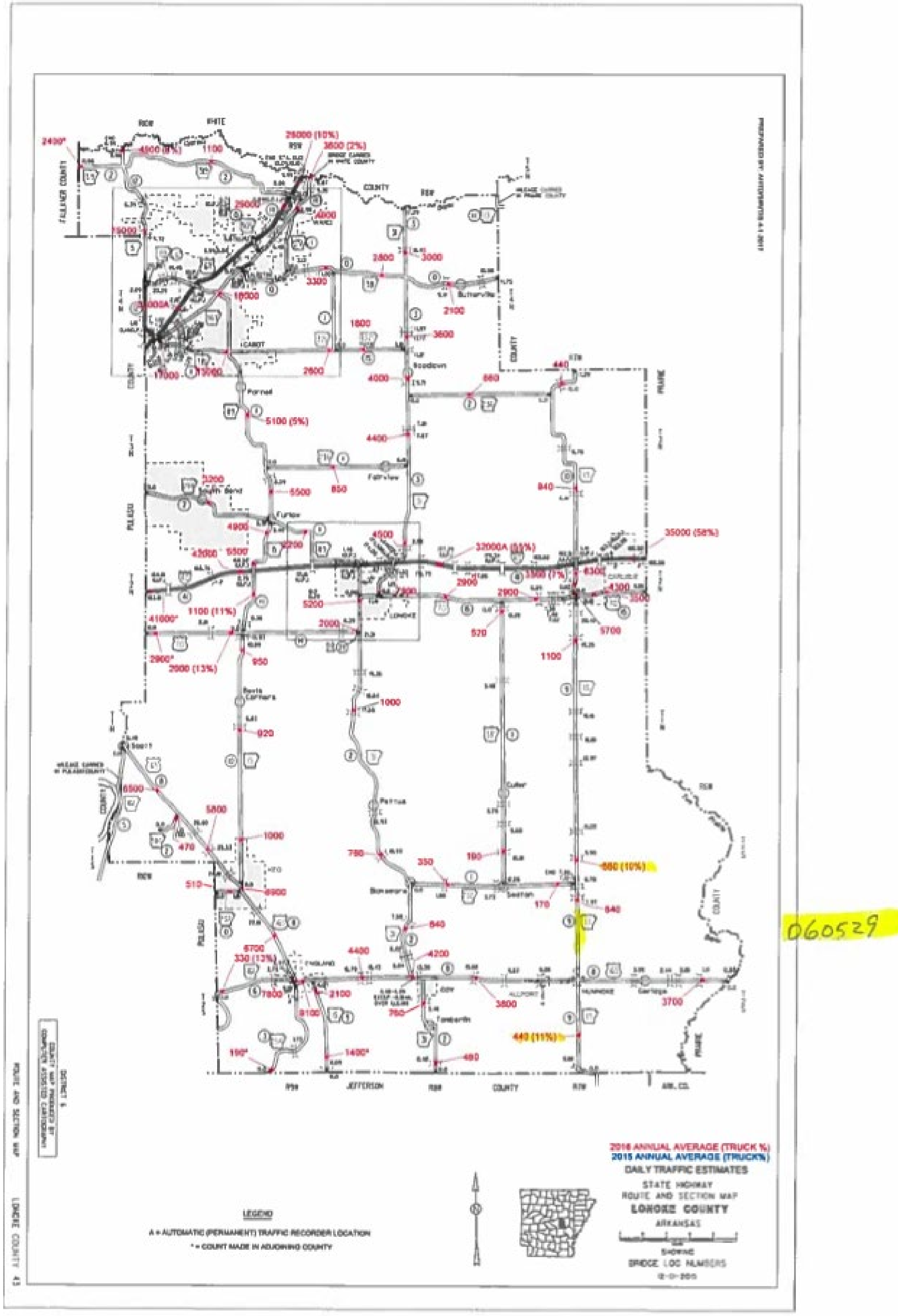


Figure 11. Annual Average Daily Traffic Map

3.3 Data Analysis

After the data collection, refinement, and analyses, it was obvious the first step to site selection would be creating a "shortlist" of the sites, based solely on amount of longitudinal cracking in the latest field data survey using the graph in Figure 9. The initial shortlist was then further refined with the consideration of pavement structure and traffic data, and mapped. Figure 12 is the geographic distribution of this shortlist, color coded by categorized longitudinal cracking.

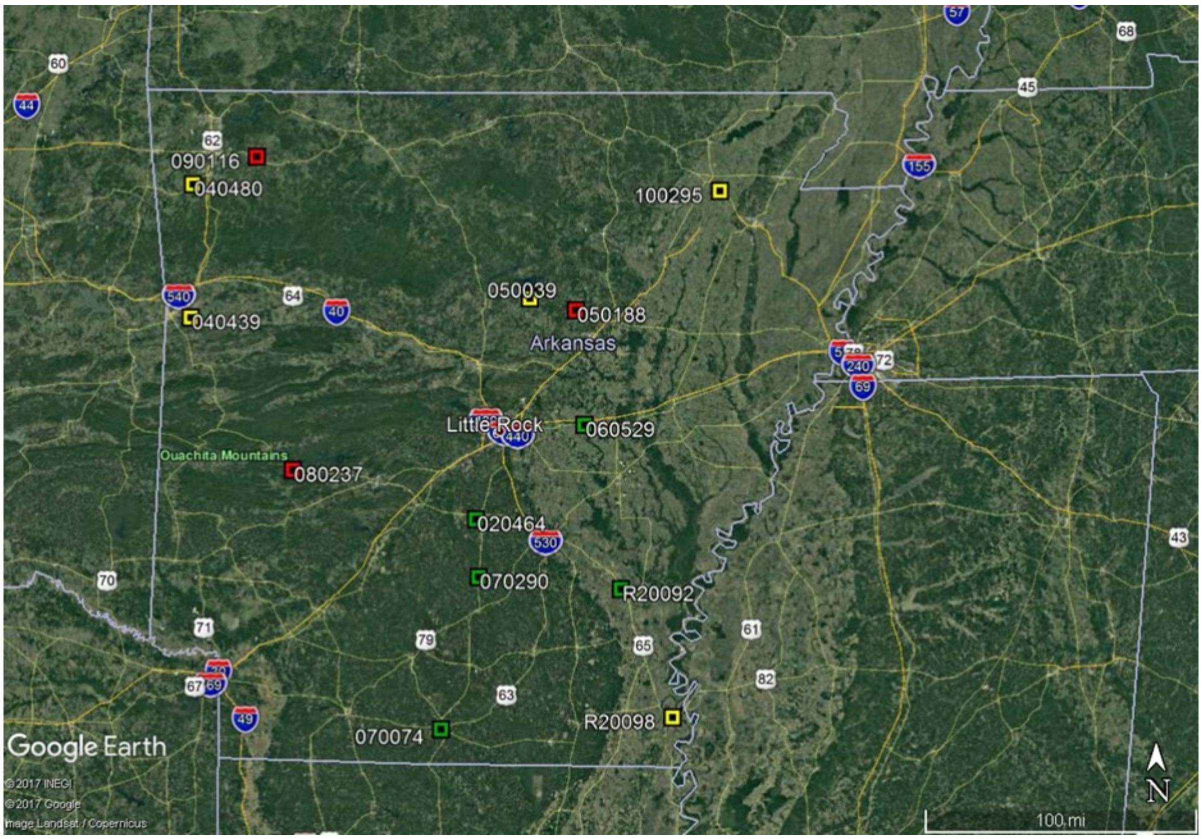


Figure 12. First Shortlist Job Site Geographic Distribution

Observing the geographic distribution of these sites, although they were dispersed relatively throughout the state, many of the sites were close. Sites close to each other were not advisable due to the desire to isolate material behavior as it relates to pavement performance. Sites in geographic proximity to each other could feature very similar materials, and therefore essentially be the same mix with very similar material properties. Therefore, differences in early-age cracking behavior would not reflect the effect of materials, but rather a function of the traffic, the quality of the construction, and the pavement structure. So, these sites were evaluated and refined again to ensure an even geographic distribution as well as comparable pavement structure and traffic. Figure 13 shows the final shortlist, bringing the proposed thirty sites down to six for testing and analysis.

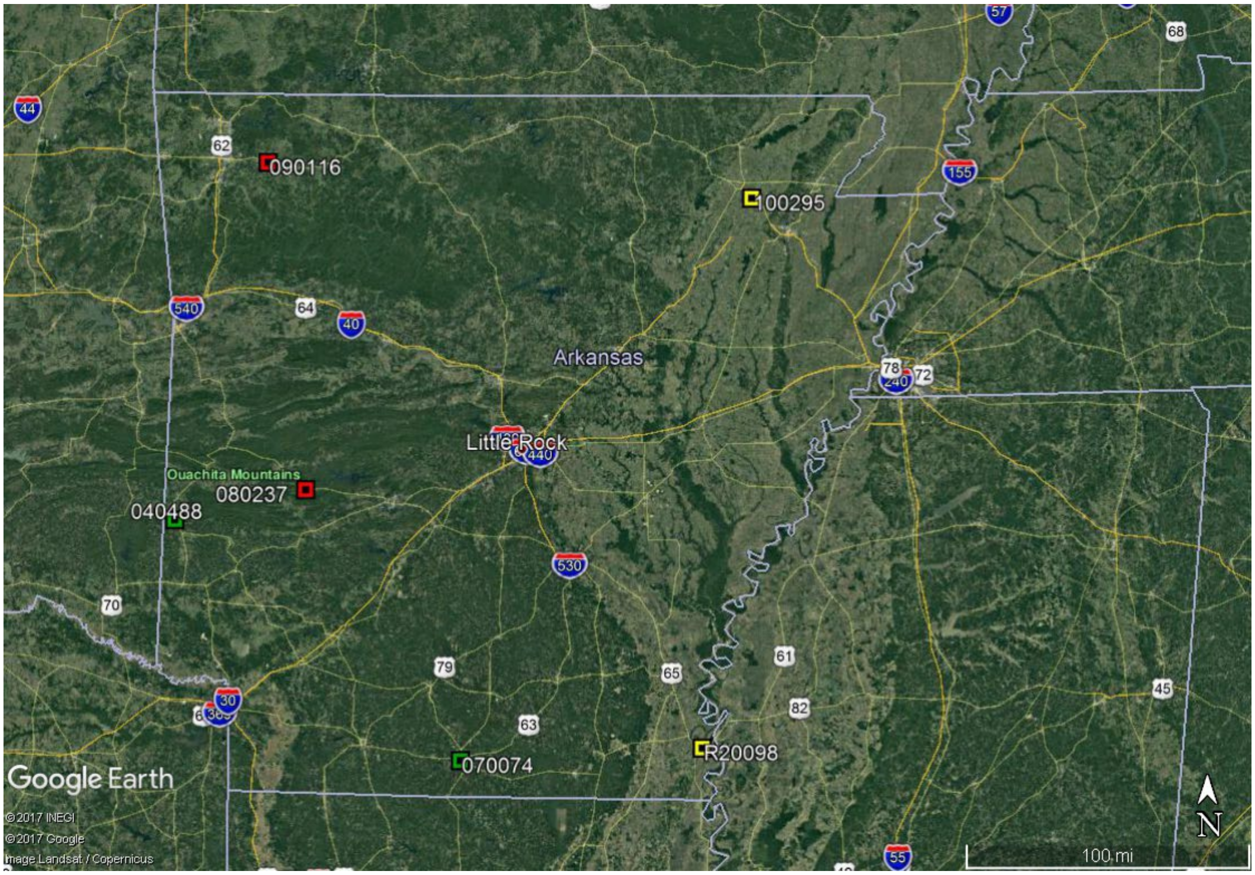


Figure 13. Geographic Distribution for Final Six Sites Selected for Analysis

Six sites were ultimately selected in part due to the ability for the lab team to recreate and test these mixes in a timely manner. The final six sites were distributed relatively evenly throughout the state, ensuring a diversity in quarried materials, mixes, and binders. The colored categories were simplified into Good, Fair, and Poor, with two of each classification being available in this subset. Good was defined as less than 5% cracking, fair defined as 5-10% cracking, and poor defined as more than 10% cracking, per ARDOT's preventive maintenance plan (ARDOT, 2018). These six sites will have their mixes recreated and tested in terms of cracking to try to predict early-age cracking; this comparison will be done by comparing the lab data to observed field performance, as recorded on the field data surveys.

The site selection procedure outlined in this text is not exclusive to TRC 1802. Rather, the procedure will be a reference and a guideline in future projects in which situational site selection is called for. The procedure establishes a base line by which general metrics and parameters can be addressed and considered, while still offering flexibility to be modified to suit the needs of the project. The site selection to get to these six sites was iterative, reaching past the original dataset by consulting outside sources to get the most appropriate and robust dataset to make an informed decision. By doing so, site selection can become a more concrete and qualitative approach methodology based on the data itself, rather than personal interpretation.

CHAPTER FOUR
DETERMINATION OF LABORATORY AGING PROTOCOLS

One salient question related to implementation of mix-design performance testing is the amount and protocol for "aging" the mixture in the laboratory, ostensibly to 'mimic' aging which occurs on in-service pavements. Many laboratory aging protocols have been proposed; some protocols are not practical for mix design (and potentially for QC/QA) due to, among other considerations, the excessive time spent in the oven. For many laboratory 'performance' tests, prediction of actual field performance is less desirable than the ability to (relatively) rank or evaluate the cracking resistance of various laboratory mixes. This paper describes the effects of different aging protocols on the discrimination potential of the Illinois Semi-circular Bending Test (SCB-IL/I-FIT).

4.1 Determination of Aging Protocols for Performance Testing

Introduction

Although the Superpave mixture design system does not include any 'performance' related tests of asphalt mixtures (it is a volumetric design process), many agencies seek to minimize the rutting potential of mixtures by incorporating a rutting-related test which provides an indication of resistance to shear and/or permanent deformation. However, comparatively few agencies have implemented such a test to assess the cracking resistance of mixtures. Interestingly, the Superpave volumetric design system is known to produce mixes that may be too lean, which in turn causes cracking problems (Bonaquist, 2014). In recent years there have been several laboratory cracking tests developed to overcome this difficulty, in order to "balance" the mixture design.

Zhou et al. (2016) identified a series of desirable characteristics for laboratory cracking tests which includes, among other features, good correlation with field cracking performance. To maximize this feature, significant effort is made to test laboratory samples under conditions that are as close as possible to the ones found in the field; this has given rise to the development of laboratory aging protocols designed to mimic mixture changes caused by field climate factors. Currently, the aging protocol that is used by most state agencies is contained in AASHTO R30. It is generally believed that the short-term oven aging method per this standard accurately represents construction conditions; nevertheless, it is also generally agreed that the standard is not sufficiently severe when mimicking long-term effects (Braham, Buttlar, Clyne, Marasteanu, & Turos, 2009; Kim et al., 2018; Newcomb et al., 2015). Thus, researchers have proposed alternative "long-term" aging protocols by varying oven temperature, oven time, and mixture condition (loose versus compacted) (Chen, Yin, Turner, West, & Tran, 2018; Kim et al., 2018).

It is vital to identify the overall objective of any laboratory testing program used to consider "performance" during the mixture design process. For an objective related to *predicting* performance (e.g. cracking) over time, it is appropriate to test specimens which have been laboratory-aged to a point which approximates mixtures in the field for a given time period. However, for an objective related to *screening* mixtures in the context of maximizing cracking resistance – in other words, applying an "index" type of testing/evaluation – the necessity of exactly mimicking effects of field aging in the laboratory aging protocol is not as vital. In such cases, a key feature of the aging protocol used its discrimination potential – that is, the ability of the aging/testing system to differentiate 'good' from 'bad' mixes, and levels of performance in between.

For this project, the overall objective for the aging/testing system is related to *screening* mixtures, to evaluate the potential for cracking resistance. Specific objectives related to the information presented here include:

- Evaluate the effects of different aging protocols on the discrimination potential of a relatively simple cracking test: I-FIT.
- Propose an aging protocol for performance testing (cracking, using I-FIT) of asphalt mixtures during the mixture design process in Arkansas.

A total of six road projects in Arkansas are included in the overall study; this chapter reports on the results from three of those projects. The field cracking performance on each project has been monitored since its construction (approximately 10 years on average). The cracking test used is the room-temperature, semi-circular bend test developed by the University of Illinois, commonly referred to as "I-FIT". Four aging protocols are applied: NCHRP Report 871 (5 days at 95C), NCAT top-down cracking program (8 hours at 135C), and AASHTO R 30 short-term for volumetric analysis (2 hours at compaction temperature) and long-term for mechanical testing (preconditioning loose mix 4 hours at 135C plus 5 days at 85C). The discrimination potential resulting from each aging protocol was evaluated by comparing field and laboratory performance using engineering judgment, and statistical tools such as ANOVA and Tukey's tests.

Background of Aging

There are two main causes of aging in asphalt mixtures. During construction, loss of volatile components and oxidation are predominant, and progressive oxidation occurs during the working life of the placed pavement (Bell, AbWahab, Cristi, & Sosnovske, 1994; Newcomb et al., 2015). Progressive oxidation rates depend, among others, on the climate of the region; hotter climates affecting pavement structures the most (Bell et al., 1994). Volumetric factors such as total air voids and interconnectivity of such voids also play a role on aging rates; and while binder content in general does affect aging as well (Newcomb et al., 2015), aging rates have been found to be independent of the type and source of the binder (Kim et al., 2018). Aggregate characteristics can also affect aging rates due to its influence in the binder film thickness during mix design (Newcomb et al., 2015).

The primary effect of aging on mixtures is hardening or stiffening. Even though stiff mixes produced from prolonged aging perform well for rutting, the same mixes are particularly susceptible to cracking and moisture damage (Bell et al., 1994). In the laboratory, the construction stage in the field has been historically simulated using the short-term oven aging (Bell et al., 1994; Newcomb et al., 2015). Laboratory long-term oven aging, on the other hand, attempts to simulate effects of climate during the working life of the placed pavement. The primary variables used to control laboratory aging are oven temperature, time spent in the oven, and the condition of the mix while aged (loose versus compacted). For instance, extended mixing, or aging of loose mixes in general, has been proven to reach the highest levels of aging than conditioning of compacted mixes (Bell et al., 1994).

Currently, AASHTO R30 is the most utilized aging protocol for mix design and performance testing, and while it is believed that the current short-term oven aging method per AASHTO R 30 accurately represents construction conditions, there is general agreement that its long-term oven aging underestimates aging levels reached in the field. It has been found that AASHTO R30 actually simulates about 2 years of field performance rather than 7 to 10 years as the original research by Bell et al. (1994) suggested. Another limitation of AASHTO R 30 is that it prescribes a unique combination of oven temperature and time (Newcomb et al., 2015) regardless of the actual climate of the region, making the standard a poor candidate for performance prediction. Research has found, however, that even though this method is not suitable for field performance prediction purposes, it can be utilized to assess the relative cracking resistance of asphalt mixtures (Braham et al., 2009).

The wide range of limitations of the long-term oven aging per AASHTO R 30 has motivated significant research to seek an alternative aging protocol that more accurately simulates field deterioration due to climate factors. In this paper, two newer long-term aging protocols are studied: NCHRP project 09-54 (Report 871) (Kim et al., 2018), and the protocol proposed by NCAT for their top-down fatigue cracking experiment (Chen et al., 2018). They were chosen due to their strong theoretical foundation and/or their suitability for mix design implementation.

NCHRP Project 09-54 (Report 871)

NCHRP project 09-54 (Report 871) developed an aging model that can accurately predict chemical (carbonyl area) and rheological [shear modulus (G*)] properties of the binder. After comparing different aging protocols, NCHRP project 09-54 proposed using loose mix at 95C with a duration that depends on the climate region, the initial binder rheological properties, and the pavement layer depth. The researchers found that although long-term aging at 135C significantly decreases the aging time, it causes chemical changes in the binder that don't occur in the field. In addition, aging compacted mixes, as required per AASHTO R30, causes specimen distortion as well as aging gradients that makes the characterization of small specimens in the laboratory difficult.

For brevity, details related to the determination of specific aging protocols using the NCHRP 9-54 procedure are not given here. It is noted that one advantage of this system is that the aging time can be tailored for a given expected failure mechanism. For instance, at 6 mm depth, the method can be used to predict surface deterioration such as top-down fatigue cracking, while 50 mm depth can be used for performance prediction of bottom-up cracking. A depth of 20-mm is recommended of viscoelastic characterization. As expected, the closer to the pavement surface, the longer the laboratory aging time. On the other hand, a major complication in using this equation is that the Enhanced Integrated Climatic Model (EICM) is required to estimate the temperature of the pavement hour by hour during the desired 'design' life of the pavement. Kim et al. (2018) provide laboratory aging duration maps, which have been developed to overcome this cumbersome process. In Arkansas, for instance, a time of 5 days at 95C in the oven represents 4 years of field aging 6-mm deep the pavement structure.

NCAT Top-Down Fatigue Cracking Experiment

NCHRP Project 09-52 (Newcomb et al., 2015) defines the cumulative degree-days (CDD) as "the sum of daily high temperature above freezing for all the days being considered from the time of construction to the time of core sampling". Based on findings by Shen, Wu, Zhang, Mohammad, and Muhunthan (2017), Chen et al. (2018) reported that top-down fatigue cracking usually appears at around 70,000 CDD. This amount of CDD (70,000) is reached at different times for different climates. In the case of Arkansas, 70,000 CDD are generally reached in 4.5 years. In addition, it was found that 5-day aging at 95C better simulates a pavement that has been subjected to 70,000 CDD. This was done by comparing the change in the high-performance grade (HPG) of the binder measured in the direct shear rheometer (DSR) in field cores with laboratory aged mixes. Other mix properties compared in the study were G* and phase angle, and carbonyl area. They concluded that 8-hour aging at 135C leads, on average, to the same change in the measured properties as 5-day aging at 95C without significantly affecting the oxidation hardening relationship between both protocols. Thus, 8-hour aging at 135C was proposed for convenience. No modified binder was included in the study.

Illinois Semi-circular Bending Test (SCB-IL/I-FIT)

The I-FIT was developed by the University of Illinois in 2015 (Al-Qadi et al., 2015). The output of the test is the *Flexibility Index* (FI), which is defined as the ratio of fracture energy, or the area underneath the force-displacement curve, to the slope of the inflection point at the back slope of such curve, multiplied by some scalar factor. The fracture energy term is a measure of the overall resistance of the material to cracking, while the slope accounts for the level of brittleness or 'flexibility' of the material. Figure 14 and the equation which follows represent the concept of the flexibility index.

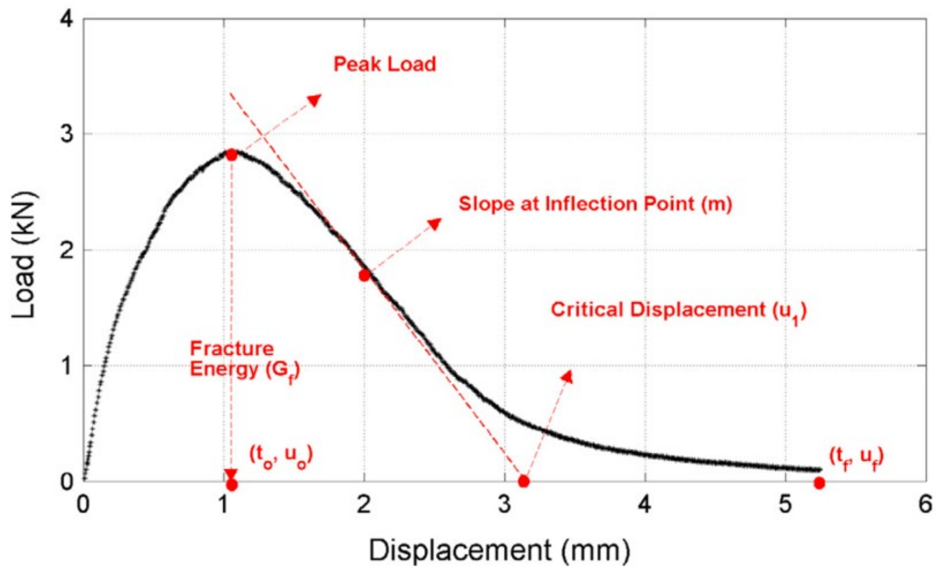


Figure 14. Flexibility Index Diagram Representation
from (Ozer, Al-Qadi, Lambros, et al., 2016)

$$FI = A \cdot \frac{G_f}{|m|}$$

where:
FI = Flexibility Index
G_f = Fracture Energy (Area underneath the displacement-load curve)
m = Slope at the inflection point at the back of the curve
A = Scalar factor (taken as 0.01)

The form of the index was inspired by the equations for the rate of crack growth derived by Zdenek and Pere (1988). Thus, the index has a strong correlation with crack growth velocity. The I-FIT is performed on a semi-circular bending (SCB) specimen. The specimen is cut from a Superpave Gyratory Compactor (SGC) cylinder; it is 50 mm thick and has a notch 15 mm deep. It is run at room temperature (25 °C) and a loading rate of 50 mm/min (AASHTO, 2017). This combination of temperature and loading rate is able to better predict and discriminate cracking performance compared to other lower temperature and/or faster loading rates (Ozer, Al-Qadi, Lambros, et al., 2016). In the original study, no aging protocol was recommended (Al-Qadi et al., 2015), nor does AASHTO TP 124 (2017) specify an aging protocol for the I-FIT test.

Several projects have studied the correlation that exist between the index and field cracking performance. Most of the studies have focused on measuring the discrimination potential of the test; that is, the strength of the index to differentiate cracking resistance among mixes that show close performance levels, as opposed to simply separating poor performing mixes from those that perform well. One study (Ozer, Al-Qadi, Singhvi, et al., 2016) illustrates the differences between fracture energy and flexibility index in characterizing the same mixtures versus their RAP content (assuming that higher RAP content results in lower field performance). Flexibility index proved superior in differentiating the various levels of RAP inclusion.

Materials and Methods

Three in-service pavements are represented in this analysis. The selection process for these projects was based primarily on the extent of cracking observed in periodic visual distress surveys of the pavements over an eight-to-ten year period after construction. Additional selection considerations included similarities in traffic level and structural characteristics of the pavement cross-section. Table 7 provides a summary of asphalt surface mixture type, material sources and characteristics, traffic, and observed cracking. The "rating" of the road is based on the total amount of cracking. All mixes are 12.5 mm surface mixes.

Table 7. Field Project Characteristics

SITE	Asphalt Grade	Traffic AADT (2016)	Traffic %Truck (2016)	Long. Cracking (ft)	Trans. Cracking (ft)	RATING
HEB SPR	PG 64-22	5100	10	382	194	Fair
JBRO	PG 70-22	11000	16	405	148	Fair
HIND	PG 70-22	11000	10	887	217	Poor

Original materials from construction were not available for this study. To recreate the mixes for the study, aggregates were collected from the asphalt plants which produced the original mixes; in all cases, the aggregate sources and types were the same, and measured specific gravities and absorption capacities were similar to those reported in the original mix designs. Matching specific binder sources with the original materials proved unfeasible; therefore, a single source of binder was used for all recreated mixes in this study – but with matching the PG grade of the binder used during construction.

The volumetric properties of the recreated mixes closely approximated those reported in the job-mix formulas for the original mixtures. In some cases, the gradation of the recreated mixes was adjusted to better match volumetric properties, particularly VMA. Table 8 presents material and volumetric properties of the mixes. It is noted that all performance-testing specimens were compacted to 7.0±0.5 percent air voids. Compacted and cut specimens for the I-FIT test were conditioned in a water bath for 24 hours at 25°C prior to testing.

Table 8. Job Mix Formula Properties

	Aggregate Absorption (%)	Binder Content, P_b (%)	Air Voids, P_a (%)	Voids in Mineral Aggregate, VMA (%)	N_{des}
HEB SPR	1.8	5.7	4.5	14.9	75
JBRO	1.2	5.0	4.5	14.4	100
HIND	1.9	5.0	4.5	14.8	100

Four aging protocols for performance testing were included in the study:

1. **AASHTO R 30 short term** for volumetric testing
   - *Justification*: although not intended for performance testing, its convenience makes it a good option for mix design and potential QC/QA.
   - *Description*: aging loose samples for two hours at compaction temperature, with stirring every 60 minutes.
   - *Comments*: Loose mix thickness required: 25 to 50 mm; relatively small pans can be used without occupying significant oven space. Since the sample is aged at compaction temperature, it can be compacted immediately after aging.

2. **AASHTO R 30 long term** for mechanical testing
   - *Justification*: used by most highway agencies for performance testing.
   - *Description*: aging compacted samples 5 days at 85°C, with preconditioning of loose samples for 4 hours at 135°C, stirring every 60 minutes (short term aging for mechanical testing).

- *Comments*: After aging loose samples for 4 hours at 135°C, extra time is required for the samples to reach compaction temperature before compaction. After samples are compacted and cut (if required), they are aged for 5 days at 85°C. This process requires several days to complete, but can be accomplished within 'traditional' workday hours.

3. **NCAT** aging protocol (used for a top-down cracking project)
   - *Justification*: A new aging protocol proposed by the National Center for Asphalt Technology, which appears to be gaining popularity
   - *Description*: aging loose mix for 8 hours at 135°C.
   - *Comments*: Even though the original NCAT study does not mention any preconditioning prior to such 8 hours at 135°C, the study used plant produced mixes. Therefore, this study does a 4-hour-at-135°C preconditioning to mimic mix deterioration due to plant production, leading to a total aging time of 12 hours at 135°C, plus the time required for the mix to reach compaction temperature. This extended continuous time makes it difficult to finish the mixing and compaction process in one 'traditional' work period.

4. **NCHRP Report 871**
   - *Justification*: The aging models that led to the development of this protocol are intended to be incorporated in future versions of Pavement ME.
   - *Description*: aging loose mix for 4 days at 95°C, with preconditioning of loose samples for 4 hours at 135°C, stirring every 60 minutes (short term aging for mechanical testing per AASHTO R 30). This combination of time and temperature was chosen since it mimics a similar field deterioration as the one proposed by NCAT (around 4 years in Arkansas).
   - *Comments*: after the short-term conditioning for mechanical testing, this protocol requires placing the loose mix in larger pans so that the thickness of the loose mix does not goes over the NMAS; thus, using significant oven space over a relatively long period of time (four days).

Aging Results

Table 9 shows the I-FIT results, representing the average I-FIT values, based on four physical I-FIT tests (two 150mm diameter discs were cut from the center of a compacted sample; each disc was subsequently split into two semi-circular test specimens, for a total of four tests). Two primary analyses are presented – the relative effect of long-term aging, and the discrimination potential of aging protocols.

Table 9. I-FIT Test Results

Site	Field Cracking Rating	Average I-FIT Value for Aging Protocol Shown			
		AASHTO R30 ST	NCAT	AASHTO R30 LT	NCHRP 871
HEB SPRG	Fair	10.6	2.5	3.1	0.7
JBRO	Fair	8.5	3.1	3.2	2.4
HIND	Poor	4.0	1.4	1.1	0.9

As expected, the effect of long-term aging on I-FIT results is substantial. Figure 15 and Table 9 demonstrate that I-FIT values for all long-term aging protocols are significantly lower than those representing short-term aging. Perhaps less expected is that there are no significant differences between I-FIT values – *for a given site* – among any of the long-term aging protocols; in other words, **all** long-term aging protocols (for a given site, with the single exception of NCHRP 871 for HEB SPRG, resulted in I-FIT values that are (statistically) the same. This suggests that the selection of a long-term aging protocol rests on its discrimination potential related to field performance and ease-of-use.

Figure 16 and Table 11 examine the discrimination potential of the aging protocols, relative to field performance. The Tukey sub-groupings of I-FIT results for a given aging protocol (reading each column in Table 10 vertically) indicate that the short-term aging protocol offers a high degree of discrimination among the sites tested; however, in the overall context of the project, the HEB SPRG and JBRO sites are considered to exhibit similar field performance. Among the long-term aging protocols, the AASHTO R30 and NCAT methods appear to discriminate between field sites in accordance with the relative judgement of field performance.

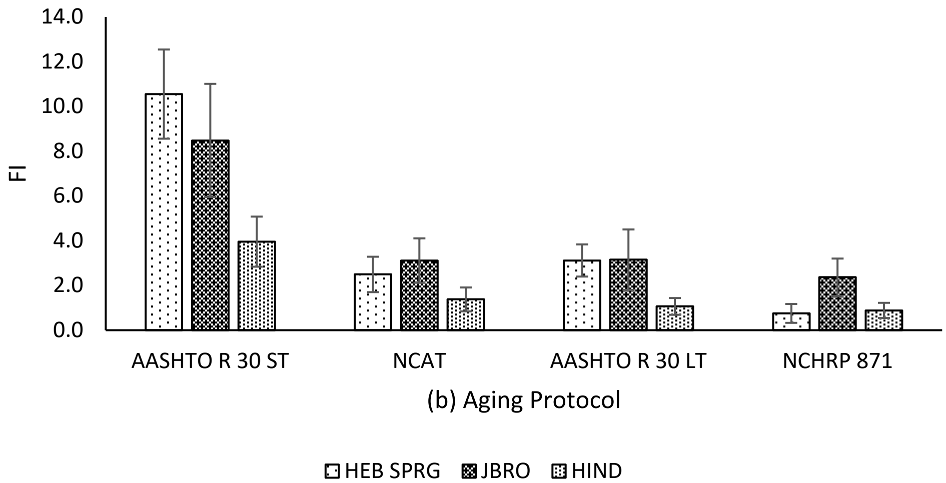


Figure 15. I-FIT Test Results, by Aging Protocol

Table 10. Tukey Sub-Groupings of I-FIT Results – Aging Protocols

Site	Tukey Sub-Grouping for Field Site Shown		
	HEB SPRG	JBRO	HIND
AASHTO R30 ST	A	A	A
AASHTO R30 LT	B	B	B
NCAT	B	B	B
NCHRP 871	C	B	B

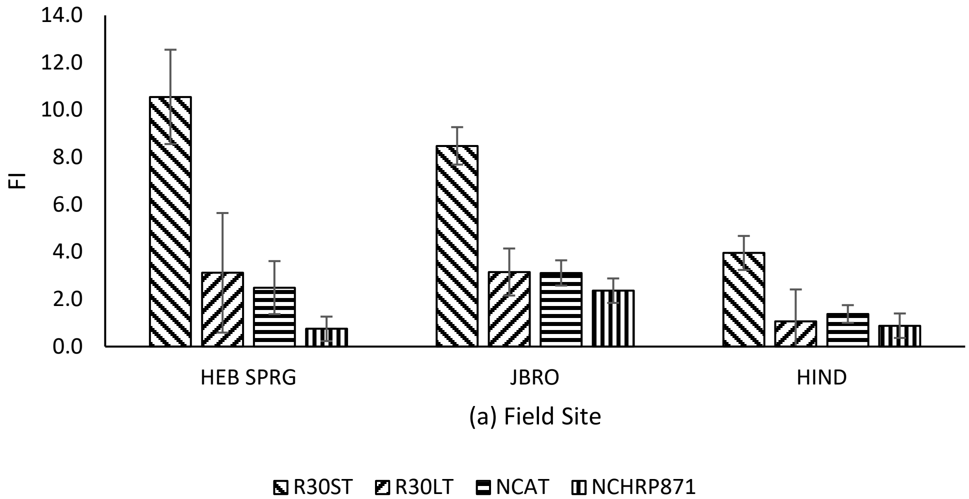


Figure 16. I-FIT Test Results, by Field Site

Table 11. Tukey Sub-Groupings of I-FIT Results – Field Sites

Site	Field Cracking Rating	Tukey Sub-Grouping for Aging Protocol Shown			
		R30 ST	**R30 LT**	**NCAT**	**NCHRP**
HEB SPRG	Fair	A	A	A	B
JBRO	Fair	B	A	A	A
HIND	Poor	C	B	B	B

Aging Conclusions

The two objectives of this study reported here include an evaluation of the discrimination potential of various aging protocols on the I-FIT test, and the selection / recommendation of an aging protocol for implementation of a cracking-related performance test for asphalt mixture design in Arkansas. Based on the somewhat-limited results to date, the following observations and conclusions are offered:

1. For a given mix/project site, all long-term aging protocols exhibited no significant differences among themselves In addition, all long-term-aged I-FIT results were significantly different from short-term-aged results.

2. For the long-term aging protocols used, the AASHTO R30 and NCAT methods appear to discriminate between mixes consistent with relative field performance.
   The limitations of the study, as presented here, are notable.

Arkansas seeks to use cracking-related performance testing in a “go/no-go” decision point process during mix design, rather than for *predicting* the potential extent of cracking on a given project. This approach emphasizes the importance of the discriminatory power in a given test method. The data examined to date suggest that the AASHTO R30 short-term aging protocol provides such discriminatory ability. In addition, two long-term aging protocols – AASHTO R30 and NCAT – also exhibit discriminatory power consistent with field performance. Given multiple protocols which provide the necessary discrimination, selecting a protocol for design may consider ease-of-use. Certainly, the short-term aging protocol offers the greatest ease-of-use; it is not only significantly shorter (time-wise) than long-term aging, it is also used for volumetric analyses of mixtures – this gives the mix designer a single protocol for all specimens in the laboratory. If a long-term aging protocol is desired, the NCAT protocol offers a significantly shorter time period for specimen preparation and provides the user with a more robust selection process for aging parameters, compared to AASHTO R30.

The preliminary recommendation for Arkansas is to continue to examine I-FIT results representing specimens prepared using AASHTO R30 short-term aging, and using the NCAT long-term aging protocol. Final selection will be based on results from field sites featuring multiple levels of cracking performance.

CHAPTER FIVE

SELECTION OF CRACKING TEST ACCEPTANCE CRITERIA

5.1 Semi-Circular Bending Test (SCB) and Illinois Flexibility Index Test (I-FIT)

The Semi-Circular Bending Test (SCB) and Illinois Flexibility Index Test (I-FIT) have been investigated as potential tests to implement in Arkansas to analyze cracking resistance during the mixture design process of asphalt pavements. I-FIT is a product of the Illinois Center for Transportation (ICT) and Illinois Department of Transportation (IDOT). I-FIT was developed as a method and protocol that can rank Asphalt Concrete (AC) mixtures based on their cracking resistance (Al-Qadi et al. 2015). This chapter discusses research conducted to evaluate the potential of the SCB and I-FIT tests to analyze the cracking resistance of AC mixtures in the state of Arkansas.

Chapter 4 included a brief introduction to the SCB/I-FIT test. A more detailed discussion is included here. The I-FIT protocol was developed by ICT and IDOT (Al-Qadi et al. 2015). The SCB test is the physical test where a semicircular specimen is tested using a SCB fixture placed in a servo-hydraulic or pneumatic AC testing machine (AASHTO TP 124). A line load is placed on the sample at 50 mm/min until failure occurs. Figure 17 is an example of a specimen being tested as well as SCB specimen dimensions. The testing is conducted at an intermediate temperature of 2 0
can be collected from the field as cores (AASHTO TP 124). Once the specimens are obtained, they are cut into two 50 mm disks which are to have 7 percent air voids. These disks are cut in half to form two semi circles, and a notch is cut into the center of each semi circle.

I-FIT is a method of analyzing of data collected during the SCB test. The I-FIT protocol was developed to evaluate an asphalt mixture's overall resistance to cracking-related damage (Al-Qadi et al. 2015; Ozer et al. 2016a; Braham et al. 2016). The test was intended to be used at the mix design and production levels (Braham et al. 2016). If deemed an acceptable test for Arkansas it would be used in that capacity.

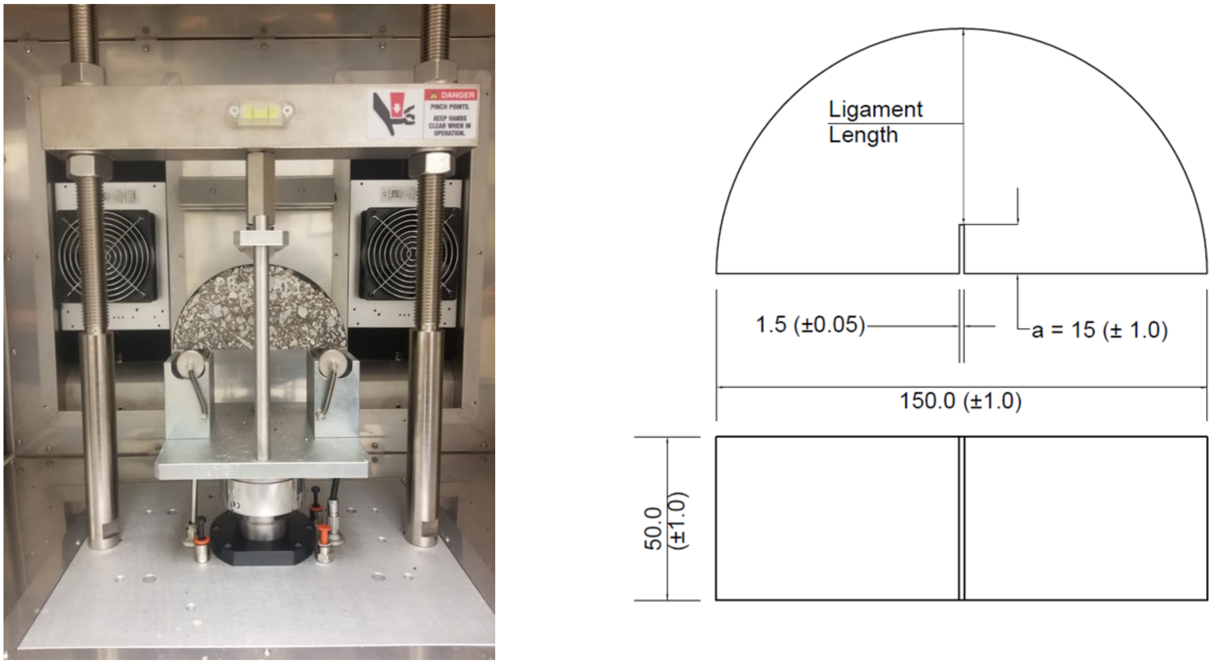


Figure 17. SCB and I-FIT Test

The main result of I-FIT is a Flexibility Index (FI). This index is a function of fracture energy (G_{fa}) reported as joules/m^2 and the absolute value of the post-peak slope at the inflection point (|m|) reported as kN/mm. The variable (A) in the equation below is a unit conversion factor and scaling coefficient (Ozer et al. 2018).

$$\text{FI} = \frac{G_{fa}}{|m|} x\ A$$

Where: Gfa = Apparent Fracture Energy

m = Slope at Post-Peak Inflection Point

A = Unit Conversion Factor and Scaling Coefficient

According to the work-of-fracture method (Hillerborg, 1985: Bazzant, 1996), fracture energy is the area under the load-displacement curve until the specimen is broken. Figure 18 is an example of the load displacement curve created. The area corresponds to the work done by load (P) on the load-point deflection (u). Assuming that all of the work of the load *P* is dissipated by the crack formation and propagation, this work would correspond to fracture energy. The method determines fracture energy, or more accurately, apparent fracture energy, because not all energy may be dissipated at the crack tip, as follows:

$$G_{fa} = \frac{1}{b(D-a)}\left[\int_{0}^{u_0} P_1(u)du + \int_{u_o}^{u_{final}} P_2(u)du\right]$$

Where: $P_1(u)$ and $P_2(u)$ = fitting equations before and after the peak, respectively;

u_0 = displacement at the peak;

u_{final} = final displacement that can be selected as the displacement at a cut-off load value where the test is considered at an end (usually taken as 0.1 kN).

If desired, the load-displacement curve can also be extrapolated to calculate the remaining area under the tail part of the curve, which is generally less than 5% of the total area. (Ozer et al. 2018)

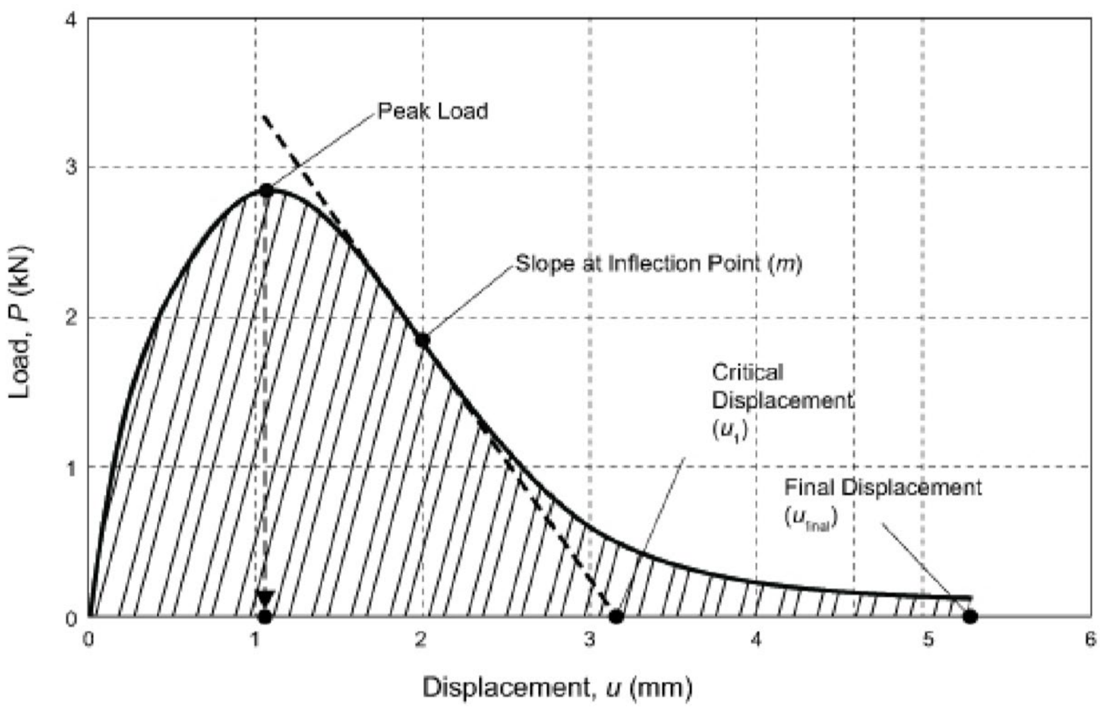


Figure 18. I-FIT Load Displacement Curve

Site Selection

ARDOT maintains a database compiling field performance data as part of its "Next 25" program. Though it is identified as having 25 projects, the database is a compilation of 40 sites across the state, including 32 asphalt pavements and 8 concrete pavements. The primary data considered from the Next 25 program was the field distress surveys. Chapter 3 details the selection of specific sites used for the study.

Mixtures

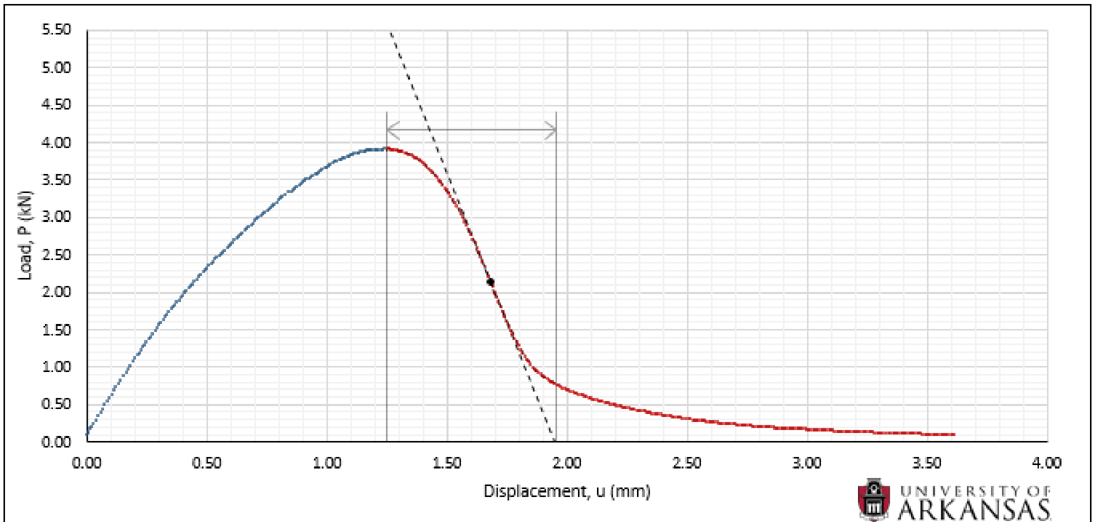
ARDOT provided the original Job Mix Formulas (JMF) for the sites selected to test. The JMF provided asphalt mixture designs for all courses (Base, Binder, and Surface). However, only the 12.5 mm surface course is being considered in this study. It was noted that the aggregate obtained would have been from a different part of the quarry (the roads selected were paved as long as 10 years ago) and that the rock may not have the same properties as the aggregate originally used. Thus, aggregate specific gravities were conducted on the rock in accordance to AASHTO T 84 and AASHTO T 85 once the aggregates were blended. Specific gravity data indicated the aggregate to be similar to the properties shown in the original mix design. The sites selected either used an unmodified PG 64-22 or polymer modified PG 70-22 binder. It is to be noted that the binder used for this research was donated from a single source, rather than attempting to obtain binder from the supplier of each specific mixture placed in the field. Laboratory aging protocols used in the study are described in Chapter 4. Table 12 summarizes the comparison of the percent air voids and VMA of the JMF compared to what was recreated in the lab.

Table 12. Mixes Used in the I-FIT Study

Site	Air Voids (%)		VMA (%)	
	JMF	**Recreated Mix**	**JMF**	**Recreated Mix**
Hindsville	4.5	4.5	14.8	14.9
Judsonia	4.5	4.7	14.8	15.0
Jonesboro	4.5	4.8	14.4	14.2
Heber Springs	4.5	4.0	14.9	14.5
Pine Bluff	4.5	4.9	14.7	14.8
DeQueen	4.5	4.5	15.8	14.8

Testing Results

All SCB/I-FIT tests were performed in accordance with AASHTO TP-124. Figures 19 and 20 provide examples of I-FIT test results.



RESULTS		
Peak Load =	3.9160	kN
Displacement at Peak Load (u_0) =	1.2405	mm
Work of Fracture (W_f) =	5.4047	J
Ligament Area ($Area_{lig}$) =	0.0030	m^2
Fracture Energy (G_f) =	1830.3	J/m^2
Post-peak slope (m) =	-7.99	kN/mm
Critical Displacement (u_1) =	1.9486	mm
Flexibility Index (FI) =	**2.29**	

SPECIMEN DIMENSIONS		
Radius (r) =	73.70	mm
Notch Length (a) =	15.10	mm
Thickness (t) =	50.39	mm

ILLINOIS FLEXIBILITY INDEX TEST (I-FIT)
UNIVERSITY OF ARKANSAS
Laboratory of Pavements and Materials

Test Date:	Wednesday, October 24, 2018
Project ID:	TRC 1802
Specimen ID:	Heber Springs NCAT
Technician:	N Parnell

File Name: UARK_SCB-IFIT_v9.4 HS_NCAT1_T1.xlsm

Figure 19. Example of I-FIT Results using the NCAT Long-term Aging Protocol

Figure 19 is a Heber Springs long term oven aging specimen and Figure 20 is a Heber Springs short term oven aging specimen. It can be noted by looking at the Figures that, although fracture energy may be similar, the FI’s can be vastly different. The fracture energy of the two specimens differs by less than 1 J/m^2. However, the FI indices differ by 10. This is due to post-peak slope. The steeper slope (Figure 19) produced the lower the flexibility index. The steeper slope indicates a brittle mixture that failed immediately after peak load was achieved.

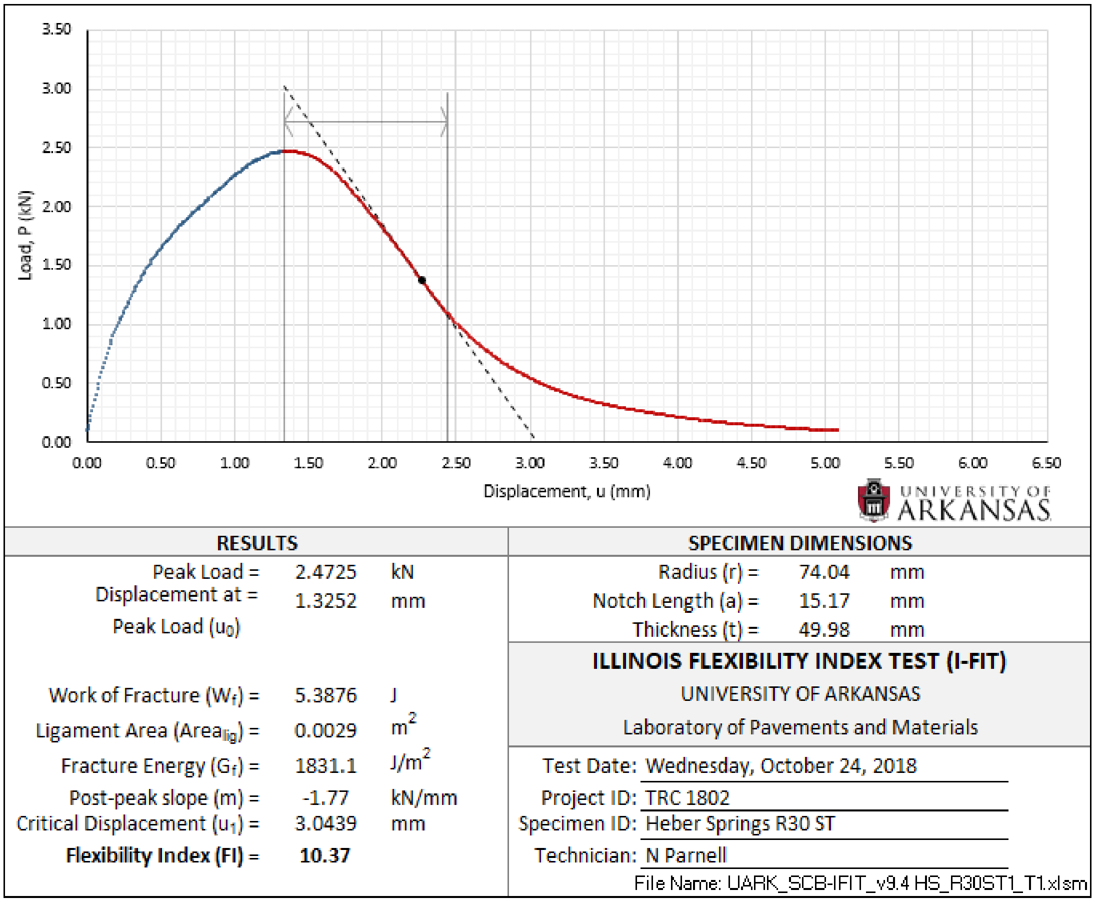


Figure 20. Example of I-FIT Results using the AASHTO R30 Short-term Aging Protocol

Table 13 and Figure 21 summarize the average FI averages for each site and aging protocol. Recall that "good" is defined as <5% cracking, "fair" as 5-10% cracking, and "poor" as >10% cacking (ARDOT, 2018).

Table 13. Flexibility Index Averages

	Short Term Aging	Long Term Aging
Hindsville	3.96	1.38
Judsonia	4.66	0.94
Jonesboro	8.48	3.11
Heber Springs	10.55	2.49
Pine Bluff	2.61	0.64
DeQueen	2.14	0.57

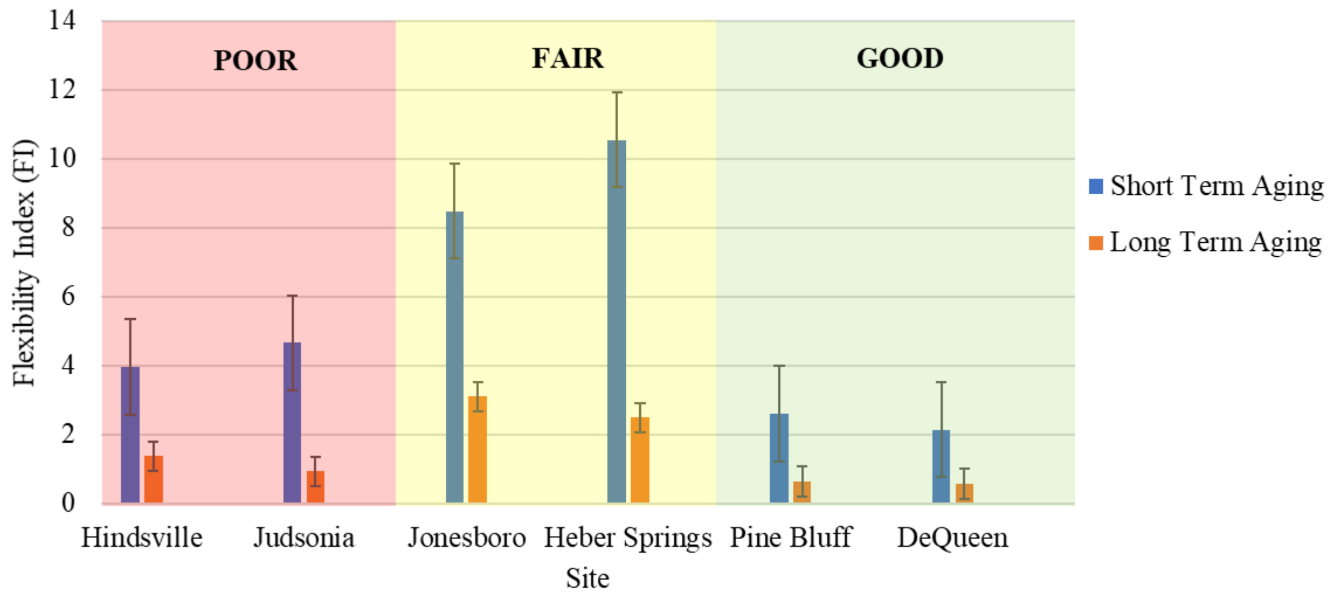


Figure 21. Flexibility Index Averages

There is consistency between the two sites within each field performance-level characterization. In addition, the effect of long-term aging is evident in the results. It is also notable that sites characterized as "fair" performing exhibited higher FI values than those characterized as "poor" performing. However, it is surprising that the two sites characterized as "good" performing did not exhibit FI values comparably higher than the other sites. One possible explanation relates to the peak load achieved during the I-FIT test for these specimens. Table 14 provides a compilation of the average peak loads achieved for each site and aging protocol.

Table 14. Average Peak Load (kN) Obtained During I-FIT Tests

Poor	**Hindsville**	**Judsonia**
Short Term Aging	3.00	3.24
Long Term Aging	4.00	4.85
Fair	**Jonesboro**	**Heber Springs**
Short Term Aging	2.74	2.62
Long Term Aging	3.75	3.97
Good	**Pine Bluff**	**DeQueen**
Short Term Aging	3.26	3.95
Long Term Aging	4.46	4.91

It is hypothesized that specimens representing the two field sites characterized as ‘good’ performing (Pine Bluff and DeQueen) have yet to reach this peak load in the field. The I-FIT test measures primarily crack propogation, not crack initiation. The specimens created in the labe are in a sense ‘pre-cracked’ because of the saw cut notch. It is possible that pavements at these sites have not yet reached crack initiation.

A series of statistical analyses were completed using the FI data. The findings confirm observations of the data, including:

- The differences between FI average values representing short-term and long-term mixture aging are significant.
- The differences between FI average values representing sites with the same field performance characterization are not significant.
- The differences between FI average values representing “poor” and “fair” field performance sites are significant. However, the differences between “poor” and “good” are not significant.
- The differences in the *variability* of test results between sites with the same field performance characterization are significant for short-term aging, but are not significant for long-term aging; this is observed across the range of field performance characterization. In other words, for a given grouping, e.g. “fair’ performing sites Jonesboro and Heber Springs, the variability of FI data for short term aging differs between sites – but does not differ for long-term aging between sites.

The results from this research compare favorably to results from research conducted in other states. A study from NCAT tested 7 different mixtures with FI values ranging between 0.4 and 10.4 (Moore 2016). A study from Missouri analyzed field cores of Superpave mixtures and FI values ranged from 0.14 to 4.98 (Butler et al. 2018). The long term oven aging results from this study are similar to those values ranging between 0.57 and 3.11.

Summary – I-FIT Testing

The primary goal of this part of the study was to determine if I-FIT could be implemented in the State of Arkansas to characterize cracking susceptibility of an asphalt concrete mixture during the mixture design process. It is important to note that the FI value – as well as the testing results from any 'index' type cracking test – is in fact an *index* and not a predictor. The FI should not be used as factual but as an *estimator* of how an asphalt mixture might perform. Recommendations stemming from the I-FIT study include:

- Additional testing should be completed for sites that would be characterized as "good" performing (field cracking performance). This is necessary to confirm and validate recommended FI values for mixture design.
- While both the AASHTO R30 and NCAT laboratory aging protocols could be used for specimen preparation, it appears that short-term aging for mechanical testing (AASHTO R30) results in adequate discrimination among various mixtures. Given the relative ease of implementation (shorter time frame, fewer stirring cycles), the R30 protocol is recommended.
- An initial acceptance value for Flexibility Index (FI) of 5 or greater is reasonable, pending additional mixture testing across Arkansas, and would not impact existing mix designs drastically. This value is comparable, but perhaps a bit lower, to published values from across the U.S. Additional validation testing will serve to refine the acceptance value.

5.2 IDEAL-CT Test

Many agencies seeking 'balanced' mix design procedures have focused on 'index type' tests (with some semi-fundamental property characterization basis) – that is, tests which provide discrimination of potential resistance to cracking (go/no-go) rather than true predictive behavior regarding the expected extent of cracking. Tests such as the Illinois flexibility index (I-FIT), the cracking resistance index (CRI), and the indirect tensile asphalt cracking test (IDEAL-CT) are popular approaches.

The IDEAL-CT stands out among other alternatives due to its simplicity in terms of both specimen fabrication and test execution: specimens do not require cutting/sawing of any kind, and displacement can be measured at the loading ram with no additional linear variable displacement transducers (LVDTs) needed. Complete, stand-alone testing packages to perform the IDEAL-CT, comparatively speaking, are economically reasonable relative to test systems capable of performing more fundamental tests. However, several agencies and small contractors could benefit from a more inexpensive equipment option.

The overall objective of this study is to evaluate the applicability/suitability of the IDEAL-CT performed on a basic Marshall testing frame. The Marshall stability/flow test has been used by many agencies around the world for decades; thus, its potential application for the IDEAL-CT is compelling, particularly outside the U.S. Specific objectives include:

- Quantifying the variability of IDEAL-CT testing parameters measured using a Marshall test frame, including load rate, as well as the repeatability of the test itself.
- Comparing IDEAL-CT outputs to those of from the I-FIT and CRI among different mixes.
- Proposing mathematical expressions to describe the relationship between the cracking tolerance index (CT_{Index}), the flexibility index (FI) and the CRI.

The Indirect Tensile Asphalt Cracking Test (IDEAL-CT).

The IDEAL-CT is performed according to ASTM standard D8225 (ASTM, 2019). It is an indirect tension test (IDT) that records displacement data from the loading ram. It is performed at room temperature, and utilizes 150 mm-diameter by 62 mm-tall compacted cylindrical specimens. Sawing is not required, and temperature conditioning can be done using a water bath. Similar to the I-FIT, it accounts for both the fracture energy and material brittleness by incorporating the slope of the back of the curve into the index. Zhou et al. (2017) addressed the limitation of the I-IFT in defining an inflection point for the back of the curve, and proposed an artificial inflection point at 75% of the peak load. The inflection point at the back of the curve is located at 75% (P_{75}) of the peak load (P_{100}), where the slope is computed using a straight line from P_{65} to P_{85}, as seen in Figure 22. P_{65} and P_{85} are points at 65% and 85% of P_{100} respectively and represent a confidence interval of 95.4% of the location of the mathematical inflection point as computed per AASHTO TP 124. The primary output of the IDEAL-CT is the cracking tolerance index (CT_{Index}) as presented in the equations that follow Figure 22.

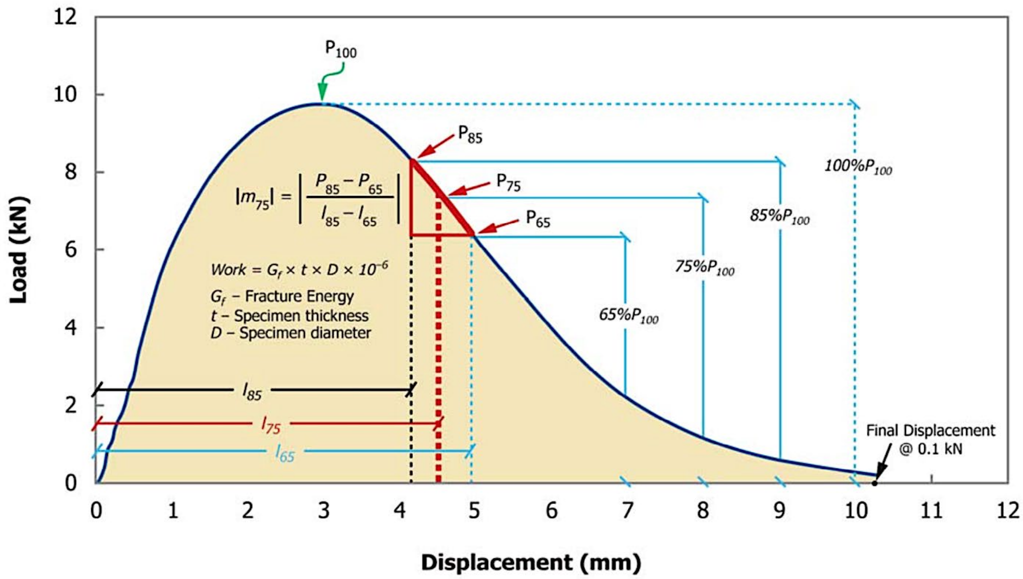


Figure 22 Typical Load-displacement Output from the IDEAL-CT *(Zhou, 2017)*

$$G_f = \frac{W_f}{D \times t} \times 10^6$$

$$|m_{75}| = \left|\frac{P_{85} - P_{65}}{l_{85} - l_{65}}\right|$$

$$CT_{Index} = \frac{t}{62} \times \frac{l_{75}}{D} \times \frac{G_f}{|m_{75}|} \times 10^6$$

where,

G_f = failure energy (Joules/m^2);

W_f = work of failure (Joules) – area underneath the load-displacement curve; computed through the quadrangle rule (*6*)

D = specimen diameter (mm);

t = specimen thickness (mm);

$|m_{75}|$ = absolute value of slope at P_{75} at the post-peak stage (N/m);

l_{65} = displacement (mm) corresponding to 65% of the peak load at the post-peak stage;

l_{75} = displacement (mm) corresponding to 75% of the peak load at the post-peak stage;

l_{85} = displacement (mm) corresponding to 85% of the peak load at the post-peak stage;

P_{65} = 65% of the peak load (kN) at the post-peak stage;

P_{85} = 85% of the peak load (kN) at the post-peak stage;

CT_{Index} = Cracking tolerance index.

Both the I-FIT and the IDEAL-CT are performed with a constant load rate of 50 mm/min. However, the IDEAL-CT does not require a pre-load stage while the I-FIT does. Both tests have been found to have strong correlation with pavement performance (Moore, 2016; Zhou, 2017).

Materials and Methods. This study is part of a larger research project focused on evaluating cracking tests for inclusion in performance-related mixture design for the state of Arkansas. Three in-service pavements are represented in this analysis. The selection process for these projects was based primarily on the extent of cracking observed in periodic visual distress surveys of the pavements over an eight-to-ten year period after construction. Additional selection considerations included similarities in traffic level and structural characteristics of the pavement cross-section. Table 15 provides a summary of asphalt surface mixture type, material sources and characteristics, traffic, and observed cracking. The "rating" of the road is based on the total amount of cracking. All mixes are 12.5 mm surface mixes.

Table 15. Field Project Characteristics

Mix/ Site	Asphalt Grade	Traffic AADT (2016)	Traffic %Truck (2016)	Long. Cracking (ft)	Trans. Cracking (ft)	Rating
M1	PG 70-22	11000	10	887	217	Poor
M2	PG 64-22	5100	10	382	194	Fair
M3	PG 70-22	7900	19	0	0	Good

Original materials from construction were not available for this study. To re-create the mixes for the study, aggregates were collected from the asphalt plants which produced the original mixes; in all cases, the aggregate sources and types were the same, and measured specific gravities and absorption capacities were similar to those reported in the original mix designs. Matching specific binder sources with the original materials proved unfeasible; therefore, a single source of binder was used for all re-created mixes in this study – but with matching the PG grade of the binder used during construction.

The volumetric properties of the re-created mixes closely approximated those reported in the job-mix formulas for the original mixtures. In some cases, the gradation of the recreated mixes was adjusted to better match volumetric properties, particularly VMA. Table 16 presents material and volumetric properties of the mixes. It is noted that all performance-testing specimens were compacted to 7.0±0.5 percent air voids.

Table 16. Job Mix Formula Properties

Mix/Site	Binder Content, P_b (%)	Air Voids, P_a (%)	Voids in Mineral Aggregate, VMA (%)	N_{des}
M1	5.0	4.5	14.8	100
M2	5.7	4.5	14.9	75
M3	4.8	4.5	14.7	100

Specimens prepared for the I-FIT were obtained from three Superpave® gyratory compactor (SGC) pills. Four SCB specimens were cut from the middle of each pill; making a total twelve specimens for the I-FIT. In case of the IDEAL-CT, only four SGC specimen were tested per mix type. No sawing was required for the IDEAL-CT. Short-term oven aging for mechanical testing per AASHTO R 30 was applied to all mixes in order to maximize the discrimination power of the tests.

I-FIT specimens were tested on an asphalt materials performance test (AMPT) in accordance with AASHTO TP124. CRI values were obtained from the I-FIT tests. IDEAL-CT specimens were tested on a standard Marshall testing frame (purchased before 1975); data acquisition for the test was accomplished using in-house-assembled instrumentation with a sampling rate of least 40 data points per second. Figure 23 shows the IDEAL-CT testing setup. The target load rate on the Marshall press was set to 50 mm/min; actual load rates are reported later. It is interesting to note that the IDEAL-CT could not be performed on the AMPT due to load-capacity limitations of the equipment.

Results

As mentioned, limitations of load capacity made impossible the evaluation of the IDEAL-CT on the AMPT. The capacity for quasi-static testing on the AMPT to which this laboratory has access is 10 kN. However, the IDEAL-CT testing requires a minimum load capability of 25 kN. Figure 24 clearly shows the large difference between an I-IFT and an IDEAL-CT load output for the same mix; this behavior that is expected due to differences in SCB and IDT failure modes. The Marshall test frame is more than adequate for typical IDEAL-CT loads; this was one of the motivations for this study.

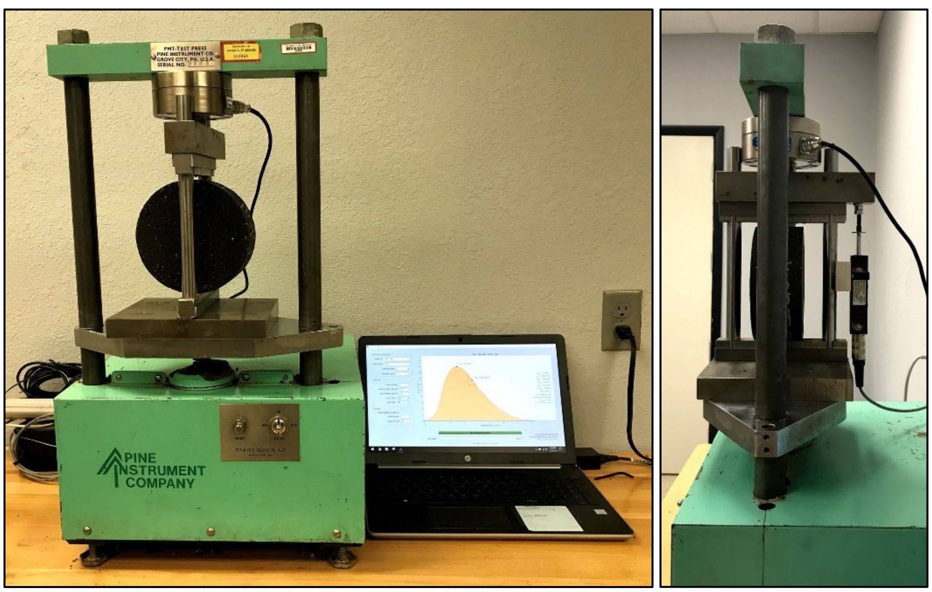


Figure 23. IDEAL-CT Performed on the Marshall Testing Frame

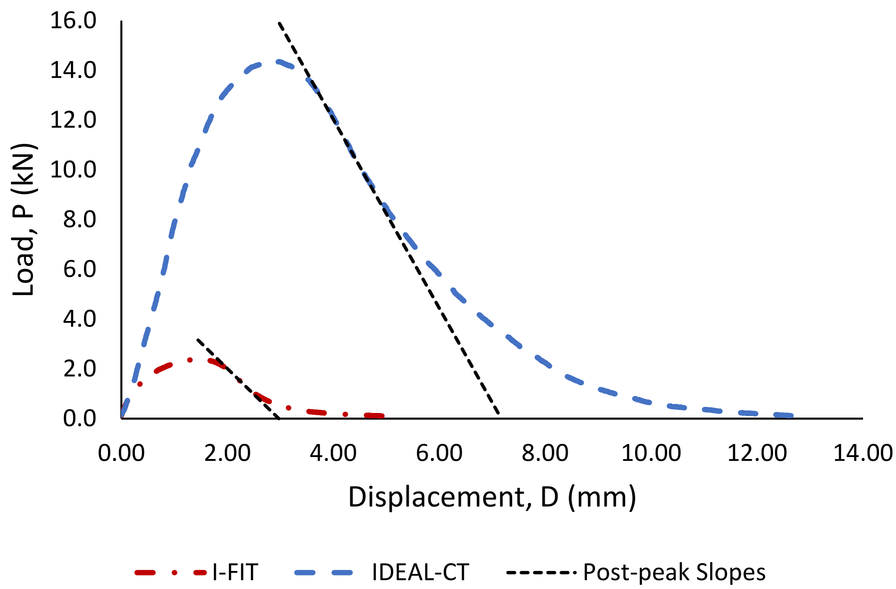


Figure 24. I-FIT and IDEAL-CT Load-displacement History for the Same Mix (M2)

A concern involving the utilization of older equipment is the reliability on testing parameters such as load rate. Figure 25 shows that the load rate measured on the Marshall frame is both significantly different and more variable than that for the AMPT. The AMPT produces a constant load rate that meets test standards (50 mm/min ± 2 mm/min). It is noted that Figure 5 represents the results from only one mix [M2]; however, the observations are consistent across multiple specimens and mixtures. The loading rate for this particular Marshall frame appears to accelerate at the beginning of the test and decelerate towards the end, with minimum rates of 42 mm/min and maximum rates of 57 mm/min.

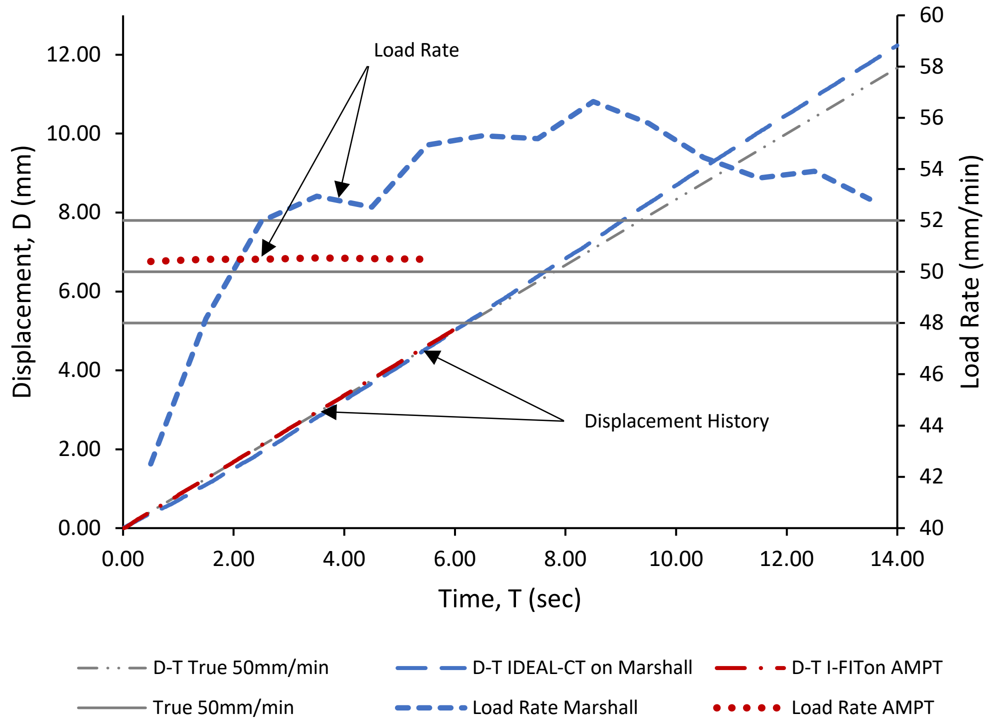


Figure 25. Load Rate for AMPT and Marshall Frame - Mix M2

Table 17 presents average index values and variability, as expressed by the coefficient of variation (COV). Flexibility indices obtained from the AMPT resulted in similar variability to CT_{Index} values obtained from the Marshall frame, despite the uneven loading rate observed for the Marshall. The CRI presents the lowest variability among all tests.

The three mixes reported here were judged to have different levels of performance in-service (Table 15). Each of the indices appear to discriminate between the mixes, in terms of cracking resistance. However, each index reports 'poorer' laboratory performance for mix M3, which was judged to exhibit 'good' in-service performance compared to mixes M1 and M2. The authors have noted that in the context of the overall study, mixture M3 – indeed – does not exhibit the same performance in the laboratory as the field. The overall study includes IFIT testing for six mixtures; the *FI* and *CRI* correctly discriminate among the other five mixes. A key take-away here is that the CT_{Index} appears to discriminate between mixtures in the same manner as the IFIT-based tests – despite the uneven loading rate displayed by the Marshall frame.

Table 17. Average Indices and Coefficient of Variation (COV)

Averages			
Mix/Site	**FI**	**CRI**	**CT_{Index}**
M1	4.0	516	28.4
M2	10.6	744	83.1
M3	2.6	426	24.9
COV (%)			
Mix/Site	**FI**	**CRI**	**CT_{Index}**
M1	28	8	18
M2	19	13	23
M3	27	8	9

Figure 26 shows the relationships between the indices evaluated in this study: CT_{Index} *FI*, and *CRI*. The relationship between CT_{Index} and *FI* appears to be similar to that between CT_{Index} and *CRI*. Preliminary mathematical relationships between the indicies are expressed in the Equations that follow. Additional testing is necessary to more firmly establish the veracity of the equations.

$$CT_{Index} = 0.2041 \times CRI - 69.2045$$
$$CT_{Index} = 7.7160 \times FI + 1.4244$$

Conclusions

Many agencies are investigating the use of a cracking-related laboratory performance test as a feature of performance-engineered asphalt mixture design. In considering the implementation of such a test, consideration should be given to both the accuracy/suitability of the test and the practical implications of equipment costs, ease of use, and other logistical issues. This study evaluated the applicability of the IDEAL-CT performed using a standard Marshall testing frame, compared specifically to the SCB/IFIT testing protocol.

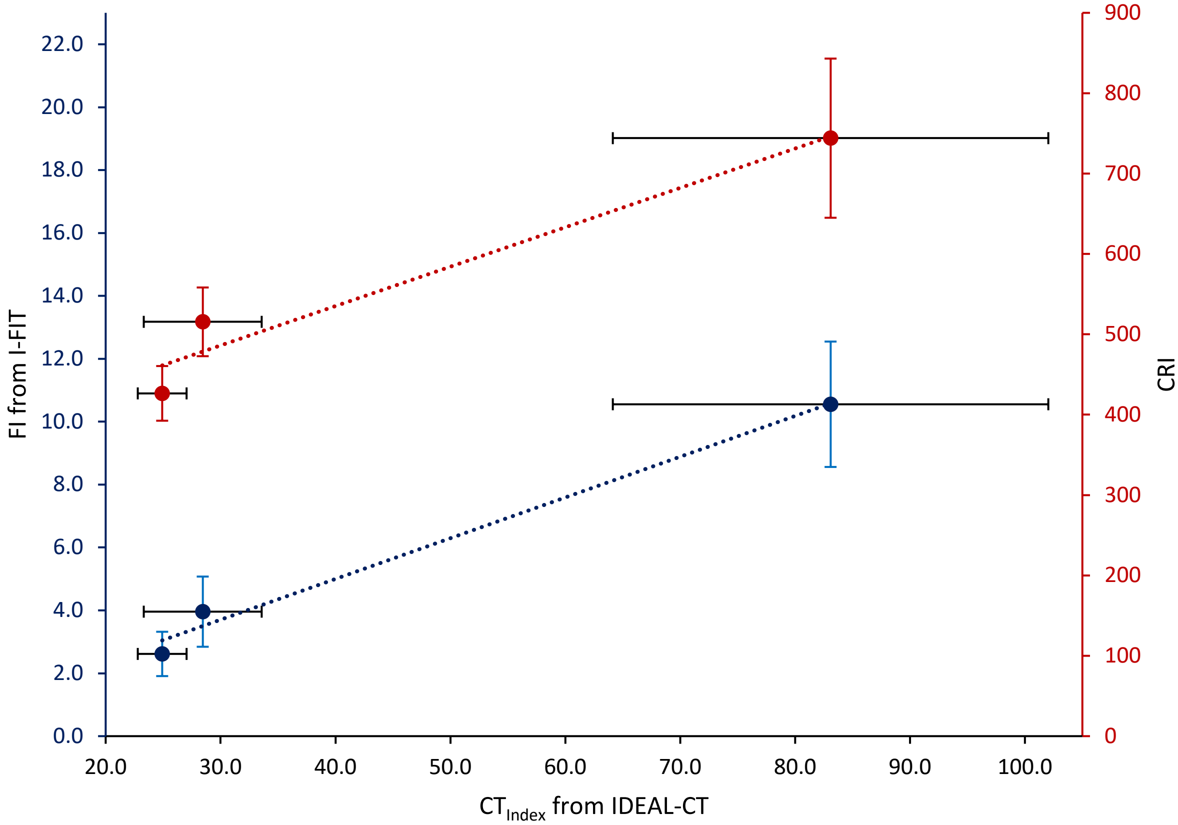


Figure 26. CT_{Index}, FI and CRI Relationships

Based on the results shown here, the following observations and conclusions are offered:

- Despite the large non-constant (and out of specification) loading rate on the Marshall testing frame, IDEAL-CT test results obtained using the frame compared favorably to IFIT test results performed on a more advanced AMPT test system.
    - The IDEAL-CT CT_{Index} generally exhibited less variability (as expressed by coefficient of variation) than the IFIT *FI*, but more variability than the IFIT *CRI*.
    - The IDEAL-CT appears to discriminate cracking-related behavior in the laboratory in a manner consistent with the IFIT test.
- There appears to be a mathematical relationship between the IDEAL-CT CT_{Index} and both the *FI* and *CRI* cracking indicies from IFIT. However, additional testing is needed to confirm such a relationship.

This study clearly indicates the potential for using the IDEAL-CT as a cracking-related 'index' type test for performance-engineered asphalt mixture design. From a practicality perspective, the IDEAL-CT offers attractive advantages: (1) potentially significant reduction in equipment costs – particularly if a given laboratory possesses a suitable load frame, e.g. a Marshall test setup; (2) significant reduction in test specimen preparation, by eliminating the need for precise sawing. From a mixture design perspective, the IDEAL-CT appears to offer similar/comparable test results to the SCB/IFIT system.

CHAPTER SIX

IMPLEMENTATION OF RESULTS

INTRODUCTION

This project sought to develop/adapt and implement a 'cracking test' for asphalt mixture design, to use in conjunction with the current APA rutting test (ARDOT Test Method 480) to shift mixture design in Arkansas to a performance, rather than volumetric, basis.

Specific project objectives included:

1. Document the current state-of-the-practice concerning asphalt cracking tests, to include: testing specifications and protocols; state Departments of Transportation policies, procedures, and specifications related to the implementation and use of these tests; agency experiences with implementation; and any other pertinent information.

2. Identify, develop, and/or adapt laboratory tests related to cracking resistance, for implementation into current mixture design procedures. Provide testing specifications (in AASHTO format) as necessary.

3. Develop initial mixture acceptance criteria for recommended cracking tests.

4. Suggest a framework for procedures to validate recommended acceptance criteria.

5. Provide recommendations for changes to ARDOT's *Standard Specifications for Highway Construction* and/or *Roadway Design Plan Development Guidelines* necessitated by the implementation of cracking tests.

The project compared laboratory cracking results to field cracking performance for mixtures placed in Arkansas. By necessity, the mixtures used in the study were "re-created" in the laboratory using aggregates from the hot-mix asphalt plants which originally supplied the field mixtures; however, while the binder *grade* was matched to that used originally, only one *source* of binder was used in the laboratory study. The study addressed specimen preparation (particularly oven aging protocols), test execution, and interpretation of test results.

RECOMMENDATIONS

It is recommended to use the Indirect Tensile Cracking Test at Intermediate Temperature (IDEAL-CT) for assessing cracking resistance during asphalt mixture design. The IDEAL-CT test method is detailed in ASTM D8225-19. The selection of the IDEAL-CT test is based on two factors: (1) the ability to discriminate between asphalt mixtures exhibiting a variation of performance; and (2) the relative ease (compared to other cracking-related tests) of specimen preparation and testing. The IDEAL-CT test requires no sawing/cutting of test specimens. It does not require specialized testing equipment, outside of an appropriate data collection system; the specified load is applied through a compression test frame with a modified-Lottman breaking head fixture (which many laboratories currently possess). Many laboratories should be able to use an existing Marshall compression frame – after demonstrating the cross-head movement meets the specified 50±2 mm/min loading rate.

Details concerning test specimen preparation, laboratory testing protocols, and mixture acceptance criteria follow.

Test Specimen Preparation

Test specimens for the IDEAL-CT are prepared using the Superpave Gyratory Compactor (SGC), following the guidelines in ASTM D8225-19. Specific notable specimen preparation details include:

- Specimen height: 62 ± 1 mm *(compacted specimen height)*
- Specimen air voids: 7 ± 0.5%
- Specimen aging: AASHTO R30, short-term aging for mechanical testing
    - Mixture condition: loose (single layer, 25 mm – 50 mm thick)
    - Oven temperature: 135C
    - Aging time: 4 hours
    - Agitation: stir loose mix every 60 minutes during aging period

The recommended mixture aging protocols are based on an extensive aging experiment completed as part of TRC1802.

Laboratory Testing Protocols

The laboratory test for assessing cracking resistance is detailed in ASTM D8225-19. Specific notable testing details include:

- Specimen conditioning: temperature chamber or water bath
  - Temperature: 25 ± 1C
  - Time: 2 hr ± 10 min.
- Test loading rate: 50 ± 2 mm / minute
- Test data: applied load and load-line displacement
  - Sampling rate: minimum of 40 data points per second
  - Stopping point: applied load drops below 100 N, after reaching peak load
- Number of Test Replicates 3 (minimum)

Mixture Acceptance Criteria

The result of the IDEAL-CT test is the *Cracking Tolerance Index* (CT_{index}). The CT_{index} is calculated using the following equations (Equations 4 and 3 in ASTM D8225-19 respectively).

$$CT_{index} = \frac{t}{62} \, x \, \frac{l_{75}}{D} \, x \, \frac{G_f}{|m_{75}|} \, x \, 10^6$$

where:

CT_{index} = cracking tolerance index

G_f = failure energy (Joules/m^2)

$|m_{75}|$ = absolute value of the post-peak slope m_{75} (N/m)

l_{75} = displacement at 75% of the peak load, after the peak (mm)

D = specimen diameter (mm)

t = specimen thickness (mm)

(Note: 10^6 is a scale factor)

$$G_f = \frac{W_f}{D \times t} x\ 10^6$$

where:

G_f = failure energy (Joules/m^2)

W_f = work of failure (Joules) *(area under the load-displacement curve)*

D = specimen diameter (mm)

t = specimen thickness (mm)

Figure 27 illustrates the concepts and quantities associated with the cracking tolerance index.

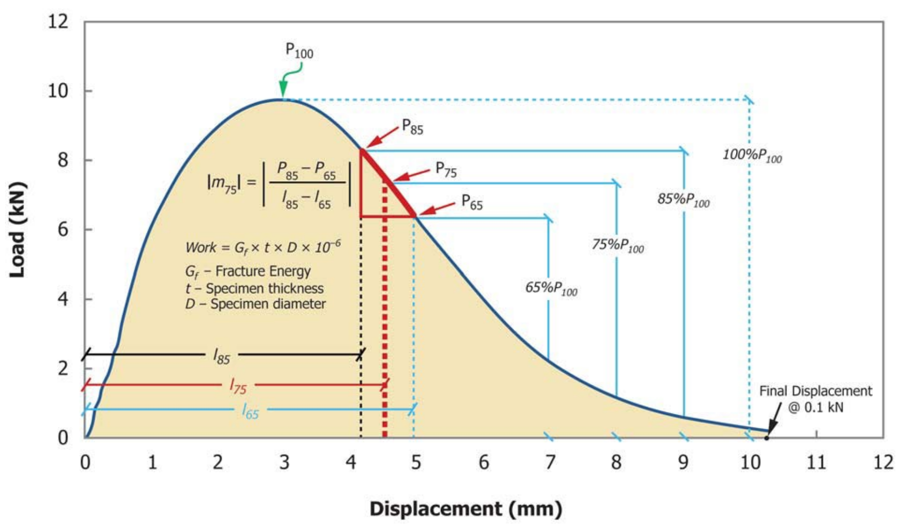


Figure 27. Load-Displacement Data Generated by IDEAL-CT Test (ASTM D8225-19)

Sites were evaluated initially using the Illinois Flexibility Index Test (IFIT). The IFIT protocol is an intermediate-temperature cracking test which is based, somewhat, on estimating the fracture properties of an asphalt specimen (particularly, crack propagation). A number of agencies have implemented, or are considering, the IFIT system for evaluating asphalt mixtures. Mixes representing eight Arkansas sites were tested using IFIT; five of those sites were also tested using IDEAL-CT. Figure 27 shows the relationship between the CT_{index} (from IDEAL-CT) and Flexibility Index (FI, from IFIT).

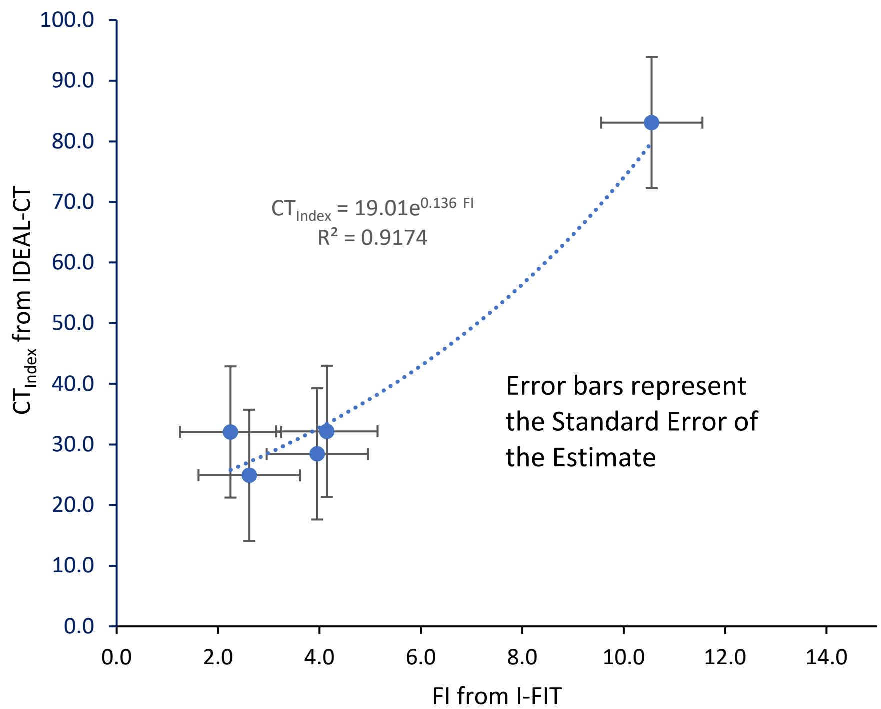


Figure 27. Relationship Between IDEAL-CT and I-FIT Results

Figure 27 illustrates that a useful relationship exists between the results from these two testing protocols. Indeed, the two tests discriminate the cracking-related behavior of asphalt mixes in the same manner. Table 18 summarizes the test results and estimates of CT_{index} and FI for mixes evaluated in this project.

Notable items related to Table 18 include:

- The CT_{index} and FI values "rank" the original six project sites in the same manner, e.g. higher values for Heber Springs and Jonesboro, intermediate values for Hindsville and Judsonia, and lower values for Pine Bluff and DeQueen. The consistency with which the two testing methods assess mixtures provides a basis for selecting/recommending a test method based on ease-of-implementation considerations.
- The CT_{index} identifies the Heber Springs and Jonesboro mixtures as exhibiting better cracking resistance than the Hindsville and Judsonia mixtures – which matches the assessment of field cracking performance. However, the laboratory test results indicate the

Pine Bluff and DeQueen mixes to have less cracking resistance than all other mixes – which does not match the assessment of field cracking performance.

- The variability of CT_{index} test results increases as the average value of CT_{index} increases; in other words, higher average CT_{index} values exhibit higher variability. This is consistent with most studies of cracking tests.

Table 18. Summary of Cracking Test Results

Site *Job Number*	Field Cracking Performance	CT_{index}			FI	
		Average	Std. Dev.	Calculated (from FI)[b]	Average	Std. Dev.
Hindsville *090116*	Poor	28	5.9		4.0	1.12
Judsonia *050188*	Poor			36	4.7	1.40
Heber Springs *050039*	Fair	83	21.9		10.6	1.99
Jonesboro *100295*	Fair			60	8.5	2.53
Pine Bluff *R20092*	Good	25	2.5		2.6	0.70
DeQueen *040488*	Good			26	2.2	0.61
Russellville[a]	N/A	32	9.2		2.2	1.02
Preston[a]	N/A	32	5.5		4.1	2.05

[a]Sites/mixtures added to study late; were not selected using field cracking performance data
[b]Not physically tested using IDEAL-CT; values calculated using equation shown in Figure 2

As discussed in the project final report, the research team speculates that field cracking data recorded at the Pine Bluff and DeQueen sites represents pavements which had not yet experienced peak load levels sufficient to initiate cracking. In addition, given that the laboratory testing performed in this study used a single source of binder (by necessity), the role of the specific binder(s) used at these two field sites on the field cracking performance is not clear. The research team is encouraged by the ability of the CT_{index} to differentiate between the two sites with "Poor" field cracking performance and "Fair" field cracking performance.

Based on the data generated in this study, the criteria for asphalt mixture acceptance based on cracking resistance is:

$$CT_{index} \geq 50$$

The recommended minimum CT_{index} corresponds (using the relationship shown in Figure 27) to an IFIT flexibility index (FI) value of 7. This value is comparable to agencies reporting cracking-related criteria based on IFIT.

IMPLEMENTATION ACTIVITIES

The findings of TRC1802 are implementable. Successful implementation of project findings will result in asphalt mixtures placed in the field being less susceptible to early-age cracking. Reducing early cracking – in effect, eliminating early maintenance and extending the life-cycle of the pavement – will provide potentially substantial cost savings to ARDOT. The sections which follow provide details related to implementation.

Specification Changes

There is little to no evidence that the majority of asphalt pavement cracking, particularly early-age cracking, initiates at the bottom of the asphalt layer system (so-called "bottom-up" cracking), outside of localized base/subgrade failures. Thus, the scope of TRC1802 was limited to surface mixes – those having a nominal maximum aggregate size (NMAS) of 9.5 mm and/or 12.5 mm (it is noted that all mixes tested in TRC1802 featured NMAS of 12.5 mm). Accordingly, changes related to the implementation of a cracking test for mixture design will target Sections 404.1(b), 404.04, and 407 (Tables 407.1 and 407.2). Recommended changes to these sections follow.

Section 404.1(b)

Starting with the sentence immediately preceding the bulleted listing of exceptions to AASHTO M323; recommendations are shown in ***highlighted boldface****:*

> The mix design will be designed in accordance with the volumetric mix design procedures contained in AASHTO M 323, its referenced standards, and the exceptions below:
>
> - PG 64-22 and PG 70-22 mixes will be designed using 4.5% air voids;
> - the fine aggregate angularity will be determined in accordance with AASHTO T 304 using the aggregate blend specific gravity of the minus No. 8 (2.36 mm) sieve through plus No. 100 (0.15 mm) sieve material;

- if any part of an ACHM Binder Course or an ACHM Base Course is within four inches (100 mm) of the pavement surface, the binder or base course lift shall comply with the angularity requirements for the top four inches (100 mm) of pavement;
- the gyratory compactor used in design, quality control, and acceptance testing must be a type evaluated by a Superpave Center and must meet the testing protocols for gyratory compactors. Gyratory compactors shall be calibrated in accordance with AASHTO T 312 and the manufacturer's recommendations. Documentation of calibration shall be made available to the Engineer upon request.
- the Voids in Mineral Aggregate (VMA) ranges will be as shown in Tables 405-1, 406-1, 407-1, or 407-2, as appropriate;
- the minimum requirement for one fractured aggregate face will be 98% and 80% for two fractured faces;
- wheel tracking test results will be determined using ARDOT Test Method 480.
- **Cracking test results will be determined using ASTM D8225-19.**
- water sensitivity will be determined using ARDOT Test Method 455A. Copies of ARDOT Test Methods are available from the Department.

Section 404.04

The table/listing of test methods; recommendations are shown in ***highlighted boldface****:*

Property	Test Method(s) (NOTE 1)
Aggregate Gradation	AASHTO T30, ARDOT 460, or AASHTO T308 1 per 750 metric tons (750 tons) Minimum
Asphalt Binder Content (NOTE 4)	ARDOT 449/449A or AASHTO T308
Stability	AASHTO T245
Air Voids (AV) (NOTE 2)	AASHTO T269
Voids in Mineral Aggregate (VMA)	ARDOT 464
Density – Maximum Theoretical	AASHTO T209
Density (Field)	AASHTO T166 or ARDOT 461
Water Sensitivity (NOTE 3)	ARDOT 455A
Wheel Tracking Test	ARDOT 480
Cracking Test	**ASTM D8225-19**

Section 407

Tables 407.1 and 407.2, immediately below specifications for the wheel tracking test; recommendations are shown in ***highlighted boldface****:*

Wheel Tracking Test	Design Gyration	Maximum Rut
(8000 cycles, 100 psi, 64°C)	75 & 115	0.315 in. (8.000 mm)
	160	0.197 in. (5.000 mm)
	205	0.197 in. (5.000 mm)
Cracking Test	**CT_{index} ≥ 50** ***(average of 3 replicates)***	

Evaluation of Implementation Impact(s)

When considering a significant change to material specifications, it is appropriate to conduct a study (or data collection activity) to gauge the impact of the change; two questions are pertinent: (1) how will the new specification impact current practice? (2) how will the new specification impact the performance of roadway pavements? Recommendations regarding each of these questions follow.

Impact on Current Practice

The addition of new criteria for asphalt mixture design will affect each mixture currently approved for use in Arkansas. Each mixture will require testing to ensure the mix meets the new specification. A preliminary recommendation of CT_{index} ≥ 50 is given for the criteria related to cracking, based on the results of TRC1802 – but it is prudent to conduct testing to gauge the number of current mixtures which might be adversely affected by this criteria. Ideally, this testing program would also include plant-produced mixes, to assess the degree to which laboratory-created specimens represent the mixture that is produced in the field. To summarize, an initial testing program would include the following:

- ArDOT-approved asphalt mixtures designs for 12.5 mm and 9.5 mm surface mixes
- Laboratory mixed (at design binder content), laboratory-compacted IDEAL-CT specimens
  - 3 replicates; 62±1 mm height; 7±0.5% air voids; AASHTO R30 short-term aging;
- Plant mixed (at design binder content), laboratory-compacted IDEAL-CT specimens
  - 3 replicates; 62±1 mm height; 7±0.5% air voids; no additional oven aging

Pavement performance data from projects featuring mixtures failing the proposed criteria should be examined for indications of early-age cracking and/or extensive cracking. In addition, pavements exhibiting early-age and/or extensive cracking – but whose asphalt mixtures do not fail the proposed criteria – should be noted and further evaluated to ascertain the probable causes of cracking.

The combination of laboratory testing and field performance data is necessary to validate the proposed criteria – and if needed, to adjust the criteria to realistically reflect field performance. It is anticipated that such a data collection effort could take up to six months or more.

Impact on Future Pavement Performance

Due to the inherent inconsistency associated with "re-creating" asphalt mixtures, it is imperative that a laboratory testing/new pavement monitoring program be initiated to assess the potential *future* impacts of implementing the findings of TRC1802. For a given time period, i.e. one to two construction season(s), asphalt mixtures approved for use on upcoming projects should be tested for laboratory cracking performance; pavements subsequently constructed using these mixtures should be intentionally monitored for performance under traffic. Such monitoring will provide the pavement management data necessary to assess the effect on pavement life-cycle attributable to the implementation of a laboratory cracking test. Ideally, pavements should be monitored for a period of ten years post-construction; beyond that period the cumulative effects of traffic and climate likely render the effects of a mix-design cracking test to be marginal.

Certainly, the concept of requiring a cracking test during asphalt mixture design is sound. An effective cracking test will provide a 'lower limit' for acceptable binder content of the mixture – in other words, it would work to ensure mixtures are not designed "too lean". The effectiveness of the test, however, rests with the mixture acceptance criteria implemented. Thus, it is vital that the recommended acceptance criteria be validated through field performance studies.

REFERENCES

AASHTO. (2016). "Standard method of test for theoretical maximum specific gravity (Gmm) and density of hot mix asphalt (HMA)." T 209-12, Washington, D. C.

AASHTO. (2017). "Standard method of test for specific gravity and absorption of fine aggregate." T 84-13, Washington, D. C.

AASHTO. (2017). "Standard method of test for bulk specific gravity (Gmb) and density of compacted asphalt mixtures using automatic vacuum sealing method." T 331-13, Washington, D. C.

AASHTO. (2017). "Standard method of test for specific gravity and absorption of coarse aggregate." T 85-14, Washington, D. C.

AASHTO. (2017) "Standard specification for Superpave volumetric mix design." M 323-17, Washington, D. C.

AASHTO. (2017). "Standard practice for Superpave volumetric design for asphalt mixtures." R 35-17, Washington, D. C.

AASHTO. (2017). "Standard Specification for Performance-Graded Asphalt Binder." M 320-17, Washington, D. C.

AASHTO (2017). "Appendix C: 1993 AASHTO Design Method" *Geotechnical Aspects of Pavements Reference Manual | Bridges and Structures,* Federal Highway Administration, 2017, www.fhwa.dot.gov/engineering/geotech/pubs/05037/ac.cfm/.

AASHTO (2018). AASHTO TP 124 Determining the Fracture Potential of Asphalt Mixtures Using Semicircular Bend Geometry (SCB) at Intermediate Temperature. American Association of State Highway and Transportation Officials, Washington, D.C., 2018.

AASHTO (2018). AASHTO R 30 Mixture Conditioning of Hot Mix Asphalt (HMA). American Association of State Highway and Transportation Officials, Washington, D.C., 2018.

Al-Qadi, I. L., H. Ozer, J. Lambros, A. El Khatib, P. Singhvi, T. Khan, J. Rivera-Perez, and B. Doll. (2015). Testing protocols to ensure performance of high asphalt binder replacement mixes using RAP and RAS (FHWA-ICT-15-017).. Illinois Center for Transportation, Illinois Department of Transportation.

ARDOT (2017). "Appendix A: Pavement Design Criteria" *Roadway Design Plan Development Guidelines | Roadway Design Division,* Arkansas Department of Transportation, 2017, www.arkansashighways.com/roadway_design_division/Roadway%20Design%20Plan%20Development%20Guidelines.pdf.

ARDOT (2018). "ARDOT Preventive Maintenance Plan," signed by Federal Highway Administration and Arkansas Department of Transportation, December 2018.

ARDOT. (2019). "Test method for determining rutting susceptibility using a loaded wheel tester (LWT)." 480-07, Little Rock, AR

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

Bazant, Z. P. (1996). "Analyzis of work-of-fracture method for measuring fracture energy of concrete." Journal of Engineering Mechanincs, Vol. 122, No.2, pp. 138-144.

Bell, C. A., AbWahab, Y., Cristi, M., & Sosnovske, D. (1994). Selection of laboratory aging procedures for asphalt-aggregate mixtures SHRP-A-383: Strategic Highway Research Program.

Bonaquist, R. (2014). Impact of Mix Design on Asphalt Pavement Durability. Transportation Research E-Circular(E-C186).

Braham, A. F., Buttlar, W. G., Clyne, T. R., Marasteanu, M. O., & Turos, M. I. (2009). The Effect of Long-Term Laboratory Aging on Asphalt Concrete Fracture Energy.

Braham, Andrew., Underwood, Shane B. (2016). "State of the art and practice in fatigue cracking evaluation of asphalt concrete pavements." Version 1.0

Butler, William G., Meister, Jim., Jahangiri, Behnam., Majidifard, Hamed., (2018). "Performance Characteristics of Modern Recycled Asphalt Mixes in Missouri, Including Ground Tire Rubber, Recycled Roofing Shingles, and Rejuvenators." University of Missouri-Columbia, Missouri Department of Transportation.

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. Paper presented at the Transportation Research Board 97th Annual Meeting, Washington DC, United States.

Goad, Stacy Denise. (1998). Superpave Mixture Design Implementation in Arkansas. Masters thesis. University of Arkansas, Fayetteville

Hall, Kevin D., Braham, Andrew F. (2017). "TRC1802 Performance-based asphalt mixture design (PBD) for Arkansas." Department of Civil Engineering, University of Arkansas, Fayetteville

Hall, Kevin D., Castillo-Camarena, Elvis A., Parnell, Nathan. (2019). "Effects of Aging Protocols of a Cracking Test in Asphalt Mixtures." Department of Civil Engineering, University of Arkansas, Fayetteville

Hall, K. D., E. A. Castillo-Camarena, and N. Parnell (2019). Effects of Aging Protocols on the Discrimination Potential of a Cracking Test in Asphalt Mixtures, ASCE Transportation and Development Institute (T&DI) International Conference on Highway and Airfield Pavements, Chicago, IL, 2019.

Hillerborg, A. (1985). "The theoretical basis of a method to determine the fracture energy of concrete." Materials and Structures, Vol 18, No.4, pp. 291-296.

Kaseer, F., F. Yin, E. Arámbula-Mercado, A. Epps Martin, J. S. Daniel, and S. Salari. (2018). Development of an Index to Evaluate the Cracking Potential of Asphalt Mixtures Using the Semi-Circular Bending Test. Construction and Building Materials, Vol. 167, 2018, pp. 286–298. https://doi.org/10.1016/j.conbuildmat.2018.02.014.

Kim, Y. R., Castorena, C., Elwardany, M., Rad, F. Y., Underwood, B. S., Gundla, A., . . . Glaser, R. R. (2018). NCHRP Report 871: Long-Term Aging of Asphalt Mixtures for Performance Testing and Prediction. Washington, D.C.: Transportation Research Board.

McCarthy, L. M., J. Callans, R. Quigley, and S. V. Scott, III (2016). “Performance Specifications for asphalt mixtures.” National Cooperative Highway Research Program, Synthesis No. 492., Washington, D. C.

Moore, Nathan D., (2016). “Evaluation of Laboratory Cracking Tests Related to Top-Down Cracking in Asphalt Pavements.” Masters Thesis. Auburn University

Newcomb, D., Martin, A. E., Yin, F., Arambula, E., Park, E. S., Chowdhury, A., . . . Coleri, E. (2015). NCHRP Report 815: Short-term laboratory conditioning of asphalt mixtures. Washington, D.C.: Transportation Research Board.

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(Supplement C), 390-401. doi:https://doi.org/10.1016/j.conbuildmat.2016.03.144

Ozer, H., Al-Qadi, I. L., Singhvi, P., Khan, T., Rivera-Perez, J., & El-Khatib, A. (2016). Fracture Characterization of Asphalt Mixtures with High Recycled Content Using Illinois Semicircular Bending Test Method and Flexibility Index. Transportation Research Record: Journal of the Transportation Research Board, 2575, 130-137. doi:10.3141/2575-14

Ozer, Hasan. Al-Qadi, I. L. (2018). “Development and implementation of the Illinois Flexibility Index Test: A protocol to evaluate the cracking resistance of asphalt mixtures.” Transportation Research Circular, (E-C237)

Richey, Allyson. (2017). “Using historic pavement performance data to identify and select sites for asphalt mixture design studies.” Honors thesis. University of Arkansas, Fayetteville

Shen, S., Wu, S., Zhang, W., Mohammad, L., & Muhunthan, B. (2017). NCHRP Report 843: Long-Term Field Performance of Warm Mix Asphalt Technologies. Washington, D.C.: Transportation Research Board.

TxDOT (2005). “Third Quarter Flexible Pavement Design Frequently Asked Questions” *Construction and Material Tips | Pavement Design Manual,* Texas Department of Transportation, 2005, https://ftp.dot.state.tx.us/pub/txdot-info/cst/tips/flex_pave_des_faq.pdf.

Zdenek, P. B., & Pere, C. P. (1988). Effect of Temperature and Humidity on Fracture Energy of Concrete. Materials Journal, 85(4). doi:10.14359/2127

Zhou, F., Newcomb, D., Gurganus, C., Banihashemrad, S., Park, E., Sakhaeifar, M., Lytton, R. (2016). “Experimental design for field validation of laboratory tests to assess cracking resistance of asphalt mixtures.” National Cooperative Highway Research Program, NCHRP 9-57., Washington, D. C.

Zhou, F., S. Im, L. Sun, and T. Scullion (2017). Development of an IDEAL Cracking Test for Asphalt Mix Design and QC/QA. Road Materials and Pavement Design, Vol. 18, No. sup4, 2017, pp. 405–427. https://doi.org/10.1080/14680629.2017.1389082.