RESEARCH

Keeping Utah Moving

Utah Department of Transportation - Research Division
4501 South 2700 West - P.O. Box 148410 - SLC,UT 84114-8410

Report No. UT-20.13

BALANCED ASPHALT CONCRETE MIX PERFORMANCE IN UTAH
PHASE IV: CRACKING INDICES FOR ASPHALT MIXTURES

Prepared For:

Utah Department of Transportation
Research & Innovation Division

Submitted By:

The University of Utah,
Department of Civil and Environmental Engineering

PEPG Consulting Engineering

Authored By:

Kevin VanFrank, P.E.
Pedro Romero, Ph.D., P.E.

Final Report
September 2020

DISCLAIMER

The authors alone are responsible for the preparation and accuracy of the information, data, analysis, discussions, recommendations, and conclusions presented herein. The contents do not necessarily reflect the views, opinions, endorsements, or policies of the Utah Department of Transportation or the U.S. Department of Transportation. The Utah Department of Transportation makes no representation or warranty of any kind, and assumes no liability therefore.

ACKNOWLEDGMENTS

The authors acknowledge the Utah Department of Transportation (UDOT) for funding this research and the following individuals from UDOT on the Technical Advisory Committee for helping to guide the research:

- Scott Andrus
- Howard Anderson
- Steve Anderson
- David Stevens
- Scott Nussbaum

The authors also acknowledge the following individuals and their work. Most of the information contained in this report is the result of their hard work. This includes the following:

- Abu Sufian Mohammad Asib, Ph.D. candidate, University of Utah
- Tim Biel, PEPG
- Faramarz Safazadeh, Ph.D., PEPG
- Michael VanMilligen, PEPG
- Clark Allen, Utah Department of Transportation
- Mike White, Utah Department of Transportation

The authors also acknowledge the financial contribution from the Mountain-Plains Consortium.

TECHNICAL REPORT ABSTRACT

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UT-20.13	N/A	N/A

4. Title and Subtitle	5. Report Date / 6. Performing Organization Code
BALANCED ASPHALT CONCRETE MIX PERFORMANCE IN UTAH, PHASE IV: CRACKING INDICES FOR ASPHALT MIXTURES	5. Report Date: September 2020 6. Performing Organization Code: N/A
7. Author(s) Kevin VanFrank (PEPG) and Pedro Romero (Univ. of Utah)	8. Performing Organization Report No. N/A
9. Performing Organization Name and Address (Prime Consultant) The University of Utah Department of Civil and Environmental Engineering 110 South Central Campus Drive Salt Lake City, UT 84112	10. Work Unit No. 5H08406H 11. Contract or Grant No. 19-8519
12. Sponsoring Agency Name and Address Utah Department of Transportation 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410	13. Type of Report & Period Covered Final Report Nov. 2018 to Sep. 2020 14. Sponsoring Agency Code PIC No. UT17.101

15. Supplementary Notes

Prepared in cooperation with the Utah Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration

16. Abstract

The Utah Department of Transportation (UDOT) has been seeking a test to balance its asphalt mixes against the stiffness encouraged by the Hamburg Wheel Tracking test and the use of recycled asphalt pavement (RAP). Without this balance, the over-stiffening of mixes has led to premature cracking. Two configurations of cracking tests have been studied previously, and conclusions about them can be found in Phases I, II and III of this series. These reports are available on the UDOT website.

In this study, the IDEAL-CT test as described in ASTM D 8225 is investigated compared to the I-FIT cracking index test on the basis of between- and within-lab repeatability and precision. It was found that, even though the IDEAL-CT test is not as fundamentally based as the I-FIT test, in samples built using the same procedures, the repeatability and precision were relatively the same. It was shown that if the samples in the IDEAL-CT test were prepared from larger samples and cut to size, repeatability was generally higher but below a tolerable threshold. If samples were built according to the ASTM procedure (62 mm uncut), the variability increased above tolerable values. If samples were compacted to a height of 75 mm, the repeatability was within tolerance standards.

It is recommended that IDEAL-CT be favored over I-FIT to evaluate mixtures for potential intermediate-temperature cracking due to the much simpler sample preparation and equivalent repeatability. It is also recommended that field-produced mixes be studied to determine sample geometry and threshold specification values.

17. Key Words	18. Distribution Statement	23. Registrant's Seal
Asphalt Concrete, Pavement Cracking, Fatigue Cracking, Flexibility Index, Semi-Circular Bend, Recycled Asphalt Pavement, IDEAL-CT	Not restricted. Available through: UDOT Research & Innov. Division 4501 South 2700 West P.O. Box 148410 Salt Lake City, UT 84114-8410 www.udot.utah.gov/go/research	N/A

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	41	N/A

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

UNIT CONVERSION FACTORS

Units used in this report and not conforming to the UDOT standard unit of measurement (U.S. Customary system) are given below with their U.S. Customary equivalents:

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)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		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)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		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. (Adapted from FHWA report template, Revised March 2003)

LIST OF ACRONYMS

AASHTO	American Association of State Highway and Transportation Officials
ANOVA	Analysis of Variance
ASTM	American Society of Testing and Materials
BBR	Bending Beam Rheometer, refers to AASHTO TP-125-16
CoV	Coefficient of Variation – Percent Ratio of the Standard Deviation to the Mean
DCSE	Dissipated Creep Strain Energy
DCT	Disc-Shaped Compact Test
DOT	Department of Transportation
FE	Fracture energy
FHWA	Federal Highway Administration
FI	Flexibility Index, refers to AASHTO TP-124-16
Gmm	Maximum theoretical specific gravity, refers to AASHTO T-209
HWTD	Hamburg Wheel Tracking Device
IDEAL-CT	Indirect tensile asphalt cracking test, refers to ASTM D 8225
IDT	Indirect Tensile Test
I-FIT	Illinois Flexibility Index Test, refers to AASHTO TP-124-16
MOI	Manual of Instruction
NDES	Designed number of gyrations
NMAS	Nominal Maximum Aggregate Size
PG	Performance Grade, refers to AASHTO M-320
RAP	Recycled Asphalt Pavement
SCB	Semi-Circular Bending
SD	Standard Deviation
SGC	Superpave Gyratory Compactor, refers to AASHTO T-312
UDOT	Utah Department of Transportation

EXECUTIVE SUMMARY

The Utah Department of Transportation (UDOT) has been seeking a test to balance its asphalt mixes against the stiffness encouraged by the Hamburg Wheel Tracking Test and the use of recycled asphalt pavement (RAP) and paying for the binder as part of the mix (making UDOT mixes drier). Without this balance, the over-stiffening of mixes has led to premature cracking. Two configurations of cracking tests have been studied previously, and conclusions about them can be found in Phases I, II and III of this series. These reports are available on the UDOT website.

In this study, the IDEAL-CT test as described in ASTM D 8225 is investigated compared to the I-FIT cracking index test on the basis of between- and within-lab repeatability and precision. It was found that, even though the IDEAL-CT test is not as fundamentally based as the I-FIT test, in samples built using the same procedures, the repeatability and precision were relatively the same. It was shown that if the samples in the IDEAL-CT test were prepared from larger samples and cut to size, repeatability was generally higher and below a tolerable threshold. If samples were built according to the ASTM procedure (62-mm uncut), the variability increased above tolerable values. If samples were compacted to a height of 75 mm, the repeatability was within tolerance standards.

It is recommended that IDEAL-CT be favored over I-FIT to evaluate mixtures for potential intermediate-temperature cracking due to the much simpler sample preparation and equivalent repeatability. It is also recommended that field-produced mixes be studied to determine sample geometry and threshold specification values.

1.0 INTRODUCTION

1.1 Background

Pavements are the Utah Department of Transportation's (UDOT's) largest and most expensive asset. Within its current practice, UDOT uses aggressive rutting and stripping testing to qualify asphalt mixes for use in highway construction. This practice was in response to the typical high-temperature and load-related distresses found in pavements. In Utah, as well as in other states, this has generally resolved rutting issues, but has led to a detrimental effect on cracking and raveling behavior in the pavements. Furthermore, in an attempt to resist rutting, increase recycling efforts, and save costs on materials, mixes now contain more Recycled Asphalt Pavement (RAP) and less asphalt binder, both virgin and total. This one-dimensional approach has been recognized as a challenge to be addressed within the mix design process, and the Department has been looking for practical tests to provide a performance balance and increase mix durability.

Binder testing alone is not adequate to predict pavement performance since it only evaluates one of the components, thus, mix performance testing is becoming increasingly important. Mixture testing has been implemented to address rutting, and low-temperature cracking is being addressed with mixture testing using the bending beam rheometer (BBR); therefore, one of the remaining major distresses in asphalt pavements is intermediate-temperature cracking (both top-down and bottom-up). Building a mix to avoid both rutting and cracking requires a balance of priorities since these behaviors are often in direct conflict.

As part of previous research efforts, the SCB I-FIT (AASHTO TP124) was recognized by UDOT as an appropriate test to measure the flexibility index (FI) of asphalt mixtures at intermediate temperatures. The research indicated that this parameter can identify mixtures that might show poor performance in terms of fatigue cracking, once placed in the field. Furthermore, it has been recognized that adoption of this, or a similar test, would lead to a more balanced mix design; one that considers low-, intermediate-, and high-temperature performance. However, a preliminary study conducted regarding the variability of test results and the reproducibility between different labs of the FI indicated that, in some cases, the variability of the test results, quantified as the standard deviation divided by the mean, was as high as 33 percent for an average of 4 samples. Furthermore, one of the labs that participated in the study showed consistent bias in the results. Based on the analysis of these results, it was concluded that there are still some variables that need to be evaluated before this test can be adopted as a specification with any level of confidence, especially if different labs will be involved in the sample preparation and testing. It is suspected that some of the variability encountered during testing is the result of variables that extend beyond the expected material variability. Discussion amongst the different labs and analysis of the data indicated that sample fabrication and analysis of data beyond the break (i.e., post-peak load) need to be refined for this test to be useful.

Meanwhile, an alternative geometry for intermediate-temperature testing was developed by Texas A&M as part of an NCHRP project. This test, called the IDEAL-CT is based on similar conceptual considerations as the SCB I-FIT, but offers certain advantages over the SCB I-FIT test in terms of specimen fabrication as it requires no cutting or notching and has simplified the determination of certain parameters, specifically the post-peak-load slope which was part of the variability encountered in the SCB I-FIT. It can be argued that if the IDEAL-CT test can provide equivalent results as the SCB I-FIT, then it has a better chance of being adopted.

1.1.1 Problem Statement

The adoption of an asphalt mixture performance-related test requires an understanding of both the mechanical ability of the test as it relates to its performance and the practicality of conducting the test in a reasonable manner. Previous phases of this research have evaluated the mechanics of the SCB I-FIT but have done very little to look at the practicality and potential for adoption of this test. Issues regarding the variability of the results coupled with the requirements for sample preparation and analysis need to be addressed before a realistic path for adoption is followed. Furthermore, following the development of the IDEAL-CT test as an alternative to the SCB I-FIT, a review-and-contrast of both tests is needed.

1.2 **Objectives**

This research represents the fourth phase of an effort to address asphalt mixture overall performance. The specific objectives of the research conducted as part of this work are:

1. Summarize existing literature on parameters that affect the variability of intermediate-temperature cracking parameters obtained from the SCB I-FIT and the IDEAL-CT tests.
2. Evaluate if the IDEAL-CT, requiring less sample preparation, provides a similar ability to differentiate asphalt mixtures as the SCB I-FIT.
3. Determine if sample preparation (cutting, compaction) can be simplified without affecting the variability of the test results.
4. Determine a procedure to evaluate the data so that reasonable variability is obtained.

1.3 Scope

This research report documents the following tasks

1. Perform a literature review on the SCB I-FIT and the IDEAL-CT in terms of their ability to predict asphalt mixture-cracking performance, their variability and their practicality.
2. Compare the variability obtained from the IDEAL-CT in reference to the SCB I-FIT.
3. Analyze the effect of sample preparation in terms of compaction density and sample manipulation (e.g., cut faces).
4. Set up a consistent methodology to collect and analyze data including evaluation of parameters so that the variability of the test is accounted for.

The material used in this study will be locally sourced based on previous work and availability. The designs will be modified to contain no RAP. The samples will be mixed and compacted in the three participating labs.

1.4 Outline of Report

This report is a continuation of the work previously described in the following research reports:

- *Development of Methods to Control Cold-Temperature and Fatigue Cracking for Asphalt Mixtures* (Report No. UT-10.08) by Romero et al. (2011);

- *Using the Bending Beam Rheometer for Low-Temperature Testing of Asphalt Mixtures* (Report No. UT-16.09) by Romero (2016);

- *Intermediate Temperature Cracking in HMA: Phase I Semi-Circular Bending (SCB) Practicality Evaluation* (Report No. UT-17.01) by VanFrank, et al. (2017);

- *Balanced Asphalt Concrete Mix Performance Phase II: Analysis of BBR and SCB-I-FIT Tests* (Report No. UT-17.21) by Romero and VanFrank; and

- *Balanced Asphalt Concrete Mix Performance in Utah, Phase III: Evaluation of Field Materials Using BBR and SCB-I-FIT Tests* (Report No. UT-19.15) by Romero and VanFrank

While some information is repeated in this report for clarity and ease of reading, most of the theoretical background has been omitted as it has already been presented in those other reports. Readers are encouraged to read the previous reports.

This report is divided into the following chapters:

- Introduction
- Literature Review and Research Methods
- Testing Variability, Sample Preparation, Data, and Evaluation
- Conclusions, Recommendations, and Implementation

2.0 LITERATURE REVIEW AND RESEARCH METHODS

2.1 Overview

Materials characterization has long been an integral part of the pavement design process. For almost 100 years, different test procedures have been developed in an effort to select the best combination of asphalt binder, aggregate, and air voids. Some of the test procedures are based on empirical relations while others are based on mechanic-based models; however, given the complex behavior of asphaltic materials and the effects of rate dependency, stress state, aging, temperature, etc., most tests end up combining both mechanic and empirical approaches. All material characterization used in representing asphalt materials must, by necessity, be a compromise between rigor and practicality (Barksdale et al., 1970).

Perhaps one of the most influential procedures for asphalt mixture design originated in the late 1930s when Bruce Marshall developed the Marshall method of mix design. His approach combined volumetric measurements with a circumferential compression test in which the peak load (referred to as stability) and the vertical displacement at peak load (referred to as flow) were measured and used as a basis for design (Leahy and McGinnis, 1999). The idea of combining a volumetric measurement with a mechanical test was later revised in the 1990s as part of the Strategic Highway Research Program. However, the type of testing suggested by SHRP turned out to be too complicated for routine use; instead different tests have been adopted by different state highway agencies to address specific pavement distresses.

Many highway agencies, including Utah DOT, have adopted the Hamburg Wheel Tracking Device (HWTD) to aggressively address high-temperature distress problems (i.e., rutting); more recently, UDOT has also made significant progress in using the BBR for mixture testing to address low-temperature distress problems (i.e., thermal cracking). Therefore, the next logical step is to address intermediate-temperature distress. The net goal is for all these tests to improve the probability that the combination of asphalt binder, aggregate, and other ingredients results in longer-lasting pavements.

2.2 Tests for Intermediate-Temperature Cracking of Pavements

Over the last 20 plus years, many cracking tests have been developed. During NCHRP project 9-57: Experiment Design for Field Validation of Laboratory Test to Assess Cracking Resistance of Asphalt Mixtures, several cracking tests were identified by the panel members as viable tests. These were (Zhou et al. 2016):

- Bending beam fatigue test (BBF)
- Texas Overlay Tests (OT)
- Disk-shaped compact tension (DCT) test
- Two versions of the indirect tensile test: the creep and strength test (IDT-CST) and the IDEAL-CT,
- Three version of the semi-circular bend (SCB) test: AASHTO TP 105 (Minnesota version), ASTM D 8044 (LSU version), and AASHTO TP-124 (Illinois version)

A summary of some of these tests, adapted from the NCHRP 9-57 report as provided by Zhou (Zhou et al. 2019) is shown in Table 2-1.

2.2.1 Semi-Circular Bend Testing Mode

Testing asphalt concrete mixtures using the semi-circular bend mode comes out of the limitation of making cylindrical samples either from the Superpave Gyratory Compactor (SGC) or from cores. In previous research UDOT evaluated both the Louisiana method and the Illinois method and determined that the Illinois method provided for adequate results with greater simplicity. The background for these tests can be found in previous UDOT reports (Report No. UT-17.01); therefore, only a small summary of the Illinois method will be provided here.

The Illinois method uses a value called the flexibility index (FI) to discriminate between asphalt concrete mixtures' potential performance. The FI is determined from a monotonic test performed on a semi-circular sample cut from an SGC cylinder and having a pre-cut notch to simulate a crack. The value is calculated by measuring the fracture energy, defined as the area under the load-displacement curve divided by the crack propagation area, and dividing this value by the absolute value of the post-peak slope at the inflection point.

Table 2-1 Summary of Cracking Tests

(Zhou et al., 2019)

Cracking Tests		Test Limitations and Equipment Cost
DCT		• Specimen prep: 3 cuts, 1 notch, and 2 holes. • Instrumentation: glue 2 studs, mount 1 clip gauge. • Equipment cost: $49,000.
SCB-AASHTO TP105		• Specimen prep: 3 cuts and 1 notch. • Instrumentation: glue 3 studs, mount 1 extensometer + 1 clip gauge. • Testing: 30 min. • Equipment cost: $52,000.
SCB-Louisiana Transportation Research Center		• Specimen prep: 9 cuts and 3 notches. • Testing: around 30 min. • Equipment cost: less than $10,000.
SCB-Illinois		• Specimen prep: 3 cuts and 1 notch. • Equipment cost: $10,000–$18,000.
IDT-CST		• Specimen prep: 2 cuts. • Instrumentation: Glue 8 studs, mount 4 extensometers. • Testing: 1–2 h. • Equipment cost: more than $50,000.
OT		• Specimen prep: 4 cuts, glue specimen to bottom plates. • Testing: 30 min–3 h. • Equipment cost: $40–50,000.
BBF		• Specimen prep: large slab, 4 cuts. • Instrumentation: glue 1 stud and mount 1 linear variable differential transformer.. • Specimen testing: 1 h to days. • Equipment cost: more than $100,000.
IDEAL-CT		• No cutting, notching, drilling, gluing, or instrumentation. • Test completion within 1 min. • Repeatable (or low variability) with COV<25%. • Practical for routine uses in DOTs and contractors' laboratories. • Low-cost test equipment (<$10,000). • Sensitive to asphalt mix composition. • Cracking performance-related.

Work by Al-Qadi and others, including UDOT, have demonstrated that the FI can distinguish between asphalt mixtures with different asphalt binder content, different RAP content, and different aging process. Field comparisons as well as comparisons with the Federal Highway Administration's Accelerated Loading Facility have shown that the FI correlates to some significant degree with field performance. (Al-Qadi et al., 2017; Ozer et al., 2016; Ozer et al. 2018). However, Romero and VanFrank showed that the geometry of the SCB I-FIT configuration seems to have a significant effect on the FI results and the overall variability of the results. Factors such as notch depth and notch thickness were specifically mentioned as being difficult to control between the labs (Report No. UT-17.21).

2.2.2 Indirect Tensile Testing Mode

Testing asphalt concrete using the indirect tensile mode has been the subject of research for many years. This is thanks to simple geometry (a cylinder) and the loading configuration (diametral compression). However, the simplicity in testing is often offset by the complex stress state that occurs during loading, making the analysis a highly sophisticated process. For example, work by Buttlar and Roque during the Strategic Highway Research Program resulted in a system capable of predicting thermal cracking in asphalt pavements (Buttlar and Roque, 1994).

In many applications, the sophisticated analysis needed to predict the mechanical response of asphalt concrete is not warranted and many simplifications are often made. An example of such simplifications is the IDEAL-CT test. The IDEAL-CT test is promoted on its simplicity with no instrumentation, cutting, gluing, drilling, or notching of the specimen required (Zhou, 2019). While this negates the mechanical aspect of the results, it provides an index that is said to have good correlation with field cracking performance. Such simplicity, while chided by academia, is embraced by practitioners who point to similarities with the Marshall mix design procedure.

2.3 **Practical Implications on Testing Methodologies**

As mentioned in the previous section, various laboratory tests have been developed to evaluate asphalt mixture potential to resist fatigue cracking. It can be argued that, for a test to be of any use, the minimum requirements are:

- Have a good correlation with observed (or measured) field performance,
- Have reasonable sensitivity to asphalt mix composition (binder type, binder content, aggregate gradation, etc.),
- Be repeatable and reproducible, with a low coefficient of variation within lab and between lab.

Besides the requirements listed above, to have a chance of being adopted, it is desirable that the test also meets the following requirements:

- Be simple to run with minimum sample preparation, cutting, or instrumentation,
- Be easy to run and analyze, requiring minimum training,
- Be reasonably priced and use equipment that is available,
- Be efficient in terms of material required and time to completion.

The relative importance of each of these factors is obviously debatable and very subjective. Nonetheless, completely ignoring any of these factors will result in a test not being selected by either highway agencies or practitioners. Therefore, a significant amount of effort is dedicated in this chapter to discuss the factors that influence the adoption of the SCB I-FIT and the IDEAL-CT.

2.3.1 Sample Preparation and Testing Procedures

Sample preparation for asphalt mixtures is limited by the ability of the material to be molded into a given shape (cylinder or square prism) with the desired properties (e.g., air void content). That means that, regardless of the test being discussed, the first step in sample preparation is the mixing and compaction process. Aggregates and asphalt binder are blended following specific recipes (job mix formula) that represent that material that will be placed in the field. The components are heated and placed in a mold to be compacted following standard procedures. Once the cylindrical sample is made, it can be tested directly or cut based on the specific test as described next.

2.3.1.1 I-FIT

To prepare the samples for I-FIT test, the gyratory compactor cylinder is cut resulting in two cylindrical samples with a height of 50 mm. Each cylindrical sample is then cut in half along its diameter resulting in four semi-circular specimens (two from each original 50-mm sample). A notch is then cut at the middle of the specimen to simulate a crack. All of the cuts are done using a cutting jig specifically designed for this test. Once the specimens are dry, they are conditioned at a temperature of 25 degrees C for at least two hours prior to testing. Testing consists of placing the specimen in a computer-controlled frame and loading it in compression at a specified rate (50 mm/min) as determined from the loading-head displacement. This setup is shown in Figure 2-1.

Figure 2-1 I-FIT Configuration

(Al-Qadi et al., 2019)

During the test, the computer data acquisition collects the load and the head displacement as a function of time, and then a specific software uses those values to calculate the area under the load-displacement curve using a numerical function. The calculated area is divided by the crack propagation area to arrive at a value referred as fracture energy. In the development of the test, it was observed that the post-peak segment of the load-displacement curves was very sensitive to changes in characteristics such as binder type and content, as well as RAP content. Therefore, the developers incorporated a shape parameter in the form of the slope of the post-peak curve into the calculations. The absolute value of the slope at the inflection point of the post-peak portion of the curve is also numerically determined by the software. The FI is calculated as the fracture energy divided by the absolute value of the slope, all multiplied by a unit conversion factor and scaling coefficient (Al-Qadi et al., 2019). This is shown schematically in Figure 2-2.

Given that four specimens are obtained from a single gyratory compactor cylinder, if all samples meet the air void requirement of 7 ± 0.5, then all four values are averaged and used to calculate the FI, otherwise corrections need to be made or samples need to be eliminated from the analysis. A more detailed discussion on this issue can be found in previous UDOT reports.

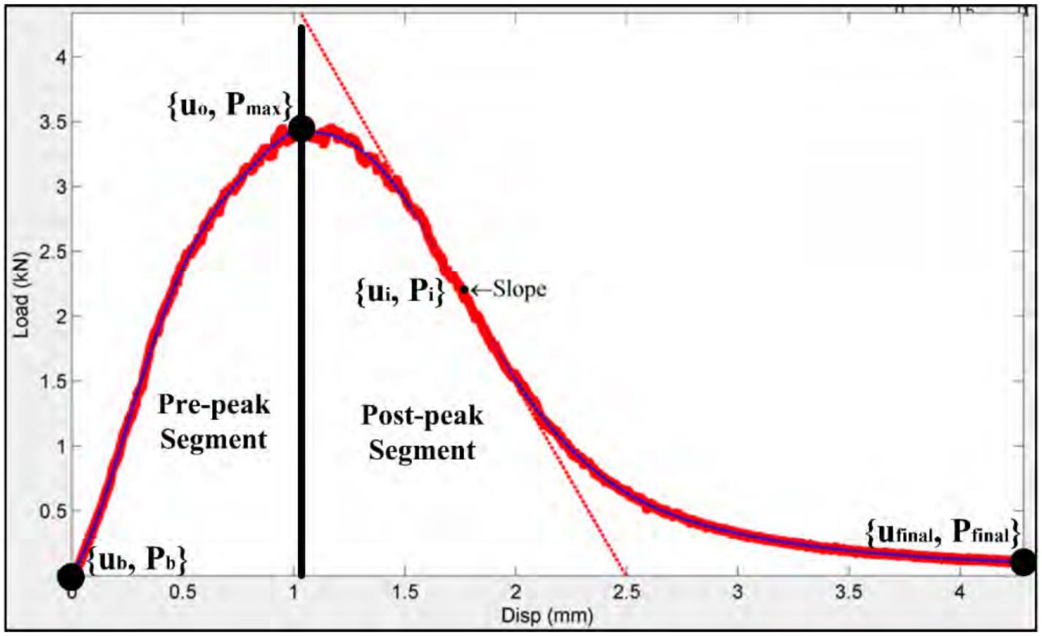

Figure 2-2 Key parameters of a load-displacement curve obtained from the I-FIT

(Al-Qadi et al., 2019)

2.3.1.2 IDEAL-CT

The IDEAL-CT is similar to traditional indirect tensile strength tests and parallels the I-FIT in the analysis process. The cylindrical specimens are prepared using the SGC that can be compacted to a specific height or cut to the desired height. If compaction is done to the proper height, then no cutting is required. This issue will be explored in more detail in Chapter 4.

In theory, the test can be run using specimens of any diameter or thickness; however, no documentation has been found to support other dimensions. Instead, the developers propose the standard diameter of 150 mm and a height of 62 mm to match the same dimensions used in Hamburg wheel tracking testing (Zhou et al., 2019). The test setup and typical results are shown in Figure 2-3.

The same inputs of force and load-head displacement are used in the analysis. However, given that this test does not have a notch or crack and the stress state is fairly complex, it is unclear how much of fracture theory actually applies. Nevertheless, the developers propose what they referred to as a cracking tolerance index (CT_{index}) with larger values relating to better cracking performance.

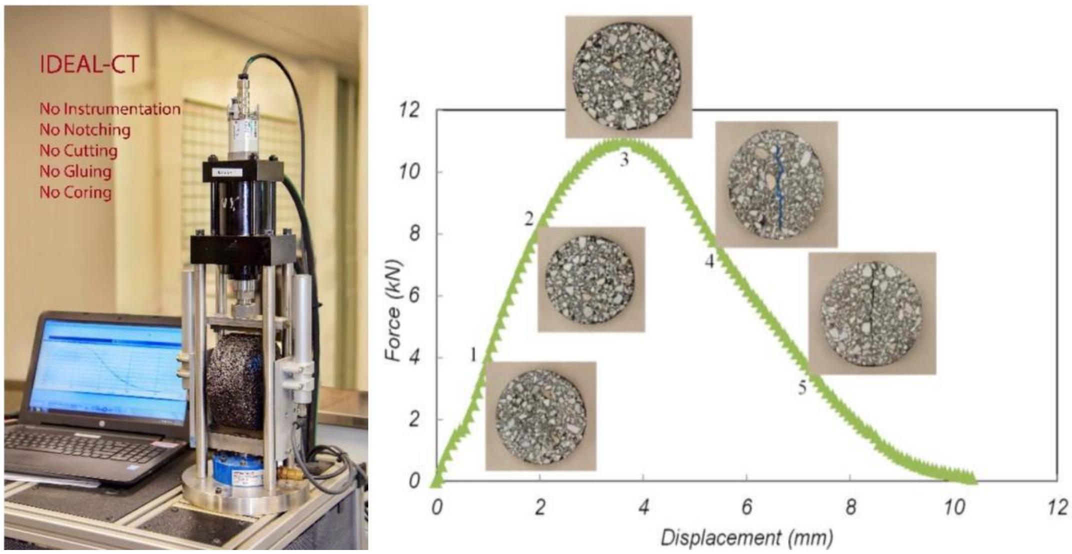

Figure 2-3 IDEAL-CT setup and typical results

(Zhou et al., 2019)

The CT-index is calculated based on the following equation:

$$CT_{index} = \frac{G_f}{\left|P/l\right|} x \left(\frac{l}{D}\right) \qquad \text{Equation 2-1}$$

In Equation 2-1, G_f represents the area under the load-versus-head-displacement curve divided by the area of the cracking face (thickness times diameter), referred to by the developers as fracture energy. The parameter |P/l| is the absolute value of the slope of the load-head displacement curve referred to as a 'modulus' parameter (P is the load and l is the displacement at a specific location). There is general agreement that the |P/l| parameter, or more appropriately $\left|\frac{\Delta P}{\Delta l}\right|$, should be determined at the inflection point of the load-head displacement curve; however, the developers recommend that for simplicity the value can be calculated between 85 and 65 percent of the peak load so this parameter is often written as $|m_{75}|$. The parameter l/D is referred to as a 'strain' tolerance parameter (the deformation tolerance under a load). The l is the head displacement and D is the diameter of the sample. For consistency, l is also determined at 75% of the maximum load. These values are represented in Figure 2-4. Equation 2-1 can be written as Equation 2-2.

$$CT_{index} = \frac{G_f}{|m_{75}|} x \left(\frac{l_{75}}{D}\right) \quad \text{Equation 2-2}$$

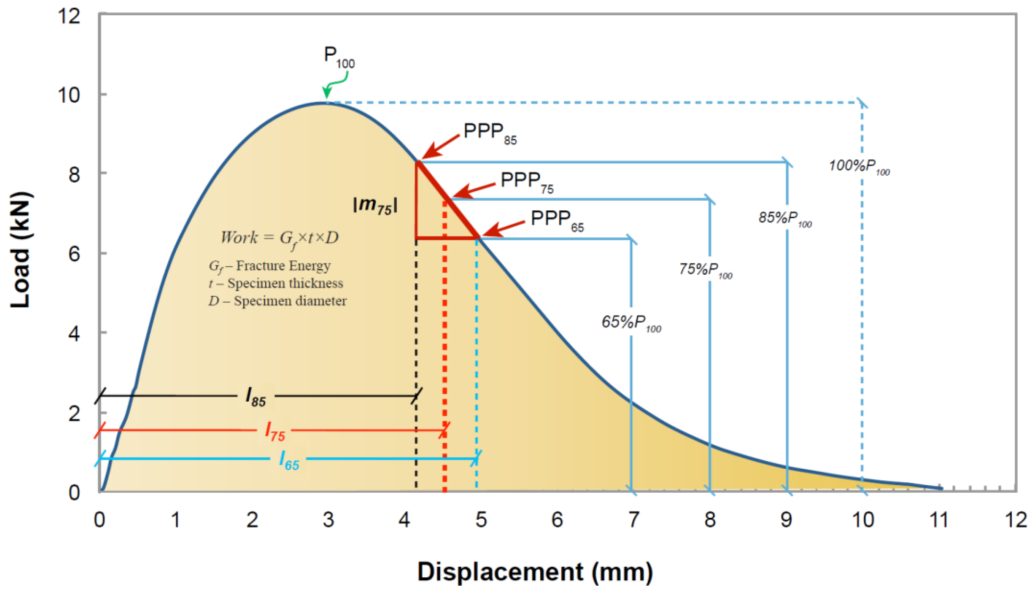

Figure 2-4 Illustration of the slope calculations.

(Zhou et al., 2019)

Similarly, to the FI, the CT_{index} is automatically calculated by the software. Most calculations assume specimens that are 62 mm thick and 150 mm in diameter. For other thickness, t, Equation 2-3 should be used. It is not clear if the results are valid for specimens with other diameters.

$$CT_{index} = \frac{t}{62} x \frac{G_f}{|m_{75}|} x \left(\frac{l_{75}}{D}\right) \quad \text{Equation 2-3}$$

The number of replicate specimens used to arrive at a CT_{index} depends on how many specimens were fabricated. Given the specimen geometry, twice the amount of material is used when compared to the FI for the same number of specimens.

2.3.2 Industry Input

One of the primary objectives of research related to asphalt materials is to develop some form of specification that can identify and likely eliminate or discourage the use of potentially troublesome mixtures. To ensure production of a durable yet economical product, asphalt mixture suppliers are also interested in performing the test during their mixture design process.

A meeting was held in which comments were requested from local industry representatives in Utah. They were asked if they would rather use the already-fabricated asphalt samples used for volumetric verification where the height is generally around 115 mm, or if they would rather compact a new sample to height, which by test procedure is 62 mm. The overwhelming comment was that the selected test for intermediate-temperature cracking should not require cutting of samples due to industry-cited liability issues regarding the use of cutting saws during routine operations.

In response to their input, it was decided by the UDOT Technical Advisory Committee for this research that compacting specimens to height, rather than cutting them from volumetric pucks, should be investigated. If no significant discrepancies are found (e.g., higher variability), the compaction to height should be the standard. This issue is investigated in Chapter 3.

2.4 **Summary**

For almost 100 years, many asphalt mixture tests have been developed; all of them share the goal of trying to identify mixtures with potential for early failure. As technology has advanced, more complex tests and analysis are possible. However, there has always been the need to balance rigor with practicality. The IDEAL-CT test seems to target the issue of practicality while still relying on some mechanic-based concepts. The relative simplicity of this test has captured the attention of the industry but, before adopting the test, it is necessary to evaluate the variability of the results.

Review of the literature and input from industry indicates that the IDEAL-CT test is a feasible candidate to control intermediate-temperature performance of asphalt mixtures.

3.0 TESTING VARIABILITY, SAMPLE PREPARATION, DATA AND EVALUATION

3.1 Overview

Testing of asphalt mixtures must consider the variability of the results. If the variability in mix performance tests is too high, then the results are problematic. Even if a test can be shown to follow an expected trend that might relate to performance, the prediction can be lost if both precision and accuracy are in doubt. A measurement is considered precise if all results are close to the same value; precision of a test can be quantified with parameters such as standard error or coefficient of variation. Of course, precision alone is of little value if the actual property value is not measured (i.e., the measurement is not accurate). Unfortunately, given the nature of material testing, accuracy is not easily quantified and practitioners often look at casual trends as a way to determine if the test is accurate.

Previous results have shown that results from fracture tests can be considered normally distributed with a small tendency to skew to the high value; as such, the researchers will accept measures of normally distributed values as valid. Readers are encouraged to read Report No. UT-19.15 for more details.

The coefficient of variation, defined as the standard deviation divided by the mean and stated as a percent, can be used to quantitatively evaluate the precision of the test and, more specifically, evaluate the repeatability of a test in a single lab with a single operator. Based on experience with asphalt mixtures testing, a coefficient of variation less than 25% is desirable to assure that the results are meaningful for acceptance testing.

To evaluate between-laboratory variability, a more sophisticated method must be employed. Since three laboratories contributed to the results, it is necessary to look at both the variation between labs as well as within labs. A single variable ANOVA evaluation will help to see whether all of the data comes from the same population. To make this analysis meaningful, all factors except the laboratory were kept as close to each other as possible. For this evaluation, the calculated F value must be smaller than the F-crit. value at the 0.05 confidence interval to conclude that the labs are getting the same answer.

3.2 **Mixture Design**

Two asphalt mixtures were used, the mixtures are representative of material used by UDOT in the Salt Lake valley but were modified to eliminate RAP as it was felt that RAP would introduce an additional variable. These are also variations of the same mixtures used in previous studies. Every attempt was made to use the same material as described in previous reports; unfortunately, the research was limited by the availability of material. The characteristics of these two mixtures are described in more detail in Report No. UT-17.21 so only a short description is presented here.

3.2.1 Gradation and Binder Content

Two aggregate gradations and two asphalt binders were selected for this study. Aggregate gradations for mixes A and B are shown in Table 3-1, and mix properties are shown in Table 3-2.

Table 3-1 Aggregate Gradation

	Mix Aggregate Gradations Percent Passing							
	19mm	12.5 mm	9.5 mm	4.75 mm	2.36 mm	1.8mm	0.30mm	0.075mm
Mix	3/4	1/2	3/8	#4	#8	#16	#50	#200
A	100	89.0	80.0	48.0	28.0	17.0	10.0	7.1
B	100	97.0	87.0	45.0	28.0	21.0	12.0	5.8
1% Lime								

Table 3-2 Mixture Characteristics

Mix Designation	Geology Description	Aggregate Bulk Specific Gravity, G_{sb}	Aggregate NMAS, mm	Design Gyrations	Binder Content, %
A	Hard Limestone	2.679	19.5	100	4.6 (m) 12.3 (v)
B	Quartzite and Granite	2.668	12.5	75	5.3 (m) 14.1 (v)

(m) by mass, (v) by volume

Laboratory samples were prepared in accordance with AASHTO TP-124 and ASTM D 8225 and mixed and compacted at the temperatures designated by the binder manufacturer. The mix was aged at compaction temperature for two hours, which is the standard practice at UDOT for all performance testing. The samples were compacted to height to achieve target air voids of 7.5 ±0.5%. Three different laboratories were involved in the study: UDOT Central Lab, PEPG, and University of Utah.

Samples for the I-FIT were tested in a previous study, and the results are described in a separate report (Report No. UT-19.15); samples for the IDEAL-CT study are described in Table 3-3.

Table 3-3 Study Matrix for IDEAL-CT testing

				Lab		
Test	Configuration	Mix	Binder	UDOT	PEPG	UofU
IDEAL-CT	115mm cut to 62 mm	A	64-34	X	X	X
		B		X	X	X
		A	70-28	X		
		B		X	X	X
	62mm uncut	A	64-34	X		
		B		X	X	X
		B	70-28	X	X	X
	75mm uncut	B	64-34	X	1/3	1/3

A full test set consisted of three sets of three pucks, totaling 9 pucks per lab, per condition.

3.3 SCB Flexibility Index Variability I-FIT

The Flexibility Index based on the Illinois version of the SCB test was the subject of a previous study (Report No. UT-19.15). In that study, the within-lab variability and the between-lab reproducibility were investigated. Based on statistical evaluations, the following was found:

- The study showed that even though there is a slight skewness towards the high end, the results from the test can be considered normally distributed, thus descriptive statistics can be used.

- The comparison of results between 3 labs indicated that, while the results are reproducible within each lab on repeated days, it is possible that a bias is introduced by a lab. Thus, it is important to verify on a regular basis that all labs are getting statistically similar results.
- The study revealed that at least 8 samples should be tested to obtain an average that represents the actual value within 20%. This requires compaction of 2 gyratory pucks.
- A coefficient of variation (CoV) between 20% - 30% was observed for samples cut from one puck. When comparing the average of four samples cut from one puck to another similar puck, the CoV was around 11%.
- There was little advantage found in performing the tests at a slower loading rate; however, more testing is recommended.

This summary is provided as a reference; readers interested in the details of the study are encouraged to read Report No. UT-19.15.

3.4 IDEAL-CT Index Variability

3.4.1 IDEAL-CT Results

To determine whether the test results from the I-FIT and IDEAL-CT tests were similar, Coefficient of Variation (CoV) for within-lab and ANOVA “F” values needed to be determined. Two aggregate blends and two binder grades were tested. Although the I-FIT binders were the same grade, the sources were different. The results are tabulated in Table 3-4. Note that not all mixes were tested in all labs. Analysis of the data indicates that, even though it has been shown that these fracture data are generally normally distributed, there is a tendency to exhibit single large values (i.e., high value outliers). About half of the time, this is the case with these results.

During testing it was observed that 62-mm, uncut samples, required high numbers of gyrations to compact to height and target density. This condition prompted the researchers to add a 75-mm-tall puck to the Mix B sample set to determine if compaction was easier. The trial was successful with compaction not exceeding 60 gyrations.

Table 3-4 IDEAL-CT Test Results

IDEAL CT Index Results								
	Lab			Lab			Lab	
UDOT	PEPG	UofU	UDOT	PEPG	UofU	UDOT	PEPG	UofU
Mix A PG 64-34 62mm Cut			Mix A PG 64-34 62mm Uncut					
161.9	146.7*	93.8	173.0		93.8			
137.8	321.2	131.3	198.1		131.3			
150.8	285.3	132.8	127.5		132.8			
164.3	382.4*	134.6	165.1		134.6			
110.5	306.6	146.0	154.6		146.0			
122.8	283.1	133.7	157.8		133.7			
147.5	298.6	181.1	77.8		181.1			
167.4	341.3	149.4	104.4		149.9			
170.3	316.4		94.7		276.2			
Mix A PG 70-28 62mm Cut								
199.1								
185.7								
203.2								
150.0								
133.5			* values are more than two standard deviations from the mean					
184.6								
171.0								
165.0								
161.4								
Mix B PG 64-34 62mm Cut			Mix B PG 64-34 62mm Uncut			Mix B PG 64-34 75mm Uncut		
531.7	515.6	280.9	173.0	274.5	227.8	439.7	404.3	305.1
320.0	464.4	464.6	198.1	337.3	556.2*	371.5	264.7	373.8
377.9	395.8	319.0	127.5	284.9	396.1	488.2	320.0	433.2
485.2	375.8	310.8	165.1	289.5	203.9	373.3		520.2
265.1	255.8	385.9	154.6	92.0	379.2	767.8*		456.5
316.9	305.6	518.4	157.8	491.6	331.1	447.7		558.0
309.0	484.6	285.9	77.8	526.4	253.5	330.0		362.9
937.3*	368.2	451.1	104.4	377.2	337.1	389.1		354.9
694.1*	437.1	529.5	94.7	596.7	382.2			393.8
217.6								
321.5								
343.7								
Mix B PG 70-28 62mm Cut			Mix B PG 70-28 62mm Uncut					
412.7	685.4	466.5	368.9	843.9	378.7			
409.8	1142.2	368.6	675.9	1181.1	429.6			
420.7	904.6	435.0	369.2	699.0	566.9			
464.5	1101.8	682.3	374.6	1209.4	529.7			
618.0	718.2	773.4	481.0	168.3	490.8			
491.4	880.7	618.3	397.2	249.3	605.1			
414.2	547.6	471.3	395.7	264.7	573.6			
457.3	1479.0	399.8	591.3	608.0	469.4			
639.8	781.3	591.9	589.1	665.4	916.3			

3.4.2 Evaluation of IDEAL-CT Results

3.4.2.1 Cracking Test Normal Distribution

In the previous study (Report No. UT-19.15), it was noticed that although the fracture index, FI, results are normally distributed, there is a tendency toward single-incidence high values. This is also noted in this data. To obtain reasonably repeatable results, the researchers discarded values falling two standard deviations above the mean. A comparison of CoV values before and after this modification are given in Table 3-5.

Table 3-5 Within-Lab Variability Modified by Removal of High Values

	Original and Adjusted Variability, Within Lab							
			UDOT		PEPG		UofU	
Mix	Sample Cond.	Binder Grade	Original	Adjust.	Original	Adjust.	Original	Adjust.
A	62 mm Cut	64-34	14.1%		21.6%	10.3%	33.5%	17.6%
	62 mm Uncut		28.9%					
A	62 mm Cut	70-28	13.2%					
B	62 mm Cut	64-34	48.7%	27.2%	21.2%		25.2%	
	62 mm Uncut		28.9%		42.6%		24.0%	
	75 mm Uncut		13.5%		21.3%		19.7%	
B	62 mm Cut	70-28	18.4%		31.1%		25.9%	
	62 mm Uncut		25.1%		58.6%		31.4%	28.1%

As expected, the CoV is greatly reduced by the adjustment.

3.4.2.2 Within-Lab and Between-Lab Results

With values greater than two standard deviations removed, the within- and between-lab results are shown using the CoV to compare within-lab results and the ANOVA single-factor test to evaluate whether the values created in different labs are derived from the same population (i.e., are the means and the variation consistent with one normal distribution or multiple distributions?). The results of these comparisons are given in Table 3-6.

Table 3-6 Variability Within and Between Labs

	Variability Within and Between Lab(s)					
Mix (Binder)	**Geometry**	UDOT	PEPG	UofU	Variability	ANOVA
A (64-34)	62 mm Cut	14.1%	10.3%	17.6%	Lower	Different
	62 mm Uncut	28.9%		33.5	Higher	NA
A (70-28)	62 mm Cut	13.2%				NA
B (64-34)	62 mm Cut	27.2%	21.2%	25.2%	Medium	Same
	62 mm Uncut	28.9%	42.6%	24.0%	Highest	Different
	75 mm Uncut	13.5%	21.3%	19.7%	Lowest	Same
B (70-28)	62 mm Cut	18.4%	31.1%	25.9%	Lower	Different
	62 mm Uncut	25.1%	58.6%	28.1%	Higher	Same

In the ANOVA column of the above table, if the between-lab data comes from the same population, the designation is 'Same". If not, the designation is "Different".

The following observations are available from this analysis:

- The adjusted coefficient of variation is below 25% except for the 62-mm uncut samples.
- Some inter-laboratory variability exists with the 62-mm cut samples independent of mix. This leads to questions about procedure.
- 62-mm uncut samples have the highest variability, independent of mix.
- 62-mm cut samples have acceptable variability.
- 75-mm uncut samples have the lowest variability; however, there were fewer replicates done with this configuration.

3.5 Comparison of Results

The coefficient of variation from the I-FIT and the IDEAL-CT tests are listed in Table 3-7. These comparisons are done with data derived from 62-mm cut pucks and discarding values falling above two standard deviations from the mean. This procedure was used to normalize the distributions in both data sets.

Table 3-7 CoV Comparison I-FIT to IDEAL-CT 62mm Cut

Cracking Index Coefficent of Variation 62 mm Cut						
	Binder Grade	Index Test		UDOT	PEPG	UofU
Mix A	64-34	I-FIT	C of V	--	10%	14%
		IDEAL 62mm	C of V	14.1%	10.3%	17.6%
	70-28	I-FIT	C of V	--	--	13%
		IDEAL 62mm	C of V	13.2%	--	--
Mix B	64-34	I-FIT	C of V	25%	21%	27%
		IDEAL 62mm	C of V	27.2%	21.2%	25.2%
	70-28	I-FIT	C of V	26%	31%	18%
		IDEAL 62mm	C of V	18.4%	31.1%	25.9%
All specimens are cut to 62-mm height with target air voids at 7 ± 0.5%						

Several observations can be made:

- Coefficient of variation with both tests occasionally exceeds the target value of 25% in this sample configuration.
- The variability is approximately the same for both tests.
- No lab is always more (or less) variable than the others.

3.6 Observation and Discussion

The data indicates that to maintain the variability below a reasonable value, the procedure requiring a 62-mm puck compacted to 7±0.5% air voids to match the Hamburg Wheel-Tracking (HWT) device specimen should not be used. Based on the limited data available in this study, testing samples compacted to 75 mm results in lower variability. There is some indication that compacting samples to 62 mm requires a large number of gyrations (higher than N_{des}) to achieve height at the target density. Since all UDOT production mixes achieve 3.5% air void at N_{des}, it seems unlikely that the contracting community would build this harsh of a mix in practice. Both of the lab mixes were adapted from mixes containing 20% RAP using a 64-34 binder and are not verified mix designs in their present forms. It is unknown at present what the comparative variability between 62-mm and 75-mm uncut samples of production mixes would be.

When pucks are cut to height in either test, the cross section is more consistent. Less variation is present due to the comparative surface smoothness vs. a raw compacted specimen. The roughness of the surface has a smaller relative impact on the cross section of a 75-mm puck vs. a 62-mm puck. All cut specimens are derived from pucks which are at least 115 mm tall. If the particles are having difficulty orienting in a thinner specimen, the compactive effort is fracturing aggregates in a random manner creating unbound surfaces and increasing variability. It is a good idea to build both the HWT specimens and the Cracking Index specimens in the same format so that they can be randomly selected as to test.

Based on the data, both the I-FIT and the IDEAL tests result in approximately the same variability. This variability can, and should, be held below a CoV of 25%. In a normal distribution, 95.5% of all data falls within two standard deviations of the mean, and 68.3% falls within one standard deviation. This statistical information should be used to develop a process by which the validity of the data could be assessed based on 3 or 4 samples. Once enough data for these tests has been collected, a procedure can be developed in which the two closest results are averaged and then the third value is rejected if it is further away from the average than the standard deviations times a value. If this rejection is necessary, one additional sample should be prepared, tested, and the results used to calculate a new average. This places three measured values inside a standard distribution curve with known variation to ensure a CoV below an acceptable value. As more data becomes available, it will be possible to state if the need to reject a sample is a rare occurrence or not. This approach is analogous to the procedures specified in ASTM C39 for testing concrete specimens.

The IDEAL-CT is a much easier and less complicated test to run. Although one more puck must be compacted for a total of three, no diametrical cutting or notching is necessary. Once the proper mass of material for compaction to height and density is obtained, pucks can be repeatedly produced to the required accuracy. The I-FIT sample is fraught with issues of inconsistent production and requires skill to get consistency.

3.7 Conclusions

The purpose of this study is twofold. The first issue is to investigate a cracking test capable of discriminating cracking-susceptible asphalt mixes from those which are not prone to cracking. To do this, a test must be repeatable within labs and between labs. The second issue is to determine if the testing can be accomplished in a reasonable time and can be performed with tools and expertise available in the field.

The I-FIT method has proven to be repeatable within the 25% threshold set for fracture testing as desired by UDOT. It is, however, a test procedure requiring a large number of replicates and significant skill. It would probably remain a tool for mix design verification.

The IDEAL-CT test has also demonstrated repeatability within the desired threshold but also has the advantage of simplicity of sample preparation. Although the exact sample geometry has yet to be determined, multiple configurations have shown promise.

The IDEAL-CT test shows the most likelihood of implementation. More work needs to be done to resolve the sample-size issue, to determine whether the sample needs to be cut or not, to select the appropriate threshold index value, or to evaluate whether a separate index value should be used for different binder grades, and whether different index values should be used for different pavement thicknesses.

4.0 CONCLUSIONS, RECOMMENDATIONS AND IMPLEMENTATION

4.1 Summary

UDOT is seeking a cracking test to offset the issues related to implementation of the Hamburg Wheel Tracking Test (HWT), the extensive use of RAP and paying for the binder as part of the mix. The high level of stiffness induced by the presence of HWT in the mix design process has resulted in a tendency for pavements to crack prematurely.

The purpose of this research was to determine whether the recently proposed ASTM D 8044 IDEAL-CT could match the I-FIT cracking index test using within- and between-lab repeatability as a criterion. Another goal was to determine if the specimen configuration specified in the IDEAL-CT test procedure could be used. This specimen consists of a 62-mm-tall asphalt puck, compacted to 7±0.5% voids. It has been determined in Phase III of this project, that the I-FIT has some difficulty achieving repeatability due to the complexity of sample preparation and the critical means of calculating the back slope of the force/displacement curve. Since sample preparation is less complex in the IDEAL-CT test and the calculation of the index is somewhat less variable, it has been argued that the test was more likely to be adopted.

To achieve parity between the two tests, the same mix design gradation and aggregates were used in both test procedures. The identical binder was however not available for the IDEAL-CT test as was used in I-FIT; however, the performance grades of the binder were the same. This disallowed direct comparison of results but did allow for variability comparison.

IDEAL-CT tests consisted of 62-mm-tall uncut, 62-mm-tall cut from 115 mm, and 75-mm-tall uncut pucks compacted to 7.5±0.5% voids. Two aggregate gradations and two binder grades were used. Three labs did three sets of three replicates for each of the four mixes. Direct comparisons were done only on the 62-mm cut samples because I-FIT is done on cut height samples.

I-FIT samples were analyzed in Phase III of this study and the results were brought forward. IDEAL-CT test results were analyzed using the coefficient of variation to determine within-lab variation and ANOVA single variable to determine between-lab repeatability. Since the analysis method was the same on both test procedures, a comparison was made of the variability of the test methods.

4.2 **Findings**

This study's purpose was to investigate the IDEAL-CT test and to determine if it could be implemented as a test to determine cracking tendency in asphalt mixes. This implementation effort is driven by the simpler sample geometry and preparation over other cracking tests. The primary concern of this study was the coefficient of variation (Standard Deviation divided by the Mean, %) which is used as a measure of test repeatability. This value is generally higher in tests that rupture the material than in tests which leave the material intact. An ideal coefficient of variation (CoV) for a cracking test would be 15%, and anything greater than 25% would be poor repeatability. The simplest sample geometry and preparation are desired.

4.2.1 Literature Review.

Review of the literature and input from industry indicates that the IDEAL-CT test is a feasible candidate to control intermediate-temperature performance of asphalt mixtures. Sample fabrication is significantly simpler than other candidates. Equipment costs are also reduced.

4.2.2 Repeatability

Comparison of IDEAL-CT CoV (within-laboratory variability) using a 62-mm cut sample compares well with I-FIT samples of the same height cut from 115-mm pucks. The CoV values range between 15 and 25% meaning that the variability is in the acceptable range for a cracking test. Changing the sample height to 62 mm without cutting resulted in a CoV between 26 and 50% meaning that this geometry might be unacceptable. When the height was increased to 75 mm, the number of gyrations needed to reach height decreased and the CoV was between 13 and 19% resulting in the lowest variability of all tests. Between-lab repeatability was also evaluated and was better with the 75-mm geometry. It is noted that the results are observed on two mixes and that neither of the mixes are production mixes nor have they been verified under UDOT mix procedures. It is also noted that mix A is a ¾-inch mix, a hot-mix asphalt (HMA) type that is no longer used by UDOT.

4.2.3 Sample Geometry

As stated in section 4.2.2, using the study mixes, the 62-mm uncut mix geometry produces poor repeatability. Cutting samples from 115-mm-high pucks produces acceptable variability, while compacting these mixes to 75 mm produces the lowest variability.

4.2.4 Index Calculation

The cracking index, as calculated, produces adequate repeatability with proper sample geometry and preparation. As with other fracture tests, since the results are observed to have some tendency to skew to the high side of the normal distribution curve with single, very large values. A consistent procedure for rejecting results which fall more than two standard deviations from the mean is needed.

4.3 Conclusions and Recommendations

The researchers' conclusion is that intermediate-temperature tests follow material ranking trends and that, through indirect evidence and input from industry, the IDEAL-CT test has the best chance of adoption. Note that the I-FIT test has been demonstrated to be repeatable with good, well-adjusted equipment and trained technicians; it is just more complicated than the alternative.

The researchers recommend the further use of the IDEAL-CT test and the abandonment of the I-FIT test to determine cracking susceptibility in asphalt paving mixes. Also recommended is further study into the sample geometry, procedures for rejection of outliers, specification thresholds, the behavior of commercially produced mixes, and the variability introduced by different test equipment.

4.4 Limitations and Challenges

One of the major challenges in this research is that the testing has been done on an adapted, RAP-containing mix. Since the presence of RAP adds an unwanted variable to the study matrix, it was removed and the mix designs were adjusted for a virgin binder. This may have resulted in a high number of gyrations being applied to achieve height and density. The resulting fractured aggregates are completely unpredictable and probably increase variability. This property is observed in the variability difference between a 62-mm- and a 75-mm-tall puck. Testing on actual production mixes is needed.

Another challenge is the inability to directly compare material ranking. Insufficient aggregate and binder have been assembled to assure consistent test samples across a multiple year study. A library of material should be developed and maintained so that pavements built today can be tested later for criteria yet to be identified.

4.5 Implementation Plan

A number of items need to be accomplished prior to implementation of the IDEAL-CT test by UDOT. They are as follows:

- Test production mixes with 62-mm uncut and 75-mm uncut geometries to determine within-lab and between-lab variability. This should involve industry labs with the variation in equipment that brings. This is an approved UTRAC study (Phase VI).
- Test mixes from the field known to have cracked prematurely and to have survived at least two years. This will help determine specification thresholds. This is an approved UTRAC study (Phase V).
- Write a specification with procedures and a manual of instruction (MOI).
- Develop a training program.
- Purchase equipment and roll the test out to the region labs for mix design verification. The UDOT central lab may be responsible to budget for this.
- Purchase equipment and roll the test out for production testing as a “For Information” standard for a period of one year. The UDOT central lab may be responsible to budget for this.
- Adjust the specification with the information and work in an incentive/disincentive structure.

The implementation should be done over a 5-year period to give industry time to adjust.

REFERENCES

Al-Qadi, I. L., D. L. Lippert, S. Wu, H. Ozer, G. Renshaw, T. R. Murphy, A. Butt, S. Gundapuneni, J. S. Trepanier, J. W. Vespa, and I. M. Said.: *Utilizing Lab Tests to Predict Asphalt Concrete Overlay Performance*. Report No. FHWA ICT-17-20. Illinois Center for Transportation, Illinois Department of Transportation, (2017).

Al-Qadi, I., Ozer, H., and Lambros, J.: Development of the Illinois Flexibility Index Test. Transportation Research Circular Number E-C251 – Relationship between Laboratory Cracking Test and Field Performance of Asphalt Mixtures. (Sept 2019)

Barksdale, R.D. and Barenberg, E.J.: *Materials Characterization*. Highway Research Board Special Report 126: Structural Evaluation of Asphalt Concrete Pavement Systems Workshop. C.L. Monismith, ed. Austin, Tx, (Dec 1970).

Buttlar, W. and Roque, R.: *Development and Evaluation of the Strategic Highway Research Program Measurement and Analysis System for Indirect Tensile Testing at Low Temperatures*. Transportation Research Record No. 1454. pp 163-171. (1994)

Leahy, R. and McGinnis, R.: *Asphalt Mixes, Materials, Design, and Characterization*. Proceedings, Association of Asphalt Paving Technologist 75th Anniversary Historical Review (68A), pp 70-127. (1999)

Ozer, H., I. L. Al-Qadi, J. Lambros, A. El-Khatib, P. Singhvi, and B. Doll.: *Development of the Fracture-Based Flexibility Index for Asphalt Concrete Cracking Potential Using Modified Semi-Circle Bending Test Parameters*. Construction and Building Materials, Vol. 115, 2016a, pp. 390–401. (2016)

Ozer, H., I. L. Al-Qadi, P. Singhvi, J. Bausano, R. Carvalho, X. Li, and N. Gibson.: *Prediction of Pavement Fatigue Cracking at an Accelerated Testing Section Using Asphalt Mixture Performance Tests*. International Journal of Pavement Engineering, Vol. 19, No. 3, pp. 264–278. (2018)

Romero, P.: *Development of Methods to Control Cold-Temperature and Fatigue Cracking for Asphalt Mixtures*. Report No. UT-10.08, Utah Department of Transportation, 2011.

Romero, P.: *Using the Bending Beam Rheometer for Low-Temperature Testing of Asphalt Mixtures*. Report No. UT-16.09, Utah Department of Transportation, 2016.

Romero, P., and VanFrank, K.: *Balanced Asphalt Concrete Mix Performance Phase II: Analysis of BBR and SCB-I-FIT Tests*. Report No. UT-17.21, Utah Department of Transportation, 2017.

Romero, P., and VanFrank, K.: *Balanced Asphalt Concrete Mix Performance in Utah, Phase III: Evaluation of Field Materials Using BBR and SCB-I-FIT Tests*. Report No. UT-19.15, Utah Department of Transportation, 2019.

VanFrank, K., *Intermediate-Temperature Cracking in HMA: Phase I Semi-Circular Bending (SCB) Practicality Evaluation.* Report No. UT-17.01, Utah Department of Transportation, 2017.

Zhou, F., Newcomb, D., Gurganus, C., Banihashemrad, S., Park, E., Sakheeifar, M., and Lytton, R.: *Experimental Design for Field Validation of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures*. Final Project Report, NCHRP project 9-57. April 2016.

Zhou, F., Im, S., and Hu, S.: Development and Validation of the IDEAL Cracking Tests. Transportation Research Circular Number E-C251 – Relationship between Laboratory Cracking Test and Field Performance of Asphalt Mixtures. (Sept 2019)

APPENDIX A: DATA

All of the data from testing was collected using electronic data acquisition of force, displacement, and temperature sensors. The data was collected in non-proprietary CSV format as generated by the data acquisition system. Spreadsheets were used to summarize and analyze the data. The raw data, called primary data, has been preserved and archived at Zenodo (https://zenodo.org/), an international repository/archive of research outputs from across all fields of research. Zenodo is listed as conforming to the USDOT Public Access Plan (https://ntl.bts.gov/publicaccess/repositories.html). According to Zenodo's policy, data entries remain accessible forever.

The data for this study is accessible at the following link:

Romero, Pedro. (2020). Evaluation of IDEAL-CT Test in Utah for Balanced Mix Design (Version 2020) [Data set]. Zenodo.
http://doi.org/10.5281/zenodo.4035139

A README file, including the metadata/information required to repeat the research, is included along with the data in the archive. Zenodo will provide proper citation for users to incorporate the data into their publications and will have a memorandum of understanding (MOU) stating that users may not re-release the data to a third party, but direct them back to the repository.