
Research Results Digest 399

Field Validation of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures: An Experimental Design

DETAILS

26 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-37560-3 | DOI 10.17226/23608

AUTHORS

Fujie Zhou, David Newcomb, Charles Gurganus, Seyedamin Banihashemrad, Maryam Sakhaeifar, Eun Sug Park, and Robert L. Lytton; National Cooperative Highway Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

August 2016

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Responsible Senior Program Officer:
Edward T. Harrigan

Research Results Digest 399

FIELD VALIDATION OF LABORATORY TESTS TO ASSESS CRACKING RESISTANCE OF ASPHALT MIXTURES: AN EXPERIMENTAL DESIGN

This digest summarizes key findings of research conducted in NCHRP Project 09-57, "Experimental Design for Field Validation of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures," by the Texas A&M Transportation Institute, Texas A&M University, College Station, Texas. This digest is based on the project final report authored by Dr. Fujie Zhou, Dr. David Newcomb, Mr. Charles Gurganus, Mr. Seyedamin Banihashemrad, Dr. Maryam Sakhaeifar, Dr. Eun Sug Park, and Dr. Robert L. Lytton. The complete project final report and three appendixes are available to download at http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=3644.

CONTENTS

Background, 1

Objective, 1

Research Approach, 1

Selection of Asphalt Cracking Tests, 2

Refinement of the Selected Cracking Tests, 9

ILS to Define Precision of the Selected Cracking Tests, 12

Experimental Design for Field Validation of the Selected Cracking Tests, 13

Summary, 25

BACKGROUND

Cracking is a primary mode of distress that frequently drives the need for rehabilitation of asphalt pavements. There are four major modes of asphalt pavement cracking—thermal, reflection, bottom-up fatigue, and top-down—that are affected by numerous factors and interactions. In the past, volumetric mixture design gave a reasonable level of comfort for potential performance since materials were relatively consistent within a given jurisdiction. However, asphalt mix designs are becoming more and more complex with the increasing use of recycled materials, recycling agents, binder additives and modifiers, and warm-mix asphalt (WMA) technologies. These changes have altered the performance of mixtures both positively and negatively so that volumetric mix design alone is no longer sufficient for evaluating the potential behavior of asphalt mixtures. Thus, there is an urgent need to establish and implement reliable performance tests that can be used to eliminate brittle mixes or with mathematical models to predict asphalt pavement cracking.

OBJECTIVE

The objective of NCHRP Project 09-57 was to develop an experimental design (including a proposed schedule and budget) for the field validation of laboratory tests selected under this study to assess the potential for the four types of cracking of asphalt mixtures. This digest briefly describes the selected cracking tests, proposed refinements, the field validation plan for the selected tests, and associated time and cost estimates.

RESEARCH APPROACH

Information was gathered from (1) a critical review of relevant research and state mixture design practices and (2) a workshop with invited experts. A comprehensive

TRANSPORTATION RESEARCH BOARD

The National Academies of

SCIENCES · ENGINEERING · MEDICINE

experimental plan, budget, and schedule were developed for a potential future project to refine and validate the selected cracking tests for use in routine asphalt mixture design.

SELECTION OF ASPHALT CRACKING TESTS

Introduction

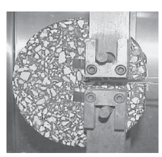
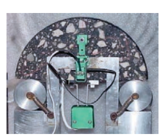
The selection of asphalt cracking tests for future field validation was a three-step process. First, relevant research and state best practices on mix designs were critically reviewed and 10 cracking tests were identified for consideration by experts participating in a Cracking Test Workshop. Next, the research team organized the workshop by (1) identifying participants, (2) conducting two cracking test webinars in which the test developers presented each test method to the workshop attendees, (3) making cracking test videos that were posted on the web and that illustrated the key steps of the cracking tests presented in the webinars, (4) developing a cracking test reference booklet tailored for the workshop participants, and (5) developing a workshop agenda. Finally, seven cracking tests were selected at the workshop held on February 11–12, 2015. The following section summarizes the cracking test selection process and the results of the workshop.

Identification of Laboratory Cracking Tests

Results of a comprehensive literature review performed to identify cracking tests suitable for routine asphalt mixture design are presented in Appendix A of the project final report. Table 1 herein summarizes the 10 candidate cracking tests identified in the literature review. Some were considered practical while others showed promise; several were either currently used by state highway agencies or were being considered for future implementation.

Selection of the 10 tests for evaluation at the workshop was based on the following considerations:

Table 1 Laboratory cracking tests.

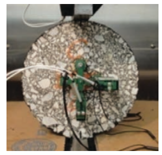
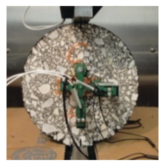
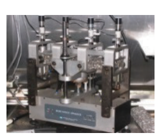
Laboratory Test						
Test Name	**Cracking Type**	**Test Standard**	**Test Configuration**	**Specimen Geometry**	**Cracking Parameter**	**Correlation to Field Performance**
DCT	Low-temperature cracking and reflection cracking	ASTM D7313 (monotonic test)		D = 6 in. T = 2 in. 2 holes D = 1 in. ND = 2.46 in.	Fracture energy	Good correlation with low-temperature cracking validated at MnROAD.
SCB	Low-temperature cracking	AASHTO TP105 (monotonic test)		D = 6 in. T = 1 in. ND = 0.6 in.	Fracture energy	Good correlation with low-temperature cracking validated at MnROAD.
	Bottom-up fatigue and top-down cracking	LTRC (monotonic test)		D = 6 in. T = 2.25 in. ND = 1, 1.25, and 1.5 in.	Energy release rate	Fair correlation to field cracking from the Louisiana Pavement Management System.

- Test variability: Cracking tests based on monotonic loading [e.g., indirect tension (IDT), semi-circular bend (SCB), disk-shaped compact tension (DCT), thermal stress-restrained specimen test (TSRST), or uniaxial thermal stress and strain test (UTSST)] generally have much lower variability with coefficients of variation (COV) often less than 15 percent. In contrast, cracking tests based on repeated loading, including the overlay test (OT), bending beam fatigue (BBF), simplified viscoelastic continuum damage (S-VECD), and direct tension (DT), may have inherently much higher COVs (>30 percent) than the monotonic tests.
- Interpretation of test results: Cracking test results can be classified as index or mechanistic-oriented parameters. Index parameters, such as fracture energy or cycles to failure, are often directly correlated to field cracking distresses; mechanistic-oriented parameters, like creep compliance from IDT and those from S-VECD and DT, must be used with cracking predictive models in order to evaluate the cracking resistance of asphalt mixtures. Index or mechanistic-oriented tests may be equally suitable for routine use as the modeling software needed for mechanistic-oriented analyses matures.
- Correlations to field performance: Regardless of whether tests use index or mechanistic-oriented parameters, or are monotonic or repeated load, any laboratory test adopted for routine use must be validated and have a good correlation with field performance.
- Test simplicity (or complexity): Technician training requirements, time for preparing and testing specimens, and difficulty in analyzing data must be weighed in the selection process. For example, IDT needs the least time for specimen preparation; in contrast, the BBF test, UTSST, S-VECD, and DT need the most time for specimen preparation. All the tests listed in Table 1 require some

Test Variability	Test Simplicity (or Complexity)	Test Sensitivity to Mix Design Parameters	Equipment Cost and Availability	Adoption by States
Low (COV = 10–15%)	Training: little time Specimen prep: 4 cuts and 2 holes Instrumentation: gluing 2 studs Testing[a]: 1–6 min. Analysis: area integration Interpretation: quick and easy (pass/fail criteria).	Asphalt binder, aggregate, RAP/RAS, and aging; insensitive to AV and P_b	Commercially available; Cost: $49,000.	Adopted by Minnesota; being considered by Colorado, South Dakota, and Montana.
Medium (COV = 20%)	Training: medium time Specimen prep: 3 cuts Instrumentation: gluing 3 studs Testing: 30 min. Analysis: area integration Interpretation: quick and easy (pass/fail criteria).	Asphalt binder, aggregate, RAP/RAS, AV and P_a	Commercially available; Cost: $52,000.	Being considered by Utah, South Dakota, Pennsylvania, and Montana.
Medium (COV = 20%)	Training: very little time Specimen prep: 4 cuts Instrumentation: none Testing: 5–10 min. Analysis: area integration and regression Interpretation: quick and easy (pass/fail criteria).	Asphalt binder, aggregate, RAP/RAS	Commercially available; Cost: $20,000.	Adopted by Louisiana; being considered by Oklahoma and New Mexico.

(*continued on next page*)

Table 1 (Continued)

Laboratory Test						
Test Name	**Cracking Type**	**Test Standard**	**Test Configuration**	**Specimen Geometry**	**Cracking Parameter**	**Correlation to Field Performance**
IDT	Low-temperature cracking	AASHTO T322: D_t and tensile strength test (monotonic tests)		D = 6 in. T = 1.5–2.0 in.	Creep compliance and tensile strength	Creep compliance and tensile strength inputs to TCMODEL. Calibrated and validated through original SHRP-I and MEPDG.
	Top-down cracking	University of Florida: M_r test, D_t test, and tensile strength test (cyclic and monotonic tests)		D = 6 in. T = 1.5–2.0 in.	Energy ratio	Validated with field cores in Florida study and confirmed at National Center for Asphalt Technology (NCAT) test track.
TSRST/UTSST	Low-temperature cracking	(monotonic test)		L = 10 in. W = 2 in. T = 2 in.	Fracture temperature	Validated with test sections during SHRP program. MnROAD test results showed moderate correlation with field performance.
Texas OT	Reflection cracking and bottom-up fatigue cracking	Tex-248-F (cyclic tests)		L = 6 in. W = 3 in. T = 1.5 in.	No. of cycles (or fracture parameters: A and n)	Good correlation with reflection cracking validated in Texas, California, and New Jersey; promising correlation with fatigue cracking validated with FHWA-accelerated loading facility (ALF) and NCAT test track.
BBF Test	Bottom-up fatigue cracking	AASHTO T321 (cyclic tests)		L = 15 in. W = 2.5 in. T = 2 in.	No. of cycles (or fatigue equation)	Correlation with bottom-up fatigue cracking historically validated.

Test Variability	Test Simplicity (or Complexity)	Test Sensitivity to Mix Design Parameters	Equipment Cost and Availability	Adoption by States
Low (COV < 11%)	Training: medium time Specimen prep: 2 cuts Instrumentation: relatively easy Testing: 1–2 hours Analysis: short and easy with data analysis software Interpretation: longer time with cracking model to predict performance.	Asphalt binder, aggregate, RAP/RAS, aging	Hydraulic test machines can be used. With test machine, more than $100,000.	AASHTO T322 is required by AASHTO-Ware.
Possibly low, similar to AASHTO T322	Training: medium time Specimen prep: 2 cuts Instrumentation: relatively easy with gauge point template Testing: 1–2 hours Analysis: easy with data analysis software Interpretation: short and easy (pass/fail criteria).	Insensitive to change in binder viscosity (Roque et al. 2010)		Being adopted by Florida.
Low (COV = around 10%)	Training: long time and intensive Specimen prep: difficult and long Instrumentation: easy and short Testing: 3–5 hours Analysis: easy and short Interpretation: quick and easy (pass/fail criteria).	Asphalt binder, aggregate, AV, P_b, and aging	Commercially available; Cost: $98,000.	Being considered by Nevada.
Relatively high (COV = 30–50%)	Training: little time Specimen prep: 4 cuts Instrumentation: none Testing: 1 min–3 hours Analysis: easy and short Interpretation: quick and easy (pass/fail criteria).	Binder, aggregate, P_b, RAP/RAS, aging, etc.	Commercially available; Cost: $46,000.	Adopted by Texas and New Jersey; being considered by Montana, Nevada, Florida, and Ohio.
Very high (COV > 50%)	Training: medium time Specimen prep: difficult and long Instrumentation: almost none Specimen testing: hours to days Analysis: easy and quick Interpretation: quick and easy (or combine with pavement analysis program to predict pavement fatigue life).	Binder, aggregate, P_b, RAP/RAS, aging, etc.	Frame (fixture) commercially available. Universal testing machine needed; could be > $100,000.	California—special pavement design; being considered by Nevada and Georgia.

(*continued on next page*)

Table 1 (Continued)

Laboratory Test						
Test Name	**Cracking Type**	**Test Standard**	**Test Configuration**	**Specimen Geometry**	**Cracking Parameter**	**Correlation to Field Performance**
S-VECD	Bottom-up fatigue and top-down cracking	AASHTO TP107 (cyclic tests) (AASHTO TP79 E* test for data analysis)		S-VECD: D = 4 in. L = 5.1 in. (E*: D = 4 in. L = 6 in.)	Fatigue equation and damage parameters (or predicted no. of cycles)	S-VECD used with MEPDG or more advanced models (LVECD and VECD-FEP++) to simulate pavement performance. Validated with FHWA-ALF test lanes and verified in North Carolina.
Direct tension	Bottom-up fatigue and top-down cracking	Texas A&M University (cyclic tests)		D = 4 in. L = 6 in.	Paris' law parameters (or no. of cycles)	Correlations with bottom-up fatigue and top-down cracking being developed under several research projects. Model and methods being validated with LTPP data.

NOTE: D = diameter; L = length; W = width; T = thickness; ND = notch depth; AV = air void; P_b = percent asphalt binder.
[a]Testing refers to the time for running the test only.

degree of specimen shaping prior to testing. Additionally, the SCB requires specimen notching, the OT gluing, and the DCT notching and drilling.

- Sensitivity to mix design parameters: Cracking tests should be sensitive to the characteristics and volumetric properties of asphalt mixtures such as binder type and content, air voids (AVs), reclaimed asphalt pavement (RAP) or recycled asphalt shingle (RAS) content, aggregate gradation, and so forth. In general, repeated loading tests are more sensitive to mix variables than monotonic loading tests.
- Other factors: Equipment availability and cost, availability of standard test methods (AASHTO, ASTM, state standard, or draft), compaction methods, and direction of loading must be considered.

Cracking Test Workshop, February 11–12, 2015

Participants

The 28 participants in the Cracking Test Workshop included the NCHRP Project 09-57 panel, members of the research team, and 14 invited experts recommended by the research team and approved by the project panel.

Cracking Test Webinars

Two half-day webinars were conducted for the workshop participants so that each cracking test

Test Variability	Test Simplicity (or Complexity)	Test Sensitivity to Mix Design Parameters	Equipment Cost and Availability	Adoption by States
Not defined	Training: very long time Specimen prep: 2 cuts and 1 coring Instrumentation: easy with a special glue jig Testing: hours to 1 day (3 more days if E* test is considered) Analysis: easy if using ALPHA-fatigue software Interpretation: quick and easy if only number of cycles is concerned [or combine with pavement analysis programs (LVECD and VECD-FEP++) to predict pavement fatigue life].	Not available	Commercially available; Cost: $97,000.	Being considered by Oklahoma, Georgia, and Pennsylvania.
Not defined	Training: very long time Specimen prep: 2 cuts and 1 coring Instrumentation: medium time and difficulty Testing: 1–2 hours Analysis: need special software Interpretation: still under development.	Model coefficients functions of AV, P_b, gradation; modulus, aging, etc.	Universal test machine needed for direct tension test; >$100,000.	Unknown

developer could present the development history, test features, lab-to-field correlation, and implementation status of each cracking test. Each presenter had 30 minutes for presentation and 10 minutes for the participants to question test developers.

Cracking Test Videos

Nine videos were produced for the 10 cracking tests; one video covered the IDT tests for both thermal and top-down cracking. The videos were a means to further assist workshop participants in understanding each cracking test and visualizing the key steps for performing it, from specimen preparation to final data analysis. Links to these cracking test videos are available at http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=3644.

Cracking Test Booklet

A 12-page cracking test summary booklet was produced specifically for the workshop, describing the key aspects of each of the 10 cracking tests. This booklet is also available at http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=3644.

Selection of Cracking Tests

The main objective of the workshop was to select tests for future laboratory and field evaluation of thermal, reflection, bottom-up fatigue, and top-down cracking. In addition to the 10 cracking tests listed in Table 1, two new cracking tests were proposed at the workshop by the participants: the SCB test in Illinois (SCB-IL) for thermal cracking

Table 2 Twelve cracking tests considered at the workshop.

Thermal Cracking Tests	Reflection Cracking Tests	Bottom-Up Fatigue Cracking Tests	Top-Down Cracking Tests
1. DCT	1. OT	1. Beam fatigue	1. IDT-Florida
2. SCB (AASHTO TP 105)	2. BBF	2. S-VECD	2. SCB-LTRC
3. SCB-IL	3. SCB-LTRC	3. Repeated tension	3. S-VECD
4. IDT (T322)	4. DCT	4. OT	4. Repeated tension
5. TSRST/UTSST		5. SCB-LTRC	5. Modified OT

NOTE: LTRC = Louisiana Transportation Research Center.

and the modified Overlay test for top-down cracking. Thus, a total of 12 cracking tests, as categorized in Table 2, were discussed at the workshop. Appendix B of the project final report documents the detailed workshop information.

A four-step process was used to select cracking tests from among the 12 candidates:

- Step 1: Determine weighting factors for seven attributes of each cracking test: (a) availability of test method, (b) test simplicity, (c) test variability, (d) sensitivity to mix parameters, (e) complexity of data analysis, (f) availability and cost of test equipment, and (g) laboratory-to-field correlation. The weighting factor ranged from 1 (least important) to 5 (most important). The weighting factors from each workshop participant were then averaged for later use in selecting cracking tests for thermal, reflection, bottom-up fatigue, and top-down cracking.
- Step 2: Determine the rating score for each cracking test under each cracking mode in terms of the seven attributes listed in Step 1. The workshop participants were divided into four groups: thermal, reflection, bottom-up fatigue, and top-down. Each group was responsible for rating the cracking tests under its cracking mode category.
- Step 3: Calculate the weighted score for each test under each cracking mode. A weighted score was determined by multiplying the scores by the attribute weighting factors. The weighted scores were then totaled for an overall ranking of the test methods.
- Step 4: Select two to three cracking tests for each cracking mode. Rank each cracking test from the highest to the lowest total score and select the first two or three tests under each cracking mode. The final candidate cracking tests selected by the workshop participants are listed in Table 3.

Summary

The final seven cracking tests selected at the workshop were (a) DCT, (b) SCB-IL, (c) SCB-TP 105, (d) SCB-LTRC, (e) OT, (f) BBF, and (g) IDT-Florida. Since the majority of these seven tests have not gone through ruggedness testing and an interlaboratory study to develop precision statements, further laboratory evaluation would be needed prior to any field validation experiment to refine the test procedures and develop their precision. A plan for this ruggedness testing and ILS (interlaboratory study) is briefly described in the next section and presented in detail in Chapter 3 of the project final report.

Table 3 Cracking tests selected at the workshop.

Thermal Cracking Tests	Reflection Cracking Tests	Bottom-Up Fatigue Cracking Tests	Top-Down Cracking Tests
1. DCT	1. OT	1. Beam fatigue	1. IDT-Florida
2. SCB-IL	2. SCB-LTRC	2. SCB-LTRC	2. SCB-LTRC
3. SCB (AASHTO TP 105)	3. BBF	3. OT*	

*OT for fatigue cracking was added later by request of the panel.

REFINEMENT OF THE SELECTED CRACKING TESTS

Ruggedness Testing

The main purpose of the proposed ruggedness testing is to identify those factors that significantly influence the cracking resistance measurements of each specific cracking test method and to estimate how closely these factors need to be controlled. Basically, the ruggedness test is a sensitivity test on variables of a test method rather than the materials under test. For a given test method, the variables may include test temperature, specimen dimensions, loading rate, and so forth. Through ruggedness testing, the sensitive test variables will be identified and the associated tolerance for each sensitive variable will be defined.

ASTM E1169-14: *Standard Practice for Conducting Ruggedness Tests* presents guidelines for ruggedness testing; it recommends that ruggedness testing be done by a single laboratory with uniform materials and precede an ILS.

Test Variables

The sensitivity of variables for each test method is the focus of the ruggedness test. These variables are features of the test and subject to control by the test method. After reviewing each test method and associated literature, the test method variables in Table 4 are proposed for inclusion in the prospective ruggedness testing.

Ruggedness Testing Experimental Design

Since many test method variables are involved, the fractional factorial Plackett-Burnam (PB) designs are often used in ruggedness testing to determine the effects of these variables. The PB designs consider two levels for each variable, and the levels chosen should be reasonably large relative to the measurement error. Thus, the high and low levels should be set at the extreme limits that could be expected to exist between different qualifying laboratories.

Table 5 shows a recommended experiment design for up to seven variables (A through G) with each variable set at two levels: (–1) for low level and (1) for high level. For five variables, Columns A, B, C, D, and F are used, for six variables, Columns A, B, C, D, F, and G. As discussed previously, the maximum number of variables for the seven cracking tests is seven. Thus, the PB designs shown in Table 5 are suitable for testing the ruggedness of the seven cracking tests evaluated in this research. The design shown in Table 5 is balanced by providing equal numbers of high and low level runs for every variable. The main effect refers to the difference between the average response of runs at the high level and the average response of runs at the low level. When the effect of a variable is the same regardless of the levels of other variables, then the main effect is the best estimate of the variable's effect.

Specifically, when transferring the PB designs shown in Table 5 to each cracking test ruggedness test, the variability of each cracking test should also be considered. In general, monotonic tests have smaller variability than repeated loading tests. Considering the variability of each cracking test, the research team determined the number of specimens required for the entire ruggedness test. Three replicate specimens are recommended for monotonic cracking tests (DCT, SCB-IL, SCB-TP 105, SCB-LTRC, and IDT-Florida) and five for repeated loading cracking tests (BBF and OT). Additionally, a safety factor of 1.5 is considered when calculating the number of specimens.

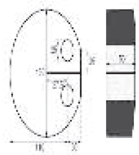
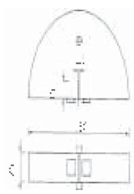
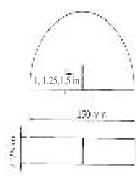
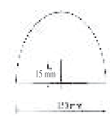
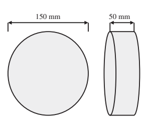
Revision of Cracking Test Methods

Statistical analysis of the results of the ruggedness testing will determine whether or not the test method is rugged with respect to the variables tested. A cracking test can be considered as rugged with regard to the variables tested if no effects are identified as statistically or practically significant, provided the experiment was carried out in the proper way. On the other hand, if some effects are practically or statistically significant, then the test method must be modified or a new specification stating acceptable ranges for the identified variables must be developed. Sometimes variable effects are statistically significant, but it is not practical to make changes. For such cases, no modification to the test method is necessary. If needed, another round of ruggedness tests may have to be conducted. After a cracking test is judged as rugged, its precision can be determined through an ILS as discussed below.

Proposed Schedule and Budget

Ruggedness testing for each cracking test must be done in one laboratory, and that laboratory must be well versed in the test procedure. Thus, multiple laboratories will likely be needed to establish the ruggedness of all of the cracking tests. At the time of this

Table 4 Proposed ruggedness testing variables for selected cracking tests.

Test Name	Test Standard	Test Configuration	Specimen Geometry	Test Variability
DCT	ASTM D7313 (monotonic test)		D = 150 mm T = 50 mm 2 holes D = 25 mm ND = 62 mm	Low (COV = 10–15%)
SCB	AASHTO TP 105 (monotonic test)		D = 150 mm T = 25 mm ND = 15 mm	Medium (COV = 20%)
	LTRC (monotonic test)		D = 150 mm T = 57 mm ND = 25.4, 31.8, 38.1 mm	Medium (COV = 20%)
	Illinois (monotonic test)		D = 150 mm T = 50 mm ND = 15 mm	—
IDT-Florida	University of Florida: D_t test and tensile strength test (monotonic test)		D = 150 mm T = 50 mm	Possibly low, similar to AASHTO T322
Texas OT	Tex-248-F (cyclic tests)		L = 150 mm W = 76 mm T = 38 mm	Relatively high (COV = 30–50%)
BBF	AASHTO T321 (cyclic tests)		L = 380 mm W = 63 mm T = 50 mm	Very high (COV > 50%)

Ruggedness Testing Variables	Standard Factor	High Level	Low Level	Tolerance
1. Specimen thickness (T)	50 mm	+5 mm	−5 mm	±2 mm
2. Crack opening displacement (COD)	1 mm/min	+5%	−5%	±2%
3. Test temperature (t)	Performance-graded (PG) low+10°C	+0.5°C	−0.5°C	±0.1°C
4. Location of holes (LH)	25 mm	+5 mm	−5 mm	±1 mm
5. Notch depth (ND)	62 mm	+3 mm	−3 mm	±1 mm
6. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen thickness (T)	25 mm	+5 mm	−5 mm	±2 mm
2. Crack opening displacement (COD)	0.03 mm/min	+5%	−5%	±2%
3. Test temperature (t)	PG low+10°C	+0.5°C	−0.5°C	±0.1°C
4. Notch depth (ND)	15 mm	+3 mm	−3 mm	±1 mm
5. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen thickness (T)	57 mm	+5 mm	−5 mm	±2 mm
2. Loading rate (LR)	0.5 mm/min	+5%	−5%	±2%
3. Test temperature (t)	25°C	+1°C	−1°C	±0.1°C
4. Notch depth (ND)	25.4, 31.8, 38.1 mm	+3 mm	−3 mm	±1 mm
5. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen thickness (T)	50 mm	+5 mm	−5 mm	±2 mm
2. Loading rate (LR)	50 mm/min	+5%	−5%	±2%
3. Test temperature (t)	25°C	+0.5°C	−0.5°C	±0.1°C
4. Notch depth (ND)	15 mm	+3 mm	−3 mm	±1 mm
5. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen thickness (T)	50 mm	+5 mm	−5 mm	±2 mm
2. Creep time (Crt)	1000 (Sec)	+5%	−5%	±2%
3. Horizontal deformation low range	0.0025	+5%	−5%	±2%
4. Horizontal deformation high range	0.019	+5%	−5%	±2%
5. Loading rate (LR)	50 mm/min	+5%	−5%	±2%
6. Test temperature (t)	10°C	+1°C	−1°C	±0.1°C
7. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen height (T)	38 mm	+1 mm	−1 mm	0.5 mm
2. Opening displacement (OD)	0.635 mm	+2%	−2%	±1%
3. Test temperature (t)	25°C	+1°C	−1°C	±0.1°C
4. Specimen width (W)	76 mm	+3 mm	−3 mm	±2 mm
5. Loading period (second)	10	+1 s	−1 s	±0.1 s
6. Air void (AV)	7%	+0.5%	−0.5%	±0.1%
1. Specimen height (T)	50 mm	+6 mm	−6 mm	±2 mm
2. Test temperature (t)	20°C	+1°C	−1°C	±0.1°C
3. Specimen length (L)	380 mm	+6 mm	−6 mm	±2 mm
4. Specimen width (W)	63 mm	+6 mm	−6 mm	±2 mm
5. Loading frequency (Fr)	10 Hz	+1 Hz	−1 Hz	0.1 Hz
6. Air void (AV)	7%	+0.5%	−0.5%	±0.1%

Table 5 Recommended experiment design for up to seven variables (ASTM E1169-14).

PB Order, Run #	A	B	C	D	E	F	G
1	1	1	1	−1	1	−1	−1
2	−1	1	1	1	−1	1	−1
3	−1	−1	1	1	1	−1	1
4	1	−1	−1	1	1	1	−1
5	−1	1	−1	−1	1	1	1
6	1	−1	1	−1	−1	1	1
7	1	1	−1	1	−1	−1	1
8	−1	−1	−1	−1	−1	−1	−1
Avg +							
Avg −							
Main Effect							

writing, it is estimated that the ruggedness testing of all seven tests will require approximately $285,000 and 12 months to complete. Since the performing laboratories must be proficient in the respective test methods, the cost estimate makes no provision for equipment purchase or testing familiarization.

ILS TO DEFINE PRECISION OF THE SELECTED CRACKING TESTS

Laboratory test results are impacted by many factors, such as operator, test equipment, calibration of the equipment, and environment (temperature, humidity, etc.). These factors vary among laboratories. The variability of the test results obtained by different operators or with different equipment will usually be greater than the variability of between-test results from a single operator with the same equipment. To quantify the variability within a single laboratory and between different laboratories, an ILS is warranted.

The main purpose of performing an ILS is to determine the precision of a test method in terms of its intralaboratory repeatability and interlaboratory reproducibility. Repeatability focuses on the variability between independent test results obtained within a single laboratory in the shortest practical period of time by a single operator with a specific set of test apparatus using test specimens taken at random from a single quantity of homogeneous material prepared for the ILS. Reproducibility deals with the variability between single test results obtained in different laboratories, each of which has applied the test method to test specimens taken at random from a single quantity of homogeneous material prepared for the ILS. The ILS plan summarized below and described in detail in Chapter 3 of the project final report was developed following ASTM E691-14: *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method.*

ASTM E691-14 requires that the ILS begin with eight or more laboratories in order to allow for attrition because the final precision statement of a test method should be developed based on acceptable test results from a minimum of six laboratories. It is highly recommended that any laboratory qualified to run the test routinely be included in the ILS. ASTM E691-14 defines *qualified* as "proper laboratory facilities and testing equipment, competent operators, familiarity with the test method, a reputation for reliable testing work, and sufficient time and interest to do a good job." For each of the selected cracking tests the project identified at least nine specific laboratories in the United States that likely meet the ILS qualifications.

ILS Experimental Design for Each Cracking Test

In the design of an ILS, an acceptable number of specimen replicates for each material must be specified in order to obtain a good estimation of the repeatability (generally the repeatability standard deviation). It is generally sound to limit the number of replicates for each material in each laboratory to a small number. When considering the repeatability of asphalt mixture cracking tests, it is recommended that the cracking tests with a monotonic load (i.e., DCT, SCB, and IDT) use three replicates, and those cracking tests with a repeated load (i.e., OT and BBF) use five replicates. On this basis, an ILS encompassing the seven selected cracking test methods would require a minimum of 1,116 specimens.

Development of Precision Statement for Cracking Tests

By following the procedures in ASTM E691-14, the test data collected in the ILS can be analyzed to produce a precision statement for each of the selected cracking tests. Availability of these precision statements will help advance the tests to the stage of practical application in routine mix design. The precision statement for each cracking test can be presented in the format of Table 6, as recommended by ASTM E691-14. The statement of precision and bias

Table 6 Precision statistics sample (ASTM E691-14).

Material	$\bar{x}$	$s_{\bar{x}}$	s_r	s_R	r	R

NOTE:
$\bar{x}$ = average of cell averages.
$s_{\bar{x}}$ = standard deviation of cell averages.
s_r = repeatability standard deviation.
s_R = reproducibility standard deviation.
r = repeatability limit, $r = 2.8S_r$.
R = reproducibility limit, $R = 2.8S_R$.

for each cracking test should at least include these specified elements:

- A brief description of the ILS procedure.
- Materials (mixtures) used.
- Number of laboratories.
- Number of test results per laboratory per material.
- Data analysis.
- A description about any deviation from test method.
- Precision statistics.

Proposed Schedule and Budget

The ILS should be conducted by one managing contractor with subcontracts or cooperative agreements with other participating laboratories. Preparation of all samples will take place in one laboratory (assumed to be the prime contractor) and the ILS testing will be done in eight laboratories (subcontractors) for each cracking test. At the time of this writing, it is estimated that an ILS of all seven tests will require approximately $972,000 and 18 months to complete. The cost estimate assumes that the participating laboratories will be compensated for their time and effort expended on the ILS.

EXPERIMENTAL DESIGN FOR FIELD VALIDATION OF THE SELECTED CRACKING TESTS

Introduction

The following sections focus on the development of experimental designs for field validation of the seven selected cracking tests and establishment of pass/fail criteria. First, the concept of the D-optimal experiment design is described, followed by a summary of specific experimental designs for thermal, reflection, bottom-up fatigue, and top-down cracking, respectively.

Chapter 4 of the project final report proposes two D-optimal experiment designs (i.e., Design 1 and Design 2) for field validation of tests for each of the four cracking types. Design 1 is the most affordable design that simply focuses on estimating all main effects. Design 2 is a more desirable design that estimates both the main effects and the two-way interaction effects, but with substantially increased requirements for field sections and materials, and thus, increased funds and time to complete the experiment. In the following sections of this digest, discussion is limited to Design 1 for each cracking type experiment.

It is important to emphasize that **the objective of the four experiment designs is to validate the cracking tests** associated with the cracking types, **not to study the mechanisms** of the different types of cracking. The experiment designs are configured to provide a wide range of cracking performance so that the validation may be done on sections likely to crack as well as sections less likely to crack. In this way, the cracking tests can be validated for materials and conditions intended to show poor cracking performance against those that will perform well. This will help reduce the occurrence of false positive results.

D-Optimal Experimental Design

The goal of statistical experimental design is to obtain valid data in order to achieve research objectives as efficiently as possible (e.g., by using as few field test sections as possible for a given degree of precision). Three options can be considered for developing an experimental design: full factorial design (consisting of all possible factor-level combinations), fractional factorial design, and D-optimal design. Using the full factorial design for field test sections is unrealistic for the experiments being considered in this project. The D-optimal designs and the fractional factorial designs are useful when the number of runs in the full factorial design is higher than the study can afford. Both designs choose a subset of runs from a full factorial design. Some fractional factorial designs can still be D-optimal designs although they are chosen through a different mechanism (e.g., the one-half fraction of the 2^k design, where k is the number of factors) rather than optimizing a statistical criterion. There are, however, many cases for which tabulated classical fractional factorial designs do not exist (most of the tabulated classical fractional factorial designs are only available in designs where each factor has two levels). Also, the practical constraints in the experiment may prevent the use of classical tabulated designs. For example, the maximum number of runs or test sections available may be too small to be accommodated by classical designs; in addition, some factor-level combinations may be impossible to run. It needs to be emphasized that in such cases, arbitrarily selecting a subset of the full factorial designs will not lead to a good design. The factor effects estimated by such an arbitrary design are likely to have poor statistical properties and there is no guarantee that all important factor effects can be estimated.

A D-optimal design is a computer-generated statistical experimental design that chooses the best subset of all possible factor-level combinations (e.g., binder/aggregate combinations) within the constraints of the experiment (e.g., limited number of test sections) by maximizing the determinant of the information matrix $X'X$, where X is the design matrix. D-optimal design can address the limitations of the traditional designs while enabling the estimation of all important effects (e.g., main effects and two-way interaction effects) with a considerably smaller number of test sections than required by traditional designs. For example, if there are five factors, each having four levels, the number of factor-level combinations for a full factorial design exceeds 1,000, which is prohibitive. However, only 16 test sections are needed for the D-optimal design to estimate all the main effects. D-optimal designs minimize the standard errors of the parameter estimates for a pre-specified model so that the factor effects of interest can be estimated more precisely than by using any other designs with the same number of test sections. Although more test sections lead to a more precise estimation of the factor effects, the maximum number of test sections that can be afforded will depend on the resources.

An example below illustrates D-optimal design. Suppose that there are three factors of interest, X1, X2, and X3 in the experiment, and researchers are interested in estimating all main effects under a linear model. The levels considered are:

- X1: 2 levels (L1, L2).
- X2: 3 levels (L1, L2, L3).
- X3: 4 levels (L1, L2, L3, L4).

A full factorial design with these three factors consists of 24 factor-level combinations, given in Table 7. D-optimal designs select a subset of runs from Table 7, maximizing the D-efficiency (i.e., minimizing the generalized variance of main effects' estimates) for a pre-specified number of runs.

Assume that due to resource limitations, only up to 20 runs can be afforded, and a smaller number of runs are preferred. The minimum number of runs needed to estimate all three main effects is seven in this case, as listed in Table 8. The minimum number of runs needed to estimate all three main effects and two-way interactions among them is 18, and those are listed in Table 9.

In summary, D-optimal designs can easily accommodate physical constraints in the experiment (such as the maximum number of runs that can be afforded or infeasible factor-level combinations), even in cases where classical designs cannot, and are optimal (in the sense that standard errors of the resulting coefficient estimates of the fitted model are

Table 7 Candidate set of runs for factors X1, X2, X3.

Run	X1	X2	X3
1	L1	L1	L1
2	L1	L1	L2
3	L1	L1	L3
4	L1	L1	L4
5	L1	L2	L1
6	L1	L2	L2
7	L1	L2	L3
8	L1	L2	L4
9	L1	L3	L1
10	L1	L3	L2
11	L1	L3	L3
12	L1	L3	L4
13	L2	L1	L1
14	L2	L1	L2
15	L2	L1	L3
16	L2	L1	L4
17	L2	L2	L1
18	L2	L2	L2
19	L2	L2	L3
20	L2	L2	L4
21	L2	L3	L1
22	L2	L3	L2
23	L2	L3	L3
24	L2	L3	L4

Table 8 D-optimal design with seven runs for main effects.

Run	X1	X2	X3
1	L2	L2	L4
2	L2	L3	L3
3	L1	L2	L2
4	L2	L1	L2
5	L1	L1	L3
6	L1	L3	L4
7	L2	L2	L1

NOTE: Seven is the minimum number of runs required to estimate all three main effects. Because this design is saturated, there are no degrees of freedom for error (i.e., there will be no error term for testing). If testing of parameters in addition to estimation is needed, at least one more run is needed to estimate an error term.

Table 9 D-optimal design with 18 runs for main effects and two-way interactions.

Run	X1	X2	X3
1	L1	L1	L2
2	L1	L3	L3
3	L2	L1	L2
4	L1	L2	L4
5	L1	L3	L1
6	L1	L3	L4
7	L1	L1	L3
8	L2	L1	L4
9	L2	L3	L1
10	L1	L2	L1
11	L1	L2	L3
12	L2	L2	L1
13	L2	L2	L2
14	L1	L3	L2
15	L1	L1	L4
16	L2	L3	L3
17	L1	L2	L2
18	L2	L1	L1

NOTE: Eighteen is the minimum number of runs required to estimate all three main effects and two-way interaction effects among them. Because this design is saturated, there are no degrees of freedom for error (i.e., there will be no error term for testing). If testing of parameters in addition to estimation is needed, at least one more run is needed to estimate an error term.

minimized) among all the designs that are subject to the same constraints. In this research, a series of field experimental designs for validating cracking tests was developed by utilizing D-optimal designs to accommodate resource constraints while ensuring that none of the important factor effects were confounded.

The tests selected for validation will be judged according to their sensitivities to important effects. For instance, if the test can detect differences in mixture properties such as asphalt content, void content, VMA, binder type, etc., it will be considered a good candidate for adoption. Conversely, if the mixture test cannot distinguish between brittle (poor performing) mixtures and ductile (crack resistant) mixtures, it will not be a good candidate. The D-optimal experimental designs will help identify tests that can make these distinctions while accounting for the effects of other parameters that might affect the performance.

General Factors Bearing on Field Test Section Selection

Climate is a key influence on the selection of test sections for each of the four cracking types and is discussed in detail in the relevant sections below.

It is also important to define high traffic and low traffic in order to distinguish between field sections. In this research, the definitions for high and low traffic are based on equivalent single axle loads (ESALs) and are as follows:

- High traffic ≥ 300,000 ESAL/year.
- Low traffic < 300,000 ESAL/year.

For the purposes of this research, field sections, depending on traffic loading, can be categorized into accelerated pavement testing (APT) sections, full-scale experiment test tracks, test roads, and in-service pavements. Table 10 summarizes the features, strengths, and weaknesses of each type of test section. While results may be rapidly obtained and conditions are well defined in APT test sections, they are subject to slow-moving traffic loading and are problematic with respect to aging and moisture effects. Full-scale test tracks allow for realistic loading on full-sized pavement structures, but their service lives only span one to four years, making the aging limited. In contrast, in-service pavements can well consider natural aging and moisture effects but take considerably longer for results and may present difficulties

Table 10 Four types of field experimental test facilities in the United States.

Items	APT	Full-Scale Test Tracks	Full-Scale Test Roads	In-Service Pavements
Examples	FHWA-ALF, Louisiana-ALF, CalTrans-heavy vehicle simulator (HVS), Florida-HVS, Illinois-Accelerated Transportation Loading Assembly (ATLAS), TxDOT-APT	WesTrack NCAT test track	MnROAD	Long-term pavement performance-general pavement studies/ specific pavement studies (LTPP-GPS/ SPS) sections and state DOT sections
Traffic load	Known traffic; well controlled traffic; often overloaded	Known traffic; WesTrack: 4 units of tractor/trailer—triple combinations NCAT track: four fully loaded trucks	Known traffic; real traffic	Unknown traffic (most of time); real traffic; many SPS sections equipped with WIMs
Traffic speed	Slow; around 5–12 mph	Around 40–45 mph	Real traffic and real speed (around 60 mph)	Real traffic and real speed (around 60 mph)
Test period	Several months	1–3 years	4 years	Several years to more than 15 years
Environment	Temperature is often controlled	Natural weather	Natural weather	Natural weather
Aging effect	Artificial aging can be considered, but not natural aging	Impact of short-term aging on performance is considered	Impact of short-/ medium-term aging is considered	Impact of long-term aging is addressed

in determining traffic loading accurately. However, some in-service pavements are equipped with weigh-in-motion (WIM) stations so that well-documented traffic loading is available. All four types of test sections, depending on cracking type, should be considered for validating cracking tests.

APT facilities generally are not appropriate for validating thermal cracking due to the required pavement section length and the secondary effect of traffic loading.

Experimental Design for Validating Thermal Cracking Tests

Thermal Cracking Mechanism and Influential Factors

Thermal cracking is tied to the climatic conditions of either (1) a slow temperature differential developed seasonally along with contraction and expansion cycles in very cold climates or (2) large diurnal temperature differentials in arid climates with a fast temperature differential over a short period of time. Thermal cracking occurs when thermal stress in the material exceeds its tensile strength. Traffic volume often accelerates the number of cracks and the rate of cracking deterioration. While many factors influence thermal cracking development, the key factors and their variation levels proposed for the thermal cracking experimental design are presented in Table 11.

Structure and mixture type should be considered when selecting test locations for thermal cracking. Ideally, pavement sections used for the study will be new and uncracked. Evaluating overlays over cracked subsurfaces creates the potential for mistakenly identifying reflection cracking as thermal cracking. Additionally, the study should evaluate the difference between thick and thin hot-mix asphalt (HMA) sections, as it is thought that thermal stresses are reduced with thicker pavement sections. Presently, the commonly accepted point of divi-

Table 11 Field experimental design factors identified for thermal cracking.

Key Factor	Variation Level
Climate	1) Cold regions—areas with few freeze-thaw (F-T) cycles that enter prolonged cold seasons. 2) Diurnal cycling regions—areas that can experience large daily temperature fluctuations in dry and hot regions. These large diurnal temperature ranges correspond with thermal fatigue.
Mix type	1) DGA mixture with regular PG binder (DGA_Regular PG). 2) DGA mixture with the binder having the same high end PG but one grade lower in terms of PG low end (DGA_PG-Lower). 3) SMA mixtures.
Pavement structure	1) Thick asphalt concrete (AC): > 150 mm (6 inches). 2) Thin AC: ≤ 150 mm (6 inches).
Traffic	1) High: > 300,000 ESAL/year. 2) Low: ≤ 300,000 ESAL/year.

sion between thick and thin is 150 mm (6 inches). For thermal cracking, three different types of mixes are recommended: (a) dense-graded asphalt (DGA) mixture with regular PG binder (DGA_Regular PG), (b) DGA mixture with the binder having the same high end PG but one grade lower in terms of PG low end (DGA_PG-Lower), and (c) stone matrix asphalt (SMA) mixtures.

Field Experimental Design for Validating Thermal Cracking Tests

The objective of the experimental design is to assess the main effects of study factors on thermal cracking development in the field and then to validate the capability of the selected thermal cracking tests for differentiating the performance of these experimental test sections. More test sections lead to a more precise estimation of the factor effects. However, the maximum number of test sections that can be afforded depends on the resources available. The most affordable D-optimal design that focuses on estimating all main effects is shown in Table 12.

Identification and Availability of Field Test Sections for Validating Thermal Cracking Tests

Much of the identification of specific potential test sections revolves around the climate. For thermal cracking, the traditional climate regions within LTPP provide general discrimination points between hot and cold.

In cold regions, low-temperature cracking will occur when the temperature becomes cold enough to induce a contraction in the pavement that overcomes the tensile strength. Locations for thermal cracking in cold regions will be north of 40°N in either the wet or dry LTPP zones, providing test sections that experience a more monotonic decline in temperature and limit any undesired fluctuations that might occur in sections between the freeze line used by LTPP and 40°N. Additionally, the cold climate region will end at 120°W to avoid the insulated area along the Pacific Coast.

Thermal cracking might also occur in the fall or spring in a climate that experiences large daily temperature swings, thereby creating thermal fatigue. Locations throughout the United States that experience the largest daily temperature differential are found in the hot, arid southern portion of the High Plains. This differential is often characterized by temperatures falling 22°C to 28°C (40°F to 50°F) within a single hour.

Table 13 is an extension of Table 12 with possible test sections included. The potential sections

Table 12 D-optimal experimental design for thermal cracking.

Test Section	Climate	Mixture	Structure	Traffic
1	Cold	DGA_Regular PG	Thick AC	High
2	Cold	SMA	Thin AC	High
3	Cold	DGA_PG-Lower	Thin AC	Low
4	Diurnal cycling regions	DGA_PG-Lower	Thick AC	High
5	Diurnal cycling regions	SMA	Thick AC	Low
6	Diurnal cycling regions	DGA_Regular PG	Thin AC	Low

NOTE: Six is the minimum number of test sections required to estimate all four main effects in this case.

Table 13 D-optimal experimental design for thermal cracking with possible test sections.

Test Section	Climate	Mixture	Structure	Traffic	Sections
1	Cold	DGA_Regular PG	Thick AC	High	MnROAD Cell 15, LTPP 18-A901
2	Cold	SMA	Thin AC	High	MnROAD Cell 16, MnROAD Cell 17, MnROAD Cell 18, MnROAD Cell 20, MnROAD Cell 21
3	Cold	DGA_PG-Lower	Thin AC	Low	MnROAD Cell 24 , MnROAD Cell 83, MnROAD Cell 84
4	Diurnal cycling regions	DGA_PG-Lower	Thick AC	High	LTPP 35-0501, 35-0502, 35-0503, 35-0504, 35-0505, 35-A001, 35-A002, 35-A003, 35-A062, 35-0901, 35-0902, 35-0903, 04-A901, 04-A903, 04-B901, 04-B903
5	Diurnal cycling regions	SMA	Thick AC	Low	LTPP 35-0802 (AC-10 binder)
6	Diurnal cycling regions	DGA_Regular PG	Thin AC	Low	LTPP 35-0801 (AC-10 binder)

meet only the climate and structure parameters. For example, MnROAD Cells 16–21 are not SMA, yet they are thin AC sections in a cold climate. Other potential test sections may be available from the current LTPP SPS-10 and the WMA and control HMA pavements studied in NCHRP Project 09-47A, "Field Performance of Warm Mix Technologies."

Proposed Schedule and Budget

The thermal cracking experiment would require a minimum of six test sites, three located in a cold climate and three in a diurnal cycling climate. Given the climate requirements, it is conceivable that all three cold climate sections can be located at one facility (MnROAD), and that the two high-traffic locations of the three diurnal cycling sites might be found in an LTPP SPS section; the other diurnal cycling site could conceivably be an LTPP site close by. At the time of this writing, it is estimated that the field experiment for thermal cracking using *available* MnROAD and LTPP test sections and materials will require approximately $255,000 and 15 months to complete. The budget estimate includes costs for sampling, inventory, transportation, and storage of test materials and the required laboratory testing.

Experimental Design for Validating Reflection Cracking Tests

This section briefly discusses reflection cracking mechanisms and associated influential factors and summarizes an experimental design for field validation of reflection cracking tests.

Reflection Cracking Mechanism and Influential Factors

Reflection cracking arises from the movement of cracked or jointed substrate below the pavement surface layer, which causes cracks to be transmitted through the surface layer. The movements may be horizontal due to the expansion and contraction of lower substrate segments, or they may be vertical as traffic loads pass over the crack and load is transferred from segment to segment, or they may be a combination of both. Major factors influencing reflection cracking and their variation levels are listed in Table 14.

Field Experimental Design for Validating Reflection Cracking Tests

The objective of this field experimental design is to assess the effects of study factors on the development of reflection cracking and validate the capability of the selected reflection cracking tests to distinguish among the experimental test sections on the basis of their reflection cracking performance.

Table 15 presents a D-optimal design for reflection cracking. The number of test sections was kept as small as possible in developing the experimental design while still ensuring that all important main factor effects can be estimated.

Table 14 Field experimental design factors identified for reflection cracking.

Key Factor	Variation Level
Climate	1) Steady state regions—typically moderate to hot weather to avoid thermally induced low-temperature cracking. 2) High-temperature cycling regions—areas that can experience large daily temperature fluctuations. These regions can include areas that experience fluctuations around freezing or those in regions that are dry and hot.
Existing pavement types	1) Cracked AC/granular base. 2) Cracked AC/cement-treated base (CTB). 3) Joint plain concrete pavement (JPCP) with poor load transfer efficiency (LTE) at joints. 4) JPCP with good LTE at joints.
Mix type	1) Superpave DGA mixture. 2) Performance mixture (such as SMA, rubber mixes, etc.). 3) Special crack resistant mixes (such as Strata, Texas CAM, etc.).
Asphalt overlay thickness	1) Thin AC: ≤ 50 mm (2 inches). 2) Medium AC: 50–150 mm (2–6 inches).
Traffic	High: > 300,000 ESAL/year.

Identification and Availability of Field Test Sections for Validating Tests for Reflection Cracking

Temperature cycling occurs in areas that experience high diurnal temperature ranges; areas that experience seasonal temperature variations but not large daily temperature differences between highs and lows are designated as *steady state*. Areas within the United States that experience the highest diurnal temperature ranges, i.e., the highest temperature cycling, are broadly found from west Texas and southern Colorado through Arizona and a large portion of western Nevada. Steady state areas should be located below 40°N to avoid confounding with thermal cracking.

A total of three overlay mixes and two different overlay thicknesses are considered in the experimental matrix. Additional parameters to be considered are existing pavement types and conditions. For reflection cracking, the existing pavement types include both JPCP and cracked asphalt pavements with either granular or CTB. Within the JPCP, the project should evaluate areas that have both high and low LTE. To identify the LTE at joints, initial forensic studies might be required. Falling weight

Table 15 D-optimal experimental design for reflection cracking.

Test Section	Climate	Existing Pavement Type	Mixture	Overlay Thickness	Traffic
1	Steady state	Cracked AC/granular base	DGA	≤ 50 mm (2 inches)	> 300,000 ESAL/year
2	Steady state	Cracked AC/CTB base	Special crack resistant mix	≤ 50 mm (2 inches)	
3	Steady state	JPCP with low LTE	Performance mix	≤ 50 mm (2 inches)	
4	Steady state	JPCP with high LTE	Special crack resistant mix	50–150 mm (2–6 inches)	
5	Temperature cycling	Cracked AC/granular base	Special crack resistant mix	≤ 50 mm (2 inches)	
6	Temperature cycling	Cracked AC/CTB base	Performance mix	50–150 mm (2–6 inches)	
7	Temperature cycling	JPCP with low LTE	DGA	50–150 mm (2–6 inches)	

NOTE: Seven is the minimum number of test sections required to estimate all four main effects in this case.

deflectometer (FWD) deflection data can greatly improve the understanding of LTE within JPCP. In fact, forensics prior to section selection for reflection cracking might be imperative in order to gain a needed understanding of existing pavements. This includes knowledge of the existing asphalt layer thicknesses, if they are not to be milled off. These data can be obtained through ground penetrating radar (GPR). Traffic is not directly considered in the experimental design for reflection cracking, but high-volume traffic is required for all reflection cracking sections.

Table 16 is the experimental design for the reflection cracking test with two possible sections for testing. As can be seen, this research was unable to identify any suitable existing field sections within the temperature cycling area described above. Thus, conduct of this experiment would require construction of several new test sections with concomitant costs over and above those planned for testing and analysis.

Proposed Schedule and Budget

The reflection cracking experiment would require the construction of five to seven new test sections on existing roadways. If suitable sites can be found, the construction costs will probably be approximately $500,000 each for a 500 ft (152 m) test section with 100 ft sampling areas at each end, requiring a total expenditure of $2.5 million to $3.5 million. At the time of this writing, it is estimated that the testing alone will require approximately $260,000 and 15 months to complete. The budget estimate includes costs for sampling, inventory, transportation, storage of test materials, and the required laboratory testing.

Experimental Design for Validating Bottom-Up Fatigue Cracking Tests

This section briefly discusses bottom-up fatigue cracking mechanisms and associated influential factors and summarizes an experimental design for field validation of tests for bottom-up fatigue cracking.

Bottom-Up Fatigue Cracking Mechanism and Influential Factors

Bottom-up fatigue cracking is the result of the application of repetitive traffic loads to a pavement surface, initiating a crack where a critical tensile strain occurs at the bottom of the asphalt layer; continued traffic loading eventually causes these cracks to propagate upward through the entire asphalt layer. Major factors contributing to bottom-up fatigue cracking—most importantly, inadequate structural cross-section and weak or brittle asphalt mixture—and their variation levels are listed in Table 17.

Field Experimental Design for Validating Tests for Bottom-Up Fatigue Cracking

The objective of this field experimental design is to assess the effects of study factors on the development of bottom-up fatigue cracking and validate the capability of the selected tests to distinguish among the experimental test sections on the basis of their bottom-up fatigue cracking performance.

Table 18 presents a D-optimal design for bottom-up fatigue cracking. The number of test sections was kept as small as possible in developing the

Table 16 D-optimal experimental design for reflection cracking with possible test sections.

Test Section	Climate	Existing Pavement Type	Mixture	Overlay Thickness	Sections
1	Steady state	Cracked AC/granular base	DGA	≤ 50 mm (2 inches)	
2	Steady state	Cracked AC/CTB base	Special crack resistant mix	≤ 50 mm (2 inches)	LTPP 01-0103, 01-0105
3	Steady state	JPCP with low LTE	Performance mix	≤ 50 mm (2 inches)	
4	Steady state	JPCP with high LTE	Special crack resistant mix	50–150 mm (2–6 inches)	LTPP 18-A901
5	Temperature cycling	Cracked AC/granular base	Special crack resistant mix	≤ 50 mm (2 inches)	
6	Temperature cycling	Cracked AC/CTB base	Performance mix	50–150 mm (2–6 inches)	
7	Temperature cycling	JPCP with low LTE	DGA	50–150 mm (2–6 inches)	

Table 17 Field experimental design factors identified for bottom-up fatigue cracking.

Key Factor	Variation Level
Climate	1) High temperature and moisture cycling regions—areas that experience fluctuations around freezing 2) All other regions—areas except high temperature/moisture cycling regions
Traffic	1) High: > 300,000 ESAL/year 2) Low: ≤ 300,000 ESAL/year
Mix type	1) Very good cracking resistant mix 2) Good cracking resistant mix 3) Medium cracking resistance mix 4) Poor cracking resistance mix
Pavement structure (AC thickness of all sections must be less than 6 inches)	1) AC with granular base 2) AC with CTB base
Subgrade	1) Poor 2) Good

experimental design while still ensuring that all important main factor effects can be estimated.

Identification and Availability of Field Test Sections for Validating Tests for Bottom-Up Fatigue Cracking Tests

Climate can impact bottom-up fatigue cracking of asphalt pavements in several ways. One is the change in binder stiffness combined with environmental effects that reduces the pavement's ability to resist cracking. High precipitation can cause subgrade and base saturation, weakening the support and facilitating fatigue cracking. Thus, the experimental design requires test sections in wet, warm areas as well as areas where such conditions are uncommon. The Gulf Coast is a promising location, combining high temperatures with moisture cycling. An advantage to this region is the existence of the NCAT test track in Alabama and the LTRC ALF in Louisiana. Another possibility is the FHWA ALF in Virginia, which also experiences high summer temperatures and moisture cycling, though to a lesser degree than Gulf Coast locations.

Table 19 is the experimental design for validation of tests for bottom-up fatigue cracking with the addition of potential test sections. These potential sections were chosen based on climate with traffic, structure, and subgrade considered when feasible. Mixture type was not considered in the initial evaluation of potential sections.

Table 18 D-optimal experimental design for bottom-up fatigue cracking.

Test Section	Climate	Traffic	Mixture	Pavement Structure	Subgrade
1	All others	High	Very good cracking resistance mix	AC/CTB base	Poor
2	High temperature/moisture cycling regions	High	Good cracking resistance mix	AC/granular base	Poor
3	All others	High	Medium cracking resistance mix	AC/granular base	Good
4	High temperature/moisture cycling regions	High	Poor cracking resistance mix	AC/CTB base	Good
5		Low	Very good cracking resistance mix	AC/granular base	Good
6	All others	Low	Good cracking resistant mix	AC/CTB base	Good
7	High temperature/moisture cycling regions	Low	Medium cracking resistance mix	AC/CTB base	Poor
8	All others	Low	Poor cracking resistance mix	AC/granular base	Poor

Table 19 D-optimal experimental design for bottom-up fatigue cracking with possible test sections.

Test Section	Climate	Traffic	Mixture	Pavement Structure	Subgrade	Sections
1	All others	High	Very good cracking resistant mix	AC/CTB base	Poor	
2	High temperature/ moisture cycling regions	High	Good cracking resistant mix	AC/granular base	Poor	LTPP 05-0803, 28-0805, 01-0102
3	All others	High	Medium cracking resistance mix	AC/granular base	Good	LTPP 04-A901, 04-A903, 55-0805, MnROAD-20, MnROAD-21
4	High temperature/ moisture cycling regions L1	High	Poor cracking resistance mix	AC/CTB base	Good	LTPP 01-0103, 01-0104, 01-0105
5		Low	Very good cracking resistant mix	AC/granular base	Good	LTPP 49-0803, 37-0801, 37-0859
6	All others	Low	Good cracking resistant mix	AC/CTB base	Good	
7	High temperature/ moisture cycling regions	Low	Medium cracking resistance mix	AC/CTB base	Poor	
8	All others	Low	Poor cracking resistance mix	AC/granular base	Poor	LTPP 35-0801, 35-0802, 30-0901, 30-0902, 30-0903 MnROAD-24, MnROAD-83, MnROAD-84

Proposed Schedule and Budget

The bottom-up fatigue cracking experiment may not require the construction of test sections due to the number of LTPP and MnROAD sections available for the study. It appears that the only sections that were not found were pavements with CTB in warm, moisture cycling and other climates. At the time of this writing, it is estimated that this experiment will require approximately $295,000 and 15 months to complete. The budget estimate includes costs for sampling, inventory, transportation, and storage of test materials and the required laboratory testing.

Experimental Design for Validating Top-Down Cracking Tests

This section briefly discusses top-down cracking mechanisms and factors before presenting the experimental design for validating top-down cracking tests. The research team also reviewed and documented all potential existing field test sections for validating top-down cracking tests in this section as well.

Top-Down Cracking Mechanism and Influential Factors

The mechanisms for top-down cracking are not as well defined as for the other types of cracking discussed. This is because top-down cracking has often been mistaken for bottom-up fatigue cracking in relatively thick asphalt pavements. Like bottom-up fatigue cracking, top-down cracking is considered to be primarily load related, but unlike bottom-up fatigue cracking, top-down cracking tends to propagate slowly after initiation. The major factors influencing the development of top-down cracking and their variation levels identified at the Cracking Test Workshop are listed in Table 20.

The experimental design includes factors related to mixture type—gradation and AV level—and

Table 20 Field experimental design factors identified for top-down cracking.

Key Factor	Variation Level
Climate	1) Hard freeze, low solar radiation 2) Hard freeze, high solar radiation 3) No freeze, low solar radiation 4) No freeze, high solar radiation
Mix type	1) DGA coarse mixture, high AV 2) DGA coarse mixture, low AV 3) DGA fine mixture, high AV 4) DGA fine mixture, low AV
Traffic	1) High volume (> 300,000 ESAL/year), high speed 2) Low volume (≤ 300,000 ESAL/year), local (low) speed 3) High volume (> 300,000 ESAL/year), low speed
Pavement structure	Thick AC: ≥ 150 mm (6 inch)

the speed of heavy vehicles. Coarse mixtures with higher in-place AVs (i.e., greater than 9 percent) tend to crack faster than those with finer gradations and lower void contents do. Finally, traffic speed and load were identified. Slower heavy vehicle speeds are expected to correlate with a greater tendency to crack due to the slower viscoelastic response of the material to loading.

The experimental design also includes the climatic factors of temperature and solar radiation. Top-down cracking is believed to be strongly related to the stiffness of the surface mix. Higher solar radiation in areas without freezing temperatures ages asphalt on the surface faster than it ages asphalt in areas with lower solar radiation and lower temperatures. There is a decreasing gradient to aging with layer depth, which consequently reduces crack resistance and makes it easier for top-down cracking to occur.

Material factors in the design include material composition, modulus gradient, and fracture and thermal properties of AC mixtures. Prior research found that the initiation of top-down cracking is affected by binder content, aggregate gradation, and binder-aggregate adhesion. High AVs at the pavement surface lead to greater aging and moisture damage, which increase the likelihood of top-down cracking due to a higher surface tensile stress.

The literature also suggests that top-down cracking is caused mainly by non-uniform stress distribution beneath rolling tires, including vertical, longitudinal, and transverse stresses. Thus, the key factor of traffic is split into three variation levels, to account for the fact that certain load positions have a much greater potential effect than others.

The final key factor is pavement structure with one level, as thick pavements are the structures most likely to exhibit top-down cracking. By only including thick pavements, the possibility of encountering bottom-up fatigue cracking while evaluating top-down cracking will be minimized.

Field Experimental Design for Validating Top-Down Cracking Tests

The objective of this field experimental design is to assess the effects of factors on top-down cracking development in the field and validate the capability of the selected top-down cracking tests for differentiating the performance of these experimental test sections. Table 21 presents the D-optimal design for top-down cracking.

Table 21 D-optimal experimental design for top-down cracking.

Test Section	Climate	Traffic	Mixture
1	Hard freeze, high solar	Low volume, low speed	DGA fine, high AV
2	Hard freeze, high solar	High volume, low speed	DGA coarse, high AV
3	Hard freeze, low solar	High volume, high speed	DGA fine, low AV
4	Hard freeze, low solar	High volume, low speed	DGA fine, high AV
5	No freeze, high solar	High volume, high speed	DGA coarse, low AV
6	No freeze, high solar	Low volume, low speed	DGA coarse, high AV
7	No freeze, high solar	High volume, low speed	DGA fine, low AV
8	No freeze, low solar	High volume, high speed	DGA fine, high AV
9	No freeze, low solar	Low volume, low speed	DGA coarse, low AV

Identification and Availability of Field Test Sections for Validating Top-Down Cracking Tests

Much of the challenge in identifying test sections for the top-down cracking experiment revolves around the climate factor, with the need to find suitable test sections in regions that experience high or low solar gain combined with no freeze or hard freeze. The research established the following four climatic regions:

- High solar gain/no freeze: west Texas to the Pacific Ocean below 35°N.
- High solar gain/hard freeze: lower elevations of Colorado, Utah, and Nevada, bounded by 37°N and 40°N, and 100°W and 120°W.
- Low solar gain/no freeze: east Texas, northern portions of Alabama and Georgia, and a narrow strip along the border of North Carolina and South Carolina eastward to the Atlantic Ocean.
- Low solar gain/hard freeze: eastern and Midwestern states bounded by 40°N and 100°W, and portions of North Dakota, Montana, Idaho, and Washington above 45°N and between 100°W and 120°W.

Structure is omitted from the experimental matrix because of the assumption that top-down cracking will likely take place on sections with high overall stiffness. With this in mind, it is assumed that top-down cracking sections will consist of thick asphalt sections [likely 150 mm (6 inches) or thicker] over good subgrades.

Table 22 is a re-creation of Table 21 with potential test sections listed corresponding with climate parameters, while traffic components are included when distinguishable within available data. MnROAD can provide many test sections for top-down cracking. The NCAT test track also provides an excellent source for potential top-down cracking sections in the low solar gain/no freeze climate. Other potential test sections may be available from the current LTPP SPS-10 and from the WMA and control HMA pavements studied in NCHRP Project 09-47A, "Field Performance of Warm Mix Technologies."

Proposed Schedule and Budget

The top-down cracking experiment may not require the construction of test sections due to the number of LTPP, MnROAD, and other sections available for the study. It appears that the only sections that were not found were high-volume, low-speed sections. It is likely that such pavements can be found through a survey of states. At the time of this writing, it is estimated that this experiment will

Table 22 D-optimal experimental design for top-down cracking with possible test sections.

Test Section	Climate	Traffic	Mixture	Sections
1	Hard freeze, high solar	Low volume, low speed	DGA fine, high AV	LTPP 30-0903, 30-0902, 30-0901, 49-0804
2	Hard freeze, high solar	High volume, low speed	DGA coarse, high AV	
3	Hard freeze, low solar	High volume, high speed	DGA fine, low AV	LTPP 18-A901, MnROAD-15
4	Hard freeze, low solar	High volume, low speed	DGA fine, high AV	LTPP 55-0806, 55-0805
5	No freeze, high solar	High volume, high speed	DGA coarse, low AV	LTPP 35-0501, 35-0502, 35-0503, 35-0504, 35-0505, 35-A001, 35-A002, 35-A003, 35-A061, 35-A062, 35-0901, 35-0902, 35-0903, 04-A901, 04-A903, 04-B901, 04-B903
6	No freeze, high solar	Low volume, low speed	DGA coarse, high AV	LTPP 35-0801, 35-0802
7	No freeze, high solar	High volume, low speed	DGA fine, low AV	
8	No freeze, low solar	High volume, high speed	DGA fine, high AV	LTPP 05-0803, 05-0804, 28-0805, 28-0902, 01-0101, 01-0102, 01-0103, 01-0104, 01-0105, 01-0106
9	No freeze, low solar	Low volume, low speed	DGA coarse, low AV	LTPP 37-0802

require approximately $315,000 and 15 months to complete. The budget estimate includes costs for sampling, inventory, transportation, storage of test materials, and the required laboratory testing.

SUMMARY

Asphalt mix designs are becoming more complex due to the increased use of recycled materials, recycling agents, binder additives and modifiers, and WMA technologies. These changes have altered the performance of mixtures both positively and negatively so that volumetric mix design alone is not sufficient for evaluating the potential behavior of asphalt mixtures, especially cracking behavior. Thus, there is an urgent need to establish and implement reliable performance tests that can be used to eliminate brittle mixes or provide input to models that predict asphalt pavement cracking.

The objective of this research was to develop an experimental design for field validation of laboratory tests selected under this study to assess the cracking potential of asphalt mixtures. A three-step process was taken to achieve the objective: (1) identify and select cracking tests, (2) define a plan to refine these selected cracking tests through ruggedness testing and an ILS, and (3) develop an experimental design for validation of the selected cracking tests. This digest first summarized the cracking test selection process and identified the cracking tests selected for field validation. The necessary steps for refining each cracking test method through ruggedness testing and ILS were then presented. Finally, a D-optimal experimental design approach was employed to develop field experimental designs and estimated costs and schedules to complete the experiment for each of four cracking types: thermal, reflection, bottom-up fatigue, and top-down. A summary of the budgets and timelines for the validation studies for the cracking tests are given in Table 23.

Table 23 Summary of budgets and timelines for cracking test ruggedness, ILS, and validation.

Experiment	Time, Months	Cost*, $
Ruggedness Testing	12	285,000
ILS	18	972,000
Thermal Cracking Validation	15	255,000
Reflection Cracking Validation	15	260,000
Bottom-Up Fatigue Cracking Validation	15	295,000
Top-Down Cracking Validation	15	315,000
TOTAL:		2,382,000

*Costs are exclusive of overhead charges, which are widely variable.

Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

ISBN 978-0-309-37560-3

Subscriber Categories: Materials • Pavements

These digests are issued in order to increase awareness of research results emanating from projects in the Cooperative Research Programs (CRP). Persons wanting to pursue the project subject matter in greater depth should contact the CRP Staff, Transportation Research Board, National Academies of Sciences, Engineering, and Medicine, 500 Fifth Street, NW, Washington, DC 20001.