VTRC
Virginia Transportation Research Council

We Bring Innovation to Transportation

Interlaboratory Study for the Indirect Tensile at High Temperature Test and Ideal Rutting Test

https://vtrc.virginia.gov/media/vtrc/vtrc-pdf/vtrc-pdf/25-R14.pdf

ILKER BOZ, Ph.D., P.E.
Senior Research Scientist

STACEY D. DIEFENDERFER, Ph.D., P.E.
Associate Principal Research Scientist

JHONY HABBOUCHE, Ph.D., P.E.
Senior Research Scientist

Virginia Transportation Research Council

AKSEL SEITLLARI, Ph.D., P.E.
Assistant Professor

State University of New York at Canton

Final Report VTRC 25-R14

VIRGINIA TRANSPORTATION RESEARCH COUNCIL
530 Edgemont Road, Charlottesville, VA 22903-2454
vtrc.virginia.gov

Standard Title Page—Report on State Project

Report No.: VTRC 25-R14	Report Date: February 2025	No. Pages: 43	Type Report: Final	Project No.: 121388
			Period Covered: May 2022 – November 2023	Contract No.:

Title:
Interlaboratory Study for the Indirect Tensile at High Temperature Test and Ideal Rutting Test

Author(s):
Ilker Boz, Ph.D., P.E., Stacey D. Diefenderfer, Ph.D., P.E., Jhony Habbouche, Ph.D., P.E., and Aksel Seitllari, Ph.D., P.E.

Performing Organization Name and Address:
Virginia Transportation Research Council
530 Edgemont Road
Charlottesville, VA 22903

Key Words: Rutting, asphalt mixtures, balanced mix design, indirect tensile at high temperature test, IDEAL RT, rapid rutting test, rutting tolerance index, performance criteria, precision estimates.

Sponsoring Agencies' Name and Address:
Virginia Department of Transportation
1401 E. Broad Street
Richmond, VA 23219

Supplementary Notes:

Abstract: The Indirect Tensile at High Temperature (IDT-HT) test and Ideal Rutting (IR) test were recommended for screening rut-susceptible asphalt mixtures in the Balanced Mix Design (BMD) process, based on a research study by the Virginia Transportation Research Council (VTRC). The Virginia Department of Transportation (VDOT) is in the initial stages of implementing the IDT-HT test as part of the BMD initiative for dense-graded surface asphalt mixtures with unmodified asphalt binders. However, for full implementation, additional considerations such as fine-tuning the test procedure and determining precision estimates of the test method are necessary, and this study specifically addressed these aspects for the IDT-HT and IR tests. The work included three tasks: a sensitivity assessment, an interlaboratory study (ILS), and a proficiency study. All tasks were based on evaluation of specimens from four mixtures with varying rutting potentials.

The sensitivity assessment investigated factors such as specimen conditioning environment, loading rate, loading frame, and storage on test results. Findings revealed significant differences in specimen conditioning time between water bath and environmental chamber, and water bath-conditioned specimens consistently showing lower parameter/index values. Additionally, the repeatability characteristics of both tests were not significantly affected by conditioning environment, loading rate, loading frame, or storage. Moreover, loading rate, loading frame, and storage did not significantly affect IDT-HT and IR test results, especially in the context of single-operator precision estimates. Furthermore, preliminary performance criteria were established, requiring minimum strength and RT index values of 100 kPa and 62 for IDT-HT and IR tests, respectively, for specimens conditioned in a water bath at 54.4°C. Based on the outcome of this task, a Virginia Test Method was developed for the IDT-HT test.

The ILS was conducted with 10 and 9 laboratories for the IDT-HT and IR tests, respectively, employing compacted specimens conditioned in a water bath. Some laboratories were provided with additional specimens for testing after environmental chamber conditioning to evaluate further the effect of the conditioning environments on test results. The precision estimates and statements for both tests were developed. The results also confirmed the sensitivity assessment's conclusion that different conditioning environments notably influence test results.

Proficiency testing was conducted for the IDT-HT test only, involving 31 VDOT and contractor laboratories in Virginia, and employed compacted specimens from a single mixture. The results revealed that 93.5% of the laboratories demonstrated satisfactory performance, indicating proficiency in the IDT-HT test.

The study recommends that VDOT should adopt (I) the developed Virginia Test Method for the IDT-HT test, (II) the developed precision estimates and statements for the IDT-HT test, and (III) the initial minimum strength criterion of 100 kPa for IDT-HT testing of dense-graded surface asphalt mixtures with unmodified asphalt binders in accordance with the developed Virginia Test Method.

FINAL REPORT

INTERLABORATORY STUDY FOR THE INDIRECT TENSILE AT HIGH TEMPERATURE TEST AND IDEAL RUTTING TEST

Ilker Boz, Ph.D., P.E.
Senior Research Scientist
Virginia Transportation Research Council

Stacey D. Diefenderfer, Ph.D., P.E.
Associate Principal Research Scientist
Virginia Transportation Research Council

Jhony Habbouche, Ph.D., P.E.
Senior Research Scientist
Virginia Transportation Research Council

Aksel Seitllari, Ph.D., P.E.
Assistant Professor
State University of New York at Canton

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

Charlottesville, Virginia

February 2025
VTRC 25-R14

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation. Any inclusion of manufacturer names, trade names, or trademarks is for identification purposes only and is not to be considered an endorsement.

ABSTRACT

The Indirect Tensile at High Temperature (IDT-HT) test and Ideal Rutting (IR) test were recommended for screening rut-susceptible asphalt mixtures in the Balanced Mix Design (BMD) process, based on a research study by the Virginia Transportation Research Council (VTRC). The Virginia Department of Transportation (VDOT) is in the initial stages of implementing the IDT-HT test as part of the BMD initiative for dense-graded surface asphalt mixtures with unmodified asphalt binders. However, for full implementation, additional considerations such as fine-tuning the test procedure and determining precision estimates of the test method are necessary, and this study specifically addressed these aspects for the IDT-HT and IR tests. The work included three tasks: a sensitivity assessment, an interlaboratory study (ILS), and a proficiency study. All tasks were based on evaluation of specimens from four mixtures with varying rutting potentials.

The sensitivity assessment investigated factors such as specimen conditioning environment, loading rate, loading frame, and storage on test results. Findings revealed significant differences in specimen conditioning time between water bath and environmental chamber, and water bath-conditioned specimens consistently showing lower parameter/index values. Additionally, the repeatability characteristics of both tests were not significantly affected by conditioning environment, loading rate, loading frame, or storage. Moreover, loading rate, loading frame, and storage did not significantly affect IDT-HT and IR test results, especially in the context of single-operator precision estimates. Furthermore, preliminary performance criteria were established, requiring minimum strength and RT index values of 100 kPa and 62 for IDT-HT and IR tests, respectively, for specimens conditioned in a water bath at 54.4°C. Based on the outcome of this task, a Virginia Test Method was developed for the IDT-HT test.

The ILS was conducted with 10 and 9 laboratories for the IDT-HT and IR tests, respectively, employing compacted specimens conditioned in a water bath. Some laboratories were provided with additional specimens for testing after environmental chamber conditioning to evaluate further the effect of the conditioning environments on test results. The precision estimates and statements for both tests were developed. The results also confirmed the sensitivity assessment’s conclusion that different conditioning environments notably influence test results.

Proficiency testing was conducted for the IDT-HT test only, involving 31 VDOT and contractor laboratories in Virginia, and employed compacted specimens from a single mixture. The results revealed that 93.5% of the laboratories demonstrated satisfactory performance, indicating proficiency in the IDT-HT test.

The study recommends that VDOT should adopt (I) the developed Virginia Test Method for the IDT-HT test, (II) the developed precision estimates and statements for the IDT-HT test, and (III) the initial minimum strength criterion of 100 kPa for IDT-HT testing of dense-graded surface asphalt mixtures with unmodified asphalt binders in accordance with the developed Virginia Test Method.

FINAL REPORT

INTERLABORATORY STUDY FOR THE INDIRECT TENSILE AT HIGH TEMPERATURE TEST AND IDEAL RUTTING TEST

Ilker Boz, Ph.D., P.E.
Senior Research Scientist
Virginia Transportation Research Council

Stacey D. Diefenderfer, Ph.D., P.E.
Associate Principal Research Scientist
Virginia Transportation Research Council

Jhony Habbouche, Ph.D., P.E.
Senior Research Scientist
Virginia Transportation Research Council

Aksel Seitllari, Ph.D., P.E.
Assistant Professor
State University of New York at Canton

INTRODUCTION

The Virginia Transportation Research Council (VTRC) conducted a research study aimed at exploring the use of monotonic tests (i.e., static loading tests) as screening tools for assessing the rutting potential of asphalt surface mixtures with A and D designations (Boz et al., 2023a). The A and D designations are applied to dense-graded surface asphalt mixtures with unmodified asphalt binders subject to 0 to 3 million equivalent single axle loads and 3 to 10 million equivalent single axle loads, respectively. In the study, various monotonic tests were evaluated in comparison with the asphalt pavement analyzer (APA) test, which is a specified rutting test in the balanced mix design (BMD) special provision of the Virginia Department of Transportation (VDOT). The results indicated that there was a high correlation between the two specific monotonic tests, namely indirect tensile at a high temperature (IDT-HT) and ideal rutting (IR) tests, and APA test. Additionally, the two monotonic tests provided either similar or better performance evaluation characteristics of asphalt mixtures than the APA test in terms of test repeatability, sensitivity to changes in asphalt mixture composition, and performance ranking and discrimination potential of asphalt mixtures. Moreover, the results showed that the overall rutting potential of the asphalt mixtures quantified by the monotonic tests and the APA test agreed with that of the fundamental rutting tests (Boz et al., 2023a and 2023b).

From the logistical point of view, the cost and availability of APA equipment in asphalt laboratories remain a concern. Loading frames used for the IDT-HT and IR tests are less costly to purchase and maintain than those for APA. In fact, VDOT district asphalt laboratories and contractors can use their existing loading frames, which are being used for other tests such as the indirect tensile cracking test (IDT-CT) at intermediate temperatures (ASTM D8225, 2019), to perform these monotonic tests. The time required to run APA tests during the production of asphalt mixtures also remains a significant challenge in the BMD process. Given the research findings and the fact that IDT-HT and IR tests are simpler and quicker than the APA test, the researchers concluded that the IDT-HT and IR tests can be used as an alternative to the APA test. Consequently, the researchers recommended implementation within VDOT's BMD specification for screening rut-susceptible asphalt mixtures (Boz et al., 2023a). This study also recommended specific performance criteria for the IDT-HT and IR tests, including a minimum average strength requirement of 133 kPa and a minimum average rutting tolerance (RT) index of 72. These performance criteria were established based on testing of compacted specimens from reheat plant-produced asphalt mixtures conditioned in an environmental chamber at 54.4°C. The researchers based its selection of test conditions—loading rate, specimen dimensions, and air void content—on the requirements of the IDT-CT for convenience during specimen preparation and the testing process of asphalt mixtures.

Furthermore, the researchers recommended implementation of these tests as part of the BMD process, using a hierarchal approach. In line with this recommendation, VDOT's Materials Division included the IDT-HT test and its associated criterion in VDOT's 2023 BMD provisional specification for the mix design stage to be reported for informational purposes only. The IDT-HT test is set to be included in VDOT's 2024 BMD specification, requirement for mix design, and for reporting purposes only during production.

As VDOT takes the initial steps toward full implementation of the test method, the agency must consider some additional issues. For example, VDOT must conduct a systematic evaluation to determine how variations in testing procedures or conditions can affect the results of a test method (Boz et al., 2022). Additionally, once the testing procedures with acceptable tolerance limits are established, precision estimates of a test method must be determined. These considerations are essential for establishing robust quality control and acceptance practices to ensure accuracy, consistency, and reliability of test methods, as well as the properties of materials being tested. Recognizing the necessity of considering those issues, the research study also recommended fine-tuning the testing protocols and determining the precision estimates for the monotonic tests (Boz et al., 2023a).

PURPOSE AND SCOPE

The purpose of this study was to determine precision estimates and develop corresponding precision statements for the rutting indices calculated from the IDT-HT and IR tests through performing an interlaboratory study (ILS). The ILS included laboratories within and outside of Virginia. Prior to performing the ILS, sensitivity assessment of the tests was conducted to fine-tune the testing procedures for these two tests. This effort focused on investigating the effects of several factors such as specimen conditioning environment, loading rate, loading frame (device types), and storage (lag and dwell times) on the test results. Moreover, following the ILS effort, a proficiency study on the IDT-HT test was conducted and involved VDOT and contractor laboratories in Virginia. All tasks undertaken as part of this study included the evaluation of specimens fabricated from four mixtures with different rutting potentials.

METHODS

Asphalt Mixtures

This study involved four dense-graded surface asphalt mixtures (designated as Mixture A through D) with each designed, produced, and compacted by an independent testing laboratory. The design of these mixtures was deliberately varied to encompass a spectrum of rutting potential, ranging from highly rut-resistant to rut-susceptible compositions. For this purpose, the study targeted specific APA rut depth brackets, which included: 0 to 4 mm for one mixture, 4 to 8 mm for two mixtures, and above 8 mm for one mixture, with the latter value representing VDOT's defined APA rutting performance threshold. Corresponding volumetric and gradation properties of these mixtures were measured in accordance with VDOT specifications. Additionally, in accordance with VDOT's BMD specification, these mixtures underwent the Cantabro test and the IDT-CT tests.

IDT-HT Test

The IDT-HT test involves subjecting cylindrical specimens to a constant rate of axial displacement along their diametrical plane. These specimens are preconditioned at a relatively high temperature for a specified duration prior to the testing. In accordance with the recommendations from a prior study (Boz et al., 2023a), the test is conducted at a temperature of 54.4 ± 0.5°C, using specimens measuring 150 ± 2 mm in diameter and 62 ± 1 mm in height, which are compacted to an air-void content of 7 ± 0.5%. The testing process includes the use of three replicate specimens conditioned in an environmental chamber, and these specimens are subjected to a loading rate of 50 ± 2 mm/min.

The IDT-HT test conditions previously described, except for the test temperature, were selected to match the test condition requirements for the IDT-CT, which currently are part of the BMD effort in Virginia. Once the testing is completed, the rutting potential of asphalt mixtures typically is evaluated by calculating the indirect tensile strength of the specimens. The strength is calculated as a function of the maximum load obtained from the test and the dimensions of the specimens, as shown in Equation 1. The required minimum average strength for specimens conditioned in an environmental chamber that indicates acceptable rutting resistance for surface asphalt mixtures with a "D" designation is 133 kPa (Boz et al., 2023a).

$$S_t = \frac{2000 \times P_{max}}{\pi \times t \times D} \times 10^3 \qquad \text{[Eq. 1]}$$

where

S_t = indirect tensile strength, kPa
P_{max} = maximum load, N
t = specimen thickness, mm
D = specimen diameter, mm

Ideal Rutting Test

The IR test follows a similar protocol as the IDT-HT test except that a shear fixture replaces an IDT fixture. A test specimen is situated on a u-shape “shear” fixture following conditioning in an environmental chamber at a temperature of 54.4 ± 0.5°C. Then, a constant rate of axial displacement of 50 ± 2 mm/min is applied on the diametrical plane of cylindrical specimens. Like the IDT-HT test, the IR test is also performed on three-replicate specimens measuring 150 ± 2 mm in diameter and 62 ± 1 mm in height, which are compacted to an air void content of 7 ± 0.5%. The rutting potential of asphalt mixtures from this test is quantified in terms of an RT index. Equation 2 shows the index calculation. A minimum average RT index of 72 for specimens conditioned in an environmental chamber indicates acceptable rutting resistance for surface asphalt mixtures with a “D” designation (Boz et al., 2023a).

$$RT_{index} = 6.618 \times 0.356 \times \frac{P_{max}}{t \times w} \times 10^{-5} \quad \text{[Eq. 2]}$$

where

RT_{index} = rutting tolerance index,
P_{max} = maximum load, N
T = specimen thickness, m
W = width of upper loading strip, 0.0191 m.

Sensitivity Assessment

The sensitivity assessment of the tests was undertaken before the ILS to ensure that the test methods were optimized and standardized before subjecting them to testing across multiple laboratories. This step aimed to minimize unnecessary variability, reduce the risk of method-related variations, and also promote reliable and consistent results in the ILS, ultimately contributing to the development of robust quality assurance practices for these tests. The sensitivity study considered factors expected to influence the test results and variables that can vary in real-world conditions. These considerations were based on factors gleaned from previous studies (Boz et al., 2021; Diefenderfer et al., 2023; Habbouche et al., 2021; 2022; and 2023) and the informed judgment of the researchers. However, before assessing these factors, the researchers conducted experiments to determine the time needed to condition test specimens in a water bath and an environmental chamber before testing. The times for both specimen conditioning modes were established through the temperature monitoring of dummy specimens. This effort focused specifically on these factors, which must be addressed in the following order:

1. *Conditioning Environment.* The initial study conducted by the authors involved conditioning compacted specimens at a temperature of 54.4°C in an environmental chamber before testing (Boz et al., 2023a). However, conditioning test specimens in a water bath at the test temperature proved to be a more convenient option for many laboratories. Conditioning in a water bath eliminates the need to invest in costly environmental chambers, as many laboratories already have multiple water baths suitable for conditioning IDT-HT or IR test specimens. Additionally, conditioning specimens in a water bath significantly reduces conditioning time compared with environmental

chamber conditioning, making it highly practical for production testing. A series of experiments were conducted to investigate the potential impact of the specimen conditioning environment on test results. To this end, specimens from the four mixtures underwent IDT-HT and IR tests after being conditioned in a water bath and an environmental chamber at 54.4°C. These experiments were performed on three replicate specimens at a loading rate of 50 mm/min with a servo-hydraulic machine capable of precisely maintaining the designated loading rate.

1. *Loading Rate.* Based on the ILS conducted to determine the precision estimates of the IDT-CT, researchers observed that asphalt laboratories in Virginia utilize a variety of loading frames or device types, including servo-hydraulic and screw-drive systems (Diefenderfer et al., 2023; Habbouche et al., 2021; 2022; and 2023). Additionally, some screw-drive loading frames failed to satisfy the specified loading rate of 50 ± 2 mm/min required for the IDT-CT, which is also the loading rate specified for the IDT-HT and IR tests (Boz et al., 2023a). Given the potential significance of loading rate on the test results for asphalt mixtures tested at the proposed temperature of 54.4°C, an experiment was conducted to investigate this impact across five different loading rates: 46, 48, 50, 52, and 54 mm/min. The selected range of loading rates was determined based on data obtained from the loading frames used during the ILS for the IDT-CT. This part of the study involved using two different mixtures (Mixtures B and C) to accomplish the objective. These experiments were performed on three replicate specimens using a servo-hydraulic machine capable of precisely maintaining the designated loading rates.

2. *Loading Frames.* As mentioned in 1, loading rate, asphalt laboratories in Virginia use a variety of loading frames made by a variety of manufacturers. Therefore, a need exists to investigate the potential impact of specific loading frames on the test results. To maintain robust quality assurance practices, the choice of test equipment must not affect test results. Using 54.4°C, the research team investigated three commonly available loading frames from various manufacturers and three different mixtures (Mixtures A, B, and D) each with three replicate specimens for this investigation to accomplish the objective.

3. *Lag and Dwell Times*. It is possible that compacted specimens and/or loose mixture samples can experience "storage" before being processed for testing, especially during a busy construction season. Additionally, test specimens prepared for any ILSs are inherently subjected to "storage" during delivery time to participants laboratories. To investigate the impact of various forms of storage on the test results, researchers conducted a series of experiments.

   (I) The first experiment involved studying the impact of "lag time" on the test results. Lag time refers to the period of delay, or waiting time, that occurs between sampling loose mixture and later reheating the loose mixture for compaction and testing. Simulated lag time ensued in this study by preparing samples of loose mixtures on the day of production (referred to as the non-reheat case, or Day 0). The loose mixtures were then set aside at a room temperature for reheating, compaction, and testing at 1, 3, 7, and 14 days after production. Reheating of loose mixtures was considered complete as soon as the mixtures reached compaction temperature and specimens were compacted immediately. The experiment also involved preparing, compacting, and testing of production mixtures on the day of production (i.e., Day 0 specimens). This test phase also provided an opportunity to evaluate the impact of reheating on the test results. Three replicate specimens were prepared from two mixtures (Mixtures B and D) and evaluated at 54.4°C using a servo-hydraulic machine capable of precisely maintaining the designated loading rate of 50 ± 2 mm/min.

(II) The second experiment involved studying the impact of "dwell time" on the test results. Dwell time refers to the period of delay or waiting time that occurs between specimen compaction and testing. Simulation of dwell time in this study occurred by compacting specimens from loose mixtures on Day 0 and setting them aside at room temperature for testing at days 0, 1, 3, 7, 14, 28, and 40 days after the compaction. Three replicate specimens were prepared from two mixtures (Mixtures B and D) and tested at 54.4°C using a servo-hydraulic machine capable of precisely maintaining the designated loading rate of 50 ± 2 mm/min.

Each loose mixture was conditioned for 2 hours at a compaction temperature before compaction. The specimens were compacted to a diameter of 150 ± 2 mm and a height of 62 ± 1 mm with a target air void content of 7 ± 0.5%. VTRC conducted all specimen tests except the lag and dwell time experiments; an independent laboratory performed those tests. Except for the water bath conditioning in the conditioning environment experiments, the specimens were conditioned in an environmental chamber for 3 hours and 30 minutes before testing. In all cases, the IDT-HT and IR tests times were less than 2 minutes after the test specimen was removed from the conditioning environment.

Interlaboratory Study

The ILS was executed in accordance with ASTM C802, *Standard Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials* (ASTM C802, 2014). This practice involves three fundamental steps: planning the ILS, overseeing the testing phase of the study, and analyzing the test results. For this effort, the ILS included the following steps:

- Identified for this study were nine laboratories, each with substantial experience in asphalt mixture testing. The study encompassed a VDOT district laboratory, the VTRC laboratory, contractor, university, and independent testing laboratories. Each participating laboratory received multiple sets of compacted specimens fabricated from the mixtures designed and evaluated in this study. These sets included five replicate test specimens from each mixture for IDT-HT testing following water bath conditioning. The increase in replicate specimens from three to five in this ILS aimed to provide an opportunity to develop precision estimates for a variety of specimen replicates. Additionally, IR testing included three replicate test specimens from each mixture following water bath conditioning. The decision to conduct the ILS using water bath as a conditioning environment was made after consulting the technical review panel of this study.

To evaluate the impact of conditioning environments on the test results further, two sets of three replicate test specimens from one of the four mixtures were provided to each laboratory for IDT-HT and IR testing, following conditioning in an environmental chamber. Eight laboratories participated in this task, with each mixture tested by two laboratories following environmental conditioning, in addition to the testing following water bath conditioning. Also supplied to the laboratories were two dummy (practice) test specimens to assist with training and protocol familiarization. After expressing interest, another state agency laboratory joined the ILS for the IDT-HT testing. That agency's test specimens had received only water bath conditioning.

- All participant laboratories received testing instructions and guidelines. The testing instructions included findings from the sensitivity study.

- Before testing began and as requested by the research team, the participating laboratories provided their equipment-collected raw data files. These data files underwent a quality check as a preliminary step before further analysis. Each raw data file had to contain recorded time, load, and displacement measurements obtained from the testing software. The data quality assessments occurred in general accordance with the guidelines outlined in a previous study (Habbouche et al., 2021). This assessment involved an examination of the load versus displacement curves and displacement versus time curves for each tested specimen.

- Statistical analyses of the collected data, including the data consistency assessment and determination of precision estimates, complied with ASTM C802.

Once the precision estimates were determined, precision statements were formulated in accordance with ASTM C670, *Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials* (ASTM C670, 2015).

Proficiency Testing

The proficiency testing included VDOT and contractor laboratories in Virginia, adhering to the same guidelines outlined for the ILS. Limited to the IDT-HT test, this effort focused on a single mixture, Mixture C and involved just one set of three-replicate specimens subjected to conditioning in a water bath. The primary objectives of this effort were to familiarize laboratories with the IDT-HT test, provide an opportunity to assess laboratory proficiency with the test, and identify areas for potential improvement.

A total of 31 laboratories participated in the proficiency testing. the procedure for testing collected data first assessed for quality and then analyzed to evaluate the accuracy of each laboratory's result with respect to the average result obtained from all laboratories (i.e., the interlaboratory average). This interlaboratory average served as the reference point or "true value" of the mixture. To quantify the deviation of each laboratory's result from this interlaboratory average, a critical statistical measure, known as the "z-score," was employed. The z-score provides a statistical quantification of how many standard deviations a specific laboratory's result deviates from the interlaboratory average. The z-score is calculated using Equation 3.

$$z - score = \frac{L_i - L_a}{SD} \quad \text{[Eq. 3]}$$

where

L_i = average test result of a given laboratory.
L_a = interlaboratory average of the proficiency testing.
SD = interlaboratory standard deviation of the proficiency testing.

A negative z-score implies that the laboratory's average is below the interlaboratory average, and a positive z-score indicates the opposite. An absolute value of z-scores was employed to categorize the performance of laboratories based on their z-score values. The z-score ratings were as follows:

- If z-score was less than or equal to 1, the rating was set to 5.
- If z-score was greater than 1 and less than or equal to 1.5, the rating was set to 4.
- If z-score was greater than 1.5 and less than or equal to 2, the rating was set to 3.
- If z-score was greater than 2 and less than or equal to 2.5, the rating was set to 2.
- If z-score was greater than 2.5 and less than or equal to 3, the rating was set to 1.
- If z-score was greater than 3, the rating was set to 0.

Any rating lower than three is considered a low rating as per the AASHTO Accreditation program.

The collected data as part of the proficiency study provided an opportunity to assess the validity of using the precision estimates of the IDT-HT testing based on five-replicate specimens for the IDT-HT testing based on three-replicate specimens. The "precision estimates" of the IDT-HT test using Mixture C with three replicates from the proficiency testing were calculated and compared with the precision estimates of the IDT-HT test based on five-replicate specimens from the ILS.

RESULTS AND DISCUSSION

Volumetric Properties and Gradations

Tables 1 and 2, show the aggregate gradations and volumetric properties, respectively, of the mixtures used in this study. Table 2 also presents the performance properties of the mixtures, which were determined in accordance with the test methods outlined in VDOT's BMD specification. These mixtures were designed such that the APA rut depths determined for each mixture were spread out to ensure a wider applicability of the study. Table 2 shows the mixtures had an APA rut depth range of 3.7 to 8.1 mm, with Mixture D exceeding VDOT's maximum APA rut depth requirement of 8 mm.

Table 1. Aggregate Gradations for the Mixtures

Sieve Size	Mixture A	Mixture B	Mixture C	Mixture D
	Percent Passing			
¾ in (19.0 mm)	100.0	100.0	100.0	100.0
½ in (12.5 mm)	86.3	100.0	97.9	100.0
3/8 in (9.5 mm)	76.1	94.8	88.8	94.0
No. 4 (4.75 mm)	52.9	56.7	66.0	65.9
No. 8 (2.36 mm)	29.6	35.7	39.7	44.5
No. 16 (1.18 mm)	14.5	22.7	26.3	28.8
No. 30 (600 µm)	9.6	14.1	18.5	16.9
No. 50 (300 µm)	7.8	7.8	13.4	8.0
No. 100 (150 µm)	6.8	5.2	9.3	4.4
No. 200 (75 µm)	6.1	4.3	6.2	3.4

Table 2. Volumetric and Performance Properties for the Mixtures

Mixture ID	**Mixture A**	**Mixture B**	**Mixture C**	**Mixture D**
RAP Content, %	30	30	0	0
Asphalt Binder	PG 76-22	PG 64-22	PG 64-22	PG 64-22
Volumetric Property				
NMAS, mm	12.5	9.5	12.5	9.5
Asphalt Binder Content, %	4.8	5.3	5.6	6.5
Rice SG (G_{mm})	2.525	2.498	2.721	2.564
Aggregate Bulk SG (G_{sb})	2.702	2.688	2.941	2.803
VTM, %	4.4	3.5	4.3	6.1
VMA, %	15.0	15.1	16.4	19.7
VFA, %	70.7	76.8	73.8	69.0
FA Ratio	1.27	0.81	1.11	0.52
Performance Property				
Cantabro Mass Loss at 25°C, %	5.4	5.7	2.8	5.8
CT Index at 25°C	32.7	79.0	118.8	148.3
APA Rut Depth at 64°C, mm	3.7	4.8	6.1	8.1

RAP = reclaimed asphalt pavement; PG = performance grade; NMAS = nominal maximum aggregate size; SG = specific gravity; VTM = voids in total mixture; VMA = voids in mineral aggregate; VFA = voids filled with asphalt; FA = fines to asphalt ratio; CT = cracking tolerance; APA = asphalt pavement analyzer.

Sensitivity Assessment

Conditioning Times for Specimen Conditioning Methods

Temperature monitoring of dummy specimens was conducted to determine the time needed to condition test specimens in each conditioning environment: water bath and environmental chamber. A single dummy specimen compacted to a 150 mm diameter and 62 mm height with an air-void content of 7% was placed into each conditioning environment. Each specimen was equipped with three thermocouples: one thermocouple was inserted in the middle center of the specimens, and the other two thermocouples were affixed to the center of the specimens' circular surfaces. An additional thermocouple was used to monitor the temperature of the conditioning environment itself. A data logger with a sampling rate of 12 data points per minute gathered the temperature. Worth noting is that the conditioning environments were already at a test temperature of 54.4°C prior to placing the dummy specimens in them. Additionally, the dummy specimens were conditioned in their natural state, that is, they were not covered with any materials such as plastic bags. Furthermore, before being placed in a conditioning environment, the dummy specimens were room temperature, ranging between 20°C and 25°C. Researchers conducted each experiment on separate days.

Figure 1 illustrates the results of the temperature monitoring for the dummy specimens for each conditioning environment, depicting only the data acquired from the thermocouple inserted in the middle center of the specimens to ensure clarity. As Figure 1 shows, a significant difference is evident in the readings between the two conditioning environments for the time required for a specimen to reach the target test temperature. The specimen conditioned in the water bath required a minimum of 0.8 hours (47.5 min) to reach the test temperature of 54.4°C, whereas the specimen conditioned in the environmental chamber required at least 2.9 hours (175 min). Based on this observation and added factors of safety, the specimens analyzed in this study underwent conditioning for 1 hour ± 10 min if a water bath was specified for conditioning or 3 hours and 30 min ± 10 min if an environmental chamber was specified for conditioning.

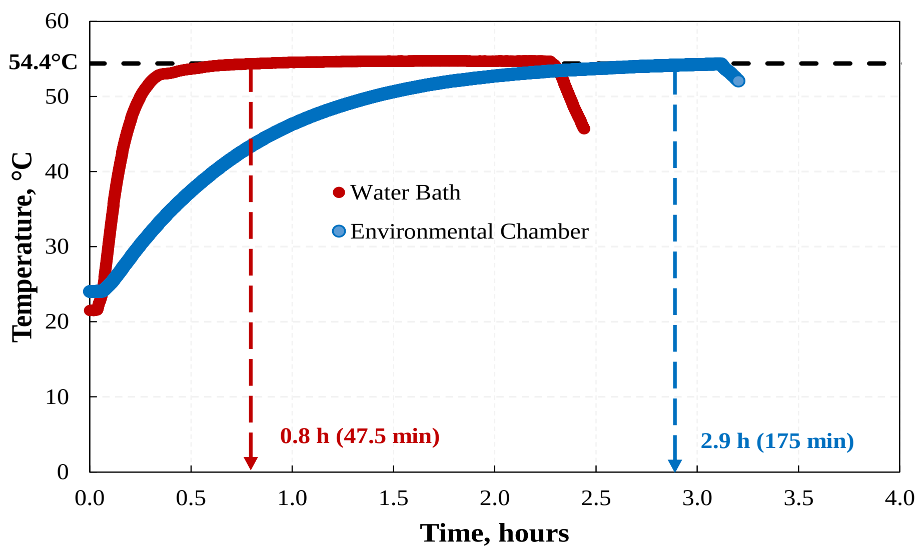

Figure 1. Conditioning Times to Reach 54.4°C in Water Bath and Environmental Chamber

Impact of Specimen Conditioning Method on the Test Results

Figure 2 presents the average strength values obtained from IDT-HT testing of the mixtures conditioned in a water bath and environmental chamber, denoted as "wet" and "dry," respectively. As shown, the mixtures conditioned in the water bath consistently exhibited lower strength values compared with those conditioned in the environmental chamber, with an approximate 25% decrease in strength. To assess the statistical significance of this observed difference, an analysis of variance (ANOVA) was conducted at a 95% confidence interval. The response variable for this analysis was strength with "mixture type" and "conditioning environment" used as factors in the model in addition to an interaction term, "mixture type*conditioning environment." However, prior to performing ANOVA, researchers verified the normality and equal variance assumptions and found results confirming the normal distribution and equal variance of the data. The equal variance finding indicated that the conditioning environment did not have a statistically significant impact on the repeatability characteristics of the IDT-HT test.

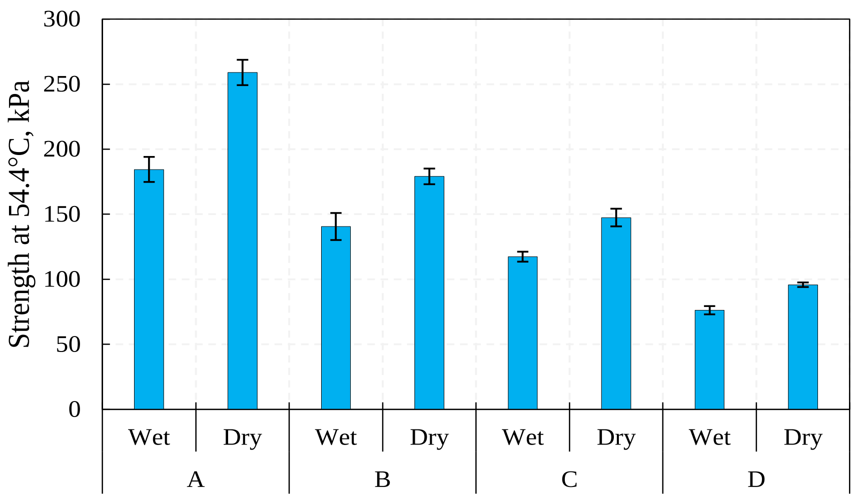

Figure 2. IDT-HT Test Results for the Mixtures Conditioned in the Water Bath (Wet) and Environmental Chamber (Dry). I-bars show ±1 standard deviation; IDT-HT = indirect tensile at a high temperature.

Table 3 presents the ANOVA statistics at a 95% confidence interval for the response variable (i.e., strength). The analysis indicated statistically significant variation in the strength parameter from the IDT-HT test based on interaction between the mixture type and conditioning environment.

Table 3. Summary of ANOVA Statistics for IDT-HT Test Results

Factor	Strength	
	DF	**p-value**
Mixture Type	3	0.000
Conditioning Environment	1	0.000
Mixture Type*Conditioning Environment	3	0.003

ANOVA = analysis of variance; IDT-HT = indirect tensile at high temperature; DF = degrees of freedom.

As shown in Table 4, pairwise comparisons using the Tukey method at a 95% confidence interval showed that, except for Mixture D, the mixtures exhibited statistically different test results when subjected to different environments (i.e., water bath or environmental chamber) for conditioning specimens prior to testing.

Table 4. Tukey Pairwise Statistical Comparisons for IDT-HT Results

Mixture Type*Conditioning Environment	Grouping
A*Dry	a
A*Wet	b
B*Dry	b
B*Wet	c/d
C*Dry	c
C*Wet	d/e
D*Dry	e/f
D*Wet	f

Interaction factors sharing the same letter in the grouping column are statistically similar. IDT-HT = indirect tensile at high temperature; Dry = environmental chamber; Wet = water bath.

The researchers initially attributed this observation to the potential impact of water at a relatively high temperature on weakening adhesive and/or cohesive bonds within the mixture medium. However, visual inspection of the cracked surfaces of the mixtures after the IDT-HT tests did not reveal the presence of adhesive damage between the aggregates and binders of the mixtures. To further evaluate the hypothesis, the authors obtained the plant-produced mixture that incorporated an antistrip agent and tested a set of specimens after subjecting them to both conditioning environments because the four mixtures in this study were laboratory prepared and compacted that did not incorporate any antistrip agents. The same observation was obtained from the plant-produced mixture. The average strength value obtained from the specimens conditioned in the water bath was lower than that of the specimens conditioned in the environmental chamber, resulting in no evidence of adhesive damage in the specimens. The drop in the strength values might have originated from potential damage within the binder medium, i.e., cohesive damage. However, anecdotal evidence of these mixtures and binders suggests no cohesive problems in the field. This observation on the test results warrants further investigation to determine the underlying mechanisms. Future research should focus on systematically evaluating the role of water-induced conditioning on binder properties and its interaction with the mixture, as well as exploring other potential factors that may contribute to the observed behavior.

The IDT-HT test results of the mixtures conditioned in both environments were compared. Figure 3 shows the correlation between the IDT-HT test results of the mixtures conditioned in the water bath and environmental chamber. The test results exhibited a strong linear correlation with the intercept set to 0. The figure also includes a data point from the plant-produced mixture, which was not used in developing the relationship shown. Although this "verification" data point lends support to the relationship, its reliability is limited, emphasizing the need for additional data to ensure robust verification. Nevertheless, referencing the established relationship and the minimum strength value of 133 kPa for surface asphalt mixtures with "A" and "D" designations conditioned in an environmental chamber in accordance with VDOT BMD specifications, a minimum strength value of 100 kPa is deemed a suitable initial threshold for surface asphalt mixtures with "A" and "D" designations conditioned in a water bath until further verification is performed.

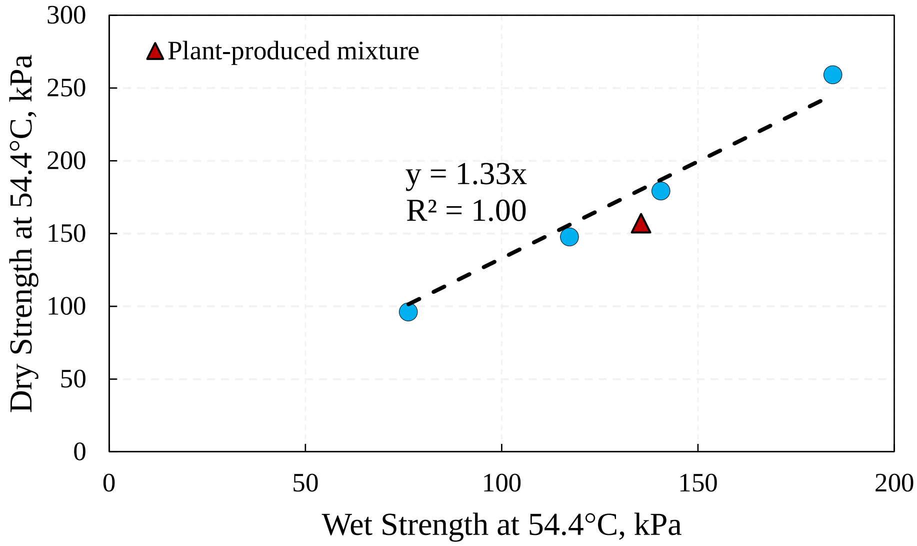

Figure 3**. Correlation Between IDT-HT Test Results for Mixtures Conditioned in Water Bath (x-axis) and Environmental Chamber (y-axis) Wet and Dry denote specimen conditioning in a water bath and environmental chamber, respectively. IDT-HT = indirect tensile at high temperature.**

IR tests replicated the analyses conducted for the IDT-HT tests. Figure 4 presents the RT index values obtained from IR testing of the mixtures conditioned in the water bath and environmental chamber, denoted as "wet" and "dry," respectively. The RT index values of the mixtures conditioned in the water bath consistently exhibited lower values, a drop in RT of approximately 14% from those of the mixtures conditioned in the environmental chamber. The normally distributed IR test results also demonstrated equal variances, indicating that the conditioning environment did not statistically impact the repeatability characteristics of the test.

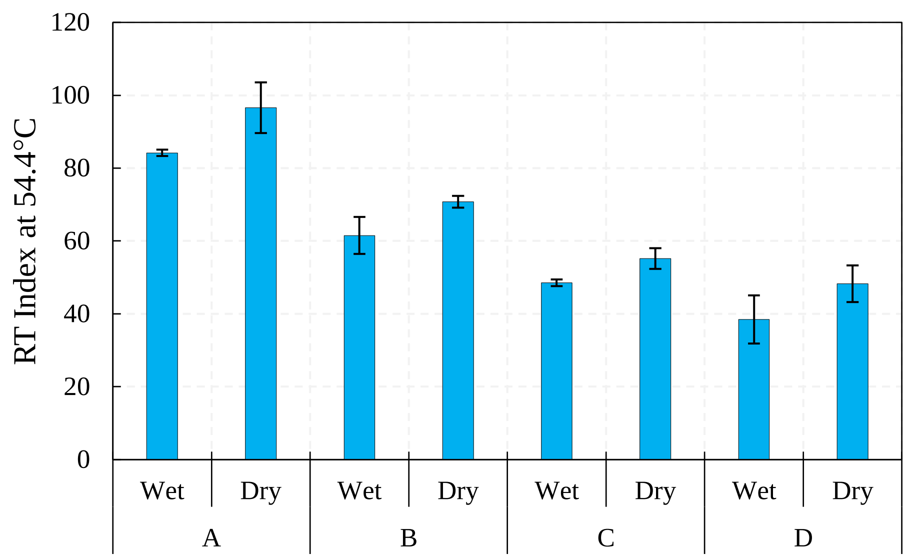

Figure 4. IR Test Results for the Mixtures Conditioned in the Water Bath (Wet) and Environmental Chamber (Dry). I-bars show ±1 standard deviation; IR = ideal rutting; RT = rutting tolerance.

Table 5 presents the ANOVA results at a 95% confidence interval for the response variable: RT index. The analysis indicated statistically significant variation in the RT index from the IR test when mixture type and/or conditioning environment changed.

Table 5. Summary of ANOVA Statistics for IR Test Results

Factor	RT index	
	DF	**p-value**
Mixture Type	3	0.000
Conditioning Environment	1	0.003
Mixture Type*Conditioning Environment	3	0.833

ANOVA = analysis of variance; RT = rutting tolerance; IR = ideal rutting; DF = degrees of freedom

As shown in Table 6, pairwise comparisons using the Tukey method at a 95% confidence interval indicated that the mixtures with statistically different results primarily belonged to sets of different mixture types subjected to conditioning either in a water bath or an environmental chamber, rather than pairs of the same mixtures conditioned in different environments. In other words, the mixtures exhibited statistically similar test results when different environments (i.e., water bath or environmental chamber) were used for conditioning specimens of the same mixture prior to testing.

Table 6. Tukey Pairwise Statistical Comparisons for IR Test Results

Mixture Type*Conditioning Environment	Grouping
A*Dry	a
A*Wet	a/b
B*Dry	b/c
B*Wet	c/d
C*Dry	c/d/e
C*Wet	d/e
D*Dry	d/e
D*Wet	e

Interaction factors sharing the same letter under the grouping title are statistically similar IR = ideal rutting; Dry = environmental chamber; Wet = water bath.

Although the RT index values were statistically similar between the mixtures conditioned in different environments, practical differences may arise when test results come from two laboratories using different specimen conditioning methods. For instance, this practice could introduce increased variability, potentially exceeding the reproducibility limits of the test method, even if the results from the two laboratories individually meet acceptable performance levels. In some instances, these differences in conditioning methods may lead to disputes, wherein mixtures with results near the performance threshold of the test method may pass for one laboratory and fail for another, despite statistically similar outcomes. Therefore, the statistical insignificance of the results of tests using different conditioning environments does not warrant including interchangeable modes of specimen conditioning environments in the test procedure. Consistent practices are crucial to mitigate the risk of method-related variations for robust quality measurement practices. In fact, the observed differences within this study between the test results of the same mixture subjected to different modes of conditioning, when evaluated with respect to the precision estimates determined later in this report, necessitate the development of a new threshold for the test methods for both IDT-HT and IR tests. This holds true regardless whether the test results are statistically significant or insignificant between the two conditioning methods.

The IR testing of the plant-produced mixture yielded the same observation as witnessed in the laboratory-produced mixtures. The specimens conditioned in the water bath had lower RT index values than those conditioned in the environmental chamber. The IR test followed that same process for assessing adhesive/cohesive failure. The observation was the same as the IDT-HT test.

Researchers compared the IR test results of the mixtures conditioned in both environments. Figure 5 shows the correlation between the IR test results of the mixtures conditioned in the water bath and environmental chamber. The test results exhibited a strong linear correlation with the intercept set to 0. The figure also includes a data point from the plant-produced mixture, which was not used in developing the shown relationship. Although this "verification" data point lends support to the relationship, its reliability is considered limited, emphasizing the need for additional data to ensure robust verification. Nevertheless, referencing the established relationship and the minimum RT index of 72 for surface asphalt mixtures with "A" and "D" designations conditioned in an environmental chamber as recommended in a previous study (Boz et al., 2023a), a minimum RT index of 62 is deemed a suitable initial threshold for surface asphalt mixtures with "A" and "D" designations conditioned in a water bath until further verification is performed.

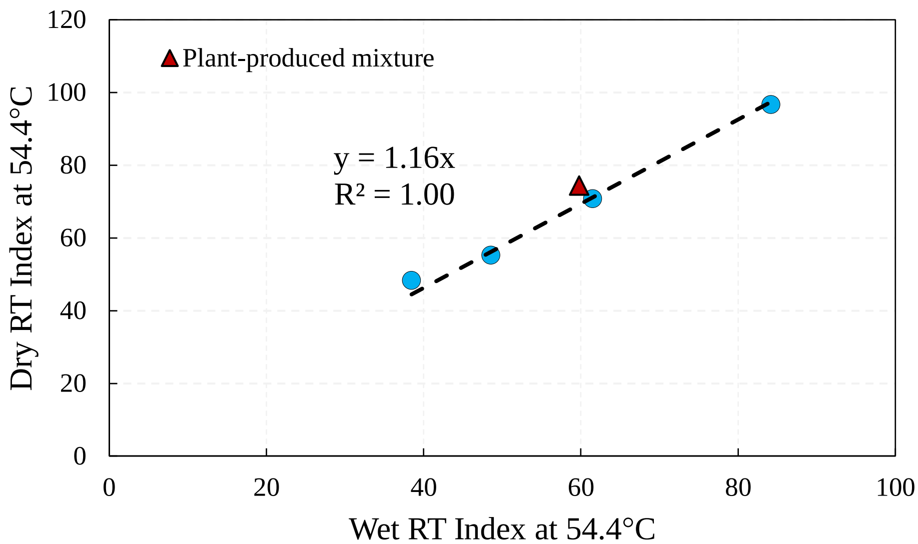

Figure 5. Correlation between IR Test Results for Mixtures Conditioned in Water Bath and Environmental Chamber. Wet and Dry denote specimen conditioning in a water bath and environmental chamber, respectively. IR = ideal rutting; RT = rutting tolerance.

Impact of Loading Rate on the Test Results

Figures 6 and 7 present the average IDT-HT and IR test results for the mixtures at each of the five loading rates, respectively. The specimens were conditioned in an environmental chamber before testing. Both figures indicate variations in the results of a given mixture for both tests across the five loading rates. The ANOVA at a 95% confidence interval was conducted to assess the statistical significance of the test results. The analyses included strength or RT index, as the response variable and factored in "mixture type," "loading rate," and an interaction term, "mixture type*loading rate." Prior to the ANOVA, the assumption of equal variance was checked and confirmed for both tests, indicating that the loading rate did not significantly impact the repeatability characteristics of the IDT-HT and IR tests. The assumption of normality was not a significant concern, given the sample size was equal to 30.

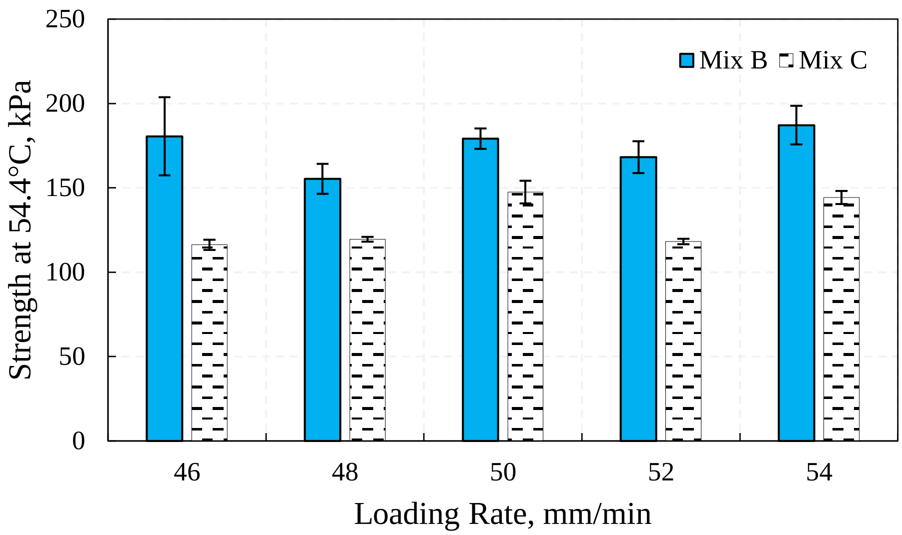

Figure 6. IDT-HT Test Results for Mixtures B and C at Five Loading Rates. I-bars show ±1 standard deviation; IDT-HT = indirect tensile at high temperature.

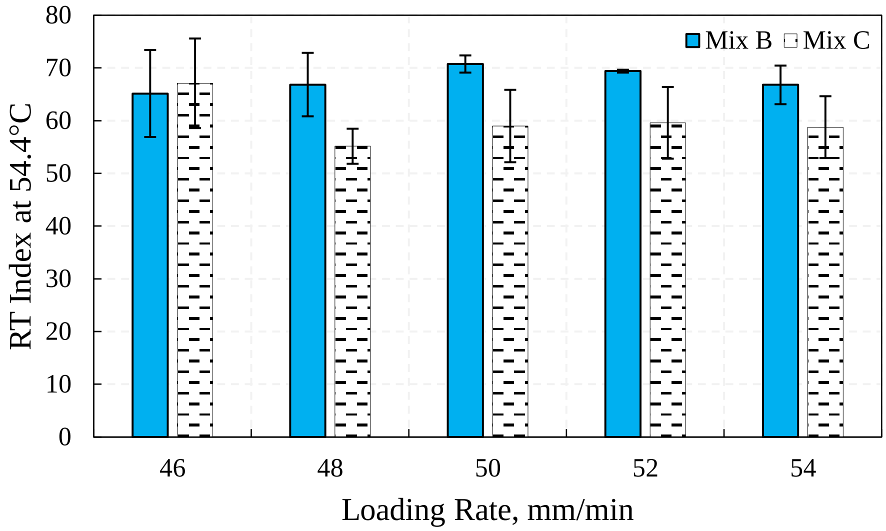

Figure 7. IR Test Results for Mixtures B and C at Five Loading Rates. I-bars show ±1 standard deviation; IR = ideal rutting; RT = rutting tolerance.

Tables 7 and 8 present the ANOVA statistics for the IDT-HT and IR test results at a 95% confidence interval, respectively. The statistical insignificance of the interaction factor (mixture type*loading rate) ($p > 0.05$) for both tests indicated that variations in the loading rate of 50 mm/min within a range of ±4 mm/min did not have a statistically significant impact on the IDT-HT and IR test results of the same mixture.

Table 7. Summary of ANOVA Statistics for IDT-HT Test Results

Factor	Strength	
	DF	**p-value**
Mixture Type	1	0.000
Loading Rate	4	0.004
Mixture Type*Loading Rate	4	0.148

ANOVA = analysis of variance; IDT-HT = indirect tensile test at high temperature; DF = degrees of freedom.

Table 8. Summary of ANOVA Statistics for IR Test Results

Factor	RT Index	
	DF	**p-value**
Mixture Type	1	0.011
Loading Rate	4	0.758
Mixture Type*Loading Rate	4	0.465

ANOVA = analysis of variance; IR = ideal rutting; RT = rutting tolerance; DF = degrees of freedom.

Impact of Using Different Loading Frames on the Test Results

Figures 8 and 9 present the average IDT-HT and IR test results for the mixtures for each of the three loading frames (referred to as LI, LII, and LIII), respectively. The specimens were conditioned in an environmental chamber before the testing. Both figures indicate variations in the results of a given mixture for both tests across the three loading frames. The ANOVA at a 95% confidence interval was conducted to determine if there was a statically significant difference in the IDT-HT and IR test results when different loading frames (devices) were used for testing. The analyses included strength or RT index as the response variable and factored in "mixture type," "loading frame," and an interaction term, "mixture type*loading frame." Before performing the ANOVA, the normal distribution and equal variance of the data for both tests were checked and confirmed, which indicated that the loading frame did not have a statistically significant impact on the repeatability characteristics of the IDT-HT and IR tests.

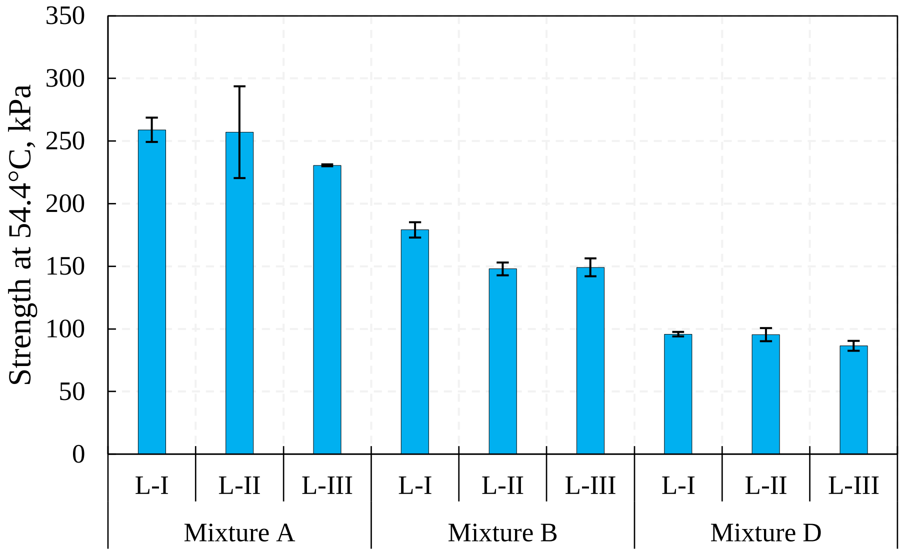

Figure 8. IDT-HT Test Results for Mixtures A, B, and D Using Three Different Loading Frames. I-bars show ±1 standard deviation. IDT-HT = indirect tensile at high temperature.

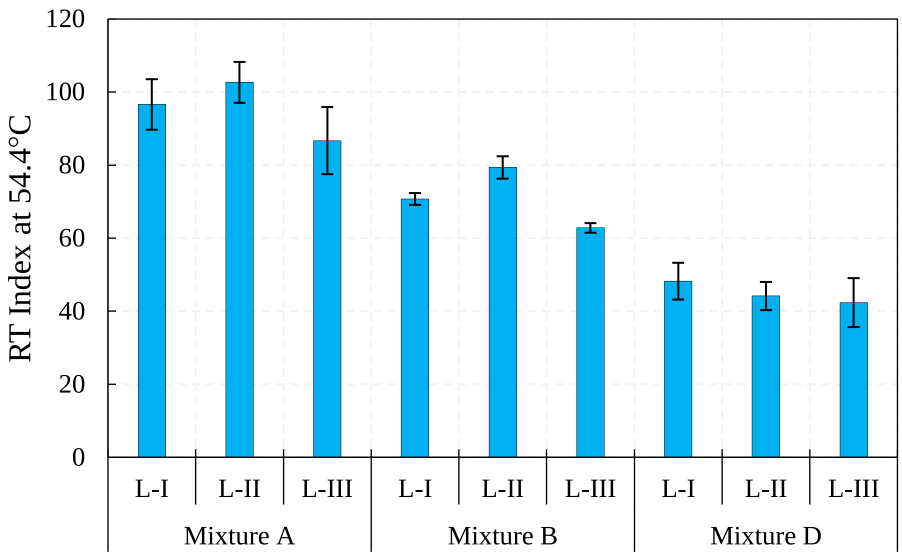

Figure 9. IR Test Results for Mixtures A, B and D Using Three Different Loading Frames. I-bars show ±1 standard deviation. IR = ideal rutting; RT = rutting tolerance.

Tables 9 and 10 present the ANOVA statistics for the IDT-HT and IR test results at a 95% confidence interval, respectively. The statistical insignificance of the interaction factor (mixture type*loading frame) ($p > 0.05$) for both tests indicated that the use of different loading frames did not have a statistically significant impact on the IDT-HT and IR test results of the same mixture.

Table 9. Summary of ANOVA Statistics for IDT-HT Test Results

Factor	Strength	
	DF	p-value
Mixture Type	2	0.066
Loading Frame	2	0.000
Mixture Type*Loading Frame	4	0.487

ANOVA = analysis of variance; IDT-HT = indirect tensile test at high temperature; DF = degrees of freedom.

Table 10. Summary of ANOVA Statistics for IR Test Results

Factor	RT Index	
	DF	p-value
Mixture Type	2	0.014
Loading Frame	2	0.000
Mixture Type*Loading Frame	4	0.322

ANOVA = analysis of variance; IR = ideal rutting; RT = rutting tolerance; DF = degrees of freedom.

Impact of Lag and Dwell Times on the Test Results

Figures 10 and 11 present the average IDT-HT and IR test results for the mixtures across different lag times, respectively. The specimens were conditioned in an environmental chamber before testing. Both figures indicate variations in the results of a given mixture for both tests across the lag times. Notably, for a given mixture, the IDT-HT and IR test results obtained on Day 0 exhibited lower strength or RT index values than those of the other days, likely indicating the impact of the reheating process on the test results. The ANOVA at a 95% confidence interval was conducted to assess the statistical significance of the test results. The analyses included strength or RT index as the response variable and considered an interaction term, “mixture type*lag time,” as a factor. Given a sample size of 30 in these experiments, the assumption of normality was not a significant concern. However, the assumption of equal variance was verified before performing the ANOVA.

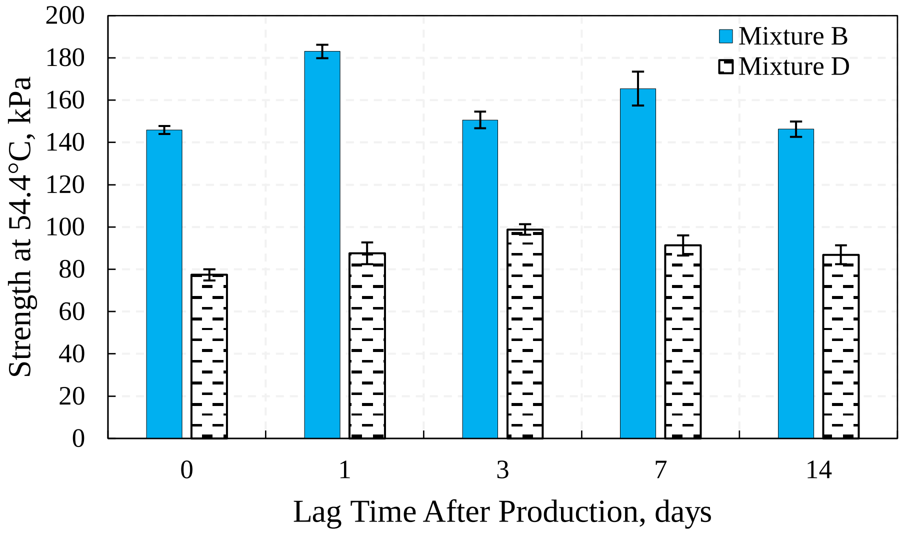

Figure 10. IDT-HT Test Results for Mixtures B and D across Different Lag Times. I-bars show ±1 standard deviation. IDT-HT = indirect tensile at high temperature.

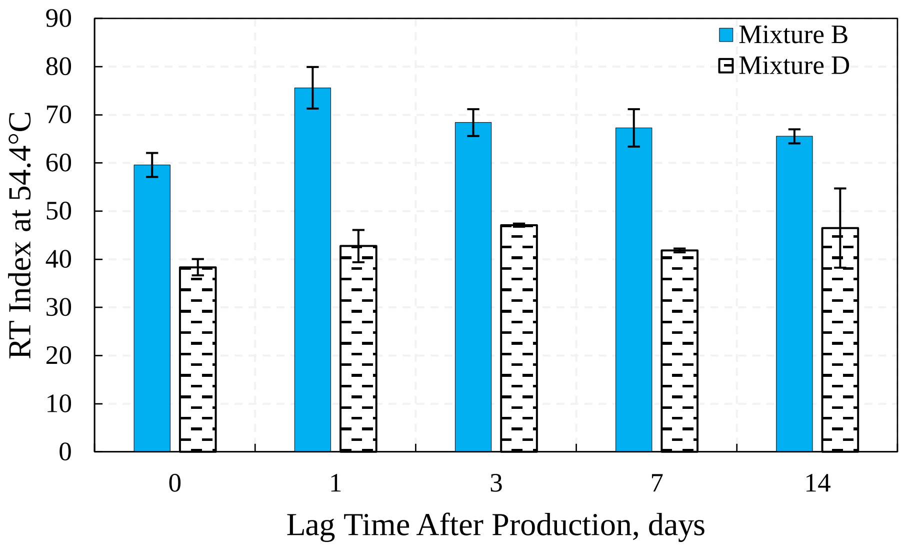

Figure 11. IR Test Results for Mixtures B and D across Different Lag Times. I-bars show ±1 standard deviation. IR = ideal rutting; RT = rutting tolerance.

The assumption of equal variance was confirmed for the IDT-HT test results, indicating that the lag time did not significantly impact the repeatability characteristics of the test. However, the ANOVA statistic of the interaction factor (mixture type*lag time) ($p = 0.000 < 0.05$) indicated a statistically significant variation in the IDT-HT test results. The pairwise comparisons using the Tukey method at a 95% confidence interval were conducted to identify the significant variations. The results of this analysis are presented in Table 11. As shown, for a given mixture, the test results significantly varied for some of the lag times. For example, for Mixture B, the test results obtained on Day 1 were statistically different from those obtained on Day 7, and they were also different from those obtained on the other days whose test results were statistically similar. On the other hand, the test results for Mixture D were statistically similar across the lag times, except for the test results obtained on Day 0.

Table 11. Tukey Pairwise Statistical Comparisons of Lag Times for IDT-HT Test Results

Mixture Type*Lag Time	Grouping
B*0	a
B*1	b
B*3	a
B*7	c
B*14	a
D*0	d
D*1	d/e
D*3	e
D*7	e
D*14	d/e

Interaction factors sharing the same letter in the Grouping column are statistically similar. IDT-HT = indirect tensile at high temperature

Although the test results obtained on Day 0 for Mixtures B and D were slightly lower in magnitude than the test results obtained on other days, they exhibited statistically similar results compared with some of the results obtained on other days. For instance, for Mixture B, the Day 0 results were statistically similar to those obtained on Days 3 and 14. Similarly, for Mixture D, the test results on Days 0, 1, and 14 showed a similar pattern of statistical similarity.

The test results exhibited "randomized" statistical similarities or dissimilarities with respect to changes in lag times. This observation may be attributed to the high repeatability characteristics observed in the dataset, resulting in a high discrimination potential among the datasets. For example, the repeatability, quantified by a coefficient of variation (COV), for these mixtures was consistently below 5.9% with an average COV of 3.5%. However, when evaluating the data in the context of the single-operator precision estimate of the test method, as presented later in the report, the results for both mixtures across the lag times are considered statistically similar.

Although acknowledging the statistically similar nature of the test results, the evident impact of reheating on the test results in terms of parameter magnitude is important, especially for results close to the performance threshold of the test method. Further investigations are warranted to understand the practical significance of reheating on the test results, particularly in a production setting.

The assumption of equal variance was not met for the IR test results. Close inspection of the data indicated that one specific data point (Day 14, Mixture D) exhibited unusually high variability, with a COV of 17.8% compared with the average COV of 4% for the rest of the data. This outlier observation likely contributed to the violation of the equal variance assumption. Consequently, there is insufficient evidence to suggest a significant impact of lag time on the repeatability characteristics of the test.

The ANOVA statistic of the interaction factor (mixture type*lag time) ($p = 0.000 < 0.05$) indicated a statistically significant variation in the IR test results. The pairwise comparisons using the Games-Howell method at a 95% confidence interval were conducted to identify the significant variations. Table 12 shows the results of this analysis. The table data infer that the test results exhibited statistical similarities across the lag times for Mixture B. Outcomes for Mixture D were the same, except for the results obtained on Day 7. Similar to the analysis performed for the IDT-HT test, evaluation of the IR test results for Mixture D in the context of the single-operator precision estimate of the test method, as quantified later in this report, are statistically similar. The observations made regarding the impact of reheating on the IDT-HT test results during production testing are also applicable to the IR test.

Table 12. Games-Howell Pairwise Statistical Comparisons of Lag Times for IR Test Results

Mixture Type*Lag Time	Grouping
B*0	a/b
B*1	a
B*3	a
B*7	a/b
B*14	a/b
D*0	c/d
D*1	c/d
D*3	b/c
D*7	d
D*14	a/b/c/d

Interaction factors sharing the same letter in the Grouping column are statistically similar: IR = ideal rutting

Figures 12 and 13 present the average IDT-HT and IR test results for the mixtures across different dwell times, respectively. The specimens were conditioned in an environmental chamber before testing. Both figures indicate variations in the results of a given mixture for both tests across the dwell times. The ANOVA at a 95% confidence interval was conducted to assess the statistical significance of the test results. The analyses included strength or RT index as the response variable and considered an interaction term, “mixture type*dwell time,” as a factor. Given a sample size of 30 in these experiments, the assumption of normality was not a significant concern. However, the assumption of equal variance was verified before performing the ANOVA.

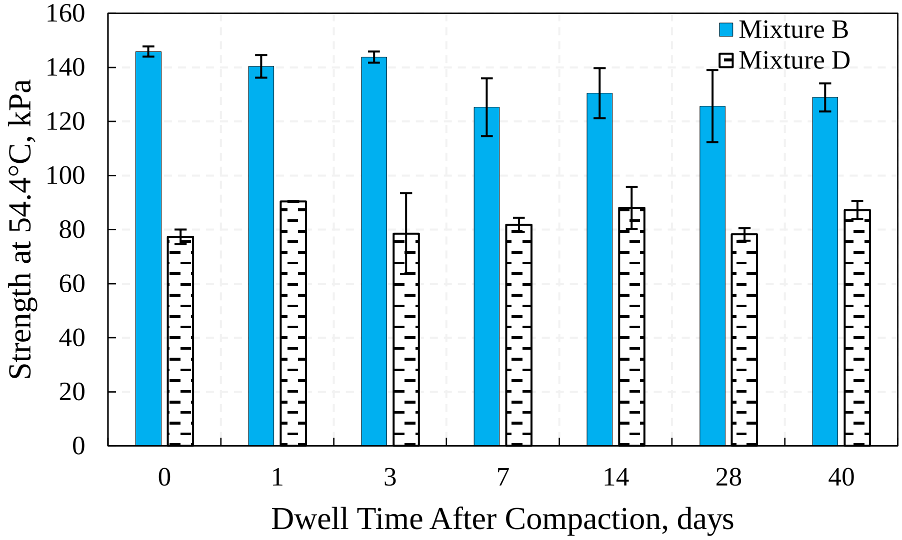

Figure 12. IDT-HT Test Results for Mixtures B and D across Different Dwell Times. I-bars show ±1 standard deviation; IDT-HT = indirect tensile at high temperature.

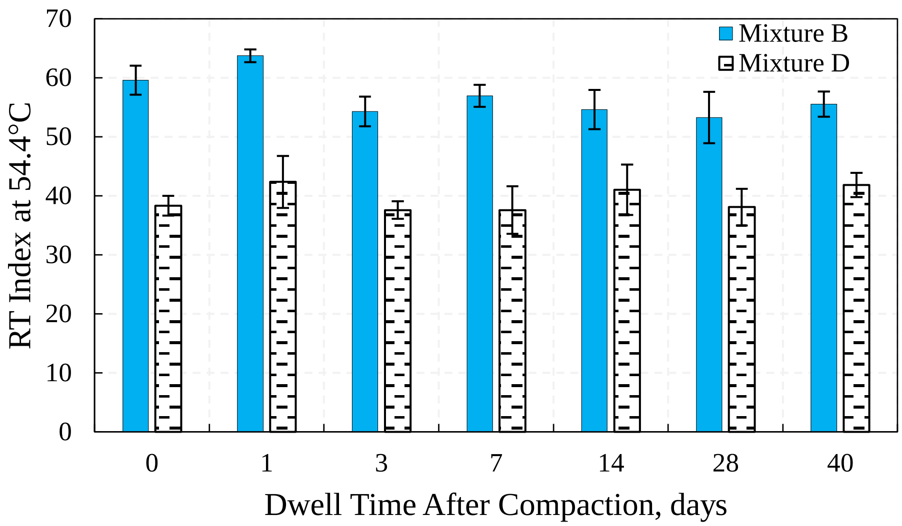

Figure 13. IR Test Results for Mixtures B and D across Different Dwell Times. I-bars show ±1 standard deviation; IR = ideal rutting; RT = rutting tolerance.

The assumption of equal variance was not met for the IDT-HT test results. Close examination of the data indicated that one specific data point (Day 3, Mixture D) exhibited unusually high variability, with a COV of 19.1% compared with the average COV of 4.5% for the rest of the data. This outlier observation likely contributed to the violation of the equal variance assumption. Consequently, there is insufficient evidence to suggest a significant impact of dwell time on the repeatability characteristics of the test.

The ANOVA statistic of the interaction factor (mixture type*dwell time) ($p = 0.000 < 0.05$) indicated a statistically significant variation in the IDT-HT test results. The pairwise comparisons using the Games-Howell method at a 95% confidence interval were conducted to identify the significant variations. Table 13 summarizes the analysis results. As the table shows, for a given mixture, the test results exhibited statistical similarities across the dwell times. This result indicates that dwell times did not have a statistically significant impact on the IDT-HT test results of the same mixture.

Table 13. Games-Howell Pairwise Statistical Comparisons of Dwell Times for IDT-HT Test Results

Mixture Type*Dwell Time	Statistical Grouping
B*0	a
B*1	a
B*3	a
B*7	a/b/c
B*14	a/b
B*28	a/b/c
B*40	a/b
D*0	c
D*1	b/c
D*3	a/b/c
D*7	b/c
D*14	b/c
D*28	b/c
D*40	b/c

Interaction factors sharing the same letter in the Grouping column are statistically similar. IDT-HT = indirect tensile at high temperature.

The assumption of equal variance met the IR test results, which suggests that no significant impact of dwell time on the repeatability characteristics of the test. The ANOVA statistic of the interaction factor (mixture type*dwell time) ($p = 0.000 < 0.05$) indicated a statistically significant variation in the IR test results. The pairwise comparisons using the Tukey method at a 95% confidence interval were conducted to identify the significant variations. Table 14 summarizes the analysis results. For a given mixture, the test results exhibited statistical similarities across the dwell times, except for Mixture B Day 1 results. This result indicates that dwell times, overall, did not have a statistically significant impact on the IDT-HT test results of the same mixture as the single data point is considered as a potential outlier.

Table 14. Tukey Pairwise Statistical Comparisons of Dwell Times for IR Test Results

Mixture Type*Dwell Time	Grouping
B*0	a/b
B*1	a
B*3	b
B*7	a/b
B*14	b
B*28	b
B*40	a/b
D*0	c
D*1	c
D*3	c
D*7	c
D*14	c
D*28	c
D*40	c

Interaction factors sharing the same letter in the Grouping column are statistically similar. IR = ideal rutting.

The findings regarding the impacts of loading rate, loading frame, and lag and dwell times on the test results are expected to remain applicable to the water bath test results. Thus, the findings from the sensitivity assessment study and the informed judgment of the authors in consultation with the technical review panel and industry were used to develop a Virginia Test Method for the IDT-HT test method, presented in Appendix A.

Interlaboratory Study

Data Quality Assessment

Before analyzing the collected data to determine precision estimates of the test methods, two distinct data quality assessments were conducted. The first assessment focused on ensuring the quality of the submitted raw data, following the guidelines outlined in a previous study (Habbouche et al., 2021). The second assessment involved evaluating the single-operator and multi-laboratory consistency of data in accordance with ASTM C802. These metrics were evaluated with respect to average and dispersion of the test results to avoid potential effects of inconsistent data on the precision of the test methods. Both assessments did not reveal any significant issues for either of the tests.

Selection of an Appropriate Statistical Parameter

The standard deviation or COV serves as the fundamental statistical parameter that underlies precision indexes in a test method. The determination of precision estimates, whether for a single operator or multiple laboratories, relies on either the standard deviation or COV. If the standard deviation remains consistent as test results vary in magnitude, it is used to establish precision statements. In cases where the standard deviation varies with changes in the magnitude of test results, precision statements are then determined using the COV, provided the coefficient remains relatively stable in response to changes in test results. When both the standard deviation and COV change with variations in test results, precision statements should be formulated based on either the standard deviation or COV relative to the magnitude of test results. Detailed guidance on developing precision statements for different scenarios can be found in ASTM C670.

The collected data were plotted to analyze the relationship between the parameters of the tests and the standard deviation as well as the COV for single-operator and multi-laboratory conditions. This analysis focused on assessing the rate of change in statistical parameters with respect to the parameters of the tests to determine the statistical form of precision statements for each test method. The data showed that COV remained nearly constant with respect to changes in both strength and RT index for both single-operator and multi-laboratory conditions, as shown in Figures 14 and 15 for IDT-HT and IR tests, respectively. Thus, the precision estimates for IDT-HT and IR tests were developed based on the COV.

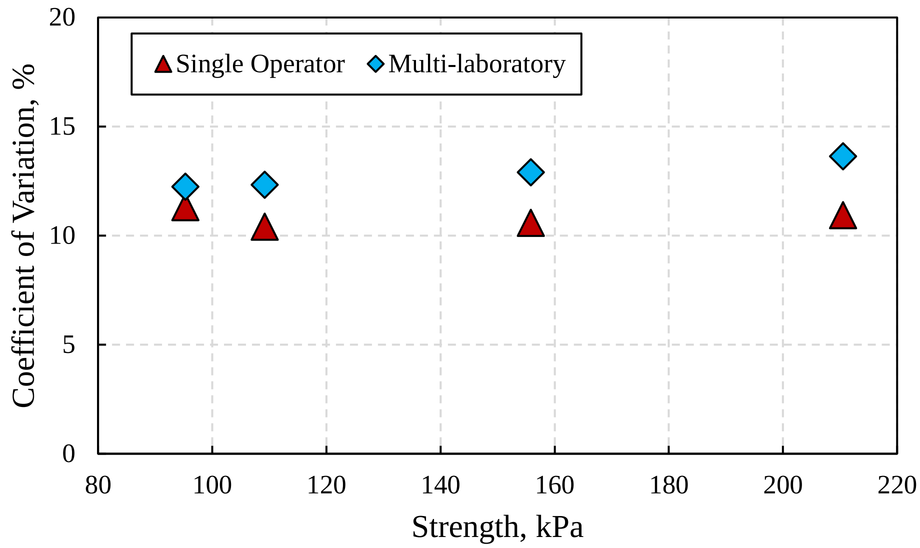

Figure 14. Variation in Coefficient of Variation over the Range of Average Strength Values of the IDT-HT Test. IDT-HT = indirect tensile at high temperature.

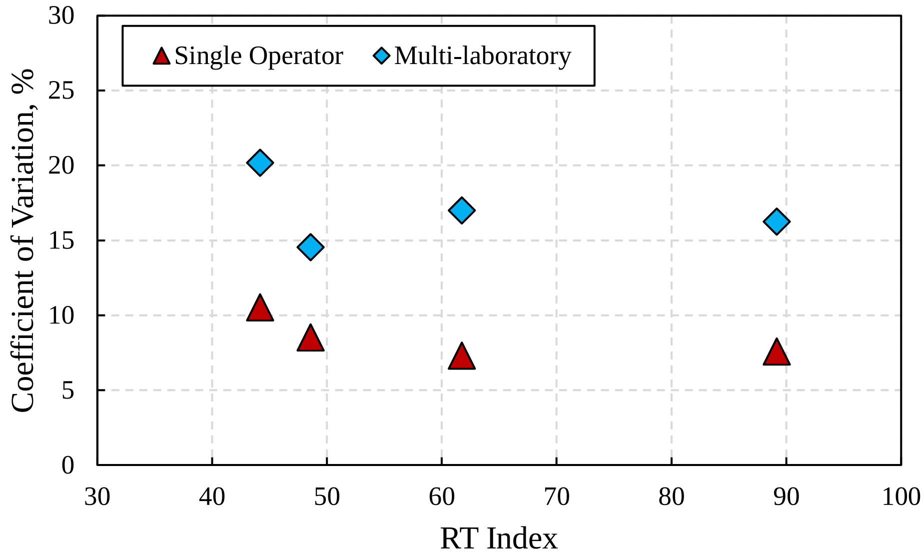

Figure 15. Variation in Coefficient of Variation over the Range of Average RT Index Values of the IR Test. IR = ideal rutting; RT = rutting tolerance.

Determination of Precision Estimates

The process for determining precision estimates for the IDT-HT and IR followed the ASTM C802 method. For a given mixture, each participating laboratory received five replicate specimens for the IDT-HT test and three replicate specimens for the IR test. The collected five-replicate datasets for the IDT-HT test were analyzed further to generate two distinct three-replicate datasets. The first three-replicate dataset was generated to represent the lowest single-operator variability condition for a given mixture for each laboratory. The second three-replicate dataset was generated to represent the highest single-operator variability condition for a given mixture for each laboratory. In other words, these datasets represent the two extreme conditions of the test variability based on three replicates. The precision estimates for the IDT-HT test were determined based on these three conditions. The ILS for the IR test was based on using three-replicate specimens.

Table 15 presents the precision estimates for the IDT-HT/Strength test under the three conditions plus the precision estimates for the IR test. Notably, the precision estimates of the IDT-HT tests generated from the five-replicate condition exhibit magnitudes that are similar to the average magnitudes of the precision estimates from the other two conditions. The use of precision estimates based on the three-replicate (extreme) conditions for the IDT-HT test may pose challenges for quality measurement practices. The precision estimates derived from conditions with the lowest single-operator variability may appear unrealistically low. For a given dataset, the multi-laboratory precision estimate is higher than the single-operator precision estimate. Moreover, precision estimates are expected to increase as the number of replicates decreases when separate sets of specimens from a given mixture are prepared with three and five replicates. In other words, the precision estimates are typically not determined by reducing the number of specimens from a dataset with initially higher replicates (specimens). However, the opposite is observed when the precision estimates are determined through the reduced number of specimens from a dataset with higher replicates. Considering these, the precision estimates from the three-replicate condition with the lowest single operator variability are expected to be lower values compared with those from the five-replicate condition, as Table 15 shows.

Table 15. Summary of Precision Estimates of the IDT-HT and IR Tests

Test/Parameter	Condition	Precision Estimate Values, COV, %	
		Single-Operator	Multi-Laboratory
IDT-HT/Strength	Five-replicate	10.8	12.8
	Three-replicate/Lowest	7.3	10.7
	Three-replicate/Highest	13.7	13.9
IR/RT Index	Three-replicate	8.5	17.0

COV = coefficient of variation; IDT-HT = indirect tensile at high temperature; IR = ideal rutting; RT = rutting tolerance.

Considering deliberate generation of the dataset to minimize variability, the multi-laboratory precision estimate for that condition (10.7%) is lower than the single-operator precision estimate of the five-replicate condition (10.8%), making the precision estimates of the three-replicate condition unrealistic. On the other hand, the precision estimates derived from conditions with the highest single operator variability may be considered relatively high. Such high variability could potentially mask compositional changes in asphalt mixtures during quality measurement practices. Therefore, the researchers decided to use the precision estimates from the five-replicate condition to evaluate asphalt mixtures with three replicates.

Development of Precision Statements

The following precision statements for the IDT-HT and IR tests were developed in accordance with ASTM C670 based on the results in Table 15.

- For the IDT-HT test:

Single-Operator Precision—The single-operator coefficient of variation is 10.8%. A test result is defined as the average of three test determinations (replicates). Therefore, the results of two properly conducted tests (each consisting of the average of three test replicates) by the same operator on the same material are not expected to differ from each other by more than 17.5% of their average. The range (difference between highest and lowest) of the three replicates used in calculating the average is not expected to differ from each other by more than 35.6% of their average.

Multi-Laboratory Precision—The multi-laboratory coefficient of variation is 12.8%. A test result is defined as the average of three test determinations (replicates). Therefore, results of two properly conducted tests by two different laboratories on specimens of the same material are not expected to differ from each other by more than 35.8% of their average.

- For the IR test:

Single-Operator Precision—The single-operator coefficient of variation is 8.5%. A test result is defined as the average of three test determinations (replicates). Therefore, the results of two properly conducted tests (each consisting of the average of three test replicates) by the same operator on the same material are not expected to differ from each other by more than 13.7% of their average. The range (difference between highest and lowest) of the three replicates used in calculating the average is not expected to differ from each other by more than 28% of their average.

Multi-Laboratory Precision—The multi-laboratory coefficient of variation is 17%. A test result is defined as the average of three test determinations (replicates). Therefore, results of two properly conducted tests by two different laboratories on specimens of the same material are not expected to differ from each other by more than 47.6% of their average.

Impact of Specimen Conditioning Method on the Test Results

As previously mentioned, eight laboratories, which had already received specimens for testing in a water bath as part of the ILS, were additionally provided with two sets of three-replicate specimens for a given mixture. These additional specimens were designated for IDT-HT and IR testing after being subjected to conditioning in an environmental chamber. The primary aim of this effort was to assess further the impact of conditioning environments on the test results, considering the various laboratories involved. Figure 16 compares the IDT-HT test results from specimens conditioned in these two distinct environments. Visual inspection of the figures shows that five of the eight laboratories had lower strength values in the wet condition compared with the dry condition. On the other hand, two laboratories exhibited dry strength values that were lower than the wet strength values. Assuming that the conditioning environment does not have a statistical impact on the test results, then the difference between the wet and dry strength values for three laboratories exceeded the single-operator precision estimate of the test, a result that represents 37.5% of the observations within this sample.

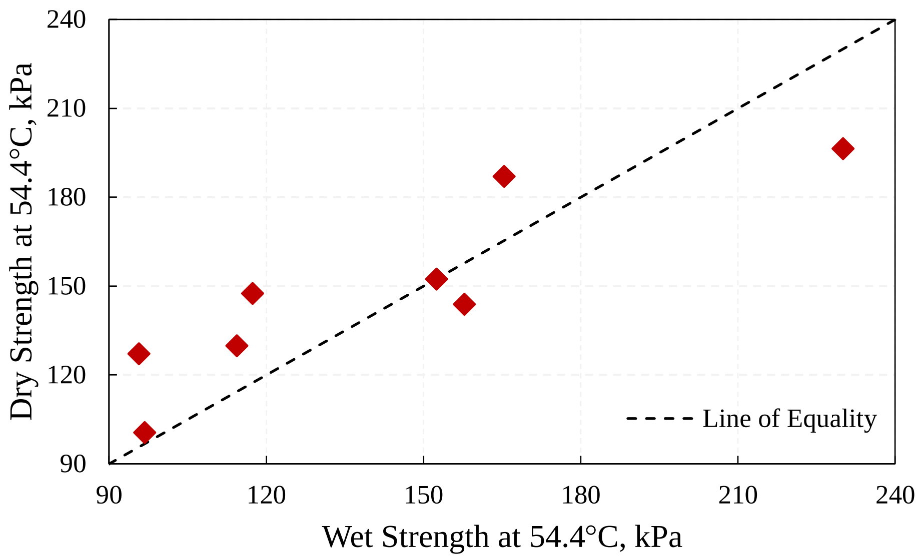

Figure 16. Comparisons between IDT-HT Test Results for Mixtures Conditioned in Water Bath and Environmental Chamber. Wet and Dry denote specimen conditioning in a water bath and environmental chamber, respectively. IDT-HT = indirect tensile at high temperature.

Figure 17 compares the IR test results from specimens conditioned in water bath and environmental chamber. Similar to the IDT-HT test, five of the eight laboratories had lower RT index values in the wet condition compared with the dry condition. Conversely, three laboratories showed lower dry RT index values than wet RT index values. The difference between the wet and dry RT index values was statistically significant for two laboratories when the results are evaluated with respect to the single-operator precision estimate of the test, which corresponds to 25% of the observations in this sample. Considering that the ILS laboratories possess substantial experience in asphalt mixture testing, the outcome from this study underlines the significance of specifying a conditioning environment in a test protocol to effectively manage and mitigate potential risks during quality measurement practices.

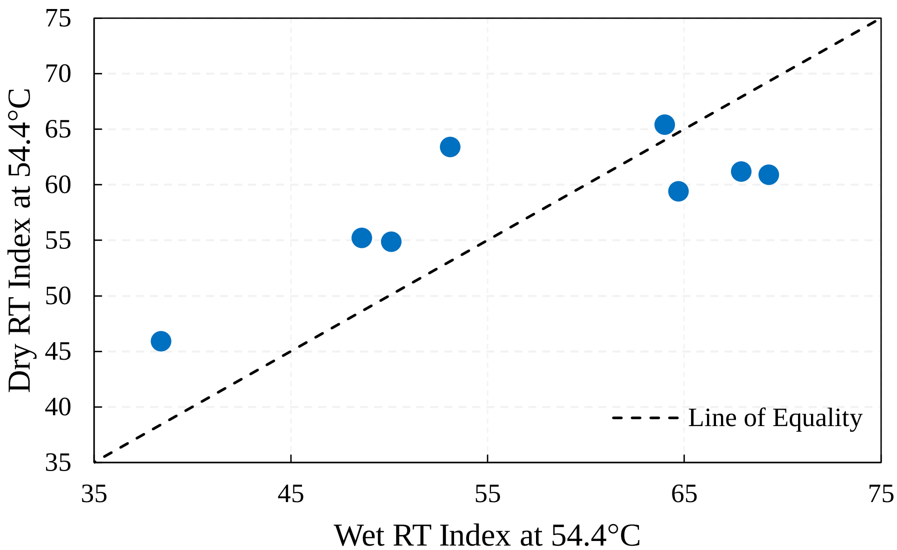

Figure 17. Comparisons Between IR Test Results for Mixtures Conditioned in Water Bath and Environmental Chamber. Wet and Dry denote specimen conditioning in a water bath and environmental chamber, respectively. IR = ideal rutting; RT = rutting tolerance.

Proficiency Testing

Figure 18 presents the z-scores for the 31 laboratories that participated in the proficiency study. Out of these 31 laboratories, 22 (corresponding to 71% of the observations) achieved z-scores within the range of -1 to 1. This range corresponds to Rating 5, the highest possible rating, indicating proficiency in conducting the test relative to the interlaboratory average. Additionally, 29 laboratories (corresponding to 93.5% of the observations) achieved z-scores within the acceptable range of -2 to 2, aligning with the rating criteria specified by the AASHTO accreditation (AASHTO Resource, 2016), which corresponds to a Rating 3 level, indicating satisfactory performance by these laboratories.

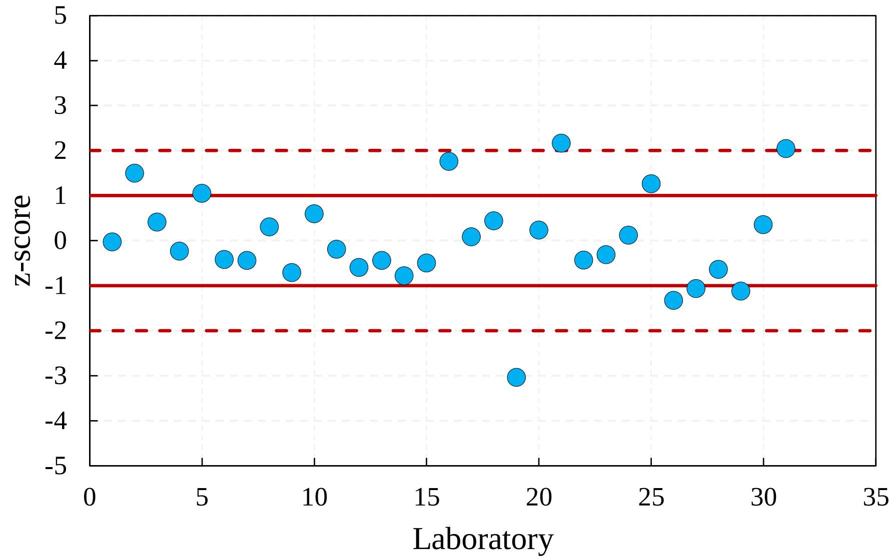

Figure 18. Z-scores of Laboratories for IDT-HT Testing Proficiency. IDT-HT = indirect tensile at high temperature. Red solid and dashed lines indicate the range for Ratings 5 and 3, respectively.

The data from the two laboratories that fell below a Rating of 3, corresponding to a z-score less than 2 and greater than -2, were investigated. However, no significant issues or anomalies were identified. In both cases, a single observation out of the three replicates was responsible for deviating the laboratory's average test result from the interlaboratory average, resulting in the lower ratings. Overall, the results indicated that most laboratories demonstrated proficiency in the IDT-HT test, which can likely be attributed to their experience with the IDT-CT test, as both tests are conducted in a similar manner. The similarity of the test protocols may have contributed to the proficiency of the participating laboratories.

The proficiency effort involved the use of three-replicate specimens from Mixture C. The collected data from this effort were analyzed in accordance with ASTM C802 to determine the "precision estimates" of IDT-HT test. The single operator and multi-laboratory precision estimates were found to be 10.3% and 12.5%, respectively. This finding served as a validation of using the five-replicate specimen precision estimates of the IDT-HT testing for the IDT-HT testing based on three-replicate specimens.

CONCLUSIONS

- *Specimen conditioning time significantly differs between different conditioning environments: water bath and environmental chamber.* Specimens conditioned in a water bath reach the target test temperature more quickly than those conditioned in an environmental chamber.

- *The conditioning environment does not significantly impact the repeatability characteristics of the IDT-HT and IR tests.*

- *Specimens conditioned in a water bath consistently exhibit lower parameter/index values compared with those conditioned in an environmental chamber, indicating a substantial* *effect of the conditioning environment on the outcomes of the IDT-HT and IR tests*. This conclusion underscores the importance of specifying and controlling the conditioning environment in a test protocol, particularly in the context of quality measurement practices, as not requiring a specific conditioning environment can significantly impact the reliability and consistency of test results.

- *The loading rate and the loading frame (device type) used do not significantly impact the repeatability characteristics of the IDT-HT and IR tests.*

- *The results of IDT-HT and IR tests are not dependent on the loading rate applied, regardless of the loading frame (device type) used.* This conclusion is valid for a loading rate range of 46 to 54 mm/min.

- *The lag and dwell times do not significantly impact the repeatability characteristics of the IDT-HT and IR tests.*

- *The lag times do not significantly impact the IDT-HT and IR test results when considering the single-operator precision estimates of the tests.*

- *Further investigations are warranted to understand the practical significance of reheating on the IDT-HT and IR test results.*

- *Dwell times do not significantly impact the IDT-HT and IR test results.*

- *The COV is an appropriate statistical parameter for evaluating the variability of IDT-HT and IR test results.*

- *Sets of three replicates of the IDT-HT test had single-operator and multi-laboratory precision estimates similar to those determined for five-replicate datasets.*

- *Preliminary performance criteria requiring minimum strength and RT index values of 100 kPa and 62 are deemed suitable for surface asphalt mixtures with "A" and "D" designation and both conditioned in a water bath at 54.4°C.* The proposed thresholds for strength and RT index values provide initial benchmarks, but further validation is necessary for robust verification.

RECOMMENDATIONS

1. *VDOT's Materials Division should adopt the Virginia Test Method presented in the appendix for the IDT-HT test.* This method prescribes the IDT-HT test to be performed following conditioning specimens in a water bath for 1 hour ± 10 minutes. It also specifies the allowable loading/deformation rate tolerance to be 50 ± 3 mm/min when performing the IDT-HT.

2. *VDOT's Materials Division should consider adopting the precision estimates and corresponding statements determined in this study for the strength parameter from the IDT-*

HT test. The recommended precision statements[1] for the strength parameter from the IDT-HT tests are as follows:

> Single-Operator Precision—The single-operator coefficient of variation is 10.8%. A test result is defined as the average of three test determinations (replicates). Therefore, the results of two properly conducted tests (each consisting of the average of three test replicates) by the same operator on the same material are not expected to differ from each other by more than 17.5% of their average.[2] The range (difference between highest and lowest) of the three replicates used in calculating the average is not expected to differ from each other by more than 35.6% of their average (see footnote 2).
>
> Multi-Laboratory Precision—The multi-laboratory coefficient of variation is 12.8%. A test result is defined as the average of three test determinations (replicates). Therefore, results of two properly conducted tests by two different laboratories on specimens of the same material are not expected to differ from each other by more than 35.8% of their average (see footnote 2).

3. *VDOT's Materials Division should adopt an initial minimum strength criterion of 100 kPa for IDT-HT testing of surface asphalt mixtures with "A" and "D" designations in accordance with the Virginia Test Method presented in the appendix.* The recommended performance threshold corresponds to specimens conditioned in water bath for 1 hour ± 10 minutes.

IMPLEMENTATION AND BENEFITS

Researchers and the technical review panel (listed in the Acknowledgments) for the project collaborate to craft a plan to implement the study recommendations and to determine the benefits of doing so. This is to ensure that the implementation plan is developed and approved with the participation and support of those involved with VDOT operations. The implementation plan and the accompanying benefits are provided here.

Implementation

With regard to Recommendations 1 and 3, VDOT's Materials Division has already adopted the Virginia Test Method presented in the appendix (designated as the Virginia Test Method - 145) and the recommended threshold in its BMD specification for 2024.

With regard to Recommendation 2, VTRC will work closely with VDOT's Materials Division to incorporate the recommended precision estimates into Virginia Test Method 145, and the single-operator precision estimate into the 2025 BMD special provisions for both mix design and production for reporting purposes only. This will be accomplished by February 1, 2026

[1] These precision statements are based on an interlaboratory study that involved 10 laboratories, four dense-graded surface asphalt mixtures with strength values ranging from 66 to 262 kPa, and three replicate tests per operator.
[2] These numbers represent the difference limits in percentage (d2s%) as described in Practice ASTM C670.

Benefits

This study developed a Virginia Test Method for the IDT-HT test and precision estimates and precision statements for the strength and RT index of asphalt mixtures determined by performing the IDT-HT and IR tests. The implementation of these findings into VDOT specifications provides unified testing standards, identifies acceptable performance levels, and establishes robust practices to validate test result consistency.

ACKNOWLEDGMENTS

The authors of this report are grateful to the following individuals who served on the technical review panel for this study: Angela Beyke (Project Champion and Assistant State Materials Engineer, VDOT Materials Division); Donald (Clyde) Landreth (Pavement Engineer, VDOT Salem District); Bryan Smith (Asphalt Pavement Field Engineer, VDOT Materials Division); Todd Rorrer (Asphalt Pavement Field Engineer, VDOT Materials Division); and Mo Zhao (Senior Research Scientist, VTRC). The authors acknowledge Mike Dudley from the Virginia Asphalt Association for his assistance in developing and executing the ILS. The contributions of the staff of the Bluegrass Testing Laboratory in designing and producing loose mixtures and test specimens are deeply appreciated. Appreciation is also extended to all participant laboratories involved in any testing in this study. The authors are also grateful to Linda Evans of VTRC for her editorial support.

REFERENCES

AASHTO Resource. *Proficiency Sample Ratings: Being Average Has Never Been So Good*, August 2, 2016. Accessed December 16, 2024. https://aashtoresource.org/university/newsletters/newsletters/2016/08/02/proficiency-sample-ratings.

ASTM International. *ASTM C670-15: Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials*. West Conshohocken, PA, 2015.

ASTM International. *ASTM C802-14: Standard Practice for Conducting an Interlaboratory Test Program to Determine the Precision of Test Methods for Construction Materials*. West Conshohocken, PA, 2014.

ASTM International. *ASTM D8225-19: Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature*. West Conshohocken, PA, 2019.

Boz, I., Coffey, G.P., Habbouche, J., Diefenderfer, S.D., and Ozbulut, O.E. A Critical Review of Monotonic Loading Tests to Evaluate Rutting Potential of Asphalt Mixtures. *Construction and Building Materials*, Vol. 335, 2022.

Boz, I., Habbouche, J., Diefenderfer, S.D., and Bilgic, Y.K. Precision Estimates and Statements for Performance Indices From the Indirect Tensile Cracking Test at Intermediate Temperature. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2021, pp. 225-241.

Boz, I., Habbouche, J., Diefenderfer, S.D., Coffey, G.P., Ozbulut, O.E., and Seitllari, A. *Simple and Practical Tests for Rutting Evaluation of Asphalt Mixtures in the Balanced Mix Design Process*. VTRC 23-R11. Virginia Transportation Research Council, Charlottesville, 2023a.

Boz, I., Habbouche, J., Diefenderfer, S.D., Coffey, G.P., Seitllari, A., and Ozbulut, O E. Evaluating the Rutting Potential of Asphalt Mixtures with Simple and Practical Tests. *Transportation Research Record: Journal of the Transportation Research Board,* 2023b. https://doi.org/10.1177/03611981231207089.

Diefenderfer, S.D., Habbouche, J., and Boz, I. Using Indirect Tensile Cracking Test Data for Quality Control and Acceptance, *Transportation Research Record: Journal of the Transportation Research Board,* 2023. https://doi.org/10.1177/03611981231170876.

Habbouche, J., Boz, I., and Diefenderfer, S.D. *Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II*. VTRC 23-R3. Virginia Transportation Research Council, Charlottesville, 2022.

Habbouche, J., Boz, I., and Diefenderfer, S.D. Impact of Device Type and Specimen Preparation on the Variability of the Indirect Tensile Cracking Test (IDT-CT) at Intermediate Temperatures, *Transportation Research Record: Journal of the Transportation Research Board*, 2023. https://doi.org/10.1177/03611981221150242.

Habbouche, J., Boz, I., Diefenderfer, S.D, and Bilgic, Y.K. *Round Robin Testing Program for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase I*. VTRC 22-R3. Virginia Transportation Research Council, Charlottesville, 2021.

APPENDIX A
VIRGINIA TEST METHOD FOR THE IDT-HT TEST

Virginia Test Method - 145

Method of Test for Determining Rutting Susceptibility of Asphalt Mixtures Using the Indirect Tensile at High Temperature (IDT-HT) Test

1. Scope

1.1. This method describes the procedures for preparing, testing, and measuring the rutting susceptibility of asphalt mixtures using the indirect tensile at high temperature (IDT-HT) test.

2. Referenced Documents

2.1. AASHTO Standards

- AASHTO R 30 Mixture Conditioning of Hot Mix Asphalt
- AASHTO R 47 Reducing Sample of Asphalt Mixtures to Testing Size
- AASHTO R 97 Sampling Asphalt Mixtures
- AASHTO T 166 Bulk Specific Gravity (Gmb) of Compacted Asphalt Mixtures Using Saturated Surface-Dry Specimens
- AASHTO T 209 Theoretical Maximum Specific Gravity (Gmm) and Density of Asphalt Mixtures
- AASHTO T 269 Percent Air Voids in Compacted Dense and Open Asphalt Mixtures
- AASHTO T 312 Preparing and Determining the Density of Asphalt Mixture Specimens by Means of the Superpave Gyratory Compactor

2.2. ASTM Standards

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

3. Summary of Test Method

3.1. A cylindrical specimen conditioned at a temperature of 54.4 ± 0.5°C is centered in the indirect tensile test fixture. A constant rate of axial displacement at 50 ± 3 mm/min is applied across the diametrical plane of the specimen using a loading frame. The peak load is measured and used to calculate the strength of the specimen.

4. Significance and Use

4.1. This method is used for the laboratory evaluation of the rutting susceptibility of asphalt mixtures. Both laboratory-produced and plant-produced asphalt mixtures can be tested and compared with a specified performance criterion.

5. Apparatus

5.1. Ovens for heating aggregate and asphalt binder.

5.2. Mixing utensils (bowls, spoon, spatula, etc.).

5.3. Equipment for determining the theoretical maximum specific gravity (Gmm) of asphalt mixtures conforming to the requirements of AASHTO T 209.

5.4. Compaction device and associated equipment for preparing laboratory specimens conforming to the requirements of AASHTO T 312.

5.5. Equipment for determining the bulk specific gravity (Gmb) of compacted asphalt mixtures conforming to the requirements of AASHTO T 166.

5.6. A caliper capable of measuring specimen height and diameter to ±0.1 mm.

5.7. A water bath capable of maintaining the test temperature of 54.4°C within 0.5°C for conditioning specimens before testing.

5.8. An indirect tensile test apparatus conforming to the requirements of ASTM D8225, Section 6.1, except that the axial loading device shall be capable of maintaining a constant deformation rate of 50 ± 3 mm/min.

6. Hazards

6.1. This standard may involve hazardous materials, operations, and equipment. This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.

7. Specimen Preparation

7.1. Number of test specimensa minimum of three cylindrical specimens with a 150 ± 2 mm diameter and a 62 ± 1 mm height for the mixtures with a nominal maximum aggregate size (NMAS) of 19 mm or smaller. All specimens shall be prepared without cutting or trimming from a larger specimen.

7.2. The test specimens for both laboratory-prepared and plant-produced mixtures shall be compacted to 7.0 ± 0.5% air voids.

7.3. Laboratory-Prepared Asphalt Mixture Specimens

7.3.1. Mixture proportions shall be batched in accordance with the desired job mix formula (JMF). The batch size shall be equal to the amount of mixture needed to achieve the required specimen dimensions and air void content for a single test specimen. A minimum of four batches are required—three for test sample fabrication and one to determine the Gmm.

7.3.2. Test samples shall be short-term aged in accordance with AASHTO R 30 at the appropriate compaction temperature recommended in Section 211.03(d)6 of the specifications for two (2) hours ± 5 min.

7.3.3. The maximum specific gravity of the test mixture shall be determined in accordance with AASHTO T 209.

7.3.4. Test samples shall be compacted to the required dimensions and air void content in accordance with AASHTO T 312.

7.3.5. Compacted specimens shall be left at room temperature (approximately 77°F [25°C]) to allow the entire specimen to cool.

7.3.6. The bulk specific gravity of the test specimens shall be determined in accordance with AASHTO T 166.

7.3.7. The air void contents of the test specimens shall be calculated in accordance with AASHTO T 269.

7.4. Plant-Produced Asphalt Mixture Specimens

7.4.1. Samples of plant-produced mixtures shall be obtained in accordance with AASHTO R 97. Enough material shall be sampled to determine Gmm in accordance with AASHTO T 209 and to fabricate a minimum of three specimens with the required dimensions and air void content.

7.4.2. Mixture samples shall be reduced to the appropriate test size in accordance with AASHTO R 47.

7.4.3. The test sample shall be placed in a container and brought to the compaction temperature. The thickness of the sample in the container shall conform to the requirements of AASHTO R 30.

7.4.4. The maximum specific gravity of the test mixture shall be determined in accordance with AASHTO T 209.

7.4.5. Test samples shall be compacted to the required dimensions and air void content in accordance with AASHTO T 312.

7.4.6. Compacted specimens shall be left at room temperature (approximately 25°C [77°F]) to allow the entire specimen to cool.

7.4.7. The bulk specific gravity of the test specimens shall be determined in accordance with AASHTO T 166.

7.4.8. The air void contents of the test specimens shall be calculated in accordance with AASHTO T 269.

8. Procedure

8.1. Measure and record the specimen height and diameter in mm at four (4) and two (2) evenly spaced locations around the circumference, respectively, to the nearest 0.1 mm.

8.2. Precondition test specimens in a water bath at a temperature of 54.4 ± 0.5°C for1 hour ± 10 minutes prior to testing. The water bath shall be at the specified temperature before the test specimens are placed in the bath. The water temperature shall be maintained at the test temperature during the course of conditioning. The test specimens shall be placed on a perforated bottom and fully submerged with at least 25 mm (1 in) of water above the specimens.

8.3. Ensure that the loading strips are parallel and contact surfaces are clean and free of debris.

8.4. Place the specimen in the fixture and ensure that the specimen is centered on its vertical diametrical plane and makes uniform contact with the loading strips.

8.5. Apply a constant rate of axial displacement at 50 ± 3 mm/min. Stop the test when the load drops below 100 N.

8.6. Complete the testing in 2 minutes or less after the specimens are removed from the water bath.

8.7. Record the time, load, and displacement at a minimum sampling rate of 40 data points per second.

9. Calculation

9.1. Calculate the IDT-HT strength using the following equation.

$$S_t = \frac{2000 \times P}{\pi \times t \times D}$$

where:

S_t = IDT-HT strength, kPa
P = maximum load, N
t = specimen height, mm
D = specimen diameter, mm.

9.2. Example calculation:

The average diameter of a specimen: 150.1 mm
The average height of a specimen: 62.1 mm
The maximum load of a specimen: 1500 N.

$$S_t = \frac{2000 \times 1500}{\pi \times 62.1 \times 150.1} = 102.5\ kPa$$

10. Report

10.1. The test report shall include the following information:

10.1.1. JMF number.

10.1.2. Laboratory name, technician name, and date of test.

10.1.3. Mixture type (e.g., surface mixture) and description (e.g. designation, NMAS).

10.1.4. Specimen type (e.g., laboratory-produced or plant-produced).

10.1.5. Average measured height and diameter of each specimen and average of all specimens to the nearest 0.1 mm.

10.1.6. Air void content of each specimen and average air void content of all specimens.

10.1.7. Test temperature to the nearest 0.1°C.

10.1.8. Axial displacement rate used in the test to the nearest mm/min.

10.1.9. Maximum load for each specimen and average maximum load for all specimens to the nearest 0.1 kN.

10.1.10. IDT strength for each specimen and average IDT strength for all specimens to the nearest 1 kPa

10.1.11. Average coefficient of variance to the nearest 0.1%.

11. Precision and Bias

11.1. The precision estimates and statements for this test method are not yet available.

11.2. No information can be presented on the bias of the procedure in this test method for measuring the tensile strength of asphalt mixtures because no material having a true reference value is available for comparison.