VTRC
Virginia Transportation Research Council

We Bring Innovation to Transportation

Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II

http://www.virginiadot.org/vtrc/main/online_reports/pdf/23-r3.pdf

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

ILKER BOZ, Ph.D.
Research Scientist

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

Final Report VTRC 23-R3

VIRGINIA TRANSPORTATION RESEARCH COUNCIL
530 Edgemont Road, Charlottesville, VA 22903-2454
vtrc.virginiadot.org

Standard Title Page - Report on Federally Funded Project

1. Report No.: FHWA/VTRC 23-R3	2. Government Accession No.:	3. Recipient's Catalog No.:

4. Title and Subtitle:
Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II

5. Report Date:
August 2022

6. Performing Organization Code:

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

8. Performing Organization Report No.:
VTRC 23-R3

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

10. Work Unit No. (TRAIS):

11. Contract or Grant No.:
118733

12. Sponsoring Agencies' Name and Address:

Virginia Department of Transportation	Federal Highway Administration
1401 E. Broad Street	400 North 8th Street, Room 750
Richmond, VA 23219	Richmond, VA 23219-4825

13. Type of Report and Period Covered:
Final

14. Sponsoring Agency Code:

15. Supplementary Notes:
This is an SPR-B report.

16. Abstract:

The indirect tensile cracking test (IDT-CT) is performed in accordance with ASTM D8225-19. This test method does not currently contain precision estimates and associated statements. This creates potential issues when test results are different among individual laboratories conducting testing on the same asphalt mixture. In 2020, Phase I of an IDT-CT interlaboratory study was conducted by researchers at the Virginia Transportation Research Council to establish precision estimates and statements for several indices associated with the IDT-CT (i.e., the cracking tolerance index, fracture strain tolerance index, strength [S_t], and cracking resistance index) through the evaluation of two asphalt mixtures. Phase I involved the evaluation of specimens fabricated and compacted by a third party laboratory and sent to participant laboratories for testing only.

The purpose of the current study (i.e., Phase II) was to build on the efforts undertaken as part of the Phase I study. The major objective of both phases was to determine acceptable variability and establish precision estimates and statements for IDT-CT results. Phase II also evaluated the impact of additional critical factors and their interactions on the IDT-CT results, such as specimen fabrication and preparation and equipment type (Phase II.1); specimen conditioning method (Phase II.2); and loading rate and data collection frequency (Phase II.3).

In Phase II.1, 24 of 50 participating laboratories submitted results (29 of 55 data sets, or about 53% of the submitted data) for both mixtures that were in full accordance with the requirements of ASTM D8225-19. As compared to Phase I.1, this was a significant decrease in the percentage of participants that submitted non-compliant data and a significant increase in the percentage of participants that submitted data that fully conformed with the requirements of ASTM D8225-19. Phase II.1 involved the evaluation of test results obtained from specimens that were fabricated by the participant laboratories from loose mixture that was produced and distributed by a third party laboratory along with detailed instructions for specimen fabrication and testing. The precision estimates for the IDT-CT indices were determined. The precision estimates for single-operator conditions were similar whether or not specimens of a given mixture were fabricated by the same laboratory. Specimen preparation introduced additional variability in the precision estimates for multi-laboratory conditions.

Phase II.2 showed that the indices tested were not dependent on the type of conditioning method used (i.e., dry vs. wet). Moreover, Phase II.3 showed that the indices were not dependent on the loading rate applied within a range of 50 ± 3 mm/min or on the frequency of data collection.

The study recommends that the Virginia Department of Transportation (VDOT) (1) include the developed precision estimates and statements in their balanced mix design specifications and adopt them for acceptance; (2) allow use of the wet conditioning method with the IDT-CT; (3) extend the allowable loading rate tolerance for the IDT-CT from 50 ± 2 mm/min to 50 ± 3 mm/min; (4) establish an annual proficiency testing program for the IDT-CT; and (5) routinely offer hands-on training and demonstrations of the laboratory tests being considered by VDOT as part of the balanced mix design initiative.

17 Key Words:
Balanced mix design, cracking, durability, performance testing, indirect tensile cracking test, cracking tolerance index, fracture strain tolerance index, indirect tensile strength, cracking resistance index, statistical analysis, device, quality, variability, reproducibility, repeatability, conditioning, impact.

18. Distribution Statement:
No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.

19. Security Classif. (of this report): Unclassified	20. Security Classif. (of this page): Unclassified	21. No. of Pages: 72	22. Price:

FINAL REPORT

INTERLABORATORY STUDY FOR THE INDIRECT TENSILE CRACKING TEST AT INTERMEDIATE TEMPERATURE: PHASE II

Jhony Habbouche, Ph.D., P.E.
Senior Research Scientist

Ilker Boz, Ph.D.
Research Scientist

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

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

August 2022
VTRC 23-R3

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 cracking test (IDT-CT) is performed in accordance with ASTM D8225-19. This test method does not currently contain precision estimates and associated statements. This creates potential issues when test results are different among individual laboratories conducting testing on the same asphalt mixture. In 2020, Phase I of an IDT-CT interlaboratory study was conducted by researchers at the Virginia Transportation Research Council to establish precision estimates and statements for several indices associated with the IDT-CT (i.e., the cracking tolerance index, fracture strain tolerance index, strength [S_t], and cracking resistance index) through the evaluation of two asphalt mixtures. Phase I involved the evaluation of specimens fabricated and compacted by a third party laboratory and sent to participant laboratories for testing only.

The purpose of the current study (i.e., Phase II) was to build on the efforts undertaken as part of the Phase I study. The major objective of both phases was to determine acceptable variability and establish precision estimates and statements for IDT-CT results. Phase II also evaluated the impact of additional critical factors and their interactions on the IDT-CT results, such as specimen fabrication and preparation and equipment type (Phase II.1); specimen conditioning method (Phase II.2); and loading rate and data collection frequency (Phase II.3).

In Phase II.1, 24 of 50 participating laboratories submitted results (29 of 55 data sets, or about 53% of the submitted data) for both mixtures that were in full accordance with the requirements of ASTM D8225-19. As compared to Phase I.1, this was a significant decrease in the percentage of participants that submitted non-compliant data and a significant increase in the percentage of participants that submitted data that fully conformed with the requirements of ASTM D8225-19. Phase II.1 involved the evaluation of test results obtained from specimens that were fabricated by the participant laboratories from loose mixture that was produced and distributed by a third party laboratory along with detailed instructions for specimen fabrication and testing. The precision estimates for the IDT-CT indices were determined. The precision estimates for single-operator conditions were similar whether or not specimens of a given mixture were fabricated by the same laboratory. Specimen preparation introduced additional variability in the precision estimates for multi-laboratory conditions.

Phase II.2 showed that the indices tested were not dependent on the type of conditioning method used (i.e., dry vs. wet). Moreover, Phase II.3 showed that the indices were not dependent on the loading rate applied within a range of 50 ± 3 mm/min or on the frequency of data collection.

The study recommends that the Virginia Department of Transportation (VDOT) (1) include the developed precision estimates and statements in their balanced mix design specifications and adopt them for acceptance; (2) allow use of the wet conditioning method with the IDT-CT; (3) extend the allowable loading rate tolerance for the IDT-CT from 50 ± 2 mm/min to 50 ± 3 mm/min; (4) establish an annual proficiency testing program for the IDT-CT; and (5) routinely offer hands-on training and demonstrations of the laboratory tests being considered by VDOT as part of the balanced mix design initiative.

FINAL REPORT

INTERLABORATORY STUDY FOR THE INDIRECT TENSILE CRACKING TEST AT INTERMEDIATE TEMPERATURE: PHASE II

Jhony Habbouche, Ph.D., P.E.
Senior Research Scientist

Ilker Boz, Ph.D.
Research Scientist

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

INTRODUCTION

Background

The Virginia Department of Transportation (VDOT) currently requires the use of the indirect tensile cracking test (IDT-CT) in its balanced mix design (BMD) specifications to evaluate the cracking resistance of dense-graded asphalt surface mixtures with A and D designations. These mixtures are usually designed to withstand traffic loads of 0 to 3 million and 3 to 10 million equivalent single axle loads, respectively. A performance criterion requiring a minimum cracking tolerance (CT) index of 70 has been recommended based on extensive mixture testing to lessen the cracking susceptibility of the asphalt mixtures (Diefenderfer and Bowers, 2019; Diefenderfer et al., 2021a; Diefenderfer et al., 2021b). This CT index was determined based on testing of compacted specimens from reheated plant-produced asphalt mixtures.

The IDT-CT is performed in accordance with ASTM D8225-19, Determination of Cracking Tolerance Index of Asphalt Mixtures Using the Indirect Tensile Cracking Test at Intermediate Temperature (ASTM, 2019a). This test method does not currently contain precision estimates and statements. This can lead to comparison concerns when different laboratories report different test results when testing the same asphalt mixture. Thus, in 2020, the Virginia Transportation Research Council (VTRC), in collaboration with the Virginia Asphalt Association, initiated the first phase of a round robin evaluation, or interlaboratory study (ILS), referred to herein as Phase I, in an effort to establish the acceptable variability of the test method. The CT index, defined in ASTM D8225-19 and currently used in VDOT's BMD specifications, and three additional performance indices (fracture strain tolerance [FST] index, strength [S_t], and cracking resistance index [CRI]) were evaluated as part of this effort (Boz et al., 2021; Habbouche et al., 2021.

Review of Phase I

Phase I of the IDT-CT ILS was a two-stage study to develop precision estimates and statements for the CT index, FST index, S_t, and CRI through the evaluation of two asphalt mixtures (Habbouche et al., 2021). Phase I involved the evaluation of specimens fabricated and compacted by a third party laboratory and sent to participant laboratories for testing only.

Stage 1, referred to herein as Phase I.1, focused on non-VDOT laboratories and included 41 participants with 46 submitted data sets. The difference between the number of participants and number of submitted data sets was due to the fact that several laboratories received more than one set of test specimens per mixture to perform testing using devices from different manufacturers. The effects of the particular device used and the loading rate on the selected IDT-CT indices were also investigated.

Stage 2, referred to herein as Phase I.2, focused on VDOT laboratories and included eight participants with eight submitted data sets. Phase I.2 was performed 1 year after the completion of Phase I.1 and used similar specimens compacted and stored in a climate-controlled setup. For consistency, all IDT-CTs were performed at 25°C on dry specimens. However, challenges related to keeping the specimens dry when conditioned in a water bath using leak-proof plastic bags were reported.

All participants in both stages of the Phase I study were asked to submit the raw data files from their equipment in addition to reporting the CT index values for each specimen. These data files were used to perform data quality checks before any further analysis was conducted to determine if the tests performed were in accordance with ASTM D8255-19. More details related to the data quality checks can be found in the Phase I study report (Habbouche et al., 2021).

In Phase I.1, 3 of 41 participating laboratories were unable to perform testing because of machine-related issues (i.e., 3 of 46 data sets). In addition, 3 of 41 participating laboratories were unable to provide the raw data from their tested specimens (i.e., 3 of 46 data sets). Further, data quality checks found that 10 of 41 participating laboratories (i.e., 10 of 46 data sets) were unable to perform the test correctly, indicating a need for training. Moreover, 14 of 41 participating laboratories (i.e., 14 of 46 data sets) performed testing using equipment incapable of meeting the loading rate requirement of 50 ± 2 mm/min in ASTM D8225-19. This resulted in only 14 of 41 participating laboratories submitting results for both mixtures in full accordance with the requirements of ASTM D8225-19 (i.e., 16 of 46 data sets, or about 35% of the submitted data). Each participant laboratory tested a set of five replicate specimens for each of the two mixtures. The research team separated the received data into two data groups (or categories) to develop the precision estimates and statements for the CT index: Category (i), a data group fulfilling all the requirements of ASTM D8225-19, including the loading rate of 50 ± 2 mm/min (i.e., 16 data sets per mixture type), and Category (ii), a data group consisting of all submitted test results including those that did not satisfy the loading rate requirement of 50 mm/min (i.e., 30 data sets per mixture type, including the 16 data sets that fulfilled the loading rate requirement of 50 ± 2 mm/min). The data were analyzed using one of two approaches: (1) an *untrimmed* approach (i.e., original data), in which the five IDT-CT index values per mixture type were included in the analysis; and (2) a *trimmed* approach (i.e., trimmed data), in which the highest and lowest IDT-CT index values were eliminated from the data set per mixture, resulting in three replicate test results for each data set. Table 1 summarizes the precision estimates resulting from Phase I.1 for the four selected IDT-CT indices. The findings also indicated that the magnitudes of the CT index, FST index, and CRI were not dependent on the type of device used. Further, the initial findings suggested that slight deviations from the loading rate of 50 ± 2 mm/min did not significantly affect the calculated IDT-CT indices.

The data collected in Phase I.2 (i.e., the results from VDOT laboratories) showed a relatively higher variability than that found in Phase I.1. This was attributable to several factors, including the relative lack of operator experience; a need for training; and potential changes in the material properties during storing, handling, and the shipping and/or testing process. The data collected in Phase I.1 and Phase I.2 also revealed that there was no statistically significant impact of 1 year of climate-controlled storage of compacted specimens on the calculated IDT-CT indices.

Phase I recommended that a follow-up study (referred to herein as "Phase II") be conducted to assess the impact of the specimen preparation process on the repeatability (i.e., single operator variability) and reproducibility (i.e., multi-laboratory variability) of IDT-CT results through an ILS program with asphalt mixtures sent to participating laboratories in the loose mixture state. The recommendation to use loose asphalt mixture as part of an ILS program was suggested so that the actual state of practice during design and production could be better reproduced. Phase I also recommended assessing the impact of testing specimens after temperature conditioning in water (wet conditioning) as compared to testing specimens under dry conditions. Further, Phase I recommended assessing the effect of loading rate on the IDT-CT results in a more controlled manner; as a factor, it was not solely varied in Phase I.

Table 1. Summary of Precision Estimates of the IDT-CT Indices in Phase I.1

IDT-CT Index	Category	Approach	Statistical Parameter Considered	Precision Estimate Values	
				Single-Operator	Multi-Laboratory
CT index	(i) 16 data sets per mix	Original data	COV	18.3%	21.3%
		Trimmed data		11.2%	15.9%
	(ii) 30 data sets per mix	Original data		20.7%	21.9%
		Trimmed data		12.8%	16.9%
FST index	(i) 16 data sets per mix	Original data	Stdv	0.56	0.58
		Trimmed data		0.31	0.43
	(ii) 30 data sets per mix	Original data		0.58	0.61
		Trimmed data		0.34	0.44
S_t	(i) 16 data sets per mix	Original data	Stdv	49.6 kPa	106.0 kPa
		Trimmed data		32.7 kPa	103.0 kPa
	(ii) 30 data sets per mix	Original data		51.5 kPa	99.6 kPa
		Trimmed data		33.1 kPa	94.3 kPa
CRI	(i) 16 data sets per mix	Original data	Stdv	0.44 1/mm*10^4	0.47 1/mm*10^4
		Trimmed data		0.21 1/mm*10^4	0.29 1/mm*10^4
	(ii) 30 data sets per mix	Original data		0.43 1/mm*10^4	0.46 1/mm*10^4
		Trimmed data		0.22 1/mm*10^4	0.30 1/mm*10^4

CT = cracking tolerance; COV = coefficient of variation; FST = fracture strain tolerance; Stdv = standard deviation; S_t = strength; CRI = cracking resistance index.

PURPOSE AND SCOPE

The purpose of this Phase II study was to build on the efforts undertaken in Phase I of the IDT-CT ILS. The major objective of both phases was to determine the acceptable variability and develop precision estimates and statements for IDT-CT results. Phase II also evaluated the impact of additional critical factors and their interactions on IDT-CT results, such as specimen fabrication and preparation, equipment type, specimen conditioning method, loading rate, and data collection frequency. The scope of the study included three major parts:

1. *Part 1, referred to herein as Phase II.1, assessed the impact of variability induced by specimen preparation.* This involved the evaluation of test results obtained from specimens fabricated by the participant laboratories from loose mixtures, which were produced and sent by a third party laboratory along with detailed instructions for specimen fabrication and testing. Phase II.1 also evaluated the impact of generating a smaller number of replicate specimens to be tested on the developed precision estimates.

2. *Part 2, referred to herein as Phase II.2, assessed the impact of specimen conditioning method (i.e., dry vs. wet) on the test results and evaluated the feasibility of performing the IDT-CT on wet specimens*.

3. *Part 3, referred to herein as Phase II.3, assessed the impact of loading rate and data collection frequency on the test results.*

METHODS

Asphalt Mixtures

The two mixtures, Mixture A and Mixture B, designed and evaluated in Phase I were also evaluated in Phase II. Mixture A was a 65-gyration 9.5 mm nominal maximum aggregate size (NMAS) Superpave mixture. Mixture A was produced using a performance grade (PG) 76-22 asphalt binder and a reclaimed asphalt pavement (RAP) content of 30%. Mixture B was a 50-gyration 12.5 mm NMAS Superpave mixture. Mixture B was produced using a PG 64-22 asphalt binder and did not contain RAP.

Two additional mixtures, Mixture C and Mixture D, were designed and evaluated as part of Phase II.2 and Phase II.3 along with Mixture A and Mixture B. Mixture C was a 65-gyration 9.5 mm NMAS Superpave mixture. Mixture C was produced using a PG 64-22 asphalt binder and a RAP content of 15%. Mixture D was a 65-gyration 9.5 mm NMAS Superpave mixture. Mixture D was produced using a PG 58-28 asphalt binder and a RAP content of 30%.

The four mixtures were designed and produced by a third party independent laboratory. Corresponding volumetric and gradation properties were reported.

IDT-CT Testing

The testing was conducted at 25 ± 0.5°C in accordance with ASTM D8225-19. The test load-displacement curve and specimen dimensions in terms of diameter and thickness were then used to calculate the CT index (Eqs. 1 and 2), FST index, S_t, and CRI (Eqs. 3 through 5).

$$CT\ index = \frac{G_f}{|m_{75}|} * \left(\frac{l_{75}}{D}\right) * (\frac{t}{62}) \quad \text{[Eq. 1]}$$

$$m_{75} = |\frac{p_{85} - p_{65}}{l_{85} - l_{65}}| \quad \text{[Eq. 2]}$$

$$FST\ index = \frac{G_f}{S_t} * 10^6 \quad \text{[Eq. 3]}$$

$$S_t = \frac{2000P_{max}}{\pi t D} * 10^3 \quad \text{[Eq. 4]}$$

$$CRI = \frac{G_f}{P_{max}} \quad \text{[Eq. 5]}$$

where

CT index = cracking tolerance index expressed in Equation 1
G_f = total area under the load-displacement curve divided by the product of the specimen thickness [t] and diameter [D], kN/mm
m_{75} = slope of interest expressed in Equation 2
p_{85}= 85% of the peak load (P_{max}) at the post-peak stage, kN
p_{75}= 75% of P_{max} at the post-peak stage, kN
p_{65}= 65% of P_{max} at the post-peak stage, kN
l_{85} = displacement corresponding to p_{85}, mm
l_{75} = displacement corresponding to p_{75}, mm
l_{65} = displacement corresponding to p_{65}, mm
FST = fracture strain tolerance expressed in Equation 3
S_t = indirect tensile strength expressed in Equation 4, kPa
CRI = cracking resistance index expressed in Equation 5
D = specimen diameter, mm
t = specimen thickness, mm.

Interlaboratory Study: Phase II.1

Phase II.1 included non-VDOT laboratories and VDOT district laboratories. The ILS was conducted in accordance with ASTM E691-19, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method (ASTM, 2019b).

The following defines four test measurement terms of interest according to ASTM E691-19 (ASTM, 2019b):

1. *Precision:* the closeness of agreements between independent test results obtained under stipulated conditions.

2. *Bias:* the difference between the expectation of the test results and an accepted reference value.

3. *Repeatability:* precision of test results from tests conducted within the shortest practical time period on identical material by the same test method in a single laboratory with all known sources of variability conditions controlled at the same levels.

4. *Reproducibility:* precision of test results from tests conducted on identical material by the same test method in different laboratories.

The experimental plan for the ILS study included the following:

- identification of qualified laboratories to participate

- full-scale execution, which included material preparation and handling, distribution, tracking of the testing progress of participant laboratories, and data inspection and examination

- statistical analysis of the generated data

- determination of precision estimates in accordance with ASTM E691-19

- development of precision statements in accordance with ASTM C670-15, Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials (ASTM, 2015); no statement on bias was made since there is no accepted reference material suitable for determining the bias in this test method

- reporting of individual results to each participant laboratory along with the study outcomes.

The requirements for each step as specified in ASTM E691-19 were met in Phase II.1 with the exception of including at least three materials representing different test levels for developing precision statements. Only two materials were included. Similar to Phase I, the reason for such an exception was to involve more laboratories rather than evaluate more materials. Supporting this, the analysis in Phase I revealed that the precision is relatively constant when compared to the average level over the range of values of interest.

Impact of Specimen Conditioning Method on IDT-CT Indices

Phase II.2 evaluated the feasibility of performing the IDT-CT on wet specimens. Currently, VDOT allows the use of the IDT-CT on only dry specimens conditioned in an environmental chamber or placed in leak-proof plastic bags in a water bath for 2 hours until the specimens reach a temperature of 25°C. In other words, the test specimens must remain dry during the conditioning process. Testing wet specimens consisted of placing specimens in a water bath for 2 hours, removing the specimens and drying them until they reached the saturate surface dry condition, and immediately performing the IDT-CT. For both methods, the IDT-CT should be completed within less than 4 minutes after the test specimen is removed from the conditioning environment.

Phase II.2 evaluated specimens fabricated and compacted by a third party laboratory and sent to participant laboratories along with detailed instructions for testing only. Five laboratories participated in Phase II.2: one independent testing laboratory (Lab 1), three VDOT district laboratories (Labs 2, 3, and 4), and the VTRC laboratory (Lab 5). Mixtures A and B were used in Phase II.2. All participants performed testing on two sets of specimens per mixture type, one set conditioned using the dry method and the other set conditioned using the wet method. All testing was performed using a similar servo-hydraulic device, Device I.

The five participant laboratories were provided with testing instructions and guidelines to ensure consistency in testing among all laboratories. These included guidelines for inspecting test specimens for any visual damages (e.g., cracks, creep, etc.) and instructions for measuring the diameter and thickness of each tested specimen. Instructions for how to condition the specimens in "dry" and "wet" conditions were also provided. All participant laboratories were asked to submit the raw data (time, load, and displacement measurements) for all tested specimens. The data quality was assessed for each specimen.

Impact of Loading Rate and Data Frequency Collection on IDT-CT Indices

Phase II.3 focused on assessing the impact of loading rate on the IDT-CT results. Phase II.3 involved the evaluation of specimens fabricated and compacted by an independent laboratory and tested by a single laboratory, in this case the VTRC laboratory. Four mixtures (Mixtures A through D) were used. Testing of all mixtures was performed using a servo-hydraulic (SH) machine (Device I) and a screw-drive (SD) machine (Device IV). Both devices were used to test each mixture. Enough specimens per mixture were fabricated to perform testing at five loading rates per machine: 46, 48, 50, 52, and 54 mm/min. Five specimens per mixture were tested at each of the considered five loading rates. All testing was performed on dry conditioned specimens at 25°C. The raw data, i.e., time, load, and displacement measurements, were processed and the data quality was assessed.

In addition, Phase II.3 attempted to evaluate the impact of data frequency collection on the IDT-CT selected indices. In that effort, a six-order polynomial function as shown in Equation 6 was fitted to the load-displacement data for each evaluated specimen. The time and displacement corresponding to when the applied load dropped to 0.1 kN after the peak load was reached were determined from the raw data. These data were used to artificially generate a load-displacement database using the fitted model of Equation 6 per evaluated specimen for each of the 12 considered frequencies: 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 Hz.

$$P = a_6 l^6 + a_5 l^5 + a_4 l^4 + a_3 l^3 + a_2 l^2 + a_1 l^1 \qquad \text{[Eq. 6]}$$

where

P = load, kN
l = displacement, mm
a_1, a_2, a_3, a_4, a_5, and a_6 = fitting parameters.

RESULTS AND DISCUSSION

Design Properties of Evaluated Mixtures

The four evaluated mixtures were produced by a third party independent laboratory. Tables 2 and 3 summarize the aggregate gradation and volumetric properties for the four mixtures, respectively. The mixtures were designed such that the CT index values determined for each mixture were spread out to ensure a wider applicability of the study to a range of CT index values from 50 to 300.

Table 2. Aggregate Gradations for Evaluated 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)	100.0	97.9	100.0	100.0
3/8 in (9.5 mm)	96.7	88.8	98.0	96.7
No. 4 (4.75 mm)	70.1	66.0	75.0	70.1
No. 8 (2.36 mm)	49.1	39.7	46.2	49.1
No. 16 (1.18 mm)	32.6	26.3	29.3	32.6
No. 30 (600 µm)	21.4	18.5	19.7	21.4
No. 50 (300 µm)	12.8	13.4	12.1	12.8
No. 100 (150 µm)	8.9	9.3	9.0	9.0
No. 200 (75 µm)	7.2	6.2	7.4	7.2

Table 3. Volumetric Properties for Evaluated Mixtures

Mixture ID	**Mixture A**	**Mixture B**	**Mixture C**	**Mixture D**
RAP Content, %	**30**	**0**	**15**	**30**
Asphalt Binder	**PG 76-22**	**PG 64-22**	**PG 64-22**	**PG 58-28**
Volumetric Property				
N_{design}, gyrations	65	50	65	65
NMAS, mm	9.5	12.5	9.5	9.5
Asphalt Binder Content, %	5.3	5.8	5.2	5.9
Rice SG (G_{mm})	2.511	2.723	2.492	2.473
Aggregate Bulk SG (G_{sb})	2.678	2.941	2.667	2.687
VTM, %	4.5	4.0	4.0	4.0
VMA, %	15.2	16.4	15.0	16.9
VFA, %	70.0	75.6	73.3	76.3
FA Ratio	1.57	1.28	1.56	1.29

RAP = reclaimed asphalt pavement; PG = performance grade; N_{design} = number of Superpave design gyrations; 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.

Interlaboratory Study, Phase II.1

Participant Laboratories

A total of 50 laboratories participated in Phase II.1. They consisted of VDOT district, VTRC, university, other DOT, contractor, and independent testing laboratories. In total, 55 pairs of buckets each for Mixture A and Mixture B were shipped to the laboratories and 55 corresponding data sets per mixture type were collected. Similar to Phase I.1, the difference between the number of participants and the number of collected data sets was due to the fact that several laboratories received more than one set of test specimens per mixture to perform testing using devices from different manufacturers.

In Phase II.1, 4 of 50 participating laboratories did not submit their testing results (i.e., 4 of 55 data sets). Four of 50 participating laboratories were also unable to provide the raw data from their tested specimens (i.e., 4 of 55 data sets). Further, 7 of 50 participating laboratories (i.e., 7 of 55 data sets) were unable to perform the test correctly or performed the tests on specimens outside the 7 ± 0.5% air-void range as specified in ASTM D8225-19. Moreover, 11 of 50 participating laboratories (i.e., 11 of 55 data sets) performed testing using equipment incapable of meeting the loading rate requirement of 50 ± 2 mm/min. This resulted in only 24 of 50 participating laboratories submitting results for both mixtures in full accordance with the requirements of ASTM D8225-19 (i.e., 29 of 55 data sets, about 53% of the submitted data). Each participant laboratory was asked to compact and test a set of at least five replicate specimens each for Mixture A and Mixture B. Figure 1 shows the breakdown of the submitted data status for Phase I.1 and Phase II.1. The significant increase in the percentage of participants submitting data fully conforming to the requirements ASTM D8225-19 is notable.

In total, 55 pairs of buckets and 10 devices (I though X) were evaluated. Device I and Device X were SH machines. and Device II through IX were SD machines. Overall, 60% and 40% of the data were collected from the SH and SD machine, respectively.

Testing Instructions and Data Quality Evaluation

Each participant laboratory received a pair of buckets, one labeled A and the other labeled B, both with the same assigned number (e.g., A1 and B1, A55 and B55). All participant laboratories were provided with testing instructions and guidelines to ensure consistency in testing among all laboratories. This included instructions to reheat and split loose mixtures, compact test specimens, and perform the IDT-CT. Moreover, participant laboratories were asked to determine the theoretical maximum specific gravity (G_{mm}), asphalt binder content by the ignition method, and gradation of the recovered aggregates from the ignition method for quality control purposes and identification of possible production errors and outliers.

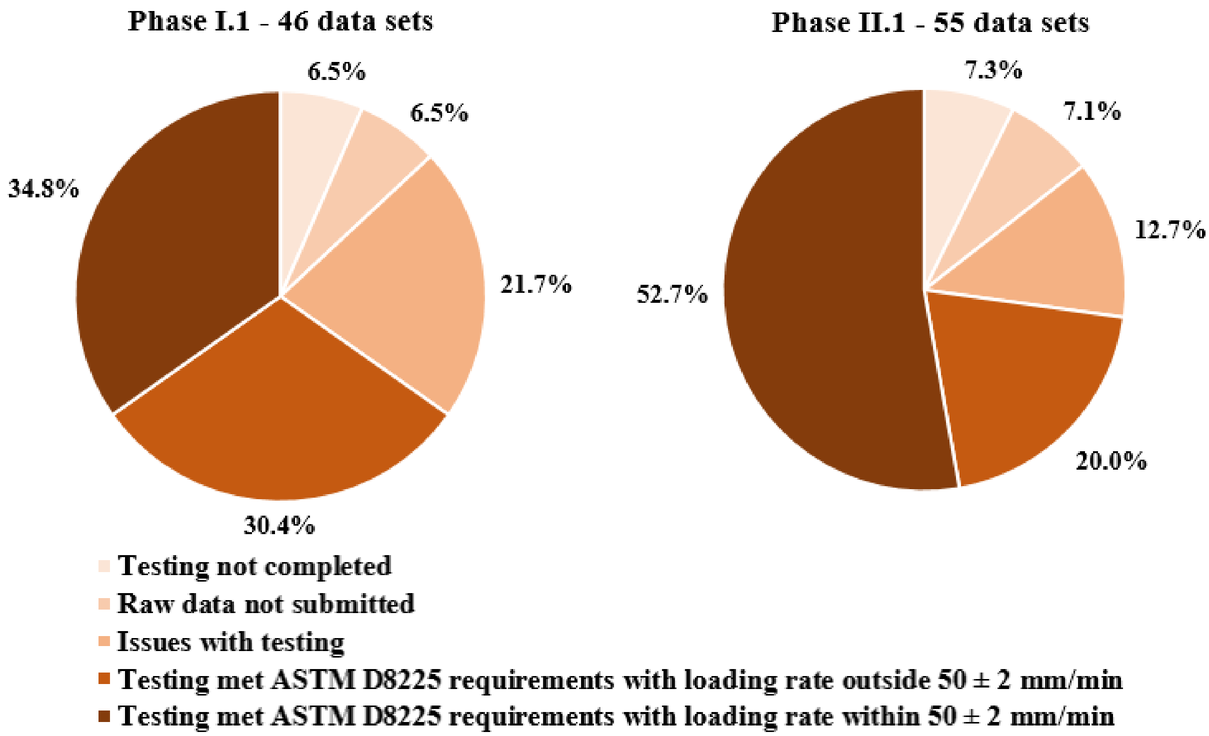

Figure 1. Status of Submitted Data for Phase I.1 and Phase II.1

All participant laboratories were provided testing instructions and guidelines summarized as follows:

- *Reheating loose mixtures.* The plastic handle must be removed from the bucket handle, and the rubber gasket must be removed from the rim of the bucket lid. The mixture bucket (with lid on top of it) should be placed in a forced draft oven set to the provided compaction temperature (i.e., 300 ± 5°F for Mixture A and 275 ± 5°F for Mixture B) for 3 to 3.5 hours prior to splitting.

- *Splitting reheated loose mixtures*. Mixtures should be split out to accommodate eight IDT-CT specimens, two G_{mm} samples, and one ignition sample. No further reheating of the mixture should be performed; all samples must be split after the first heating of the mixture bucket. All IDT-CT samples should be evenly distributed to a uniform thickness in pans. All pans should be labeled, covered with aluminum foil, and stored at room temperature until needed.

- *Compacting test specimens.* A weight of mixture must be determined so that all corresponding compacted IDT-CT specimens are within 7.0 ± 0.5%; thus, a starting weight for each mixture was provided to all participants. A trial IDT-CT specimen 150 ± 2 mm in diameter by 62 ± 1 mm in height should be compacted using the starting trial specimen weight. Once the trial specimen cooled completely, the bulk specific gravity was determined in accordance with AASHTO T 166, Method A. If the air voids of the trial specimen are not within 7.0 ± 0.5%, the starting weight should be adjusted using the Excel spreadsheet in the data form file entitled "Trial

Mass Calculation," which was previously shared with all participants, and another trial specimen should be compacted. Once the accurate weight of mixture needed to reach 7.0 ± 0.5% air voids is determined, at least five IDT-CT specimens within 7.0 ± 0.5% air voids should be compacted.

- *Performing* the *IDT-CT*. Each specimen's diameter should be determined by measuring to 0.1mm at two locations along the specimen. Moreover, each specimen's thickness should be determined by measuring to 0.1 mm at four locations around the specimen. All specimens must be dry prior to testing. All specimens must be conditioned at 25°C for at least 2 hours while the dry condition is maintained. The IDT-CT is then performed in accordance with ASTM D8225-19. Additional information such as testing date and time, data file name, CT index value, and raw data file must be reported.

- *Determining the theoretical maximum specific gravity (G_{mm}) of loose mixtures.* The G_{mm} for two replicates each for Mixtures A and B should be determined in accordance with AASHTO T 209.

- *Determining the asphalt binder content of asphalt mixtures by ignition method.* The asphalt binder content of Mixtures A and B should be determined by the ignition method in accordance with the standard that the participant's agency requires or follows (e.g., VTM 102, AASHTO T 308, etc.). No correction factors should be applied.

- *Conducting sieve analysis of recovered aggregates.* The gradation analysis of the recovered aggregates should be determined in accordance with AASHTO T 30.

All participants were asked to submit the raw data files from their equipment in addition to reporting the CT index values for each specimen. These data files were used for data quality checks before further analysis was performed to determine if the tests were performed in accordance with the requirements of ASTM D8255-19. More details related to the test data quality evaluation can be found in the Phase I report (Habbouche et al., 2021).

Cracking Test Results

The four indices considered in this study (i.e., CT index, FST index, S_t, and CRI) were evaluated using the collected and/or processed data. Each laboratory reported a minimum of five replicate measurements each for Mixture A and Mixture B. Similar to Phase I, two approaches, original untrimmed and trimmed, were used in performing data analysis. The original untrimmed approach uses all five replicate measurements of the collected data. The trimmed approach removes the minimum and maximum value of the selected index and uses the remaining three measurements for each mixture. Figure 2 presents the distribution of the individual CT index values for the original untrimmed data collected from all participant laboratories for Mixture A and Mixture B. The distributions of values for the FST index, S_t, and CRI are shown in Appendix A.

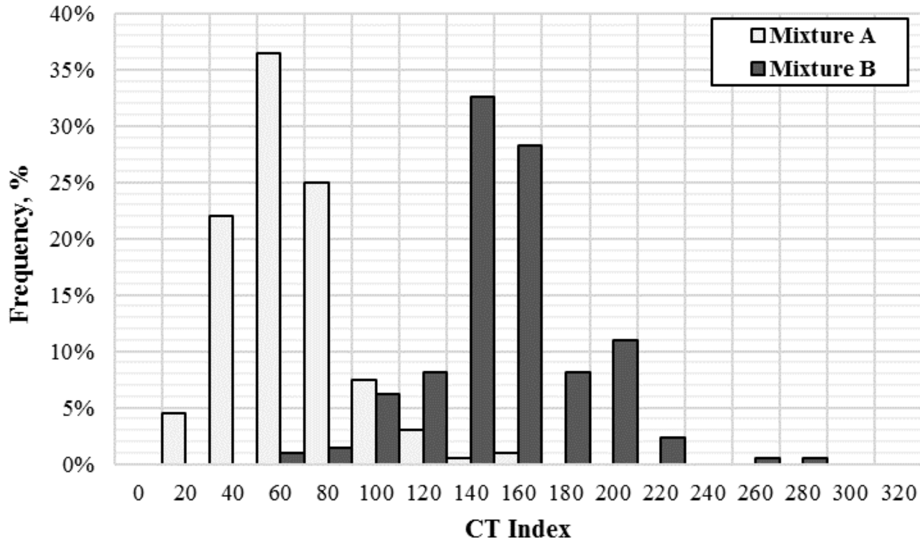

Figure 2. Distribution of CT Index Values for Mixture A and Mixture B Collected as Part of Phase II.1. CT = cracking tolerance.

Effect of Using Different Devices on IDT-CT Indices: Preliminary Assessment

Phase II.1 featured the use of 10 devices labeled Devices I through X with three participating laboratories capable of testing using multiple devices. This presented an opportunity to assess the impact of the device used on the four IDT-CT indices considered in this study. Two approaches were considered.

The first approach evaluated the data sets collected from all laboratories. An analysis of variance (ANOVA) at a 95% confidence interval was performed to determine if there was a statistically significant difference in the CT index results when different devices were used for testing. For the response variable, i.e., the CT index, the factors considered were Device; Laboratory, nested under the factor Device; and Mixture Type, nested under the factors Device and Laboratory. Table 4 presents the ANOVA statistics for the CT index using the data sets from Category (i) and Category (ii). Similar to Phase I.1, Category (i) consisted of the data sets fulfilling all requirements of ASTM D8225-19, including the loading rate of 50 ± 2 mm/min; Category (ii) consisted of the data group of all submitted test results including the ones that did not satisfy the loading rate requirement of 50 mm/min. As shown, Mixture Type was a statistically significant factor for the data sets of both categories for specific laboratory and device combinations.

Table 4. Summary of ANOVA Results for CT Index Using Original Untrimmed Data of Categories (i) and (ii)

Factor	Category (i)		Category (ii)	
	DF	**p-value**	**DF**	**p-value**
Device	6	0.000	8	0.000
Laboratory (Device)	25	0.000	36	0.000
Mixture Type (Device, Laboratory)	25	***0.000***	37	***0.000***

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; CT = cracking tolerance; DF = degree of freedom.

A pairwise comparison using the Bonferroni method indicated that, for a given mixture, a few laboratories with different device types had statistically different test results compared to the test result from the other laboratories. It was also observed that these laboratories with and without statistically similar test results shared the same device types. In other words, for example, Laboratory M with Device I fell under the statistically different test results group, but Laboratory N with Device I fell under the statistically similar test results group. This indicated that the source of statistical difference may not be the device type, but potentially the mixture and/or laboratory/operator variability. Thus, further analysis using a second approach was carried out to evaluate the impact of device type on the test results.

The second approach evaluated only the data sets of the three participating laboratories with multiple devices, Lab X, Lab Y, and Lab Z. Lab X and Lab Y were each provided two buckets per mixture to be tested using different devices by the same operator. The third participant, Lab Z, was provided with four buckets per mixture for the same reason. The analysis was performed only on the data compiled in Category (ii). Among the equipment used by the three laboratories, five devices were also identified and highlighted in the corresponding analysis: Device I (SH) (Lab X, Lab Y, and Lab Z); Device II (SD) (Lab X); Device III (SD) (Lab Y and Lab Z); Device IV (SD) (Lab Z); and Device VIII (SD) (Lab Z).

Both data analysis approaches, original untrimmed and trimmed, were considered. Figure 3, Figure 4, and Figure 5 show the average CT index of data reported by Lab X, Lab Y, and Lab Z, respectively. The average CT index results from the original untrimmed data were similar to the average CT index results from the trimmed data for all three laboratories (i.e., Lab X, Lab Y, and Lab Z). The results of a paired t-test at a 5% significance level on each pair of data points for each laboratory indicated no statistically significant differences between the CT index results before and after trimming, confirming the findings of Phase I.1 (Habbouche et al., 2021). There was a considerable decrease in the average coefficient of variation (COV) when the trimmed approach was employed. The COV for the original data with all data points of the three laboratories ranged from 5.9% to 27.9% with an average COV of 17.0%. The COV for the trimmed data ranged from 3.4% to 23.9% with an average COV of 10.9%.

An ANOVA at a 95% confidence interval was performed to determine if there was a statically significant difference in the CT index results when different devices were used for testing. For the response variable CT index, the parameters used as factors in the analysis model were Mixture Type and Device. An interaction term Mixture Type*Device was also added into the model as a factor. The assumption of normality was satisfied for the CT index data obtained from any of the three laboratories. Table 5 presents the ANOVA statistics for the CT index for the three laboratories, Lab X, Lab Y, and Lab Z, using both the original untrimmed and trimmed data. The data confirmed an expected observation that Mixture Type was a statically significant factor across the three laboratories for both original untrimmed and trimmed data. The factor Device was statistically significant for the trimmed data of Lab X and for both the original untrimmed and trimmed data for Lab Z.

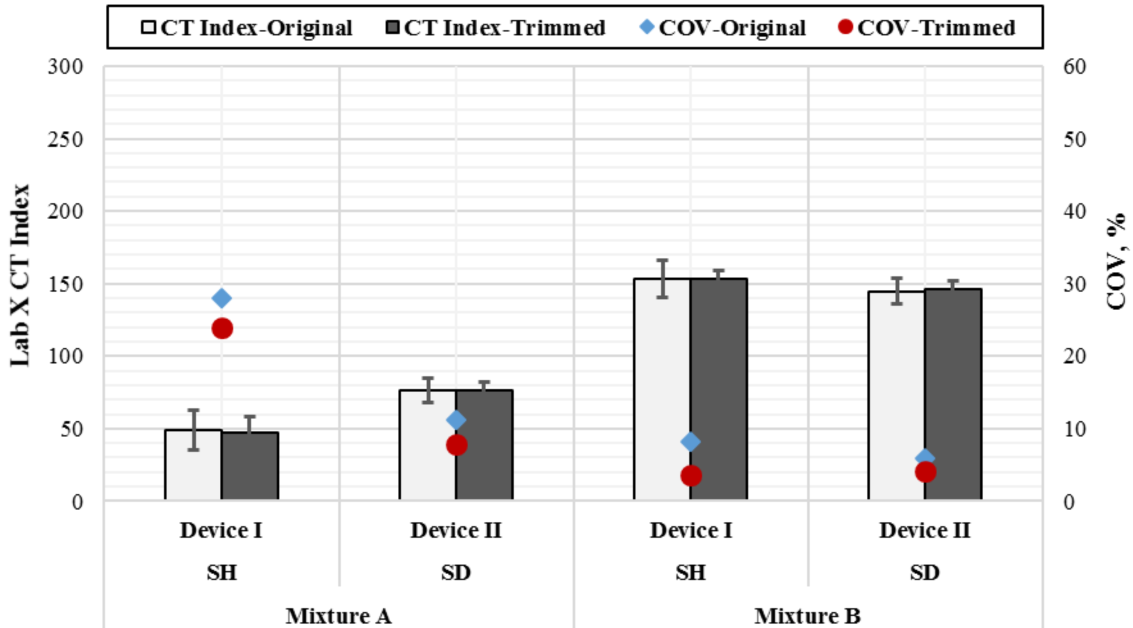

Figure 3. CT Index Values Reported by Lab X for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. CT = cracking tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

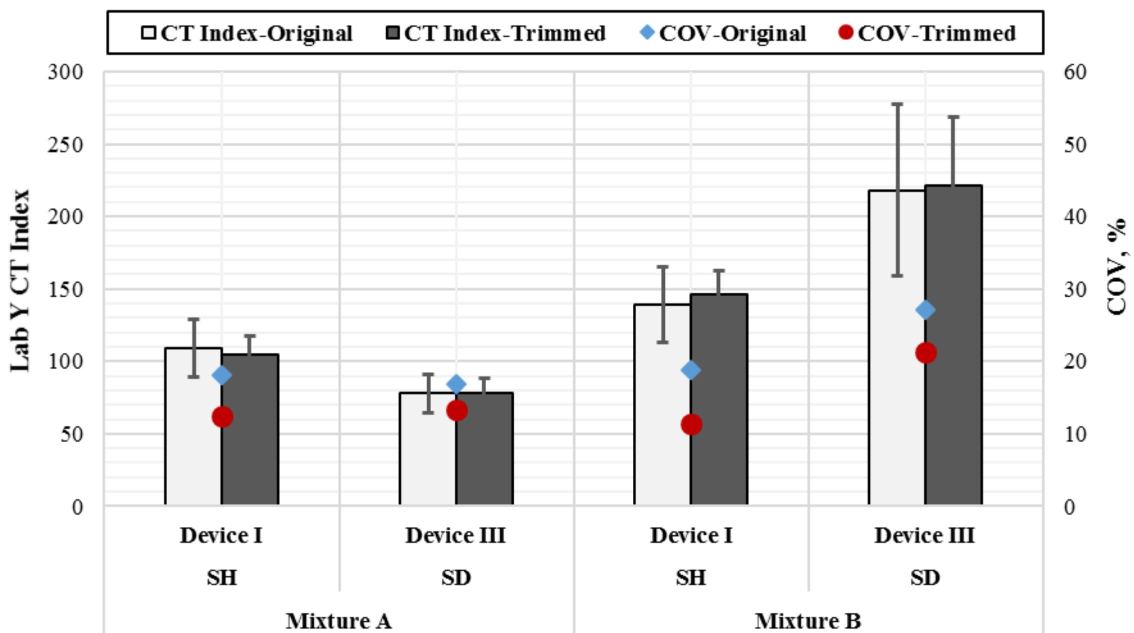

Figure 4. CT Index Values Reported by Lab Y for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. CT = cracking tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

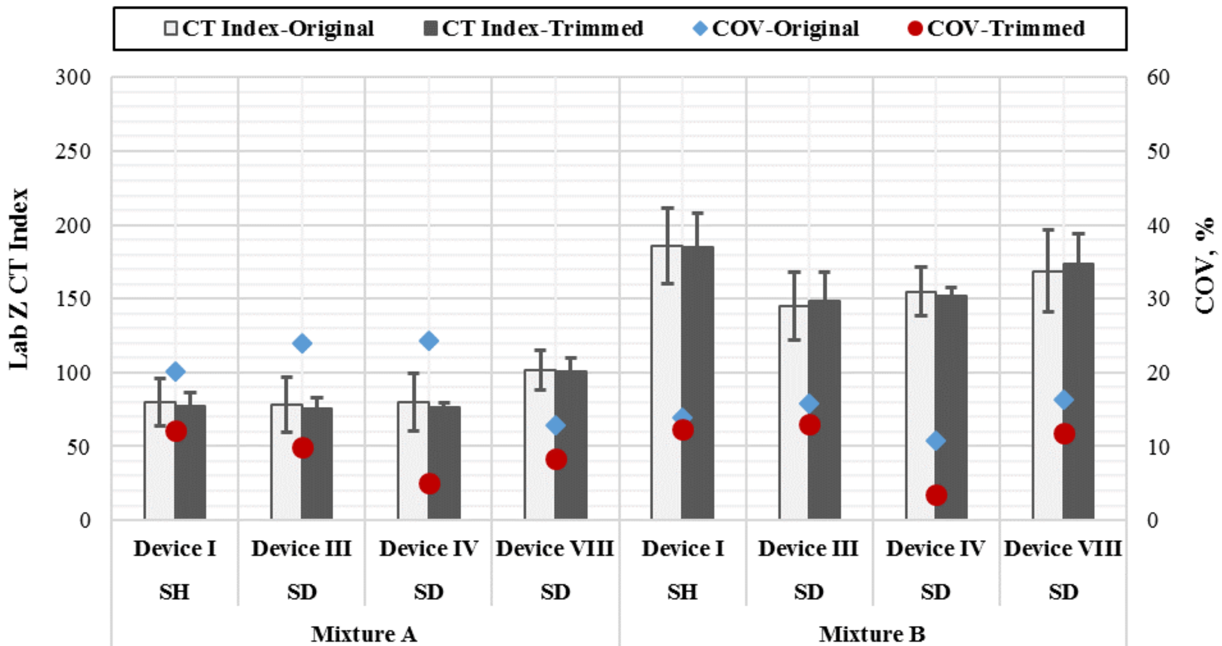

Figure 5. CT Index Values Reported by Lab Z for Mixture A and Mixture B Collected as Part of Phase II.1 Using Four Devices. I-bars show ±1 standard deviation. CT = cracking tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

Table 5. Summary of ANOVA Results for CT Index Using Original Untrimmed and Trimmed Data for Lab X, Lab Y, and Lab Z

Factor	Lab X		Lab Y		Lab Z	
	DF	**p-value**	**DF**	**p-value**	**DF**	**p-value**
Original Untrimmed Data						
Mixture Type	1	0.000	1	0.000	1	***0.000***
Device	1	0.080	1	0.142	3	***0.035***
Mixture Type*Device	1	***0.002***	1	***0.003***	3	0.131
Trimmed Data						
Mixture Type	1	0.000	1	0.000	1	***0.000***
Device	1	0.041	1	0.148	3	***0.014***
Mixture Type*Device	1	***0.003***	1	***0.010***	3	0.123

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; DF = degree of freedom.

A pairwise comparison using the Tukey method on all data subgroups indicated for Lab X that Device I and Device II resulted in statistically different CT index values for Mixture A for both original untrimmed and trimmed data. For Lab Y, Device I and Device II resulted in statistically similar CT index values for Mixture A for both original untrimmed and trimmed data. Finally, for Lab Z, the factor Device was not a statistically significant factor regardless of the mixture type and trimming. The results here suggest that device type might be a significant factor for mixtures with relatively low CT values, although further evaluation is necessary to confirm this observation.

Other IDT-CT Indices and Parameters

Similar to the CT index, evaluations were carried out to assess the impact of test device on the FST index, S_t, and CRI. The statistical analyses were limited to the original untrimmed data sets of Category (ii). The analyses are available in Appendix B. The same observations made for the CT index remain valid for the FST index, S_t, and CRI.

Development of Precision Estimates

ASTM E691-19 was used to determine the precision estimates for the four considered indices. Both approaches, original untrimmed and trimmed, were considered when the data were analyzed. Similar to Phase I.1, the data obtained as part of Phase II.1 were grouped and analyzed in two categories: Category (i) and Category (ii). Category (i) considered 29 data sets from laboratories that were able to perform the IDT-CT in full accordance with the requirements of ASTM D8225-19 including satisfying a loading rate of 50 ± 2 mm/min. Category (ii) considered 40 data sets collected from laboratories that were able to perform the IDT-CT in accordance with ASTM D8225-19 with the exception of satisfying the specified loading rate. The 40 data sets from Category (ii) included the 29 data sets from Category (i). Category (ii) was considered because small deviations from a loading rate range of 50 ± 2 mm/min appeared not to have a statistically significant impact on the four considered IDT-CT indices.

Selection of an Appropriate Statistical Parameter

In order to determine the precision estimates, the relationship between each IDT-CT index/parameter and the statistical parameters of interest for single-operator and multi-laboratory conditions were investigated for Categories (i) and (ii) of the data sets. The statistical parameters considered were the standard deviation (Stdv) and COV. Table 6 presents the variation of the Stdv and COV functions of the CT index collected in Phase I.1 and Phase II.1 for Mixtures A and B, for Categories (i) and (ii), and with the use of both the untrimmed original and trimmed approaches. The variation was evaluated by computing the Stdv and COV rate changes, and the parameter associated with the lowest absolute value rate change was selected. The Stdv or COV rate change was defined as the difference between the Stdv or COV of the CT index for Mixture B and the Stdv or COV of the CT index for Mixture A normalized by the average of the Stdv or COV of the CT index for Mixtures A and B. A positive rate indicates that the statistical parameter (Stdv or COV) increased with the increase of CT index value (as the CT index for Mixture B is greater than the CT index for Mixture A), and a negative value indicates the opposite. Figure 6 and Figure 7 illustrate the variation of Stdv and COV over the range of evaluated CT index values for data collected in Phase I.1 and Phase II.1, respectively. The data in Table 6, Figure 6, and Figure 7 indicated that the COV had a lower rate of change when compared with the Stdv for the same data sets for Mixtures A and B during both Phase I.1 and Phase II.1.

Table 6. Variation of Standard Deviation and Coefficient of Variation Functions of CT Index for Phase I.1 and Phase II.1

Phase	Category	Approach	Mix	Avg. Value	Single-Operator				Recommended Statistical Parameter	Multi-Laboratory				Recommended Statistical Parameter
					Stdv	**Stdv Rate**	**COV**	**COV Rate**		**Stdv**	**Stdv Rate**	**COV**	**COV Rate**	
I.1	(i)	Untrimmed	A	43.6	8.0	1.149	18.2	***0.005***	**COV**	10.5	0.963	24.1	***-0.258***	**COV**
			B	161.7	29.6		18.3			30.0		18.6		
		Trimmed	A	43.4	5.5	0.952	12.6	***-0.250***	**COV**	9.1	***0.605***	20.9	-0.637	**Stdv**
			B	157.1	15.5		9.8			17.0		10.8		
	(ii)	Untrimmed	A	44.8	10.6	0.907	23.7	***-0.290***	**COV**	11.5	0.861	25.6	***-0.343***	**COV**
			B	159.1	28.2		17.7			28.9		18.1		
		Trimmed	A	44.8	6.6	0.862	14.8	***-0.322***	**COV**	9.2	0.773	20.5	***-0.419***	**COV**
			B	155.7	16.6		10.7			20.8		13.4		
II.1	(i)	Untrimmed	A	63.1	15.4	0.475	24.4	***-0.398***	**COV**	24.0	***0.307***	38.1	-0.566	**Stdv**
			B	153.2	25.0		16.3			32.7		21.3		
		Trimmed	A	61.4	7.8	0.528	12.7	***-0.363***	**COV**	20.6	***0.239***	33.6	-0.640	**Stdv**
			B	152.1	13.4		8.8			26.2		17.3		
	(ii)	Untrimmed	A	65.2	15.6	0.514	23.9	***-0.332***	**COV**	23.7	***0.374***	36.4	-0.472	**Stdv**
			B	154.0	26.4		17.1			34.6		22.5		
		Trimmed	A	63.4	7.2	0.753	11.4	***-0.092***	**COV**	19.5	***0.378***	30.8	-0.489	**Stdv**
			B	153.2	15.9		10.4			28.6		18.7		

Avg. = average; Stdv = standard deviation; COV = coefficient of variation.

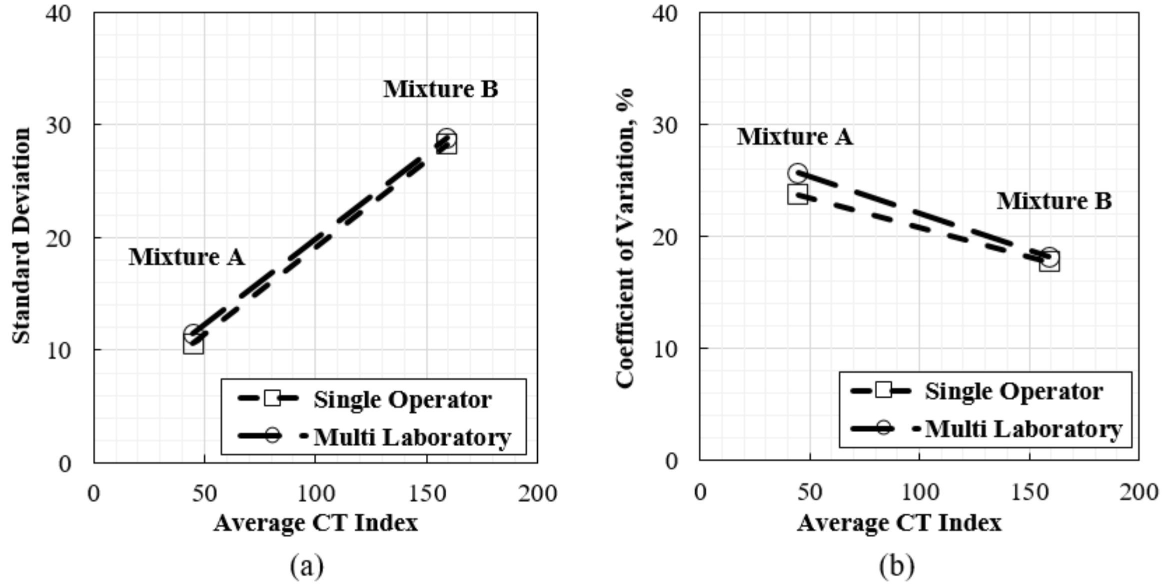

Figure 6. Variation of Statistical Parameters Over the Range of Evaluated CT index Values for Data Collected in Phase I.1: (a) standard deviation; (b) coefficient of variation.

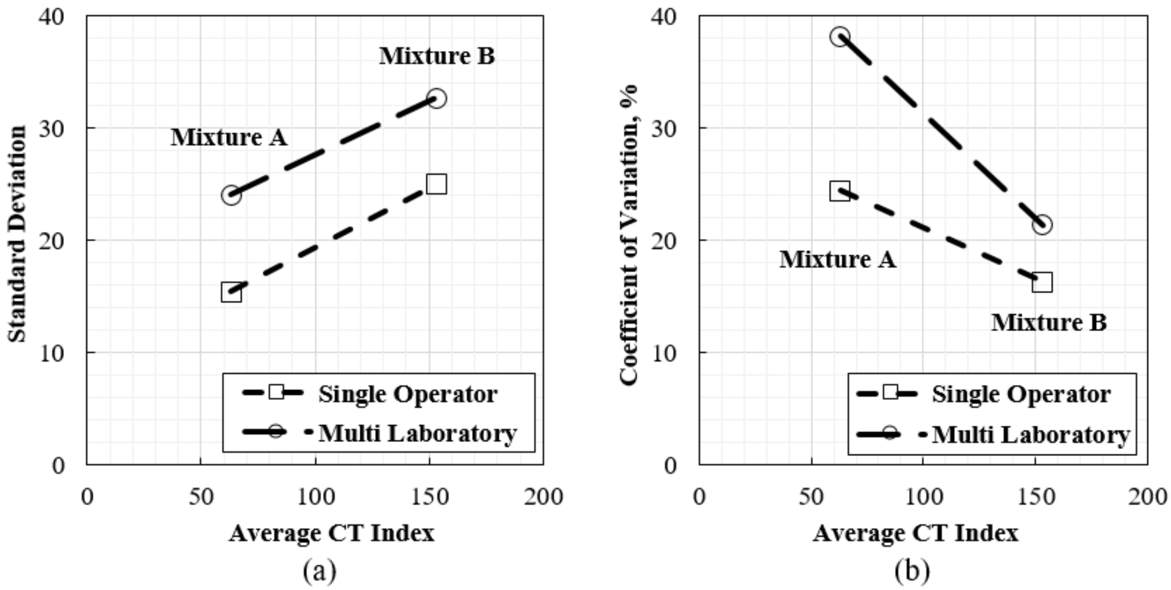

Figure 7. Variation of Statistical Parameters Over the Range of Evaluated CT index Values for Data Collected in Phase II.1: (a) standard deviation; (b) coefficient of variation.

Similar analyses were performed for the FST index, S_t, and CRI. The details of these analyses are available in Appendix C. In summary, the Stdv had an overall lower rate of change when compared with the COV for both the FST index and CRI. However, for S_t, the Stdv had a lower rate of change for the data sets of Phase I.1, and the COV had a lower rate of change for the data sets of Phase II.1.

Summary of Precision Estimates

Table 7 summarizes the precision estimates resulting from Phase II.1 for the four IDT-CT indices (i.e., CT index, FST index, S_t, and CRI). Trimming of the data resulted in a substantial decrease in Stdv and COV for both single-operator and multi-laboratory conditions. Overall, the Stdv and COV for both conditions were similar for the original untrimmed or trimmed approach regardless of the category ([i] or [ii]) and number of data sets (29 or 40) considered as part of the analysis.

Table 8 compares the precision estimates for COV for the CT index determined in Phase I.1 and those determined in Phase II.2. The precision estimates for single-operator conditions were similar for Phase I.1 and Phase II.1 across the same approach and regardless of the category. However the multi-laboratory precision estimates for Phase II.1 were greater than those for Phase I.1. This reflects the additional variability due to specimen preparation. Similar observations remain valid for the FST index, S_t, and CRI.

Table 7. Summary of Precision Estimates of the IDT-CT Indices Determined in Phase II.1

IDT-CT Index	Category and Condition		Statistical Parameter Considered	Precision Estimate Values	
				Single-Operator	Multi-Laboratory
CT index	(i) 29 data sets per mix	Original data	COV	20.4 %	29.7 %
		Trimmed data		10.7 %	25.4 %
	(ii) 40 data sets per mix	Original data		20.5 %	29.4 %
		Trimmed data		10.9 %	24.7 %
FST index	(i) 29 data sets per mix	Original data	Stdv	0.59	0.87
		Trimmed data		0.32	0.73
	(ii) 40 data sets per mix	Original data		0.59	0.86
		Trimmed data		0.34	0.72
S_t	(i) 29 data sets per mix	Original data	COV	5.6 %	11.8 %
		Trimmed data		3.2 %	11.0 %
	(ii) 40 data sets per mix	Original data		5.7 %	12.1 %
		Trimmed data		3.1 %	11.4 %
CRI	(i) 29 data sets per mix	Original data	Stdv	0.41 1/mm*10^4	0.60 1/mm*10^4
		Trimmed data		0.22 1/mm*10^4	0.50 1/mm*10^4
	(ii) 40 data sets per mix	Original data		0.40 1/mm*10^4	0.59 1/mm*10^4
		Trimmed data		0.23 1/mm*10^4	0.49 1/mm*10^4

CT = cracking tolerance; COV = coefficient of variation; FST = fracture strain tolerance; Stdv = standard deviation; S_t = strength; CRI = cracking resistance index.

Table 8. Comparison of Precision Estimates for the CT Index Determined in Phase I.1 and Phase II.1

Data Sets and Phase		Precision Estimate Values, COV, %	
		Single-Operator	Multi-Laboratory
Category (i): Data sets meeting the ASTM loading rate requirement of 50 ± 2 mm/min			
Original data	Phase I.1	18.3 %	21.3 %
	Phase II.1	20.4 %	29.7 %
Trimmed data	Phase I.1	11.2 %	15.9 %
	Phase II.1	10.7 %	25.4 %
Category (ii): Data sets *not* meeting the ASTM loading rate requirement of 50 ± 2 mm/min			
Original data	Phase I.1	20.7 %	21.9 %
	Phase II.1	20.5 %	29.4 %
Trimmed data	Phase I.1	12.8 %	16.9 %
	Phase II.1	10.9 %	24.7 %

COV = coefficient of variation.

Reduction of Specimen Replicates

With the move toward initial implementation of BMD and the continuous increase in the number of awarded contracts requiring the use of BMD mixtures, there is mutual interest from the industry and VDOT to consider generating a smaller number of replicate specimens for the IDT-CT. Currently, VDOT's BMD specifications require testing a set of five IDT-CT specimens compacted to an air-void range of 7.0 ± 0.5% during both the design and production stages. In addition, VDOT's BMD specifications currently require contractors to compact sets of five non-reheated IDT-CT specimens to be tested in their laboratories for quality control and during independent assurance sampling, sets of five non-reheated IDT-CT specimens to be tested by VDOT laboratories for quality assurance (QA).

This requirement is feasible during the mix design stage when contractors have enough time to produce sets of five IDT-CT replicate specimens fulfilling the requirements of ASTM D8225-19 in terms of diameter, height, air-void level, etc. However, the requirement is challenging during production as there is only one opportunity for success per sample. Excess sampled material cannot remain at temperature indefinitely to continue to compact acceptable specimens due to changes in the test results after excessive heating of the loose material. Further, additional sampling cannot be performed to collect material for additional specimens, as this material is considered a new sample.

This brings up several concerns. What if some of the IDT-CT specimens produced by the contractors to be used for quality control or QA did not meet the ASTM requirement of 7 ± 0.5%? What if some of the IDT-CT specimens were mishandled and damaged at any point? What if there were any machine- and/or operator-related issues during testing of the compacted specimens? Any of these situations could result in specimen sets consisting of fewer than five replicates for quality control or QA testing. Therefore, the impact on the precision estimates determined in Phase I.1 and Phase II.1 of using four or three replicates instead of five was evaluated. The minimum number of replicates was limited to three to maintain statistical soundness.

The five-replicate data sets from Phase I.1 and Phase II.1 were further evaluated using two scenarios. The first assessed combinations of four replicates randomly selected from each set of five replicates with no repetitions allowed. This resulted in five potential combinations per data set per mixture for each laboratory. The second scenario considered combinations of three replicates randomly selected from each set of five replicates with no repetitions allowed. This resulted in 10 potential combinations per data set per mixture for each laboratory. R-package was used in the analysis (R-package, 2022). Table 9 summarizes the precision estimates determined for the CT index when four and three replicates were analyzed of a set of five replicates from Phase I.1 and Phase II.1. Table 9 also compares these precision estimates to the ones already determined for the five-replicate data sets and those of three results obtained by trimming the high and low CT index values of each five-replicate data set. Precision estimates were similar for sets of five, four, or three untrimmed replicates for both categories and phases of the study. Therefore, the findings from this study remain applicable to the sets of four and three untrimmed replicates.

Table 9. Summary and Comparisons of Precision Estimates for the CT Index Using Phases I.1 and II.1 and Considering Various Numbers of Replicates per Data Set

No. of Replicates	Description	Precision Estimates, COV, %			
		Category (i)		Category (ii)	
		Single-Operator	Multi-Laboratory	Single-Operator	Multi-Laboratory
Phase I.1					
5	Data used as collected	18.3 %	21.3 %	20.7 %	21.9 %
4	Combinations of 4 of 5	18.4 %	20.7 %	20.9 %	22.2 %
3	Combinations of 3 of 5	18.3 %	20.8 %	20.8 %	22.1 %
3	After trimming	11.2 %	15.9 %	12.8 %	16.9 %
Phase II.1					
5	Data used as collected	20.3 %	29.7 %	20.5 %	29.3 %
4	Combinations of 4 of 5	20.3 %	29.4 %	20.5 %	29.1 %
3	Combinations of 3 of 5	20.3 %	29.4 %	20.5 %	29.1 %
3	After trimming	10.7 %	25.3 %	11.0 %	24.7 %

CT = cracking tolerance; COV = coefficient of variation.

Impact of Specimen Conditioning Method on IDT-CT Indices

Each laboratory reported five replicate measurements for each evaluated mixture. The obtained data were used to evaluate all four indices considered in this study: CT index, FST index, S_t, and CRI. Both approaches, untrimmed and trimmed, were used in performing data analysis.

Original Untrimmed Data

Figure 8 presents the average of five CT index values per each mixture for each of the participant laboratories. Table 10 shows the descriptive statistics of the CT index for Mixtures A and B and dry and wet conditioning modes. Figure 9 compares the average CT index value collected from the five participant laboratories for specimens conditioned using the dry and wet modes for both mixtures.

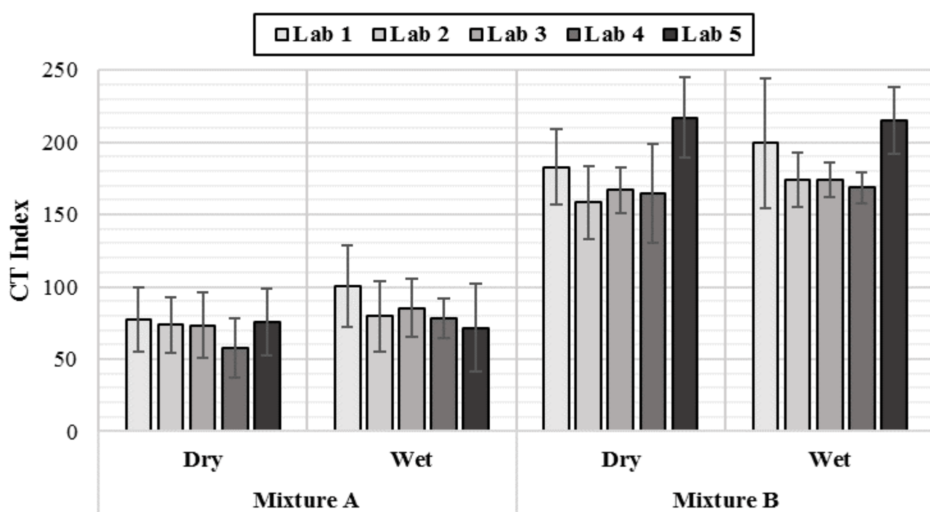

Figure 8. Average CT Index Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. CT = cracking tolerance; Lab = laboratory.

Table 10. Descriptive Statistics for CT Index Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min.	Q1	Q3	Max.	Range	IQR
Mixture A	Dry	72	38	58	83	115	76	25
	Wet	83	35	65	98	139	104	33
Mixture B	Dry	178	122	156	196	258	136	40
	Wet	186	146	169	203	254	108	35

CT = cracking tolerance; Q1 = quartile 1; Q3 = quartile 3; IQR = interquartile range.

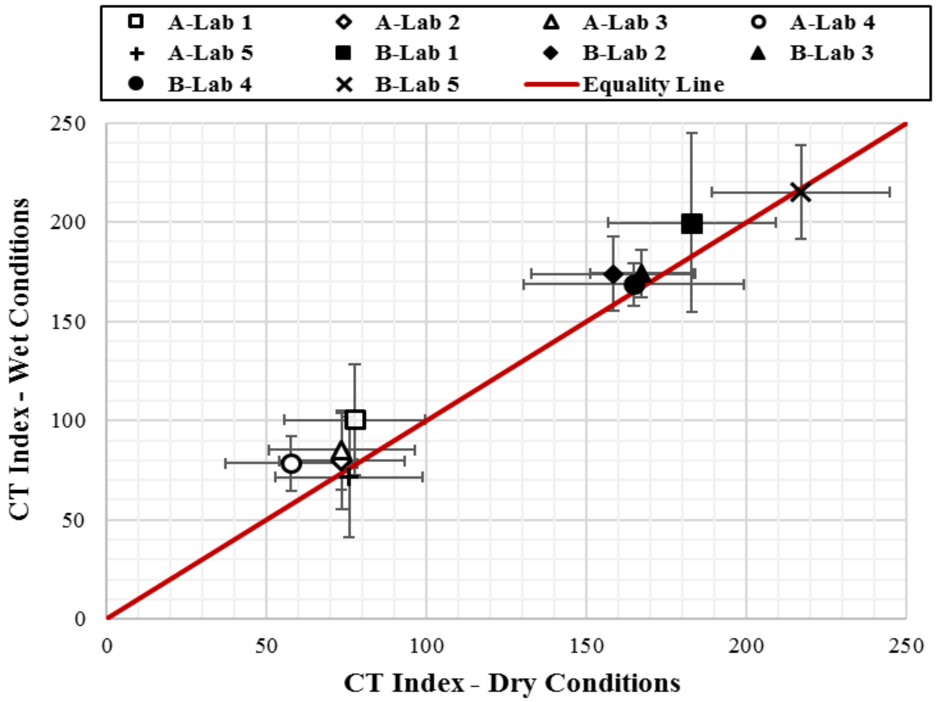

Figure 9. Dry Versus Wet CT Index Values of Original Untrimmed Data for Mixture A and Mixture B. I-bars show ±1 standard deviation. A = Mixture A; B = Mixture B; CT = cracking tolerance; Lab = laboratory.

Trimmed Data

Figure 10 presents the average of three CT index values per evaluated mixture for each of the participant laboratories after the high and low CT index values per data set were trimmed. Table 11 shows the descriptive statistics of the CT index with respect to mixture type (Mixture A and Mixture B) and conditioning mode (dry and wet). Figure 11 compares the average CT index value for specimens conditioned using the dry and wet modes for both mixtures by the five participant laboratories.

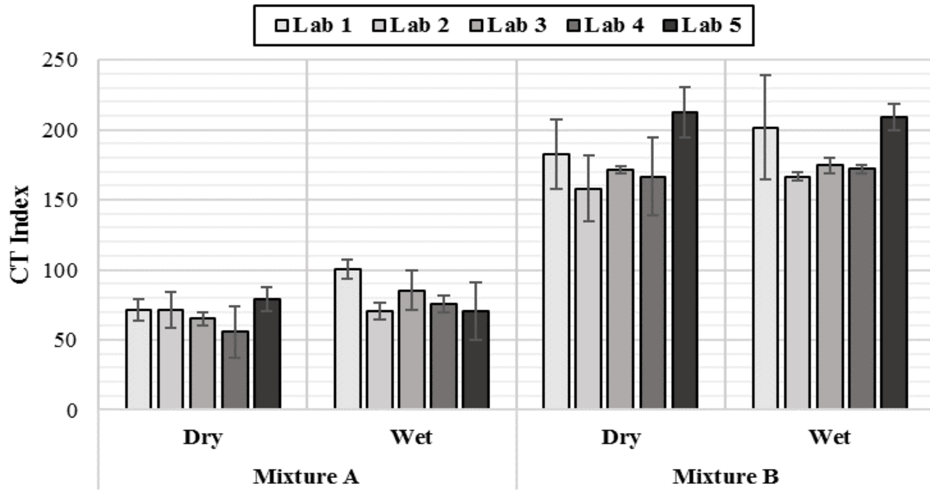

Figure 10. Average CT Index Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. CT = cracking tolerance; Lab = laboratory.

Table 11. Descriptive Statistics for CT Index Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min.	Q1	Q3	Max.	Range	IQR
Mixture A	Dry	68	44	63	78	86	42	16
	Wet	80	54	69	94	108	54	26
Mixture B	Dry	178	136	161	196	232	96	35
	Wet	185	164	170	193	245	80	24

CT = cracking tolerance; Q1 = quartile 1; Q3 = quartile 3; IQR = interquartile range.

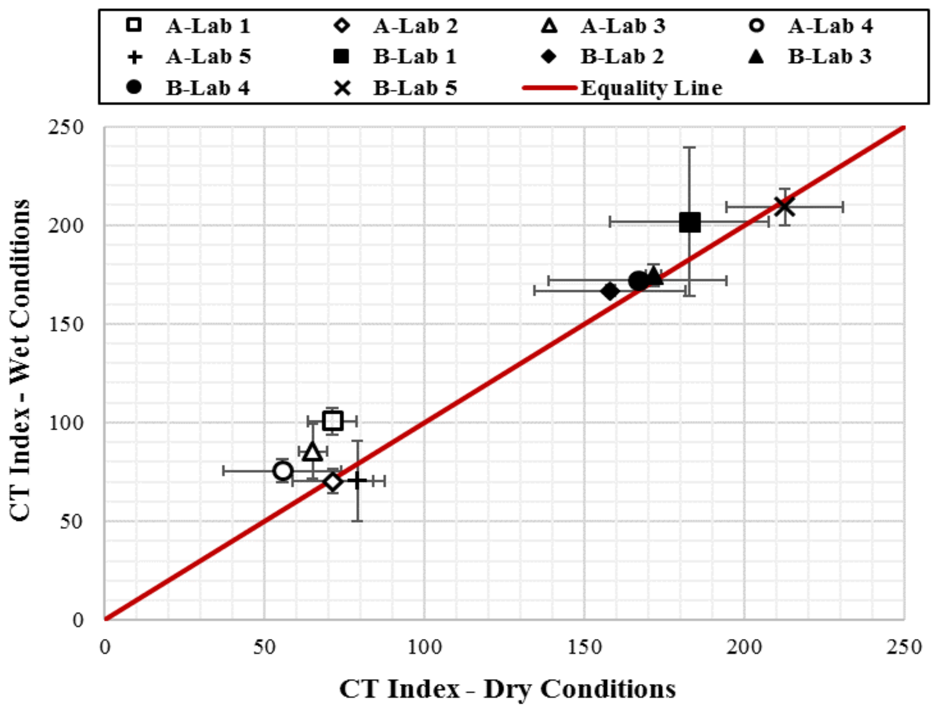

Figure 11. Dry Versus Wet CT Index Values of Trimmed Data for Mixture A and Mixture B. I-bars show ±1 standard deviation. CT = cracking tolerance; A = Mixture A; B = Mixture B; Lab = laboratory.

Statistical Analyses

For both the original and trimmed data, an ANOVA at a 95% confidence interval was performed to determine if there was a statically significant difference in the CT index results when the two different conditioning methods were used for testing. For the response variable, the CT index in this case, the parameters used as factors in the analysis model were Mixture Type, Conditioning Method, and Laboratory. Four interaction terms (Mixture Type*Conditioning Method, Mixture Type*Laboratory, Conditioning Method*Laboratory, and Mixture Type*Conditioning Method*Laboratory) were also added into the model as factors. The assumption of normality was satisfied for the dry and wet CT index data obtained from the participant laboratories regardless of the approach (untrimmed vs. trimmed). Minitab software was used in the analysis (Minitab, 2021). Table 12 presents the ANOVA statistics for the CT index. For both untrimmed and trimmed data, the test results were statistically different by conditioning method and by mixture type specific to a laboratory.

A pairwise comparison using the Tukey method, shown in Table 13, was performed to identify the statistically significant observations in the data. In the table, the observations that do not share a letter indicate a statistically significant difference in their means. Overall, regardless of the data type used (original or trimmed), the analysis indicated a statistical similarity of the CT index data collected using dry and wet conditioning methods for a given mixture and a laboratory. The statistical differences came solely from the test results for Mixture B from Lab 2 in the dry condition, and this specific observation differed from the test results of Mixture B from all other laboratories, whether or not the data were obtained in the dry or wet condition, although the test results for Mixture B from Lab 2 were the same for the dry and wet testing, as can be seen in Table 12. A similar analysis was performed in which testing data for Mixture B from Lab 2 under dry and wet conditions were removed. No statistical differences were found. Thus, the conditioning method is not a statistically significant factor for the CT index.

Table 12. Summary of ANOVA Results for CT Index of Dry and Wet Specimens Using Original Untrimmed and Trimmed Data

Factor	Untrimmed Data		Trimmed Data	
	DF	**p-value**	**DF**	**p-value**
Mixture Type	1	0.000	1	0.000
Conditioning Method	1	***0.046***	1	***0.031***
Laboratory	4	0.001	4	0.000
Mixture Type*Conditioning Method	1	0.747	1	0.509
Mixture Type*Laboratory	4	***0.006***	4	***0.017***
Conditioning Method*Laboratory	4	0.680	4	0.236
Mixture Type*Conditioning Method*Laboratory	4	0.935	4	0.789

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; CT = cracking tolerance; DF = degree of freedom.

Table 13. Summary of Tukey Pairwise Comparison for CT Index of Dry and Wet Specimens Using Original Untrimmed and Trimmed Data

Mixture Type*Conditioning Method*Laboratory	Untrimmed Data			Trimmed Data		
	N	Mean	Grouping	N	Mean	Grouping
Mixture A – Dry – Lab 4	5	57.6	a	3	55.6	a'
Mixture A – Wet – Lab 5	5	71.4	a	3	65.0	a'
Mixture A – Dry – Lab 3	5	73.4	a	3	70.3	a'
Mixture A – Dry – Lab 2	5	73.5	a	3	70.4	a'
Mixture A – Dry – Lab 5	5	75.8	a	3	71.1	a'
Mixture A – Dry – Lab 1	5	77.5	a	3	71.3	a'
Mixture A – Wet – Lab 4	5	78.4	a	3	75.4	a'
Mixture A – Wet – Lab 2	5	79.6	a	3	78.9	a'
Mixture A – Wet – Lab 3	5	85.4	a	3	85.5	a'
Mixture A – Wet – Lab 1	5	100.4	a	3	100.6	a'
Mixture B – Dry – Lab 2	5	158.3	b	3	158.0	b'
Mixture B – Dry – Lab 4	5	164.7	b c	3	166.5	b' c'
Mixture B – Dry – Lab 3	5	167.0	b c	3	166.7	b' c'
Mixture B – Wet – Lab 4	5	168.6	b c	3	171.4	b' c'
Mixture B – Wet – Lab 2	5	174.0	b c	3	171.9	b' c'
Mixture B – Wet – Lab 3	5	174.2	b c	3	174.5	b' c'
Mixture B – Dry – Lab 1	5	182.8	b c	3	182.7	b' c'
Mixture B – Wet – Lab 1	5	199.7	b c	3	201.6	b' c'
Mixture B – Wet – Lab 5	5	215.1	c	3	209.1	c'
Mixture B – Dry – Lab 5	5	217.0	c	3	212.5	c'

Lab = laboratory; N = number of elements; a, b, c = statistical grouping.

Other IDT-CT Indices and Parameters

Similar to the CT index, the analysis was carried out to assess the impact of the dry and wet specimen conditioning methods on the FST index, S_t, and CRI. The statistical analysis was limited to the original untrimmed data set. The distribution of collected data for original untrimmed data and trimmed data and the statistical analyses for untrimmed data for the FST index, S_t, and CRI are provided in Appendix D. Overall, Conditioning Method was statistically insignificant for the measured IDT-CT indices.

Impact of Loading Rate and Data Collection Frequency on IDT-CT Indices

The obtained data were used to evaluate all four indices considered in this study: CT index, FST index, S_t, and CRI. Both approaches, untrimmed and trimmed, were used in performing data analysis.

Loading Rates: Servo-Hydraulic vs. Screw-Drive Machines

Figure 12 compares the nominal loading rates to which the IDT-CT devices were set prior to testing and the exact measured loading rates computed using the data collected during testing. For the SH machine and as shown in Figure 12a, the nominal and measured loading rates were almost identical, indicating effective control of the applied load and rate of loading for the SH-based setup.

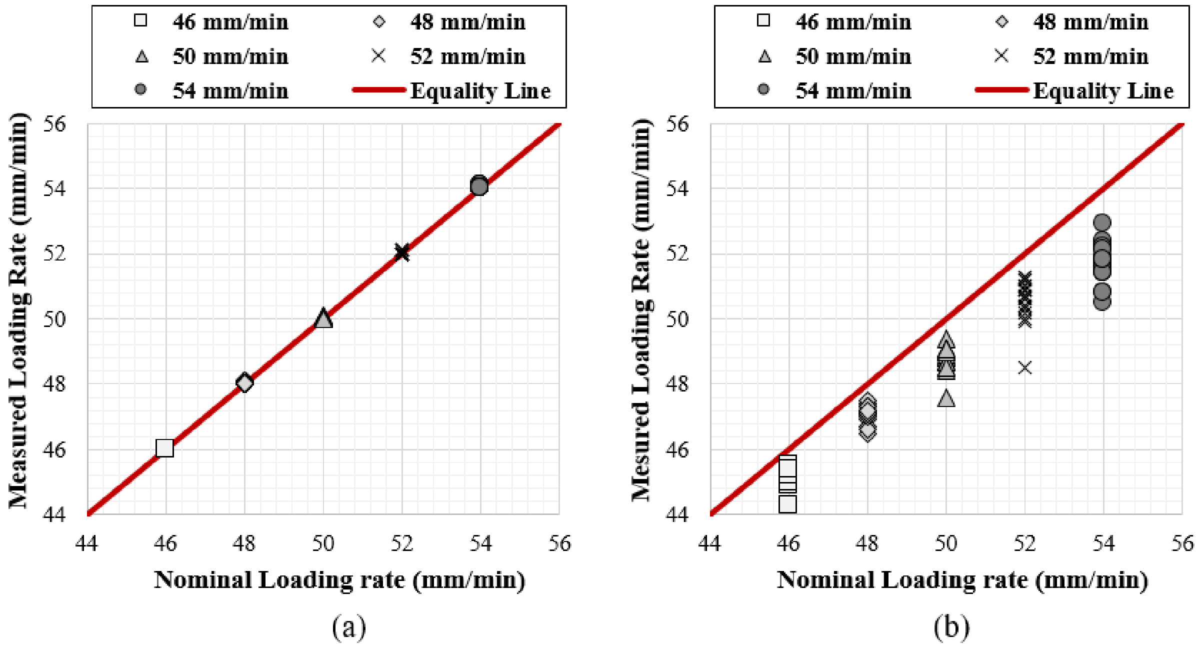

Figure 12. Comparison of Nominal and Measured Loading Rates for Evaluated Specimens of Mixtures A, B, C, and D: (a) using an SH machine; (b) using an SD machine. SH = servo-hydraulic; SD = screw-drive.

However, the measured loading rates were lower than the nominal loading rates for the SD machine, as shown in Figure 12b. Moreover, the difference in magnitude between the nominal and measured loading rates increased with the increase of the nominal maximum loading rate. Based on this, the effect of loading rate should be considered when an SD machine is used.

Analysis of Data Collected Using SH Machine

Figure 13 and Figure 14 present the original untrimmed and trimmed CT index values, respectively, for each evaluated mixture collected using an SH machine for each of the five loading rates. Table 14 presents the ANOVA statistics at a 95% confidence interval of the CT index with respect to mixture type (Mixtures A, B, C, and D) and loading rate (46, 48, 50, 52, and 54 mm/min) for both original untrimmed and trimmed data. The interaction factor Mixture Type*Loading Rate was considered in the analysis. Mixture Type was identified as a statistically significant factor ($p < 0.05$) for the original untrimmed and trimmed data. Loading Rate was identified potentially to be statistically significant for the trimmed data. A pairwise comparison using the Tukey method was performed on each of the original untrimmed and trimmed data sets using both the nominal and measured loading rates. Only the analysis performed using the nominal loading rate is shown in Table 15. In general, the observations that do not share a letter indicate a statically significant difference in their mean values, output, or index. Regardless of the combination, for each evaluated mixture, the average CT index at the five loading rates belonged to the same group, indicating that Loading Rate is not a statistically significant factor when an SH machine is used. Similar analyses were performed for the remaining IDT-CT indices. The data are provided in Appendix E. A similar conclusion was derived.

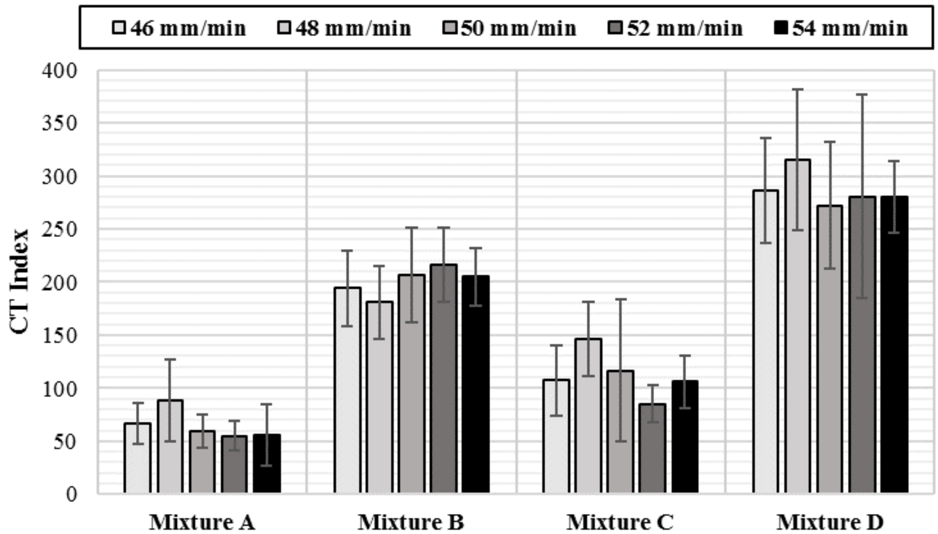

Figure 13. Average CT Index Values of Original Untrimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SH = servo-hydraulic.

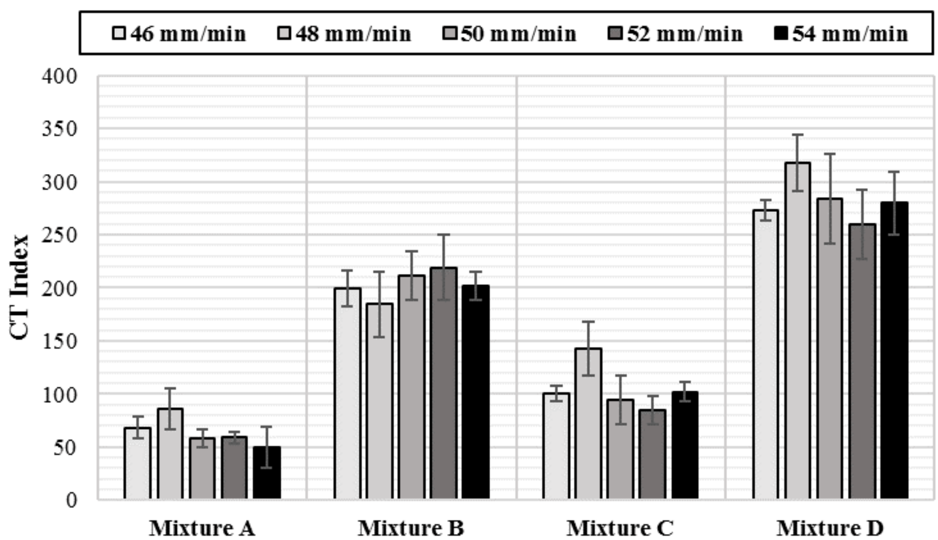

Figure 14. Average CT Index Values of Trimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SH = servo-hydraulic.

Table 14. Summary of ANOVA Results for CT Index of Mixtures A, B, C, and D Using Original Untrimmed and Trimmed Data at Five Different Loading Rates Using an SH Machine

Factor	Untrimmed Data		Trimmed Data	
	DF	p-value	DF	p-value
Mixture Type	3	***0.000***	3	***0.000***
Loading Rate	4	0.447	4	***0.029***
Mixture Type*Loading Rate	12	0.755	12	0.509

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; SH = servo-hydraulic; CT = cracking tolerance; DF = degree of freedom.

Table 15. Summary of Tukey Pairwise Comparison for CT Index of Mixtures A, B, C, and D Using Original Untrimmed Data at Five Different Loading Rates Using an SH Machine

Mixture Type*Loading Rate	Untrimmed Data		
	N	**Mean**	**Grouping**
Mixture A – 52.0	5	55.0	a
Mixture A – 54.0	5	55.6	a
Mixture A – 50.0	5	59.3	a
Mixture A – 46.0	5	66.7	a
Mixture C – 52.0	5	85.0	a b
Mixture A – 48.0	5	88.4	a b
Mixture C – 54.0	5	105.9	a b c
Mixture C – 46.0	5	107.3	a b c d
Mixture C – 50.0	5	116.6	a b c d e
Mixture C – 48.0	5	146.0	a b c d e
Mixture B – 48.0	5	181.1	b c d e f
Mixture B – 46.0	5	194.2	c d e f g
Mixture B – 54.0	5	204.8	c d e f g
Mixture B – 50.0	5	207.0	d e f g
Mixture B – 52.0	5	216.1	e f g h
Mixture D – 50.0	5	272.2	f g h
Mixture D – 54.0	5	279.9	f g h
Mixture D – 52.0	5	280.2	f g h
Mixture D – 46.0	5	286.0	g h
Mixture D – 48.0	5	315.0	h

CT = cracking tolerance; SH = servo-hydraulic; N = number of elements; a, b, c, d, e, f, g, h = statistical grouping.

Analysis of Data Collected Using SD Machine

Figures 15 and 16 present the original untrimmed and trimmed CT index values, respectively, for each evaluated mixture collected using an SD machine for each of the five loading rates. Table 16 presents the ANOVA statistics at a 95% confidence interval of the CT index with respect to mixture type and loading rate for both original untrimmed and trimmed data. In this case, Loading Rate was nested under Mixture Type. Loading Rate was not a statistically significant factor ($p < 0.05$) regardless of the mixture type and approach followed (untrimmed vs. trimmed). Similar to that described in the previous section, a pairwise comparison using the Tukey method was performed on each of the original untrimmed and trimmed data sets using the measured loading rates. Regardless of the combination, for each evaluated mixture, the average CT index at the five loading rates belonged to the same group, indicating that Loading Rate was not a statistically significant factor when an SD machine was used. Similar analyses were performed for the remaining IDT-CT indices. The data are provided in Appendix E. A similar conclusion was derived.

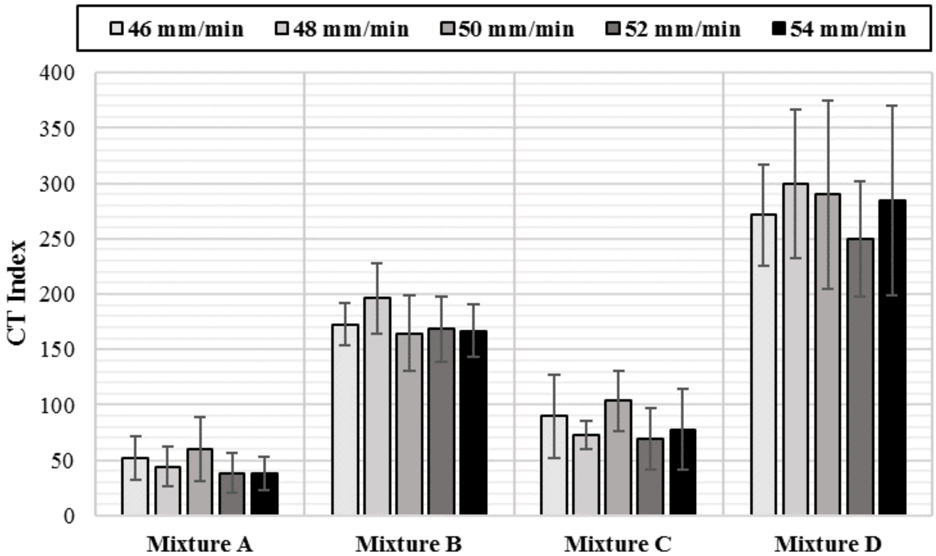

Figure 15. Average CT Index Values of Original Untrimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SD Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SD = screw-drive.

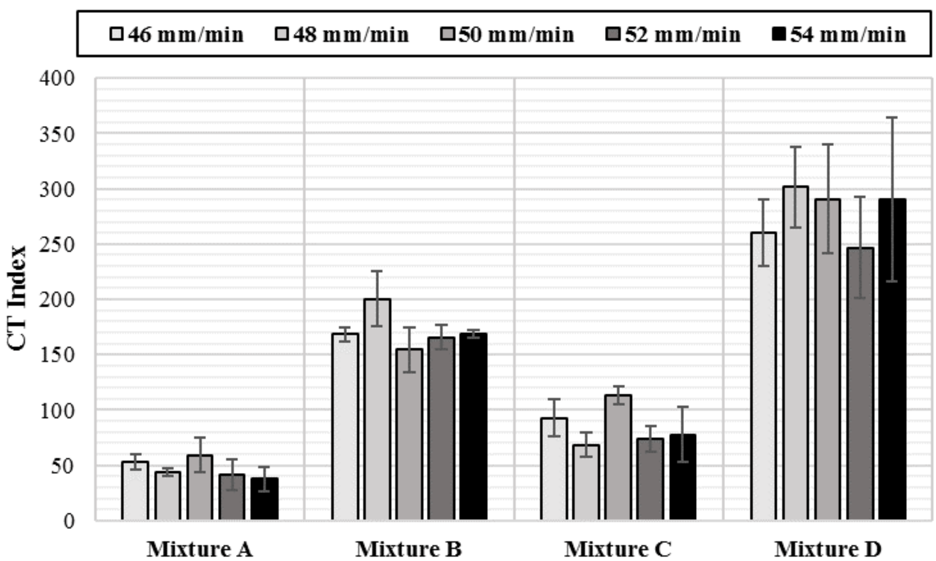

Figure 16. Average CT Index Values of Trimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SD Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SD = screw-drive.

Table 16. Summary of ANOVA Results for CT Index of Mixtures A, B, C, and D Using Original Untrimmed and Trimmed Data at Five Different Loading Rates Using an SD Machine

Factor	Untrimmed Data		Trimmed Data	
	DF	p-value	DF	p-value
Mixture Type	3	***0.000***	3	***0.000***
Loading Rate (Mixture Type)	69	0.792	46	0.514

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; CT = cracking tolerance; SD = screw-drive; DF = degree of freedom.

Impact of Device on the IDT-CT: Additional Analyses Using an Independent Data Set

The data collected as part of Phase II.3 present another opportunity to assess the impact of the device used on the considered IDT-CT indices. The data collected using an SH machine and an SD machine at a loading rate of 50 mm/min were evaluated. Table 17 presents the ANOVA statistics at a 95% confidence interval for the CT index when the factors Mixture Type and Machine Type and the interaction factor Mixture Type*Machine Type were considered. In this case, Machine Type was not a statistically significant factor for either the untrimmed general or trimmed data sets.

Table 17. Summary of ANOVA Results for CT Index Using Original Untrimmed and Trimmed Data Collected Using SH and SD Machines

Factor	Untrimmed		Trimmed	
	DF	p-value	DF	p-value
Mixture Type	3	***0.000***	3	***0.000***
Machine Type	1	0.563	1	0.505
Mixture Type*Machine Type	3	0.590	3	0.122

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; CT = cracking tolerance; SH = servo-hydraulic; SD = screw-drive; DF = degree of freedom.

Data Collection Frequency: Results and Statistical Analyses

The impact of data collection frequency on the four selected IDT-CT indices was evaluated using the data from Phase II.3. The analysis was performed using both the SH and SD machines. The analysis used the original untrimmed data collected when the IDT-CT was performed at a loading rate of 50 mm/min on Mixtures A, B, C, and D. Figure 17 shows the load-displacement data collected at a frequency of 50 Hz and the fitted model using Equation 6.

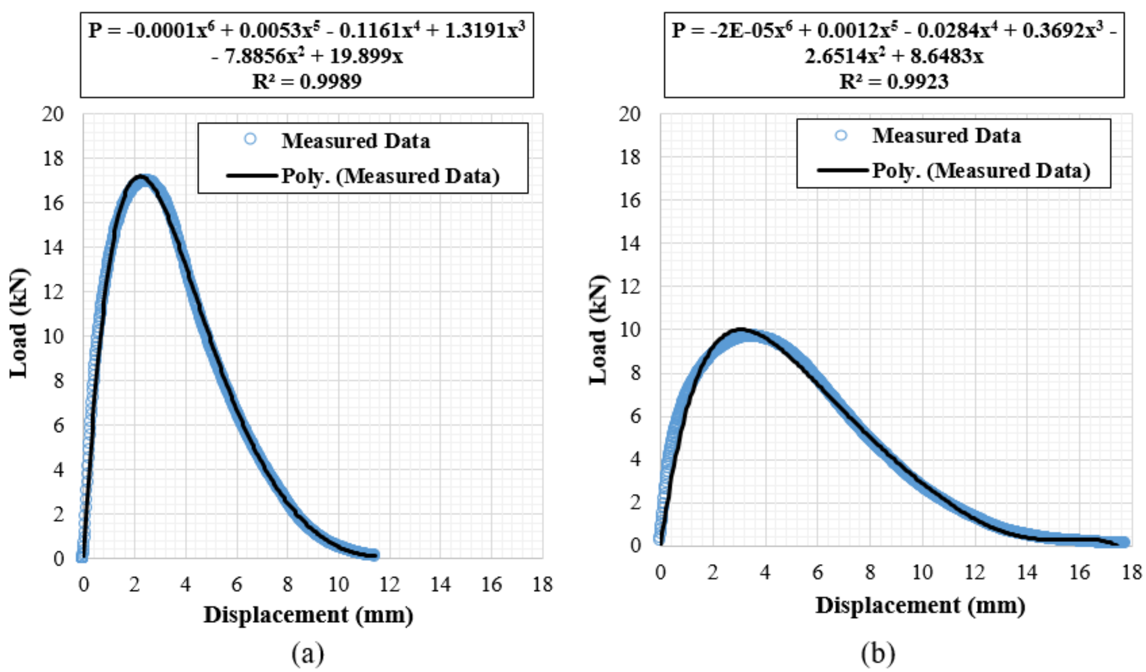

Figure 17. Examples of As-Collected and Fitted Load-Displacement Curves: (a) Mixture A; (b) Mixture D.

Figures 18 and 19 show the CT index determined for the four mixtures at 12 evaluated frequencies for data collected using the SH and SD machine, respectively. Overall, the data were statistically similar over a data collection frequency range of 1 to 100 Hz. Similar analyses were performed for the remaining IDT-CT indices. The data are provided in Appendix F. A similar conclusion was derived.

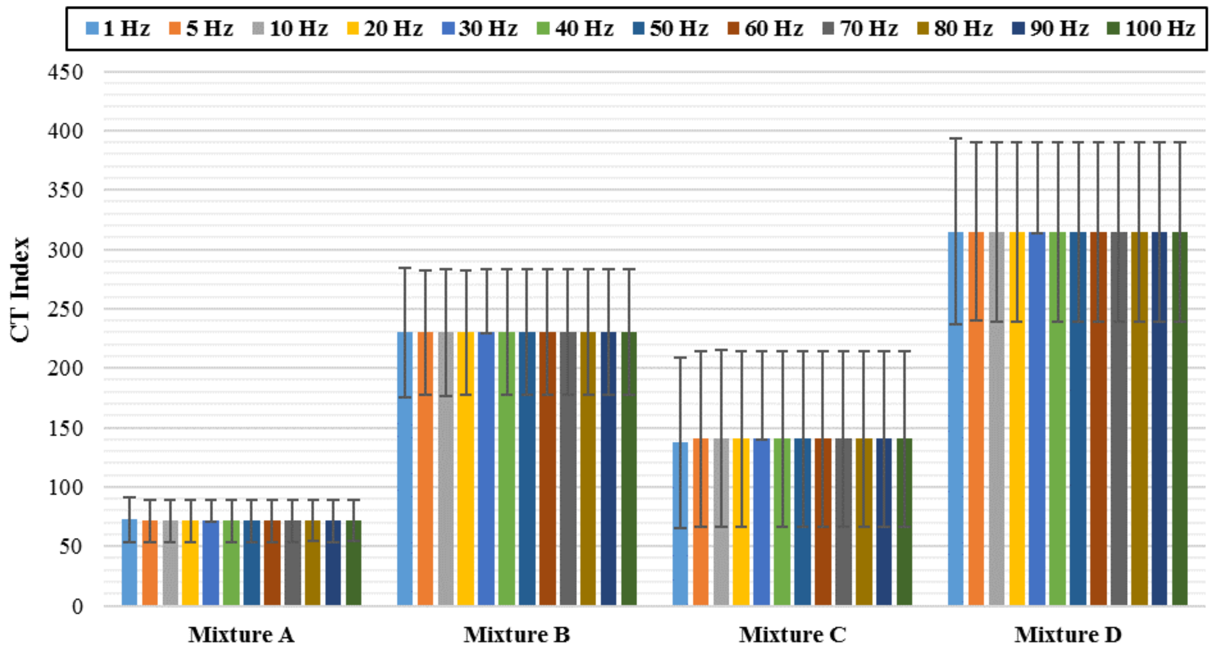

Figure 18. CT Index of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using the SH Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SH = servo-hydraulic.

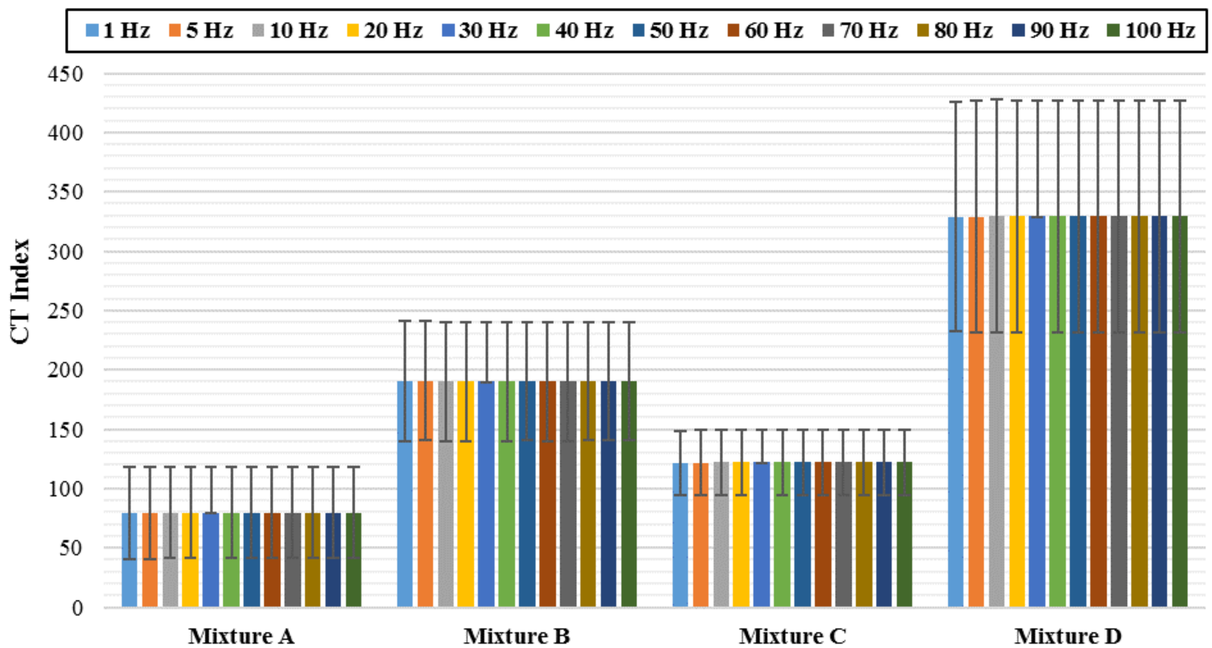

Figure 19. CT Index of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SD Machine. I-bars show ±1 standard deviation. CT = cracking tolerance; SD = screw-drive.

Precision Estimates in Design and Acceptance

The analyses performed as part of Phase I.1 and Phase II.1 confirmed that greater variability is introduced during specimen preparation for multi-laboratory conditions. Therefore, consistency in mixture processing and specimen fabrication is very important. In this section, a discussion is provided on how to apply and use precision estimates during mix design and acceptance. The layout uses the CT index since it is used in the current VDOT BMD specifications. In this discussion, the single-operator precision estimate is referred to as "SO" and the multi-laboratory precision estimate is referred to as "ML."

Mix Design Stage

During the mix design stage (D) and according to the current 2023 VDOT BMD specification, contractors must meet the requirement of a minimum CT index of 70 with a variability computed using the COV lower than SO. Prior to the determination and evaluation of the CT index, test data quality should be evaluated by the contractor. In addition, contractors must store the raw data of all tested IDT-CT specimens for possible review by VDOT at any time.

The following constitutes a guide for the industry/contractors to check both the average and variability of the data they submit for approval of the mix design by VDOT:

- *Case D.1*. The CT index threshold of 70 is met at optimum binder content (OBC) with a COV lower than SO. In this case, the mix design meets the CT index requirements for approval.

- *Case D.2*. The CT index threshold of 70 is met at OBC with a COV greater than SO. In this case, the mix design does not meet the criteria for approval. The contractor must produce, compact, and test a new set of five IDT-CT specimens. This process is repeated until both the CT index criteria are met and the COV of the evaluated set of specimens is less than SO. Then, the mix design meets the CT index requirements for approval. Mixing and matching specimens among various batches are not allowed as this negate the statistical validity of the test.

- *Case D.3*. The CT index threshold of 70 and/or a COV lower than SO were not met at OBC, so the design cannot be approved. The contractor must redesign.

Production Stage

During the production stage (P), the requirement of a minimum CT index of 70 remains valid when reheated IDT-CT specimens are being evaluated. In the case of non-reheated specimens, no thresholds are currently available. The following presents a guide for VDOT to accept a mix design during production based on the performance of reheated specimens. Similar to the requirements of Section 211.08 of the VDOT specifications (VDOT, 2018), acceptance of mixtures based on any performance testing (including the one evaluated using the IDT-CT) can be added to VDOT's QA program. This program includes the testing of production samples by the contractor and of monitor samples tested by VDOT.

For all cases, regardless of whether or not testing is performed by the contractor or VDOT, all data sets should have a COV lower than SO. In cases where the SO is not met, the reported results are invalid because of excess variability. The contractor and/or VDOT must immediately re-sample and re-test. If testing continues to fail variability, the contractor and/or VDOT should initiate an investigation to determine the reasons for the failure.

In the case of the CT index threshold of 70 not being met by the contractor or VDOT, VDOT and/or the contractor should initiate an investigation to determine the reasons for the failure.

In the case of the CT index threshold of 70 being met by the contractor and VDOT, a comparative statistical analysis should be conducted. If the COV between the data set of the contractor and that of VDOT is greater than ML, indicating a statistically significant difference, VDOT and the contractor should conduct an investigation to determine the reasons for the difference.

A Case Study: A Numerical Example

An example of applying variability requirements to CT index data using data collected from a paving contract is provided in this section. The intent of the example is to show how precision estimates can be numerically applied using data generated by the contractor during the mix design stage and data generated by the contractor and agency during the production stage.

During mix design, the CT index must be determined and reported at the design OBC (short- and long-term aged), at 0.5% above the design OBC (short-term aged), and at 0.5% below design OBC (short-term aged). The varying binder content testing is intended to provide an indication of mix design sensitivity to changes in the binder content. The data collected during the mix design stage are shown in Table 18. After each complete set of specimens was assessed, the following observations were made:

- *At OBC − 0.5%.* The average CT index of short-term aged specimens was 69, failing the criterion of 70; the COV was 37.2%, exceeding the acceptable single-operator COV of 18.3%.

- *At OBC.* The average CT index of short-term aged specimens was 137, passing the criterion of 70; the COV was 19.5%, exceeding the acceptable single-operator COV of 18.3%.

- *At OBC + 0.5%.* The average CT index of short-term aged specimens was 181, passing the criterion of 70; the COV was 37.2%, exceeding the acceptable single-operator COV of 18.3%.

- *At OBC.* The average CT index of long-term aged specimens was 36, 74% lower than the average CT index at OBC of short-term aged specimens; the COV was 18.7%, slightly exceeding the acceptable single-operator COV of 18.3%.

During production, the IDT-CT was performed by the contractor and VDOT district laboratories on non-reheated specimens, all compacted by the contractor. Moreover, loose mixtures were sampled by VDOT and VTRC during production and were tested at a later time. These specimens are referred to as "reheated." The data collected on non-reheated and reheated mixtures collected during production are shown in Table 19.

After each complete set of specimens was assessed, the following observations were made:

- *Non-reheated specimens.* The COVs of the data sets collected by the contractor and by VDOT were 7.7% and 16.9%, respectively, both meeting the acceptable single-operator COV of 18.3%. Moreover, the multi-laboratory COV between the contractor and VDOT data sets was 3.9%, meeting the acceptable multi-operator COV of 21.3% and indicating that these two data sets were statistically similar.

Table 18. Example of Variability in Data Set CT Index Values During Mix Design

Sample ID	Short-Term Aged Specimens			Long-Term Aged Specimens
	OBC − 0.5%	**OBC**	**OBC + 0.5%**	**OBC**
1	71.9	147.4	211.8	40.4
2	109.3	127.6	143.8	41.6
3	53.0	120.2	162.6	38.7
4	70.9	111.4	159.0	35.2
5	41.2	178.7	227.1	24.9
Analysis Using Original Untrimmed Data				
Average CT Index	69	137	181	36
Required Average CT index	70			N/A
Pass/Fail	Fail	***Pass***	***Pass***	N/A
COV, %	37.3	19.5	20.1	18.7
Single-Operator COV, %	18.3			
Pass/Fail	Fail	Fail	Fail	Fail
Analysis Using Trimmed Data (Optional)				
Average CT Index	65	132	178	38
Required Average CT index	70			N/A
Pass/Fail	Fail	***Pass***	***Pass***	N/A
COV, %	16.3	10.7	16.6	7.0
Single-Operator COV, %	11.2			
Pass/Fail	Fail	***Pass***	Fail	***Pass***

The underlined numbers represent maximum and minimum CT index values for each data set. OBC = optimum binder content; CT = cracking tolerance; N/A = not available; COV = coefficient of variation.

Table 19. Example of Variability in Data Set CT Index Values During Production

Sample ID	Non-Reheated Specimens		Reheated Specimens	
	Contractor	VDOT	VDOT	VTRC
1	176.8	239.7	188.3	121.6
2	198.3	237.5	124.7	110.6
3	210.1	189.7	152.5	141.5
4	195.9	237.9	183.4	150.0
5	175.9	161.0	219.8	225.6
Analysis Using Original Untrimmed Data				
Average CT Index	191	213	174	150
Required Average CT index	N/A		70	
Pass/Fail	N/A		***Pass***	***Pass***
Single-Operator COV, %	7.7	16.9	20.9	30.1
Required Single-Operator COV, %	18.3			
Pass/Fail	***Pass***	***Pass***	Fail	Fail
Multi-Laboratory COV, %	3.8		5.3	
Required Multi-Laboratory COV, %	21.3			
Pass/Fail	***Pass***		***Pass***	
Analysis Using Trimmed Data (Optional)				
Average CT Index	190	222	175	138
Required Average CT index	N/A		70	
Pass/Fail	N/A		***Pass***	***Pass***
Single-Operator COV, %	6.2	12.5	11.1	10.6
Required Single-Operator COV, %	11.2			
Pass/Fail	***Pass***	Fail	***Pass***	***Pass***
Multi-Laboratory COV, %	5.4		8.5	
Required Multi-Laboratory COV, %	15.9			
Pass/Fail	***Pass***		***Pass***	

The underlined numbers represent maximum and minimum CT index values for each data set. OBC = optimum binder content; CT = cracking tolerance; N/A = not available; COV = coefficient of variation.

- *Reheated specimens.* The COVs of the data sets collected by the VDOT district laboratory and by VTRC were 20.9% and 30.1%, respectively, both exceeding the acceptable single-operator COV of 18.3%. Most likely, the oven time required for reheating and resultant aging of the loose mixture had an adverse effect on the material and quality of the resulting data, although other factors may have also contributed. Moreover, the multi-laboratory COV between the VDOT and VTRC data sets was 5.3%, meeting the acceptable multi-operator COV of 21.3%, indicating that these two data sets were statistically similar.

CONCLUSIONS

- *The frequent hands-on training classes and demonstrations regarding the IDT-CT (and other laboratory tests considered as part of the BMD concept) between the completion of Phase I.1 and the beginning of Phase II.1 led to a significant increase in the percentage of participants submitting data fully conforming to the requirements of ASTM D8225-19.*

- *The Phase II.1 results of the testing of the IDT-CT indices suggest that device type may be a significant factor for mixtures with relatively low IDT-CT index values.* This may raise a significant concern regarding the evaluation of long-term aged asphalt mixtures using the IDT-CT. For that case, the CT index is expected to be low and the device type may have a significant impact on the test results.

- *Different statistical parameters are appropriate for evaluating the variability of different IDT-CT indices. The COV is a better metric for the CT index, and the Stdv is a better metric for the FST index, S_t, and CRI.*

- *Trimming of the data will help eliminate outliers whenever they exist*. In fact, trimming of the data results in a substantial drop in the magnitude of the selected statistical parameter for both single-operator and multi-laboratory precision estimates.

- *The precision estimates for single-operator conditions are similar whether or not the specimens of a given mixture are fabricated by the same laboratory*.

- *Specimen preparation introduces additional variability in the precision estimates for multi-laboratory comparisons.* The multi-laboratory precision estimates for Phase II.1 were greater than those for Phase I.1.

- *Sets of four and three replicates had single-operator and multi-laboratory precision estimates similar to those determined for five-replicate data sets.* This conclusion does not apply to trimmed data sets.

- *The IDT-CT indices are not dependent on the type of conditioning method (i.e., dry vs. wet) regardless of whether or not data trimming is applied.* This conclusion is based on the definitions of the "dry" and "wet" conditioning methods in this study.

- *The servo-hydraulic machine setup used in this study had effective control when a load at a given rate was applied.*

- *The results of the IDT-CT indices are not dependent on the loading rate applied, regardless of the machine type (SH or SD) and regardless of whether or not data trimming is applied .* This conclusion is valid for a loading rate range of 46 to 54 mm/min.

- *The results of the IDT-CT indices are not dependent on the data collection frequency selected.* This conclusion is valid for a data collection frequency range of 1 to 100 Hz.

RECOMMENDATIONS

1. *VDOT's Materials Division should consider adopting the precision estimates and corresponding statements determined in this study for the CT index as a sound way to monitor the repeatability and reproducibility of reported cracking performance data during both the mix design stage and production.* The recommended precision statements for the CT index are as follows:

- *For Original Untrimmed Data*

  Single-Operator Precision—The single-operator coefficient of variation was 18.3%. Therefore, results of two properly conducted tests by the same operator on the same material are not expected to differ from each other by more than 51.1%[A] of their average.

  Multi-Laboratory Precision—The multi-laboratory coefficient of variation was 21.3%. 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 59.7%[A] of their average.

  [A]These numbers represent the difference limits in percent (d2s%) as described in Practice ASTM C670.

  Note X—These precision statements are based on an interlaboratory study that involved 14 laboratories (16 data sets), two materials with CT index values ranging from 44 to 162, and five replicate tests per operator.

- *For Trimmed Data (Optional)*

  Single-Operator Precision—The single-operator coefficient of variation was 11.2%. Therefore, results of two properly conducted tests by the same operator on the same material are not expected to differ from each other by more than 31.3%[A] of their average.

  Multi-Laboratory Precision—The multi-laboratory coefficient of variation was 15.9%. 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 44.5%[A] of their average.

  [A]These numbers represent the difference limits in percent (d2s%) as described in Practice ASTM C670.

  Note X—These precision statements are based on an interlaboratory study that involved 14 laboratories (16 data sets), two materials with CT index values ranging from 44 to 162, and three replicate tests per operator.

2. *VDOT's Materials Division should allow the IDT-CT to be performed on wet specimens.* Currently, VDOT allows the use of the IDT-CT on only dry specimens conditioned in an environmental chamber or placed in leak-proof plastic bags in a water bath for 2 hours until the specimens reach a temperature of 25°C. Testing wet specimens would allow placing specimens in a water bath for 2 hours; removing the specimens and drying them to the saturated surface dry condition; and immediately performing the IDT-CT. This procedure would simplify specimen conditioning, particularly in production laboratories.

3. *VDOT's Materials Division should extend the allowable loading/deformation rate tolerance to 50 ± 3 mm/min when performing the IDT-CT.* Currently, ASTM D8225-19 requires use of

a loading apparatus capable of maintaining a constant loading/deformation rate of 50 ± 2.0 mm/min. The findings from this study (Phases I and II) and those from NCHRP Project 09-57A (Zhou et al., 2022), *Ruggedness of Laboratory Tests to Assess Cracking Resistance of Asphalt Mixtures*, have shown that the CT index values determined at a loading rate of 50 ± 3.0 mm/min remain statistically similar to those determined at a loading rate of 50 ± 2.0 mm/min.

4. *VDOT's Materials Division should establish an annual proficiency testing program for the IDT-CT.* The program should include VDOT and contactor laboratories in Virginia. The testing would entail one set of five specimens from a single mixture sent to all participant laboratories. The program would provide an opportunity to assess laboratory proficiency with the IDT-CT and identify any necessary improvements.

5. *VTRC and VDOT's Materials Division should continue organizing hands-on training classes and demonstrations of the laboratory tests being considered by VDOT as part of the BMD initiative.* The in-person BMD training classes offered in 2022 have helped familiarize a large number of agency and industry personnel with the current practices and procedures required for a successful initial implementation of the BMD concept.

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 Recommendation 1, VDOT's Materials Division has already included the single-operator precision estimate for the CT index as part of the 2023 BMD special provisions / specifications at the mix design stage. VTRC will work closely with the Materials Division to include the multi-laboratory precision estimate in the 2024 BMD special provisions / specifications and apply it as part of the acceptance program by no later than February 1, 2023.

With regard to Recommendations 2 and 3, VDOT's Materials Division will allow contractors to use the wet conditioning method as defined in this study. Moreover, the Materials Division will accept testing data collected under a loading rate controlled in the range of 50 ± 3.0 mm/min. VTRC will work closely with the Materials Division to incorporate both changes into the 2024 BMD special provisions / specifications by no later than February 1, 2023.

With regard to Recommendation 4, VDOT's Materials Division will deliberate internally on the necessity of establishing a proficiency program for the IDT-CT at intermediate temperature considering budget and resources needed. A decision is anticipated by June 2023.

With regard to Recommendation 5, VTRC and VDOT's Materials Division, with the help of the Virginia Asphalt Association and Germanna Community College through the Virginia Education Center for Asphalt Technology, organized and hosted five in-person BMD training classes in 2022. The first half of the curriculum included course work and presentations (study guides) on topics related to the BMD concept, VDOT's BMD approach, BMD tests and associated performance thresholds, and VDOT's BMD special provisions. The second half of the curriculum included hands-on training to perform BMD tests including the IDT-CT using various types of machines/devices. These BMD training classes are to continue in 2023.

Benefits

This study developed precision estimates and statements for the CT index, FST index, S_t, and CRI of asphalt mixtures determined by performing the IDT-CT at intermediate temperature. The findings of this study provide VDOT with a single-operator precision estimate for use in the acceptance of IDT-CT results for asphalt mixtures. Moreover, the findings provide VDOT with multi-laboratory precision estimates for repeatability and reproducibility to be incorporated into acceptance specifications. This work provides sound precision statements and references to determine if individual IDT-CT results from the same evaluated asphalt mixture can be considered statistically similar. Finally, this study will contribute to the data set to be used by ASTM to develop and adopt precision statements for ASTM D8225.

ACKNOWLEDGMENTS

The authors are grateful to the following individuals who served on the technical review panel for this study: Sungho Kim (Project Champion and Statewide Asphalt Program Manager, VDOT Materials Division); Robert Crandol (Assistant State Maintenance Engineer, VDOT Maintenance Division); Candice Entwisle (Assistant Asphalt Program Manager, VDOT Materials Division); Travis S. Higgs (District Materials Engineer, VDOT Salem District); Bernard L. Kassner (Senior Research Scientist, VTRC); Todd Rorrer (Asphalt Pavement Field Engineer, VDOT Materials Division); and Thomas Schinkel (District Materials Engineer, VDOT's Richmond District). The authors acknowledge Mike Dudley from the Virginia Asphalt Association for his assistance in developing and executing both phases of this IDT-CT round robin. 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.

REFERENCES

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 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, 2019a.

ASTM International. *ASTM E691-19: Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method.* West Conshohocken, PA, 2019b.

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, Issue 5, 2021.

Diefenderfer, S.D., and Bowers, B.F. Initial Approach to Performance (Balanced) Mix Design; The Virginia Experience. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2673, 2019, pp. 335-345.

Diefenderfer, S., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: Phase I: Initial Roadmap Development and Specification Verification.* VTRC 21-R15. Virginia Transportation Research Council, Charlottesville, 2021a.

Diefenderfer, S., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: 2019 Field Trials.* VTRC 21-R21. Virginia Transportation Research Council, Charlottesville, 2021b.

Habbouche, J., Boz, I., Diefenderfer, S., 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.

Minitab. Software for Statistics, 2021. https://support.minitab.com/en-us/minitab/21/. Accessed May 25, 2022.

R-package. The R Project for Statistical Computing, 2022. https://www.r-project.org/. Accessed May 4, 2022.

Transportation Research Board. *E-Circular 235: Glossary of Transportation Construction Quality Assurance Terms,* Seventh Edition. Washington, DC, 2018.

Virginia Department of Transportation. *Road and Bridge Specifications*. Richmond, 2018.

Zhou, F., Newcomb, D., and Hu, S. *NCHRP Report 987: Ruggedness of Laboratory Tests for Asphalt Mixture Cracking Resistance*. Transportation Research Board, Washington, DC, 2022.

APPENDIX A
DISTRIBUTION OF COLLECTED DATA FOR PHASE II.1

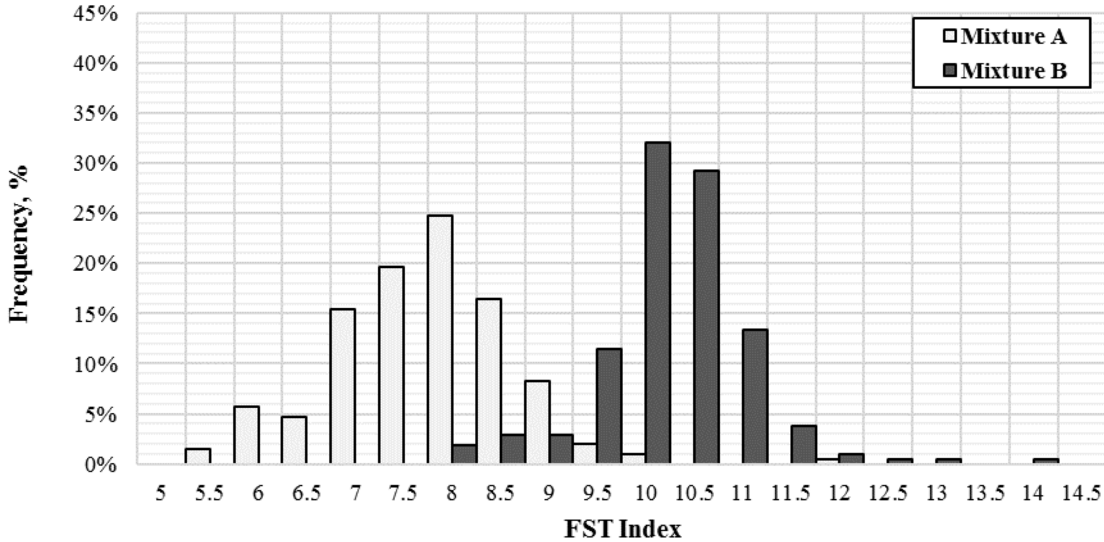

Figure A1. Distribution of FST Index Values for Mixture A and Mixture B Collected as Part of Phase II.1. FST = fracture strain tolerance.

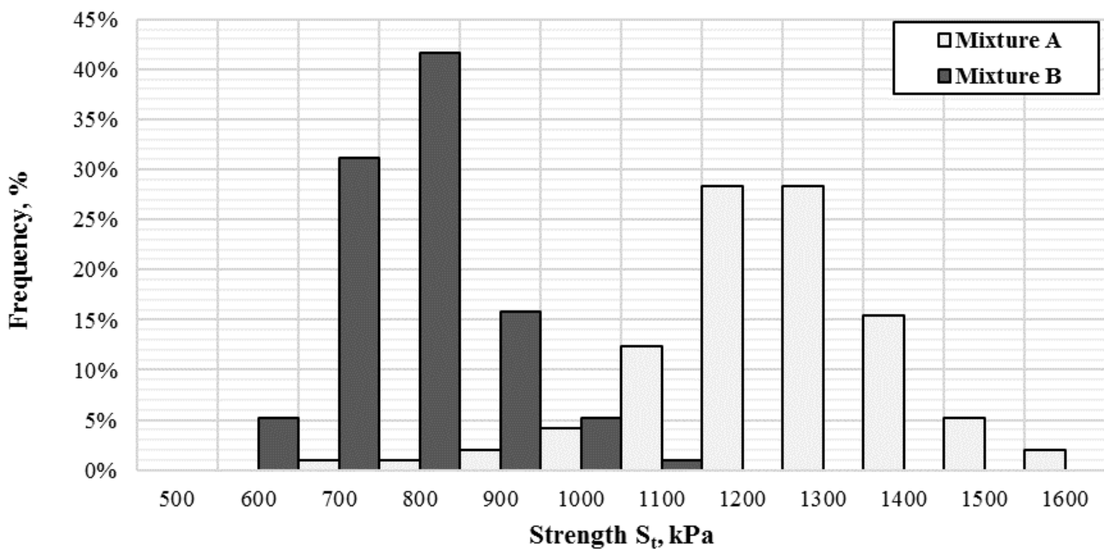

Figure A2. Distribution of S_t Values for Mixture A and Mixture B Collected as Part of Phase II.1. S_t = strength.

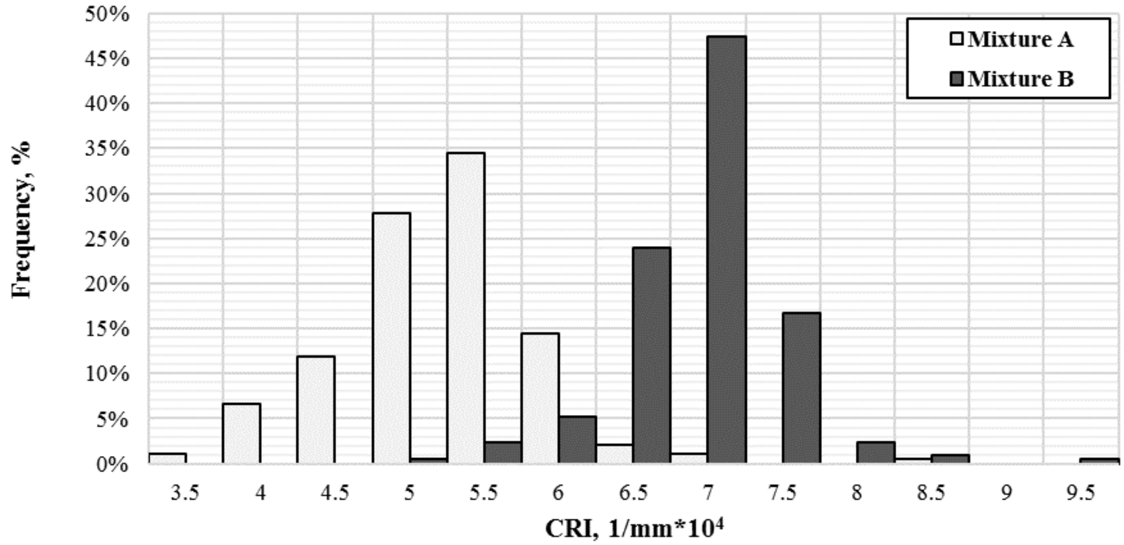

Figure A3. Distribution of CRI Values for Mixture A and Mixture B Collected as Part of Phase II.1. CRI = cracking resistance index.

APPENDIX B
PHASE II.1: EFFECT OF USING DIFFERENT DEVICES ON IDT-CT INDICES

FST Index

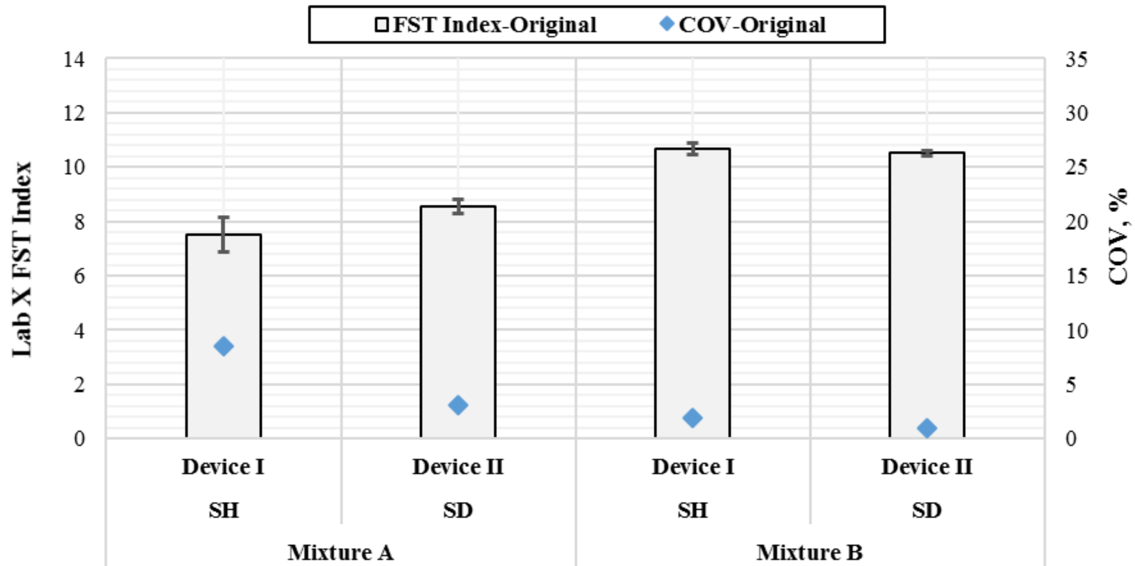

Figure B1. FST Index Original Untrimmed Values Reported by Lab X for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. FST = fracture strain tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

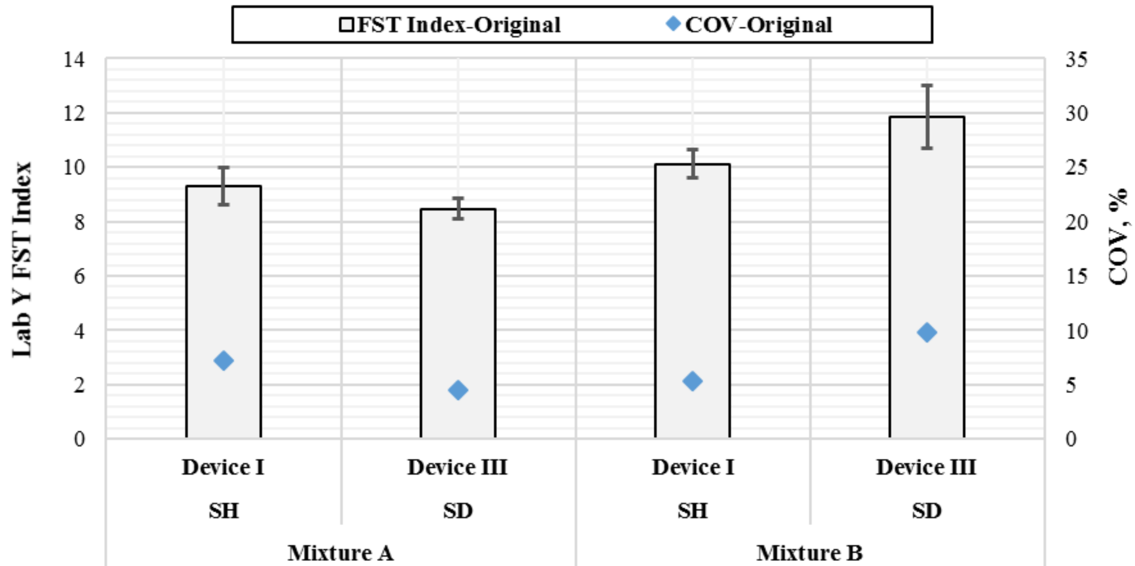

Figure B2. FST Index Original Untrimmed Values Reported by Lab Y for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. FST = fracture strain tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

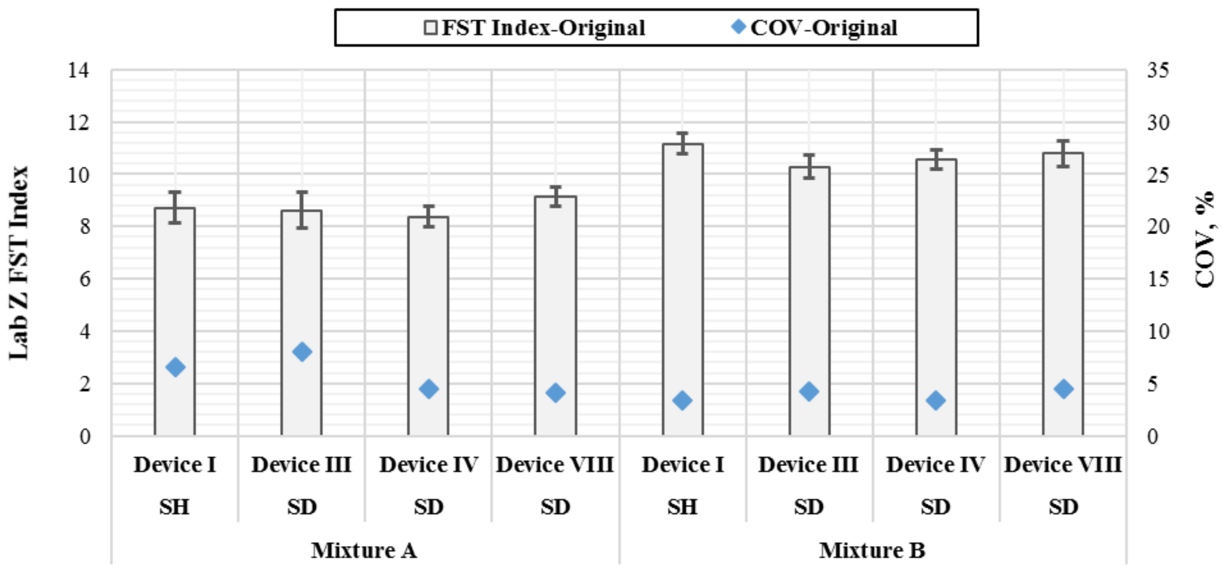

Figure B3. FST Index Original Untrimmed Values Reported by Lab Z for Mixture A and Mixture B Collected as Part of Phase II.1 Using Four Devices. I-bars show ±1 standard deviation. FST = fracture strain tolerance; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

Strength

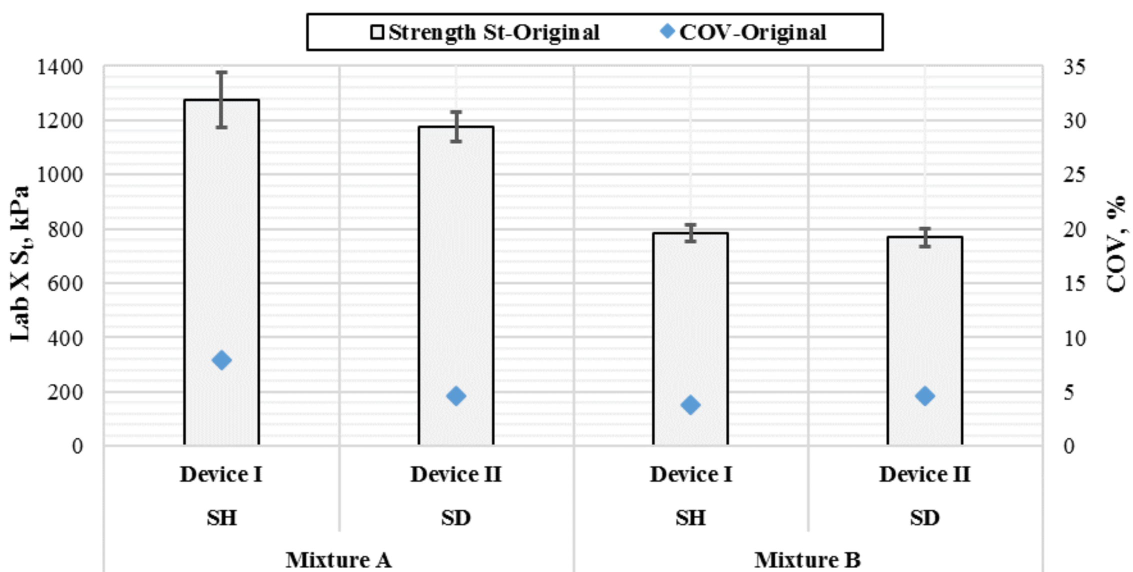

Figure B4. S_t Original Untrimmed Values Reported by Lab X for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. S_t = strength; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

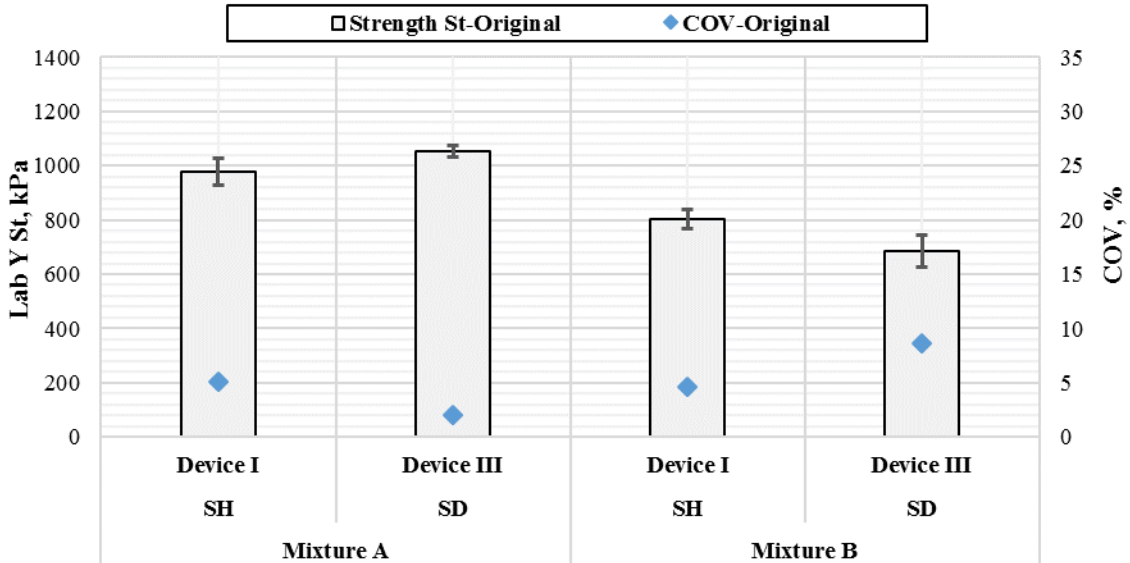

Figure B5. S_t Original Untrimmed Values Reported by Lab Y for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. S_t = strength; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

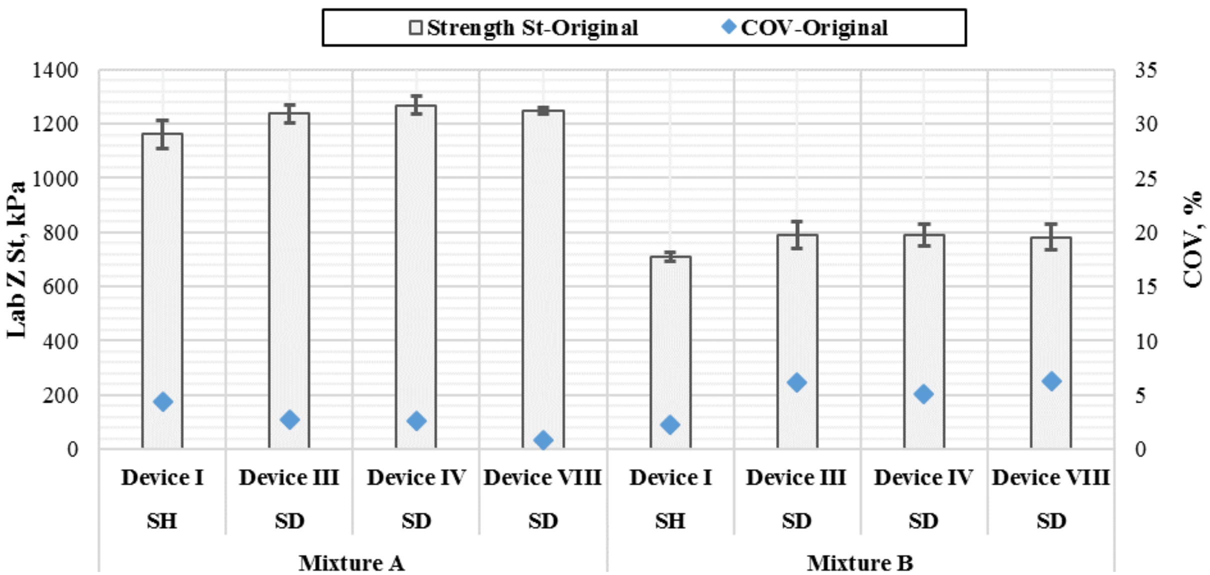

Figure B6. S_t Original Untrimmed Values Reported by Lab Z for Mixture A and Mixture B Collected as Part of Phase II.1 Using Four Devices. I-bars show ±1 standard deviation. S_t = strength; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

CRI

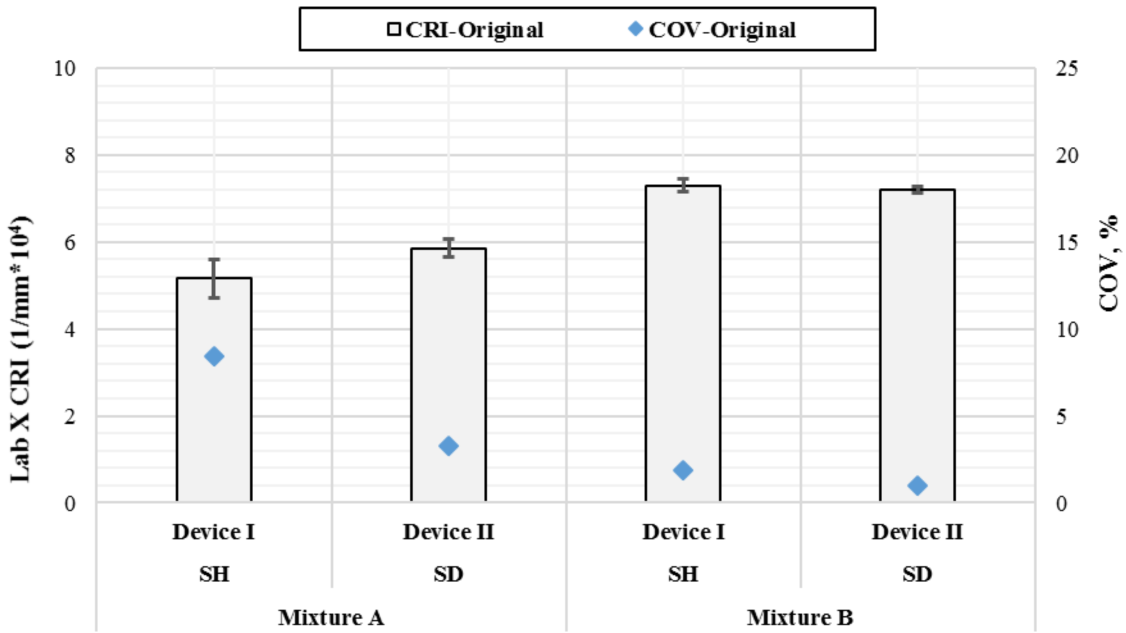

Figure B7. CRI Original Untrimmed Values Reported by Lab X for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. CRI= cracking resistance index; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

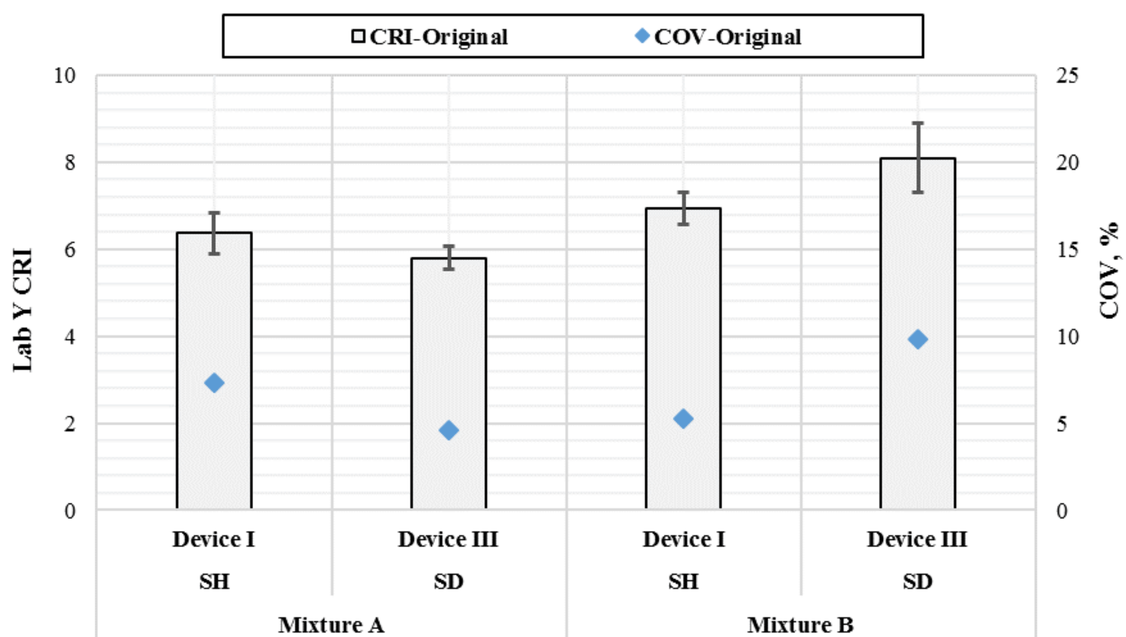

Figure B8. CRI Original Untrimmed Values Reported by Lab Y for Mixture A and Mixture B Collected as Part of Phase II.1 Using Two Devices. I-bars show ±1 standard deviation. CRI = cracking resistance index; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

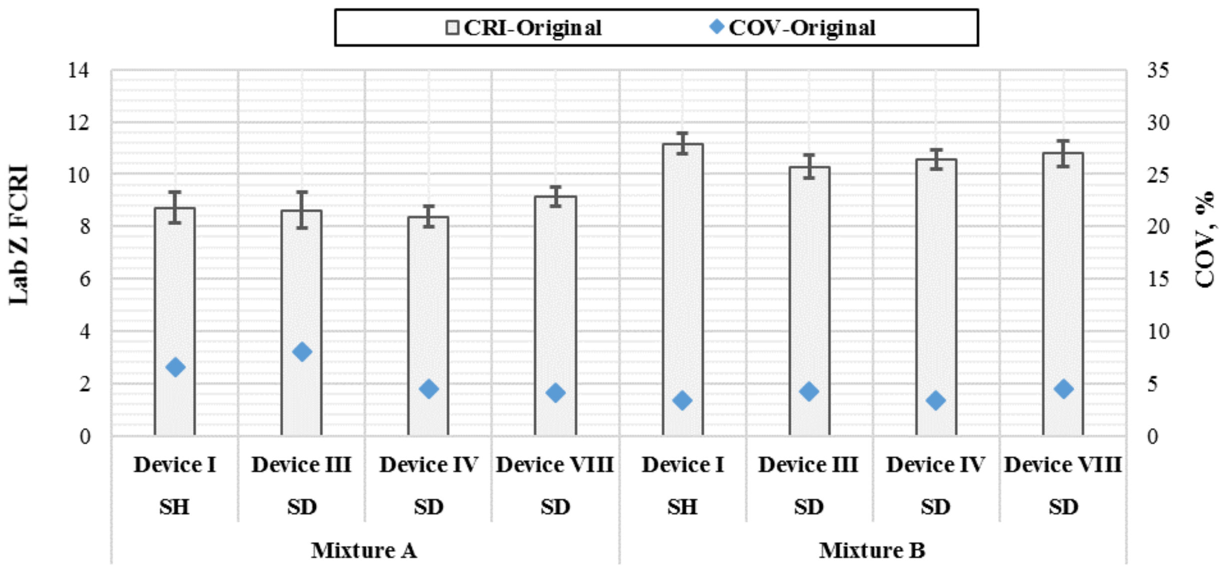

Figure B9. CRI Original Untrimmed Values Reported by Lab Z for Mixture A and Mixture B Collected as Part of Phase II.1 Using Four Devices. I-bars show ±1 standard deviation. CRI = cracking resistance index; COV = coefficient of variation; SH = servo-hydraulic; SD = screw-drive.

Statistical Analysis

Table B1. Summary of ANOVA Results for FST Index, S_t, and CRI Using Original Untrimmed and Trimmed Data for Lab X, Lab Y, and Lab Z

Factor	Lab X		Lab Y		Lab Z	
	DF	p-value	DF	p-value	DF	p-value
FST						
Mixture Type	1	0.000	1	0.000	1	***0.000***
Device	1	0.015	1	0.208	3	***0.021***
Mixture Type*Device	1	***0.002***	1	***0.002***	3	0.192
S_t, kPa						
Mixture Type	1	***0.000***	1	0.000	1	***0.000***
Device	1	***0.049***	1	0.304	3	***0.000***
Mixture Type*Device	1	0.146	1	***0.000***	3	0.796
CRI, 1/mm*10^4						
Mixture Type	1	0.000	1	0.000	1	***0.000***
Device	1	0.016	1	0.216	3	***0.020***
Mixture Type*Device	1	***0.002***	1	***0.002***	3	0.185

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; FST = fracture strain tolerance; S_t = strength; CRI = cracking resistance index; DF = degree of freedom.

APPENDIX C
SELECTION OF STATISTICAL PARAMETERS AND DEVELOPMENT OF PRECISION ESTIMATES

FST Index

Table C1. Variation of Standard Deviation and Coefficient of Variation Function of FST Index for Phase I.1 and Phase II.1

Phase	Category	Approach	Mix	Avg. Value	Single-Operator				Recommended Statistical Parameter	Multi-Laboratory				Recommended Statistical Parameter
					Stdv	Stdv Rate	COV	COV Rate		Stdv	Stdv Rate	COV	COV Rate	
I.1	(i)	Untrimmed	A	7.1	0.53	***0.107***	7.4	-0.277	**Stdv**	0.57	***0.051***	8.0	-0.336	**Stdv**
			B	10.6	0.59		5.6			0.60		5.7		
		Trimmed	A	7.1	0.27	0.258	3.8	***-0.141***	**COV**	0.43	***0.023***	6.1	-0.369	**Stdv**
			B	10.5	0.35		3.3			0.44		4.2		
	(ii)	Untrimmed	A	7.1	0.55	***0.103***	7.8	-0.294	**Stdv**	0.60	***0.049***	8.4	-0.333	**Stdv**
			B	10.5	0.61		5.8			0.63		6.0		
		Trimmed	A	7.1	0.31	***0.176***	4.4	-0.200	**Stdv**	0.44	***0.022***	6.2	-0.362	**Stdv**
			B	10.4	0.37		3.6			0.45		4.3		
II.1	(i)	Untrimmed	A	7.9	0.61	***-0.068***	7.7	-0.351	**Stdv**	0.98	***-0.253***	12.4	-0.518	**Stdv**
			B	10.5	0.57		5.4			0.76		7.3		
		Trimmed	A	7.9	0.30	***0.125***	3.8	-0.171	**Stdv**	0.83	***-0.274***	10.5	-0.545	**Stdv**
			B	10.5	0.34		3.2			0.63		6.0		
	(ii)	Untrimmed	A	8.0	0.59	***-0.017***	7.4	-0.277	**Stdv**	0.93	***-0.163***	11.6	-0.429	**Stdv**
			B	10.5	0.58		5.6			0.79		7.5		
		Trimmed	A	8.0	0.30	0.235	3.7	***-0.027***	**COV**	0.78	***-0.167***	9.7	-0.425	**Stdv**
			B	10.5	0.38		3.6			0.66		6.3		

Bold italic text refers to the lowest values among the Stdv and COV change rates. FST = fracture strain tolerance; Avg. = average; Stdv = standard deviation; COV = coefficient of variation.

Strength, S_t

Table C2. Variation of Standard Deviation and Coefficient of Variation Function of S_t for Phase I.1 and Phase II.1

Phase	Category	Approach	Mixture	Avg. Value	Single-Operator				Recommended Statistical Parameter	Multi-Laboratory				Recommended Statistical Parameter
					Stdv	Stdv Rate	COV	COV Rate		Stdv	Stdv Rate	COV	COV Rate	
I.1	(i)	Untrimmed	A	1159.3	54.0	***-0.177***	4.7	0.351	**Stdv**	122.4	-0.311	10.6	***0.226***	***COV***
			B	675.5	45.2		6.7			89.5		13.3		
		Trimmed	A	1163.2	33.6	***-0.052***	2.9	0.474	**Stdv**	117.1	-0.275	10.1	***0.251***	***COV***
			B	681.5	31.9		4.7			88.8		13.0		
	(ii)	Untrimmed	A	1160.2	55.1	***-0.138***	4.7	0.420	**Stdv**	111.3	***-0.235***	9.6	0.316	**Stdv**
			B	665.5	48.0		7.2			87.9		13.2		
		Trimmed	A	1163.8	34.1	***-0.064***	2.9	0.494	**Stdv**	104.2	***-0.210***	9.0	0.333	**Stdv**
			B	667.7	32.0		4.8			84.4		12.6		
II.1	(i)	Untrimmed	A	1272.3	69.5	-0.424	5.5	***0.036***	**COV**	136.3	-0.288	10.7	***0.179***	**COV**
			B	794.5	45.2		5.7			102.0		12.8		
		Trimmed	A	1272.5	39.3	-0.396	3.1	***0.062***	**COV**	126.2	-0.273	9.9	***0.200***	**COV**
			B	795.6	26.3		3.3			95.9		12.1		
	(ii)	Untrimmed	A	1248.8	66.1	-0.319	5.3	***0.140***	**COV**	151.7	-0.457	12.1	***0.000***	**COV**
			B	789.3	47.9		6.1			95.3		12.1		
		Trimmed	A	1248.3	36.3	-0.350	2.9	***0.098***	**COV**	145.3	-0.489	11.6	***-0.035***	**COV**
			B	789.5	25.5		3.2			88.2		11.2		

Bold italic text refers to the lowest among the Stdv and COV change rates. S_t = strength; Avg. = average; Stdv = standard deviation; COV = coefficient of variation.

CRI

Table C3. Variation of Standard Deviation and Coefficient of Variation Function of CRI for Phase I.1 and Phase II.1

Phase	Category	Approach	Mix	Avg. Value	Single-Operator Stdv	Single-Operator Stdv Rate	Single-Operator COV	Single-Operator COV Rate	Recommended Statistical Parameter	Multi-Laboratory Stdv	Multi-Laboratory Stdv Rate	Multi-Laboratory COV	Multi-Laboratory COV Rate	Recommended Statistical Parameter
I.1	(i)	Untrimmed	A	4.9	0.36	0.382	7.4	***0.000***	**COV**	0.39	0.358	7.9	***-0.026***	**COV**
			B	7.2	0.53		7.4			0.56		7.7		
		Trimmed	A	4.8	0.20	***0.049***	4.2	-0.333	**Stdv**	0.29	***0.000***	6.0	-0.400	**Stdv**
			B	7.2	0.21		3.0			0.29		4.0		
	(ii)	Untrimmed	A	4.9	0.38	0.233	7.7	***-0.139***	**COV**	0.41	0.217	8.4	***-0.168***	**COV**
			B	7.1	0.48		6.7			0.51		7.1		
		Trimmed	A	4.9	0.21	***0.091***	4.2	-0.240	**Stdv**	0.30	**0.033**	6.2	-0.362	**Stdv**
			B	7.1	0.23		3.3			0.31		4.3		
II.1	(i)	Untrimmed	A	5.4	0.42	***-0.074***	7.7	-0.351	**Stdv**	0.67	**-0.252**	12.4	-0.518	**Stdv**
			B	7.2	0.39		5.4			0.52		7.3		
		Trimmed	A	5.4	0.20	***0.140***	3.8	-0.141	**Stdv**	0.57	***-0.280***	10.5	-0.545	**Stdv**
			B	7.2	0.23		3.3			0.43		6.0		
	(ii)	Untrimmed	A	5.5	0.40	***0.000***	7.3	-0.264	**Stdv**	0.64	***-0.169***	11.6	-0.429	**Stdv**
			B	7.2	0.40		5.6			0.54		7.5		
		Trimmed	A	5.5	0.20	0.261	3.6	***0.000***	**COV**	0.53	***-0.141***	9.7	-0.410	**Stdv**
			B	7.2	0.26		3.6			0.46		6.4		

Bold italic text refers to the lowest among the Stdv and COV change rates. CRI = cracking resistance index; Avg. = average; Stdv = standard deviation; COV = coefficient of variation.

APPENDIX D
PHASE II.2: DISTRIBUTION OF COLLECTED DATA AND STATISTICAL ANALYSES FOR FST INDEX, S_t, AND CRI

FST Index

Original Untrimmed Data

Table D1. Descriptive Statistics for FST Index Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	8.4	6.6	7.8	8.8	10.0	3.3	1.0
	Wet	8.8	6.8	8.4	9.2	10.0	3.2	0.8
Mixture B	Dry	11.1	10.1	10.7	11.5	12.6	2.5	0.8
	Wet	11.4	10.7	10.9	11.6	15.9	5.2	0.6

FST = fracture strain tolerance; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

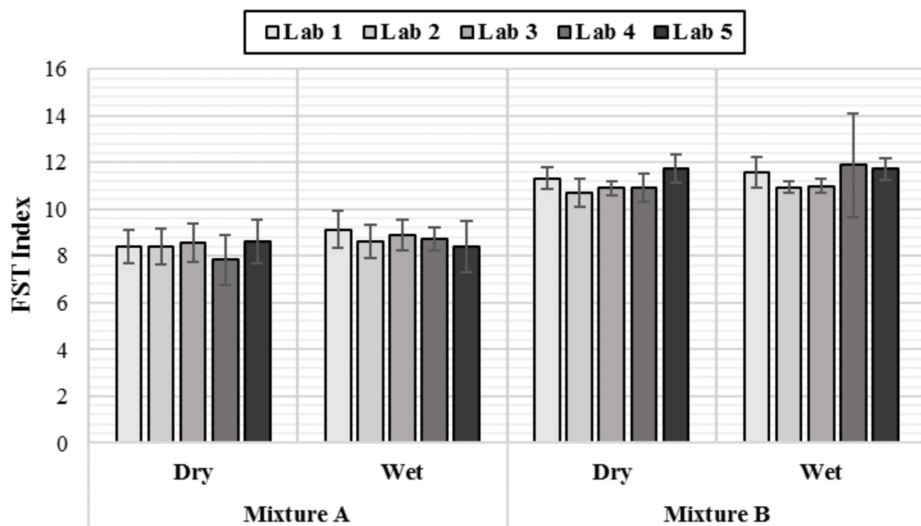

Figure D1. Average FST Index Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. FST = fracture strain tolerance; Lab = laboratory.

Trimmed Data

Table D2. Descriptive Statistics for FST Index Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	8.3	7.1	8.0	8.7	9.0	1.9	0.7
	Wet	8.7	7.8	8.4	9.0	9.8	1.9	0.6
Mixture B	Dry	11.1	10.3	10.8	11.4	12.1	1.8	0.6
	Wet	11.5	10.8	10.9	11.6	15.9	5.1	0.6

FST = fracture strain tolerance; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

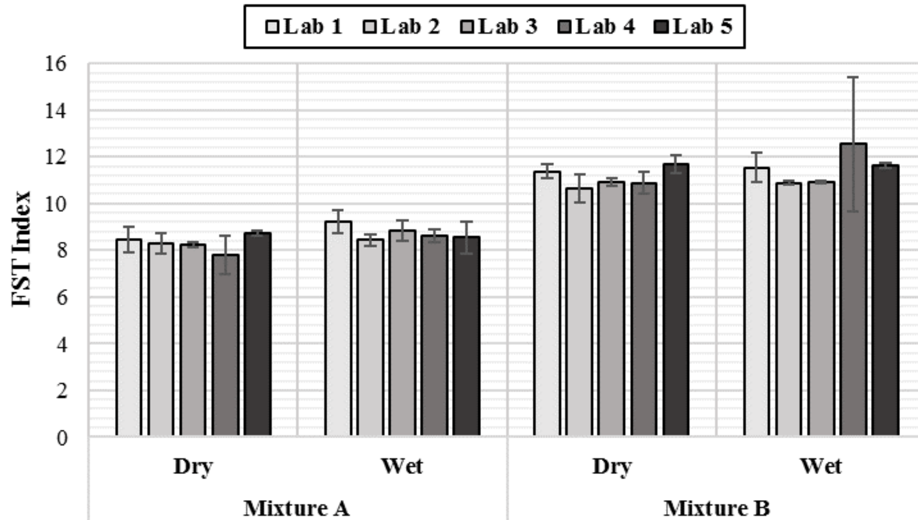

Figure D2. Average FST Index Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. FST = fracture strain tolerance; Lab = laboratory.

Strength, S_t

Original Untrimmed Data

Table D3. Descriptive Statistics for S_t Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	1245.1	1066.1	1146.9	1289.3	1514.1	448.0	142.4
	Wet	1187.3	1012.9	1151.3	1230.8	1389.2	376.3	79.5
Mixture B	Dry	725.4	552.3	660.1	776.7	902.7	350.4	116.6
	Wet	678.4	449.5	649.3	717.1	777.2	327.7	67.8

S_t = strength; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

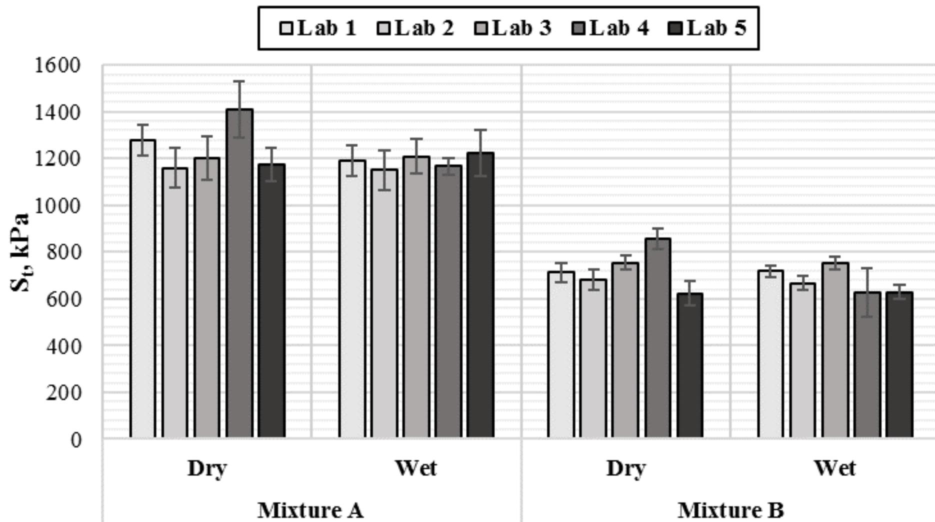

Figure D3. Average S_t Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. S_t = strength; Lab = laboratory.

Trimmed Data

Table D4. Descriptive Statistics for S_t Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	1253.3	1070.8	1176.0	1289.3	1505.8	435.0	113.3
	Wet	1191.2	1149.1	1155.2	1218.5	1259.0	109.9	63.3
Mixture B	Dry	725.9	599.4	669.0	772.2	894.2	294.8	103.2
	Wet	684.9	614.4	650.3	727.3	774.3	159.9	77.0

S_t = strength; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

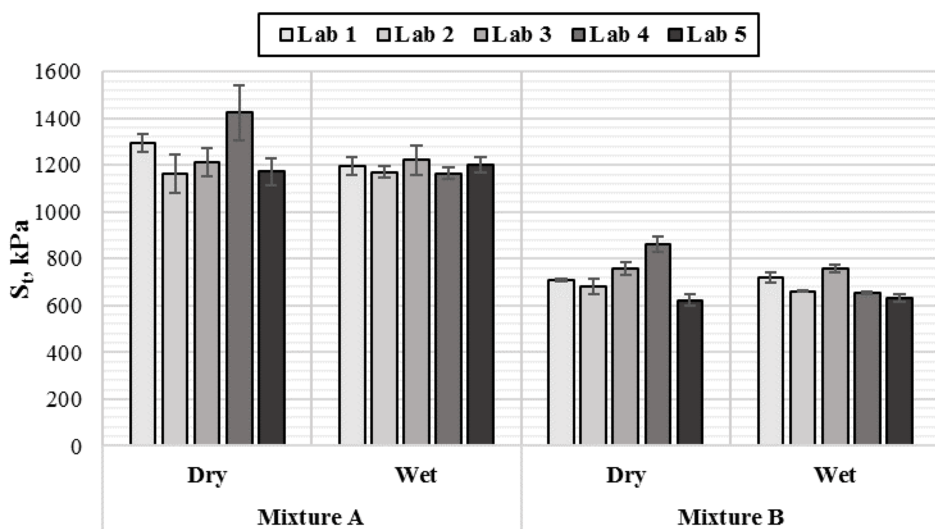

Figure D4. Average S_t Values of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. S_t = strength; Lab = laboratory.

CRI

Original Untrimmed Data

Table D5. Descriptive Statistics for CRI Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	5.7	4.5	5.4	6.1	6.8	2.2	0.7
	Wet	6.0	4.7	5.7	6.3	6.9	2.2	0.6
Mixture B	Dry	7.6	6.9	7.3	7.8	8.6	1.7	0.5
	Wet	7.8	7.3	7.4	7.9	10.9	3.6	0.5

CRI = cracking resistance index; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

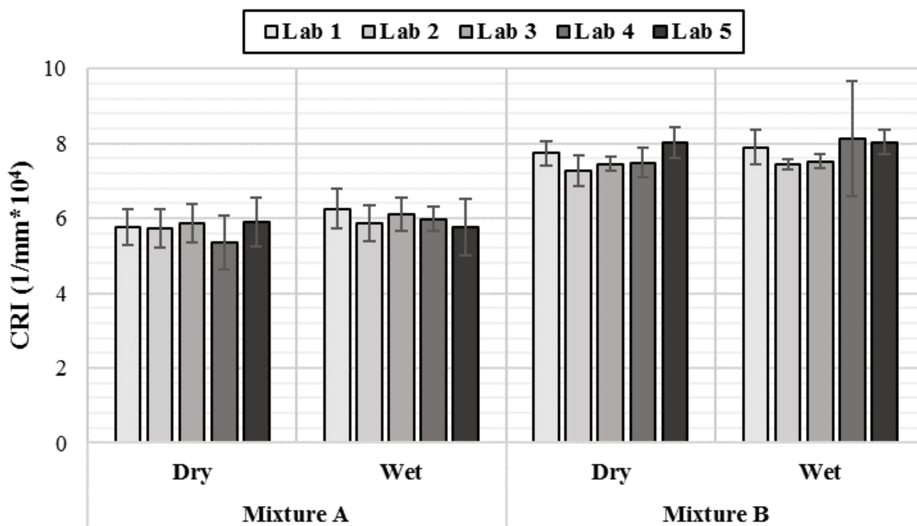

Figure D5. Average CRI Values of Original Untrimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. CRI = cracking resistance index; Lab = laboratory.

Trimmed Data

Table D6. Descriptive Statistics for CRI Values ($1/mm*10^4$) of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning.

Mixture ID	Conditioning Mode	Mean	Min	Q1	Q3	Max	Range	IQR
Mixture A	Dry	5.7	4.9	5.5	5.9	6.2	1.3	0.4
	Wet	6.0	5.4	5.7	6.2	6.7	1.3	0.4
Mixture B	Dry	7.6	7.0	7.4	7.8	8.3	1.3	0.4
	Wet	7.6	7.3	7.5	7.9	8.4	1.0	0.4

CRI = cracking resistance index; Min = minimum; Q1 = quartile 1; Q3 = quartile 3; Max = maximum; IQR = interquartile range.

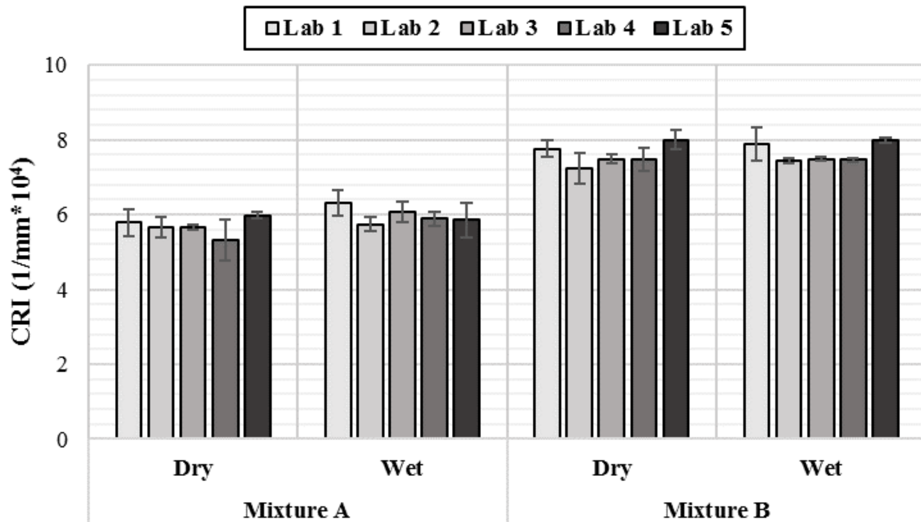

Figure D6. Average CRI Values (1/mm*10^4) of Trimmed Data for Mixture A and Mixture B After Dry and Wet Conditioning. I-bars show ±1 standard deviation. CRI = cracking resistance index; Lab = laboratory.

Table D7. Summary of ANOVA Results for FST Index, S_t, and CRI (1/mm*10^4) of Dry and Wet Specimens Using Original Untrimmed Data

Factor	FST Index		Strength, S_t		CRI	
	DF	**p-value**	**DF**	**p-value**	**DF**	**p-value**
Mixture Type	1	***0.000***	1	***0.000***	1	***0.000***
Conditioning Method	1	0.066	1	***0.000***	1	0.056
Laboratory	4	***0.049***	4	***0.000***	4	***0.026***
Mixture Type*Conditioning Method	1	0.296	1	0.451	1	0.261
Mixture Type*Laboratory	4	0.126	4	***0.040***	4	0.121
Conditioning Method*Laboratory	4	0.660	4	***0.000***	4	0.653
Mixture Type*Conditioning Method*Laboratory	4	0.673	4	0.447	4	0.651

Bold italic text indicates that the p-values were lower than 0.05. ANOVA = analysis of variance; FST = fracture strain tolerance; CRI = cracking resistance index; DF = degree of freedom.

Table D8. Summary of ANOVA Results for FST of Dry and Wet Specimens Using Original Untrimmed Data

Mixture Type*Conditioning Method*Laboratory	FST Index		
	N	Mean	Grouping
Mixture A – Dry – Lab 4	5	7.83	a
Mixture A – Dry – Lab 1	5	8.37	a
Mixture A – Wet – Lab 5	5	8.38	a
Mixture A – Dry – Lab 2	5	8.41	a
Mixture A – Dry – Lab 3	5	8.57	a
Mixture A – Dry – Lab 5	5	8.61	a
Mixture A – Wet – Lab 2	5	8.62	a
Mixture A – Wet – Lab 4	5	8.74	a
Mixture A – Wet – Lab 3	5	8.92	a
Mixture A – Wet – Lab 1	5	9.12	a
Mixture B – Dry – Lab 2	5	10.69	b
Mixture B – Wet – Lab 4	4	10.88	b
Mixture B – Dry – Lab 3	5	10.90	b
Mixture B – Dry – Lab 4	5	10.92	b
Mixture B – Wet – Lab 2	5	10.94	b
Mixture B – Wet – Lab 3	5	11.0	b
Mixture B – Dry – Lab 1	4	11.33	b
Mixture B – Wet – Lab 1	5	11.56	b
Mixture B – Dry – Lab 5	5	11.73	b
Mixture B – Wet – Lab 5	5	11.74	b

FST = fracture strain tolerance; Lab = laboratory; N = number of elements.

Table D9. Summary of ANOVA Results for S_t of Dry and Wet Specimens Using Original Untrimmed Data

Mixture Type*Conditioning Method*Laboratory	Strength, S_t		
	N	Mean	Grouping
Mixture A – Dry – Lab 4	5	1411.3	a
Mixture A – Dry – Lab 1	5	1278.5	a b
Mixture A – Wet – Lab 5	5	1223.2	b
Mixture A – Wet – Lab 3	5	1208.7	b
Mixture A – Dry – Lab 3	5	1199.9	b
Mixture A – Wet – Lab 1	5	1188.4	b
Mixture A – Dry – Lab 5	5	1175.5	b
Mixture A – Wet – Lab 4	5	1166.2	b
Mixture A – Dry – Lab 2	5	1160.1	b
Mixture A – Wet – Lab 2	5	1150.3	b
Mixture B – Dry – Lab 4	5	857.5	c
Mixture B – Dry – Lab 3	5	754.0	c d
Mixture B – Wet – Lab 3	5	752.2	c d
Mixture B – Wet – Lab 1	5	717.7	c d
Mixture B – Dry – Lab 1	5	712.1	c d
Mixture B – Dry – Lab 2	5	681.3	d
Mixture B – Wet – Lab 4	4	670.2	d
Mixture B – Wet – Lab 2	5	667.3	d
Mixture B – Wet – Lab 5	5	628.7	d
Mixture B – Dry – Lab 5	5	622.1	d

S_t = strength; Lab = laboratory; N = number of elements.

Table D10. Summary of ANOVA Results for CRI (1/mm*10^4) of Dry and Wet Specimens Using Original Untrimmed Data

Mixture Type*Conditioning Method*Laboratory	CRI (1/mm*10^4)		
	N	Mean	Grouping
Mixture A – Dry – Lab 4	5	5.35	a
Mixture A – Dry – Lab 2	5	5.73	a
Mixture A – Dry – Lab 1	5	5.75	a
Mixture A – Wet – Lab 5	5	5.76	a
Mixture A – Dry – Lab 3	5	5.86	a
Mixture A – Wet – Lab 2	5	5.88	a
Mixture A – Dry – Lab 5	5	5.91	a
Mixture A – Wet – Lab 4	5	5.98	a
Mixture A – Wet – Lab 3	5	6.12	a
Mixture A – Wet – Lab 1	5	6.26	a b
Mixture B – Dry – Lab 2	5	7.27	b c
Mixture B – Wet – Lab 4	4	7.45	c
Mixture B – Wet – Lab 2	5	7.45	c
Mixture B – Dry – Lab 3	5	7.46	c
Mixture B – Dry – Lab 4	5	7.49	c
Mixture B – Wet – Lab 3	5	7.52	c
Mixture B – Dry – Lab 1	4	7.74	c
Mixture B – Wet – Lab 1	5	7.90	c
Mixture B – Wet – Lab 5	5	8.04	c
Mixture B – Dry – Lab 5	5	8.04	c

CRI = cracking resistance index; Lab = laboratory; N = number of elements.

APPENDIX E
PHASE II.3: DISTRIBUTION OF COLLECTED DATA AND STATISTICAL ANALYSES FOR FST INDEX, S_T, AND CRI

FST Index

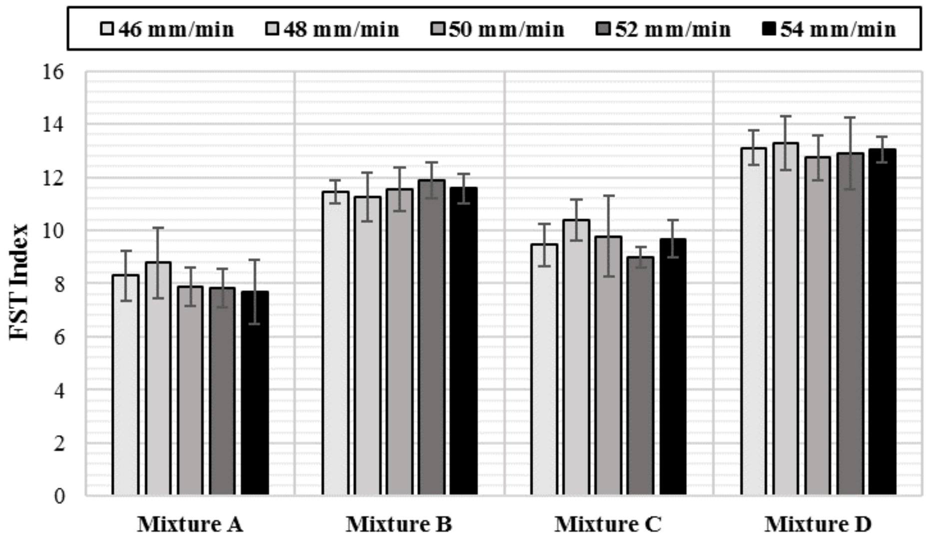

Figure E1. Average FST Index Values of Original Untrimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. FST = fracture strain tolerance; SH = servo-hydraulic.

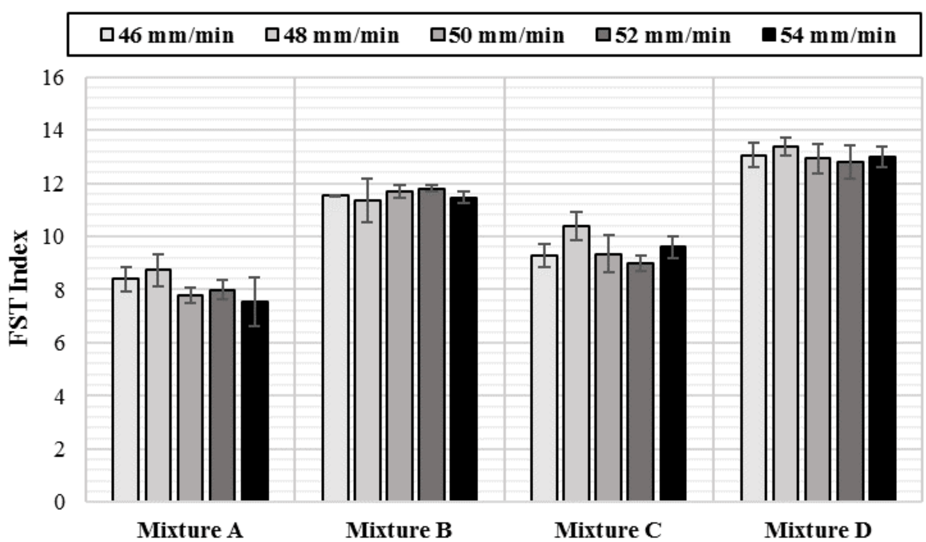

Figure E2. Average FST Index Values of Trimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. FST = fracture strain tolerance; SH = servo-hydraulic.

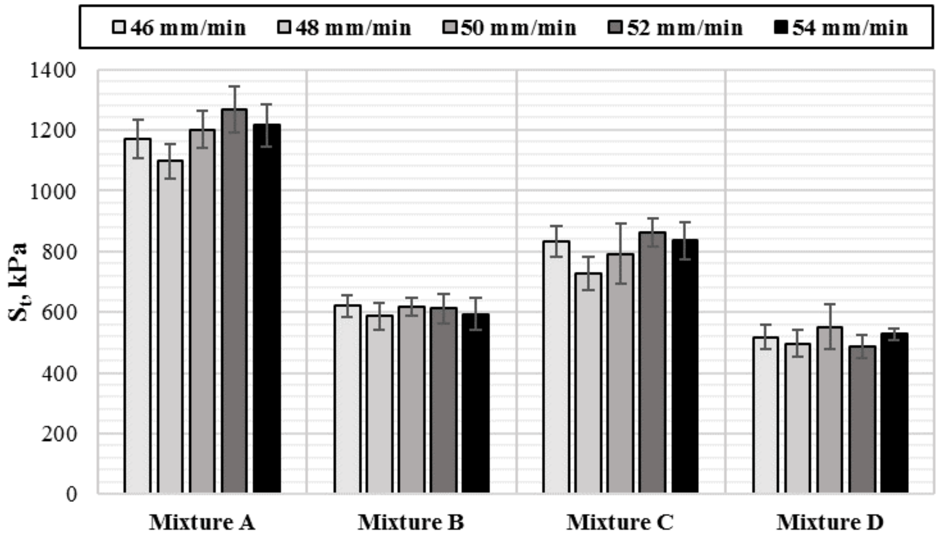

Figure E3. Average S_t Values of Original Untrimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. S_t = strength; SH = servo-hydraulic.

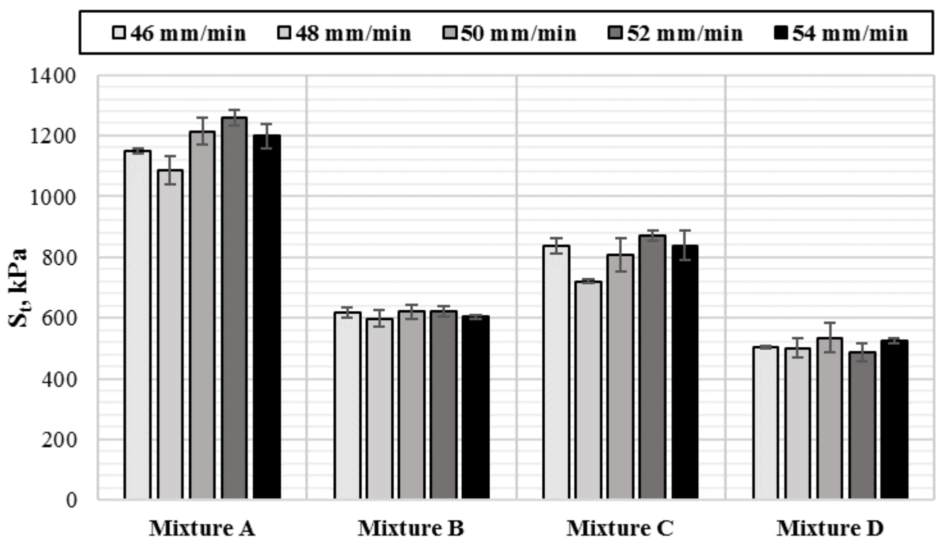

Figure E4. Average S_t Values of Trimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. S_t = strength; SH = servo=hydraulic.

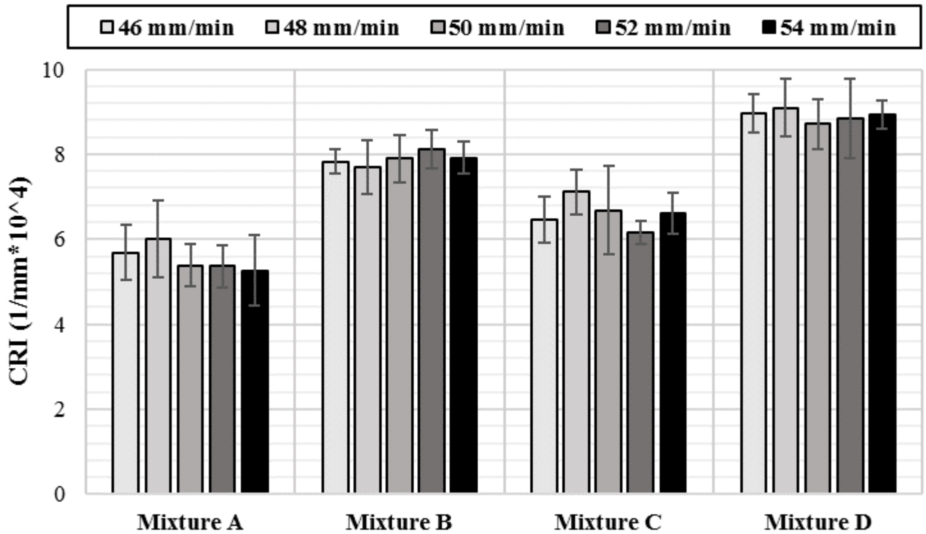

Figure E5. Average CRI Values of Original Untrimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. CRI = cracking resistance index; SH = servo-hydraulic.

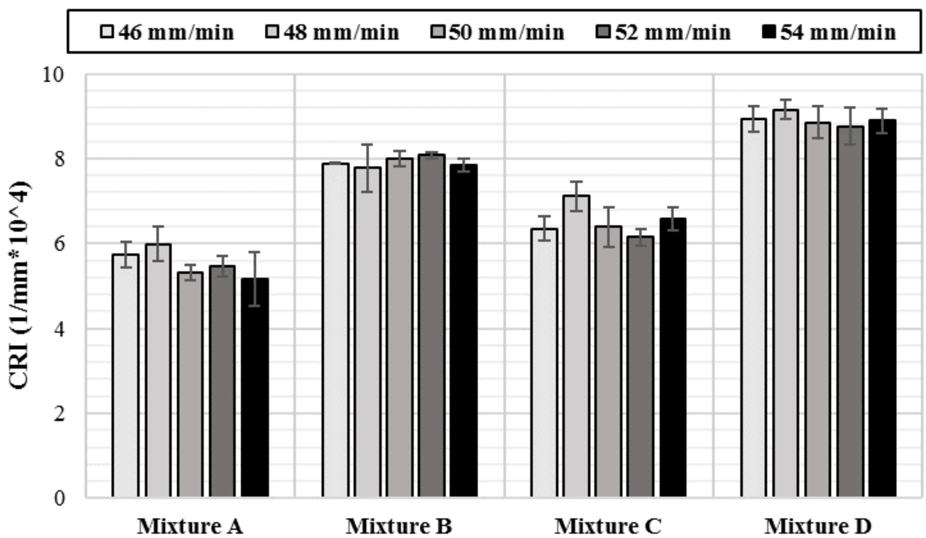

Figure E6. Average CRI Values of Trimmed Data for Mixtures A, B, C, and D at Five Loading Rates Using an SH Machine. I-bars show ±1 standard deviation. CRI = cracking resistance index; SH = servo-hydraulic.

APPENDIX F
PHASE II.3: IMPACT OF DATA COLLECTION FREQUENCY ON IDT-CT INDICES

Servo-Hydraulic Machine (Device I)

FST Index

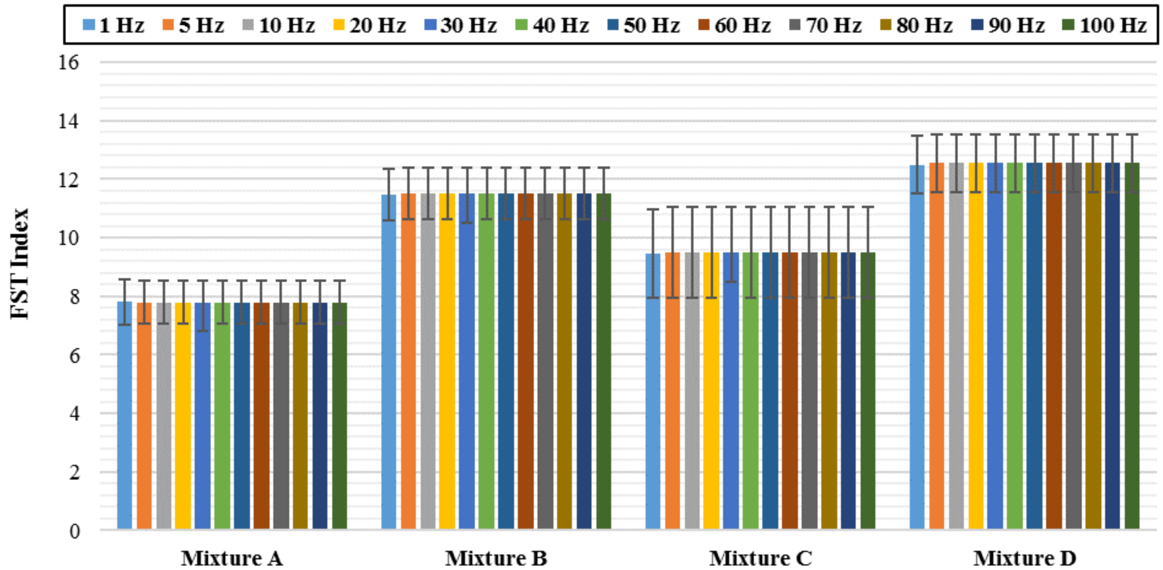

Figure F1. FST Index of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SH Machine. I-bars show ±1 standard deviation. FST = fracture strain tolerance; SH = servo-hydraulic.

Strength, S_t

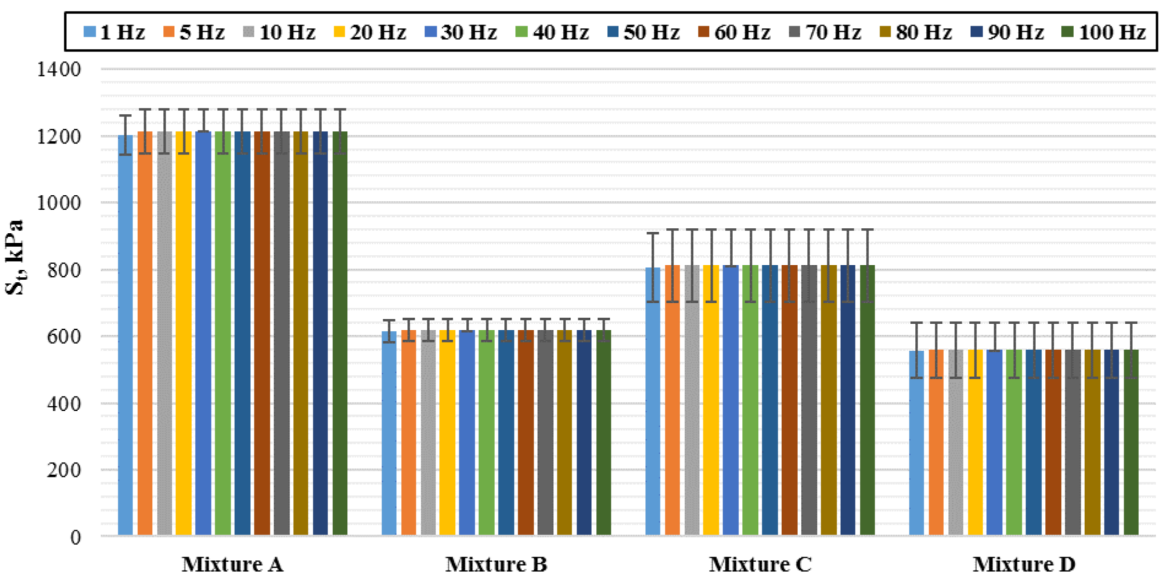

Figure F2. S_t of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SH Machine. I-bars show ±1 standard deviation. S_t = strength; SH = servo-hydraulic.

CRI

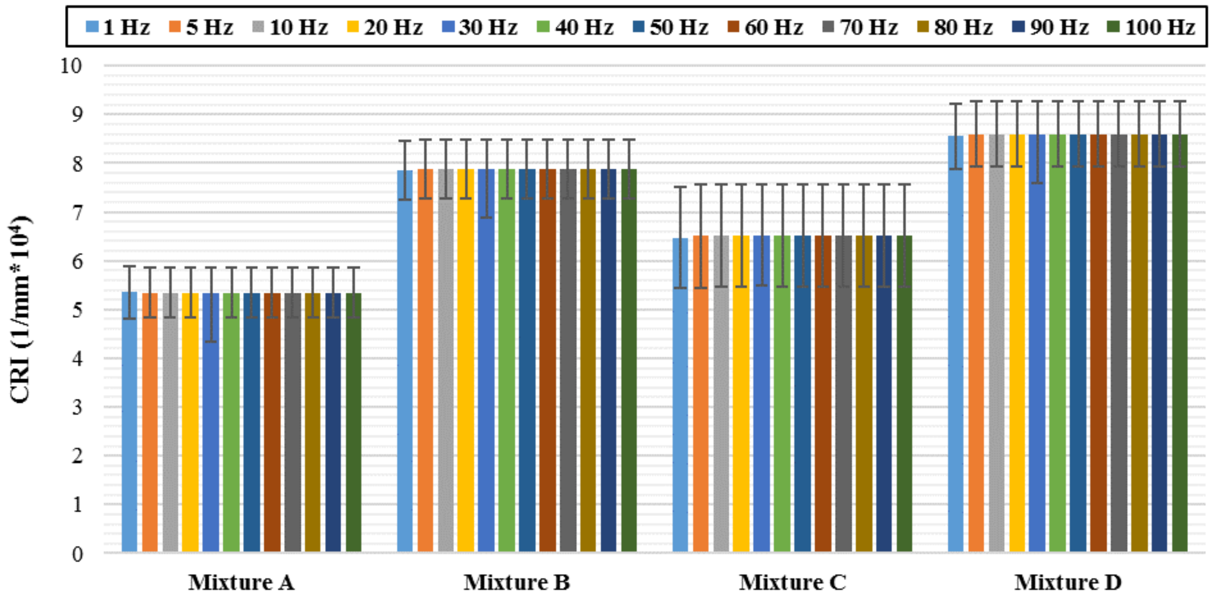

Figure F3. CRI of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SH Machine. I-bars show ±1 standard deviation. CRI = cracking resistance index; SH = servo-hydraulic.

Screw-Drive Machine (Device IV)

FST Index

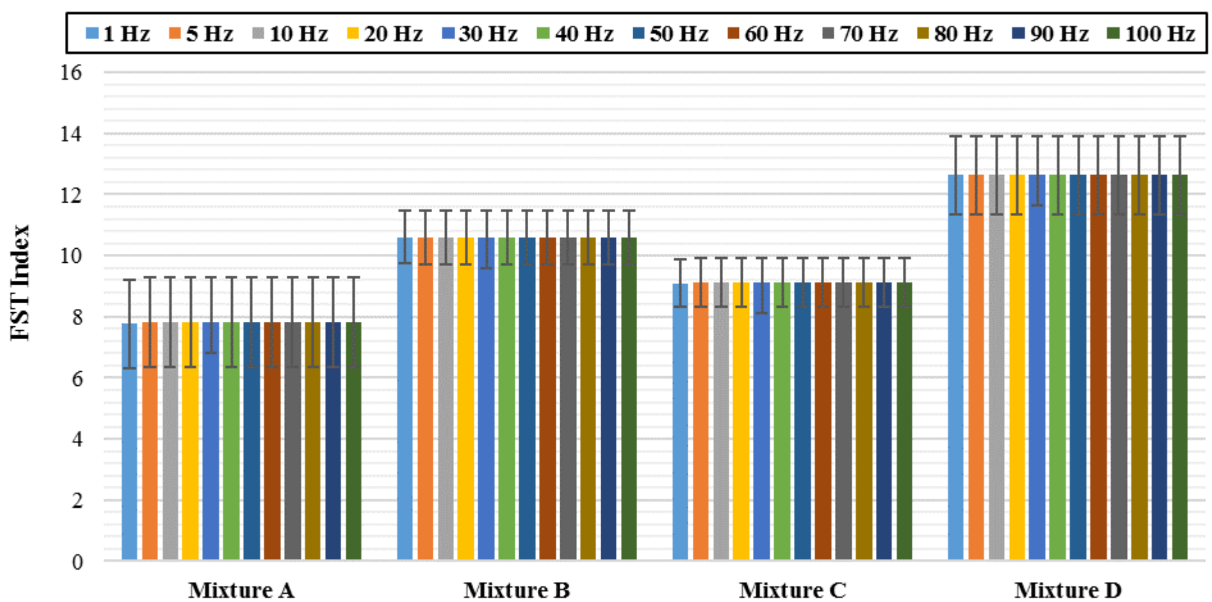

Figure F4. FST index of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SD Machine. I-bars show ±1 standard deviation. FST = fracture strain tolerance; SD = screw-drive.

Strength, S_t

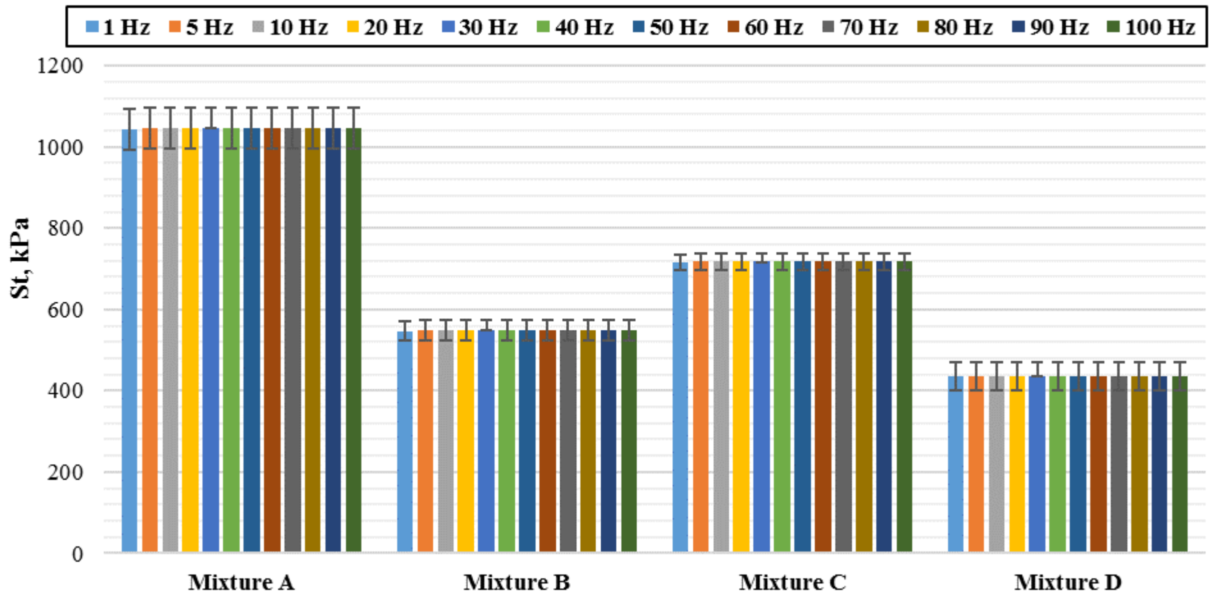

Figure F5. S_t of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SD Machine. I-bars show ±1 standard deviation. S_t = strength; SD = screw-drive.

CRI

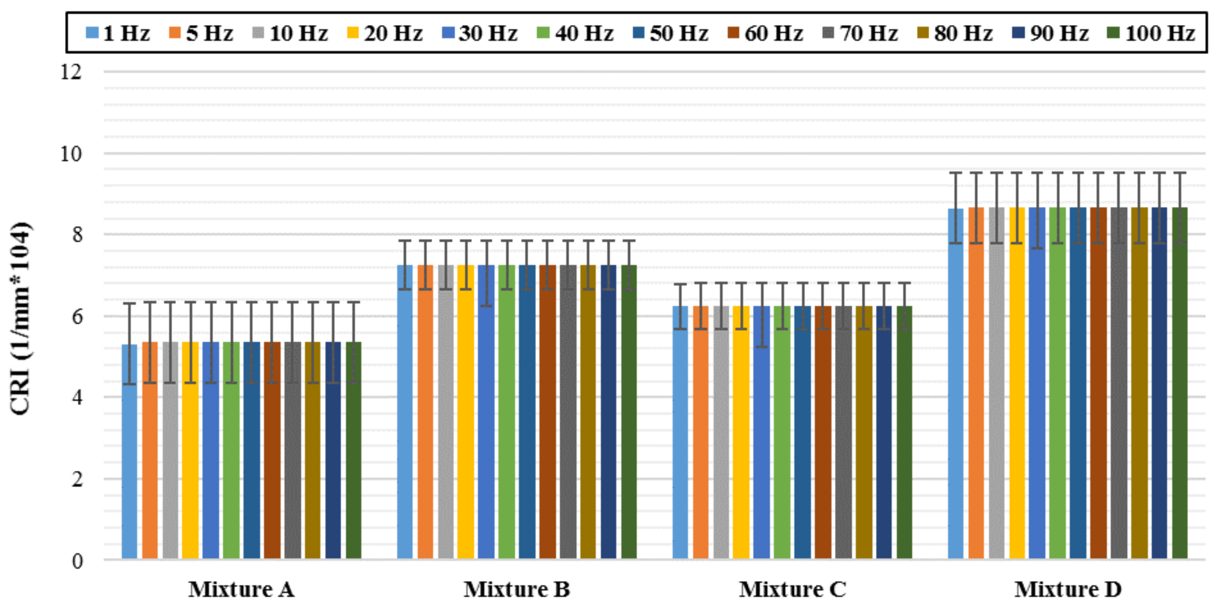

Figure F6. CRI of Mixtures A, B, C, and D Determined Over a Range of 12 Frequencies Collected Using an SD Machine. I-bars show ±1 standard deviation. CRI = cracking resistance index; SD = screw-drive.