VTRC
Virginia Transportation
Research Council

We Bring Innovation to Transportation

Ruggedness Study of Specimen Preparation and Fine-Tuning of Test Methods for IDT-CT and IDT-HT Test

https://vtrc.virginia.gov/media/vtrc/vtrc-pdf/vtrc-pdf/26-R22.pdf

ILKER BOZ, Ph.D., P.E., Senior Research Scientist
Virginia Transportation Research Council

JHONY HABBOUCHE, Ph.D., P.E., Western U.S. Regional Engineer
Asphalt Institute

ADAM TAYLOR, P.E., NCAT Engineer
National Center for Asphalt Technology at Auburn University

NATHAN MOORE, P.E., Assistant Director for Test Track Research
National Center for Asphalt Technology at Auburn University

Final Report VTRC 26-R22

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

Standard Title Page - Report on Federally Funded Project

1. Report No.: FHWA/VTRC 26-R22	2. Government Accession No.:	3. Recipient's Catalog No.:
4. Title and Subtitle: Ruggedness Study of Specimen Preparation and Fine-Tuning of Test Methods for IDT-CT and IDT-HT Test		5. Report Date: November 2025 6. Performing Organization Code:
7. Author(s): Ilker Boz, Ph.D., P.E., Jhony Habbouche, Ph.D., P.E., Adam Taylor, P.E., and Nathan Moore, P.E.		8. Performing Organization Report No.: VTRC 26-R22
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.: 124496
12. Sponsoring Agencies' Name and Address: Virginia Department of Transportation 1221 E. Broad Street Richmond, VA 23219	Federal Highway Administration 400 North 8th Street, Room 750 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 Virginia Department of Transportation utilizes the indirect tensile cracking test (IDT-CT) and the indirect tensile at high temperature (IDT-HT) test as part of its balanced mix design framework to evaluate asphalt mixture performance and support performance-based acceptance decisions. Although these tests are critical tools for assessing performance characteristics, considerable variability has been observed in test results, raising concerns about the effects of specimen preparation procedures on the repeatability and reliability of the outcomes.

This study was undertaken to systematically evaluate the influence of key specimen preparation factors on IDT-CT and IDT-HT test results and to establish clear, data-driven guidelines to standardize preparation practices. The experimental program included a ruggedness phase and a fine-tuning phase. Each phase was designed to identify and quantify the effects of critical preparation variables. Factors evaluated included mixture homogenization, splitting methods, heating duration, container type, specimen reheating, and heating processes. Laboratory testing was conducted on a range of asphalt mixtures using controlled procedures, and statistical analyses were performed to determine the significance of preparation factors on the test indices.

The results highlighted that inconsistent preparation practices can introduce practical and sometimes statistically significant variability, potentially leading to false positive or false negative performance assessments, particularly when comparing with specification thresholds such as the cracking tolerance index. Based on the study findings, a draft Virginia Test Method was developed to guide standardized preparation practices and improve the reliability, repeatability, and interpretability of IDT-CT and IDT-HT test results.

This study recommends that the Virginia Department of Transportation should (1) adopt the Virginia Test Method developed in this study as part of the standard specimen preparation procedures for Virginia Department of Transportation district laboratories and asphalt contractors to use within the balanced mix design process and (2) consider benchmarking the variability of IDT-CT and IDT-HT test results (i.e., cracking tolerance index and strength) following the adoption of the Virginia Test Method developed in this study to assess improvements in consistency and reliability.

17. Key Words: Balanced mix design, performance testing, cracking, rutting, specimen preparation, CT index, strength, variability, asphalt mixtures.		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: 49	22. Price:

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

FINAL REPORT

RUGGEDNESS STUDY OF SPECIMEN PREPARATION AND FINE-TUNING OF TEST METHODS FOR IDT-CT AND IDT-HT TEST

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

Jhony Habbouche, Ph.D., P.E.
Western U.S. Regional Engineer
Asphalt Institute

Adam Taylor, P.E.
NCAT Engineer
National Center for Asphalt Technology at Auburn University

Nathan Moore, P.E.
Assistant Director for Test Track Research
National Center for Asphalt Technology at Auburn University

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

Charlottesville, Virginia

November 2025
VTRC 26-R22

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 Virginia Department of Transportation utilizes the indirect tensile cracking test (IDT-CT) and the indirect tensile at high temperature (IDT-HT) test as part of its balanced mix design framework to evaluate asphalt mixture performance and support performance-based acceptance decisions. Although these tests are critical tools for assessing performance characteristics, considerable variability has been observed in test results, raising concerns about the effects of specimen preparation procedures on the repeatability and reliability of the outcomes.

This study was undertaken to systematically evaluate the influence of key specimen preparation factors on IDT-CT and IDT-HT test results and to establish clear, data-driven guidelines to standardize preparation practices. The experimental program included a ruggedness phase and a fine-tuning phase. Each phase was designed to identify and quantify the effects of critical preparation variables. Factors evaluated included mixture homogenization, splitting methods, heating duration, container type, specimen reheating, and heating processes. Laboratory testing was conducted on a range of asphalt mixtures using controlled procedures, and statistical analyses were performed to determine the significance of preparation factors on the test indices.

The results highlighted that inconsistent preparation practices can introduce practical and sometimes statistically significant variability, potentially leading to false positive or false negative performance assessments, particularly when comparing with specification thresholds such as the cracking tolerance index. Based on the study findings, a draft Virginia Test Method was developed to guide standardized preparation practices and improve the reliability, repeatability, and interpretability of IDT-CT and IDT-HT test results.

This study recommends that the Virginia Department of Transportation should (1) adopt the Virginia Test Method developed in this study as part of the standard specimen preparation procedures for Virginia Department of Transportation district laboratories and asphalt contractors to use within the balanced mix design process and (2) consider benchmarking the variability of IDT-CT and IDT-HT test results (i.e., cracking tolerance index and strength) following the adoption of the Virginia Test Method developed in this study to assess improvements in consistency and reliability.

TABLE OF CONTENTS

FINAL REPORT

RUGGEDNESS STUDY OF SPECIMEN PREPARATION AND FINE-TUNING OF TEST METHODS FOR IDT-CT AND IDT-HT TEST

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

Jhony Habbouche, Ph.D., P.E.
Western U.S. Regional Engineer
Asphalt Institute

Adam Taylor, P.E.
NCAT Engineer
National Center for Asphalt Technology at Auburn University

Nathan Moore, P.E.
Assistant Director for Test Track Research
National Center for Asphalt Technology at Auburn University

INTRODUCTION

Overview and Background

The Virginia Department of Transportation (VDOT) currently requires using the indirect tensile cracking test (IDT-CT), the Asphalt Pavement Analyzer rutting test, and the Cantabro mass loss test in its balanced mix design (BMD) specifications. These tests are used to evaluate the performance properties of dense-graded asphalt surface mixtures (SMs) with A and D designations, specifically for resistance to cracking, rutting, and durability, respectively (Diefenderfer and Bowers, 2019). A recent effort completed by the Virginia Transportation Research Council (VTRC) proposed adopting the indirect tensile at high temperature (IDT-HT) test as a simpler and more practical alternative to the Asphalt Pavement Analyzer rutting test (Boz et al., 2023).

The 2024 VDOT Special Provision for BMD SMs designed using performance criteria requires the following performance testing:

- A minimum cracking tolerance (CT) index of 70, determined by IDT-CT, during the mix design stage. During production, the minimum CT index is 70 for reheated specimens and 95 for non-reheated specimens.
- A maximum rut depth of 8.0 mm at 64°C, determined by the Asphalt Pavement Analyzer rutting test, during both the mix design and production stages.

- A minimum tensile strength of 100 kPa at 54.4°C, determined by the IDT-HT test, for specimens conditioned in a water bath (without sealing). This requirement applies during the mix design stage and is reported only for information during production.
- A maximum mass loss of 7.5%, determined by the Cantabro mass loss test, during both the mix design and production stages.

These performance thresholds were established based on extensive testing of compacted specimens prepared from reheated plant-produced asphalt mixtures (Boz et al., 2023; Diefenderfer and Bowers, 2019; Diefenderfer et al., 2021; Habbouche et al., 2022a).

In 2020, VTRC, in collaboration with the Virginia Asphalt Association, initiated a two-phase round-robin evaluation, or interlaboratory study, to establish the acceptable variability of the indices (e.g., CT index) from IDT-CT through the evaluation of two asphalt mixtures. Phase I of this effort involved 41 participating laboratories testing a set of five specimens per mixture fabricated and compacted by a third-party laboratory (Habbouche et al., 2021). Phase II involved 50 participating laboratories fabricating and testing a minimum of five test specimens (per mixture) using a loose mixture sent to the laboratories along with detailed instructions (Habbouche et al., 2022b). The National Center for Asphalt Technology (NCAT) and the Center for Advanced Infrastructure and Transportation at Rutgers University undertook a similar effort (Bennert et al., 2020; Taylor et al., 2022). Based on the results from the three studies, the precision estimates for single-operator conditions were found to be similar, whether or not the same laboratory fabricated the specimens of a given mixture. However, specimen preparation introduced significant additional variability in the precision estimates for multi-laboratory comparisons. For example, Virginia's effort showed a 7.4% increase in terms of coefficient of variation (COV) for the CT index, increasing the initial multi-laboratory variability of the test by nearly 35% (Boz et al., 2022; Habbouche et al., 2022b). Similarly, in the NCAT study, the COV for multi-laboratory precision for the CT index increased from 20 to 35% because of variations in specimen preparation (Taylor et al., 2022).

VTRC, NCAT, and the Center for Advanced Infrastructure and Transportation at Rutgers University pursued a similar effort to determine and develop precision estimates and statements for the rutting index (i.e., strength) calculated from the IDT-HT test (Boz et al., 2025). Similarly, specimen preparation was found to introduce significant additional variability in the precision estimates for single- and multi-laboratory comparisons.

Variability Assessment of VDOT 2024 Balanced Mix Design Database

The precision estimates for the CT index and strength were determined initially using specimens produced under highly controlled laboratory conditions. In 2024, asphalt contractors in Virginia were given the option to submit CT index results from non-reheated specimens compacted during production. As noted previously, a minimum CT index of 95 was required for these specimens. In practice, contractors sampled mixtures during production and transferred them to the laboratory, where the mixtures were immediately split and placed in an oven to prevent cooling. The specimens were then compacted and tested. Several challenges can be encountered during this process, potentially contributing to increased variability in test results. For example, achieving the target air voids of 7.0 ± 0.5% often requires adjusting pill weights, leading to differences in heating times between the first and subsequent specimens in a given set. As a result, industry stakeholders reported difficulty with consistently meeting the precision estimates for BMD tests (e.g., IDT-CT and IDT-HT) during production.

Figure 1 presents the observed variability in CT index results collected from asphalt contractors and VDOT district laboratories during the 2024 paving season. The dataset included 858 data points, each obtained from five replicate sets. Of these sets, 331 data points (38.6% of the dataset) failed to meet the single-operator COV threshold of 18.3%, and 90 data points (10.5%) did not fall within the acceptable range for single-operator variability. For IDT-CT, the acceptable range for the CT index should not exceed 71.4% of the average value within five replicate sets. Considering the current CT index threshold values specified by VDOT (i.e., 70 and 95), this acceptable range is relatively broad. Therefore, practices contributing to violations of this range and the single-operator variability limit of 18.3% should be minimized to ensure data reliability and compliance with established specifications (Habbouche et al., 2021).

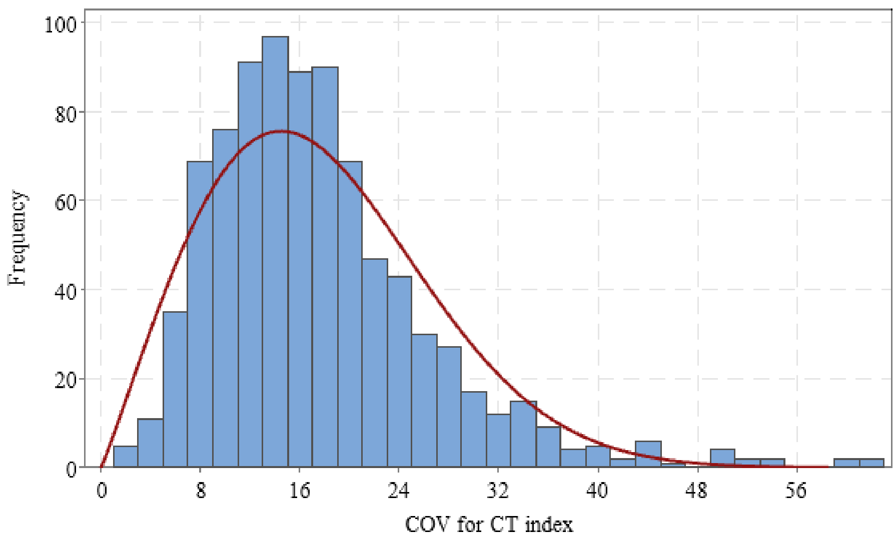

Figure 1. Variability Assessment of VDOT 2024 Balanced Mix Design Database for CT Index Determined by Indirect Tensile Cracking Test. COV = coefficient of variation; CT = cracking tolerance.

Figure 2 shows the variability in strength measurements obtained from the IDT-HT test during the 2024 paving season. This dataset consisted of 630 data points, each derived from three replicate sets. Of these sets, 188 data points (29.8%) did not meet the single-operator COV threshold of 10.8%, and 49 data points (7.8%) were outside the acceptable range for single-operator variability. For the IDT-HT test, the acceptable range for strength values should not exceed 35.6% of the average value within the three replicate sets. Although the observed variability in IDT-HT test results is less pronounced than in IDT-CT, a notable proportion of data points still fails to meet the established precision estimates. This failure underscores the necessity for consistent specimen preparation and testing protocols to enhance the reliability and repeatability of IDT-HT test results.

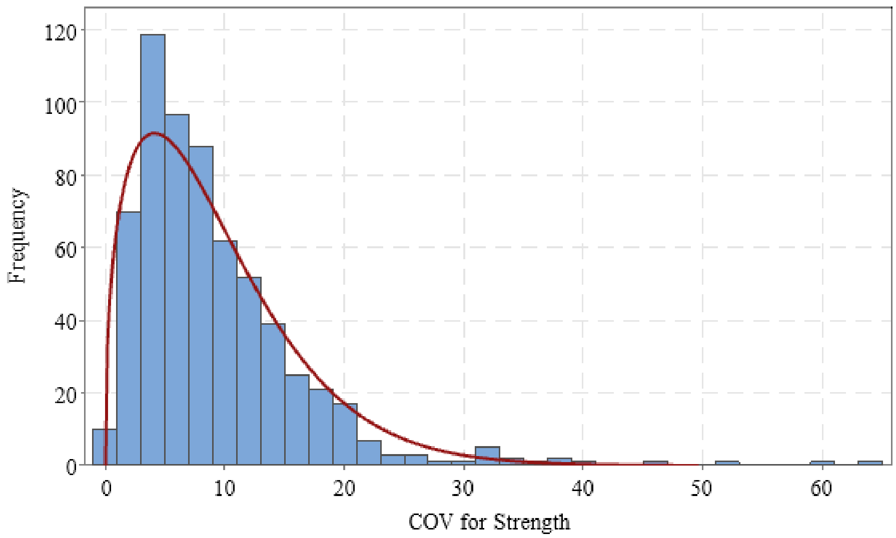

Figure 2. Variability Assessment of VDOT 2024 Balanced Mix Design Database for Strength Determined by Indirect Tensile at High Temperature Test. COV = coefficient of variation.

Moving Forward

Instead of tolerating higher (single-operator and multi-laboratory) variability, it is essential to identify the key factors that could contribute to an increase in the variability of both the CT index (from IDT-CT) and strength (from the IDT-HT test). These factors may include, but are not limited to, the method used for sampling and splitting the asphalt mixture, the type of container used for heating or reheating prior to compaction, the thickness of the sample layer within the pan or container, the duration and method of heating, and the type and calibration of ovens used. Such variables can significantly influence the consistency and reliability of test results. Addressing these factors may require stricter process controls and the development of more detailed, comprehensive procedural guidelines, particularly during production but also in the mix design phase. In the production context, such guidance should consider both non-reheated and reheated specimens to ensure improved repeatability and reproducibility of performance test results.

PURPOSE AND SCOPE

The purpose of this study was to perform a comprehensive laboratory evaluation to identify and assess key factors that necessitate stricter control and the development of more detailed guidelines during specimen preparation for the IDT-CT and IDT-HT test methods. The scope of work encompassed a literature review and stakeholder survey to identify potential sources of variability in these test methods, sampling and characterization of materials, a ruggedness evaluation of the identified factors, further assessment of selected factors through sensitivity analyses (fine-tuning studies), and the formulation of standardized specimen preparation guidelines.

METHODS

Review of the Current State of Knowledge and Practice

A comprehensive literature search was conducted to gather state-of-the-art information relevant to the objectives of this study. Researchers searched various transportation engineering databases and search engines—including Transport Research International Documentation, Transportation Research Information Services, Scopus, Catalog of Worldwide Libraries, Google Scholar, ProQuest, and Web of Science—for relevant literature.

Three documents were identified as particularly relevant for reviewing the current state of knowledge and practice:

- National Cooperative Highway Research Program (NCHRP) Report 752—*Improved Mix Design, Evaluation, and Materials Management Practices for Hot Mix Asphalt with High Reclaimed Asphalt Pavement Content* (West et al., 2013).
- NCHRP Synthesis 552—*Practices for Fabricating Asphalt Specimens for Testing in Laboratories* (Sias et al., 2020).
- National Asphalt Pavement Association IS-145 Publication—*Guide on Asphalt Mixture Specimen Fabrication for BMD Performance Testing* (Moore and Taylor, 2023).

In addition to the literature review, information on current practices related to specimen preparation during both the mix design and production stages, including procedures for reheated and non-reheated specimens, was collected through a survey. The survey was distributed to technical review panel (TRP) members, VDOT district laboratories, and Virginia asphalt contractors. The questionnaire (included in Appendix A) was designed to gather key information on specimen preparation procedures that may influence the results of IDT-CT and the IDT-HT test during both design and production phases.

Material Selection and Sampling

The experimental program for this study was divided into two major parts. Part I focuses on a study of ruggedness to identify key factors that require stricter control and the development of more comprehensive guidelines. Part II involves fine-tuning a limited number of factors selected based on the findings from Part I.

Five dense-graded asphalt SMs, labeled Mixes 1 through 5, were sampled for evaluation in Part I. All mixes were produced and placed in Virginia during the 2023 paving season. Three additional mixtures, labeled Mixes 6 through 8, were sampled for Part II and were produced and placed during the 2024 paving season.

The mixtures incorporate two nominal maximum aggregate sizes (9.5 mm and 12.5 mm), various reclaimed asphalt pavement (RAP) contents, and binders of similar performance grades (PG), sourced from different terminals with varying crude oil origins. One exception is Mix 6, which used a polymer-modified asphalt binder. The selected mixtures also include two different recycling agents (RAs) and multiple types of warm mix asphalt additives. This selection approach was intended to ensure the broad applicability of the research findings. It is important to note that the inclusion of a polymer-modified asphalt mixture in the experimental program (Mix 6) was due to the timing of Part II sampling, which occurred late in the 2024 paving season, when the availability of conventional mixtures was limited.

The specific mixtures evaluated in this study are as follows:

- Mix 1: SM-12.5A—15% RAP + PG 64S-22, sampled from the Culpeper District (S denotes standard traffic).
- Mix 2: SM-12.5D—30% RAP + PG 64S-22, sampled from the Richmond District.
- Mix 3: SM-9.5D—26% RAP + PG 64S-22, sampled from the Salem District.
- Mix 4: SM-9.5A—40% RAP + PG 64S-22 + RA1, sampled from the Northern Virginia District.
- Mix 5: SM-12.5A—40% RAP + PG 64S-22 + RA2, sampled from the Hampton Roads District.
- Mix 6: SM-12.5E—15% RAP + PG 64E-22, sampled from the Hampton Roads District (E denotes extremely heavy traffic).
- Mix 7: SM-9.5A—30% RAP + PG 64S-22, sampled from the Richmond District.
- Mix 8: SM-9.5A—30% RAP + PG 64S-22, sampled from the Fredericksburg District.

The designs of Mixes 1, 2, 4, 5, 7, and 8 were developed in accordance with VDOT's BMD special provisions. Mixes 3 and 6 were designed using the conventional Superpave methodology (VDOT, 2020).

Laboratory Characterization

Volumetric Properties and Aggregate Gradations of Mixtures

The theoretical maximum specific gravity of each mixture was determined in accordance with American Association of State Highway and Transportation Officials (AASHTO) T 209, *Standard Method of Test for Theoretical Maximum Specific Gravity (G_{mm}) and Density of Asphalt Mixtures* (AASHTO, 2020). The asphalt binder content of each mixture was determined by the ignition method in accordance with Virginia Test Method (VTM)-102, *Determination of Asphalt Content from Asphalt Paving Mixtures by the Ignition Method* (VDOT, 2013). The size distribution (gradation) of the recovered aggregate was determined in accordance with AASHTO T 11, *Standard Method of Test for Materials Finer Than 75-µm (No. 200) Sieve in Mineral Aggregates by Washing* (AASHTO, 2020), and AASHTO T 27, *Standard Method of Test for Sieve Analysis of Fine and Coarse Aggregates* (AASHTO, 2020). Loose mixtures were conditioned at the compaction temperature and then compacted to N_{design} gyrations using a Superpave gyratory compactor in accordance with AASHTO T 312, *Preparing and Determining the Density of Asphalt Mixtures Specimens by Means of the Superpave Gyratory Compactor* (AASHTO, 2019). Basic physical characteristics and volumetric parameters in terms of bulk specific gravity (G_{mb}), voids in total mixture, voids in mineral aggregate, voids filled with asphalt, fines to aggregate ratio, aggregate effective specific gravity, aggregate bulk specific gravity, absorbed asphalt binder content, effective asphalt binder content, and effective film thickness were determined.

VDOT Balanced Mix Design Performance Tests

Indirect Tensile Cracking Test

IDT-CT was conducted at 25°C and a loading rate of 50 ± 3 mm/min in accordance with VTM-143, *Determination of Cracking Tolerance Index for Asphalt Mixture Using the Indirect Tensile Cracking Test (IDT-CT) at Intermediate Temperature* (VDOT, 2023a), using specimens fabricated from reheated plant-produced mixtures. All test specimens measured 150 mm in diameter by 62 mm in height and were tested at 7.0 ± 0.5% air voids. The CT index was calculated from the test load-displacement curve collected during testing. A higher CT index value indicates greater resistance to cracking.

Indirect Tensile at High Temperature Test

The IDT-HT test was conducted by applying a constant rate of axial displacement on the diametrical plane of the test specimens. The test was conducted at 54.4°C and a loading rate of 50 ± 3 mm/min in accordance with VTM-145, *Method of Test for Determining Rutting Susceptibility of Asphalt Mixtures Using the Indirect Tensile at High Temperature (IDT-HT) Test* (VDOT, 2023b), using specimens fabricated from reheated plant-produced mixtures. The specimens were fabricated at a diameter of 150 mm and a height of 62 mm at 7 ± 0.5% target air voids. The specimens were conditioned for 60 minutes in a water bath at the test temperature before being tested. Once the IDT-HT test was completed, the indirect tensile strength (S_t) of the specimens was determined. A higher S_t value indicates higher resistance to rutting.

Ruggedness Analysis

This study identified 22 factors that may influence the preparation of specimens for IDT-CT and the IDT-HT test. The seven highest priority factors were selected for further evaluation through a ruggedness analysis (i.e., Part I). Selection was based on a combination of current industry practices, survey responses, and input from the TRP.

The ruggedness analysis of the collected data was conducted in accordance with the procedures outlined in American Society for Testing and Materials (ASTM) C1067-20, *Standard Practice for Conducting a Ruggedness Evaluation or Screening Program for Test Methods for Construction Materials* (ASTM International, 2020), and ASTM E1169-21, *Standard Practice for Conducting Ruggedness Tests* (ASTM International, 2021).

Fine-Tuning Analysis—Evaluation of Significant Factors

Selected factors identified as significantly affecting test result variability during the ruggedness analysis were further evaluated in a more detailed analysis. The objective of this phase (i.e., Part II) was to determine threshold or cutoff values at which these factors begin to affect the IDT-CT and IDT-HT test results. These determinations were made using both statistical analysis and engineering judgment. The primary goal of this step is to refine and optimize specimen preparation procedures for IDT-CT and the IDT-HT test.

Development of Best Practices and Guidelines

Based on the findings from the ruggedness analysis and the subsequent fine-tuning evaluation, researchers developed a series of best practice recommendations and comprehensive guidelines. These recommendations are intended to improve the consistency and reliability of specimen preparation procedures and reduce variability in test results obtained from IDT-CT and the IDT-HT test.

RESULTS AND DISCUSSION

State of Practice—Key Factors Identified

A total of 22 factors were identified and included in a survey distributed to TRP members, VDOT district laboratory personnel, and asphalt contractors. These factors were identified based on guidance from NCHRP Report 752, NCHRP Synthesis 552, National Asphalt Pavement Association's IS-145 publication, and the experience of the research team. Please note that "sample" indicates the asphalt mixture sampled at the plant and "specimen" refers to individual asphalt specimens that have either been split into pans or compacted to a height and target air void level. The factors, labeled F1 through F22, are defined as follows:

- **F1—Mix homogenization of mixture sample containers.**
  This factor involves determining whether it is necessary to thoroughly blend material sampled from various spots of an asphalt mixture pile (from either the mixture dumped on the ground during production or sampled from the back of a truck) using multiple buckets or containers before or during the splitting process.
- **F2—Mixture sample splitting method.**
  This factor refers to the method used to split the sample, such as the quartering method, mechanical splitter, or manual scooping.
- **F3—Container types in which mixture samples are conditioned.**
  This factor involves choosing whether to heat samples in pans or buckets for example.
- **F4—Mixture sample thickness in the container.**
  This factor could involve selecting either a standard thickness or a thickness based on nominal maximum aggregate size.
- **F5—Covering of mixture samples.**
  This practice may vary regarding whether to cover samples when heating or to leave the pans uncovered in the oven.
- **F6—Location of mixture sample when placed in the oven.**
- **F7—Stirring versus not stirring during the heating and conditioning of the mixture sample.**
- **F8—Mixture sample heating to a specific duration.**
  This factor addresses variations in heating among loose mixture samples within the same batch during specimen fabrication and determines whether the mixture sample should be heated for a specific duration, regardless of the temperature.
- **F9—Mixture sample heating to a specific temperature.**

This factor addresses variations in heating among loose mixture samples within the same batch during specimen fabrication and determines whether the mixture sample should be heated to a specific temperature, irrespective of duration.

- **F10—Oven type.**
  This factor involves using various types of ovens, such as a forced-draft oven and a nonforced-draft oven.
- **F11—Types of material transfer device used to move mixture sample from the pans to the compaction molds (e.g., funnel, scoop, others).**
- **F12—The act of transferring the mixture sample from the pans to the compaction molds.**
  This factor involves using a funnel or another device to transfer the mixture sample from the pans to the molds, scooping, or transferring the mixture sample directly from the conditioning pans into the compaction mold (with no transfer device).
- **F13—Temperature of material transfer device (e.g., funnel, scoop, others).**
- **F14—Temperature of compaction molds.**
- **F15—Reheating mixture specimen pan multiple times.**
  This factor involves allowing mixture specimens to be reheated multiple times.
- **F16—Mixture sample heating process.**
  This factor involves placing all mixture samples in the oven simultaneously (at the same time) or in stages.
- **F17—Time lapsed until compaction (lag time).**
  Lag time is the time between when the mixture is sampled and then compacted.
- **F18—Time lapsed until testing (dwell time).**
  Dwell time is the time between when specimens are compacted and tested.
- **F19—Equipment used by laboratories to fabricate performance test specimens.**
  This factor includes the type of equipment used to sample, split, and compact IDT-CT and IDT-HT specimens, such as mix splitter, gyratory compactor, compaction mold, core drying device, conditioning ovens, vacuum-bag sealing device, and others.
- **F20—Storage time of loose mixture samples.**
  This factor involves determining the maximum allowable storage time for loose asphalt mixtures.
- **F21—Cooling mechanism of asphalt mixture specimen post-compaction.**
  This factor may involve various practices, including leaving the specimens in front of a fan, using accelerated cooling systems such as air conditioners, and others.
- **F22—Number of technicians involved in fabricating specimens collected from the same batch of mixtures.**

Survey respondents were asked to rate the importance of each factor on a scale from 1 to 10, with 1 indicating “not important” and 10 indicating “highly important” in minimizing variability and achieving consistent test results. Each factor was rated under three conditions—during the mix design stage, during production for non-reheated specimens, and during production for reheated specimens. Appendix B provides a summary and analysis of the responses from the three surveys.

For instance, the identified 22 factors were grouped into four major categories:

- Category 1 included factors deemed to have the lowest importance based on responses from survey participants and the TRP. These factors were either not directly aligned with the primary objective of this study, such as F3, F6, F11, F13, F14, and F19, or were already considered in Part II of the study, such as F5.
- Category 2 encompassed factors currently addressed in other ongoing studies, either by VTRC or by other institutions at the national level. This category includes factors F17, F18, F20, and F22.
- Category 3 consisted of secondary candidate factors, which served as backup options in case certain primary factors could not be considered for laboratory evaluation. This category included factors F7, F10, F12, and F21.
- Category 4 contained the most critical factors that directly support the primary objective of this study. These factors included F1, F2, F4, F8, F9, F15, and F16. Therefore, these seven factors were selected for further evaluation in Part I of the study.

Based on analysis of the data in Appendix B and further consultation with TRP members, the following seven factors were selected for consideration as part of the ruggedness study (i.e., Part I):

- F1—Mix homogenization of mixtures' sample containers. F2—Mixture sample splitting method.
- F4—Mixture specimen thickness in the container.
- F8—Mixture sample heating to a specific duration.
- F9—Mixture specimen heating to a specific temperature.
- F15—Reheating mixture specimens multiple times.
- F16—Mixture specimen heating process.

Figure 3 presents the average ratings from TRP members and contractors for the seven selected factors when considered during the mix design stage. VDOT district laboratories did not provide ratings for this stage because contractors developed the mix designs, which they submit to VDOT for review and approval only. Figure 4 shows ratings for the seven selected factors when considered during production for non-reheated specimens, again rated by TRP members and contractors. VDOT district laboratories did not rate this category because contractors typically compact non-reheated specimens during production. Figure 5 presents ratings for the production of reheated specimens, including responses from TRP members, contractors, and VDOT district laboratories.

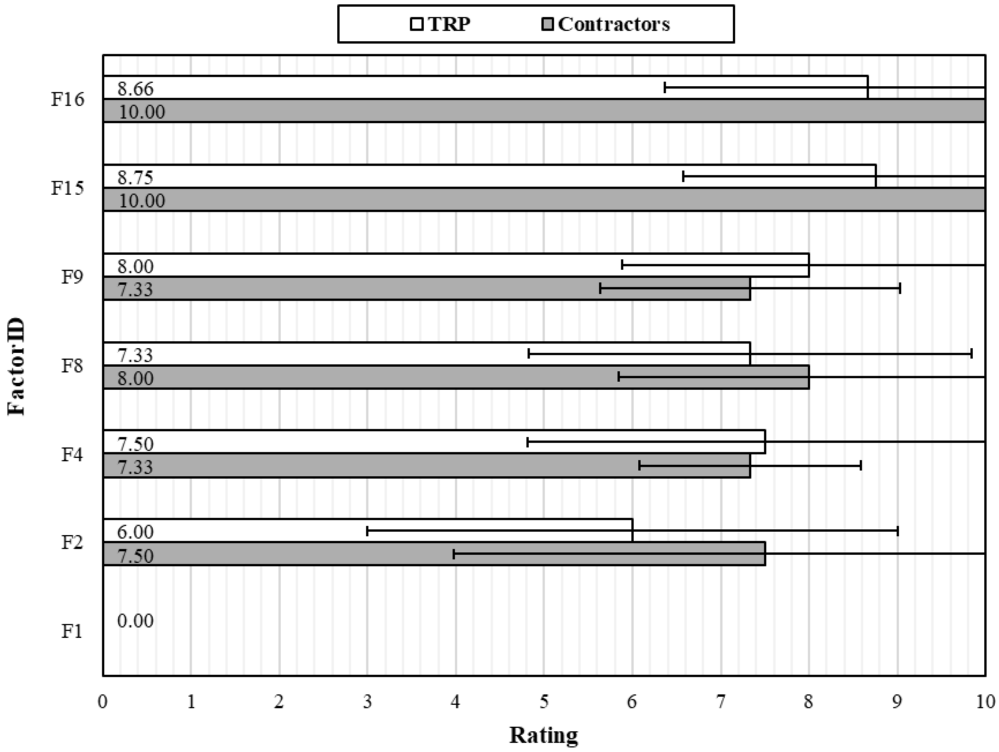

Figure 3. The Seven Selected Factors TRP Members and Contractors Rated during the Mix Design Stage. Error bars indicate plus or minus 1 standard deviation. TRP = technical review panel.

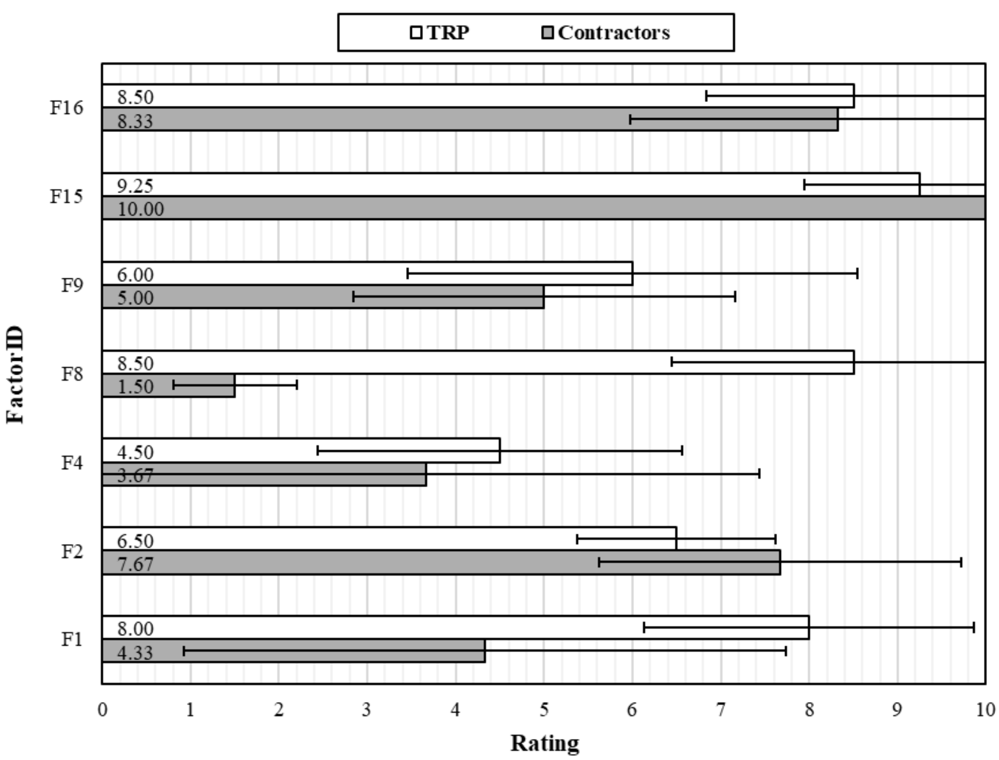

Figure 4. The Seven Selected Factors TRP Members and Contractors Rated during Production for Nonreheated Specimens. Error bars indicate plus or minus 1 standard deviation. TRP = technical review panel.

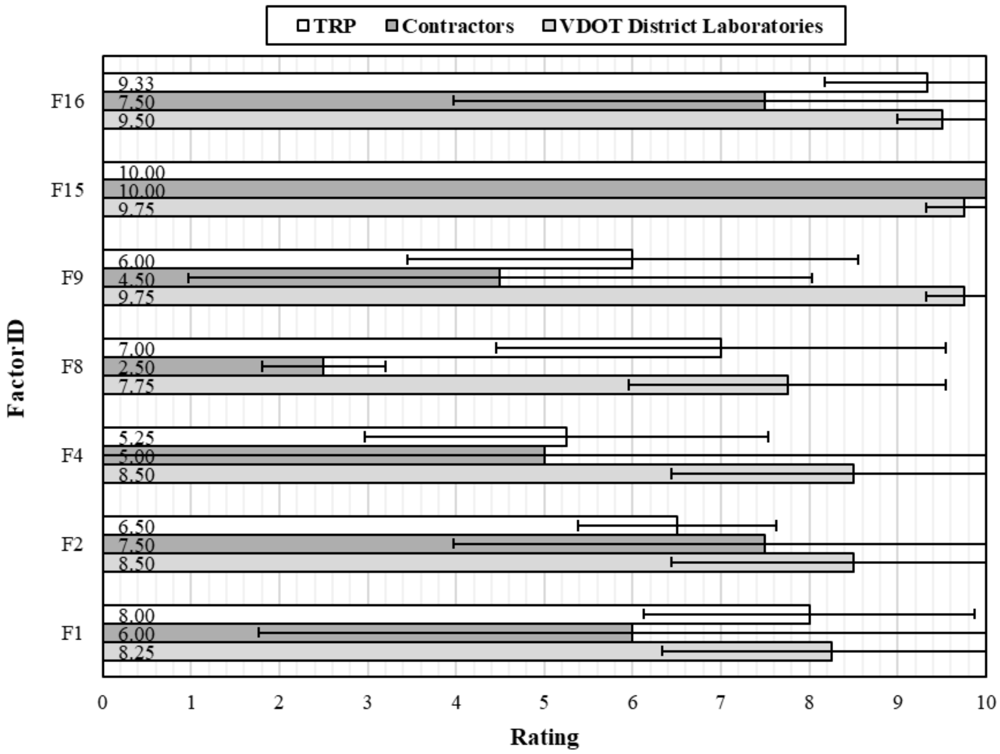

Figure 5. The Seven Selected Factors TRP Members, Contractors, and VDOT District Laboratories Rated during Production for Reheated Specimens. Error bars indicate plus or minus 1 standard deviation. TRP = technical review panel.

Laboratory Evaluation of Evaluated Asphalt Mixtures

Table 1 summarizes the volumetric properties, aggregate gradations, and performance test results for the eight asphalt mixtures evaluated in this study. Performance characteristics include Cantabro mass loss for durability, CT index for cracking measured by IDT-CT, and strength for rutting measured by IDT-HT test. All performance tests were conducted on reheated specimens prepared in the VTRC laboratory. The selected mixtures featured two nominal maximum aggregate sizes (i.e., 9.5 mm and 12.5 mm) and a range of RAP contents. The experimental program also includes mixtures incorporating two different RAs and a variety of warm mix asphalt additives. These mixtures exhibited a wide range of CT index values, including some below the threshold of 70, other mixtures near 70, and several mixtures exceeding 70. This diverse selection was designed to enhance the generalizability of the study findings across a wide range of materials and field practices.

Table 1. Volumetric Properties, Aggregate Gradation, and Selected Performance Properties for All Evaluated Mixtures

Mixture ID	**Mix 1**	**Mix 2**	**Mix 3**	**Mix 4**	**Mix 5**	**Mix 6**	**Mix 7**	**Mix 8**
Description	BMD	BMD	Non-BMD	BMD	BMD	Non-BMD	BMD	BMD
District	Culpeper	Richmond	Salem	Northen Virginia	Hampton Roads	Hampton Roads	Richmond	Fredericksburg
Composition								
RAP Content, %	15	30	26	40	40	15	30	30
Asphalt Binder	PG 64S-22	PG 64S-22	PG 64S-22	PG 64S-22	PG 58-28	PG 64E-22	PG 64S-22	PG 64S-22
Additives	WMA	WMA	WMA	WMA + RA	WMA	WMA	WMA	WMA
Property								
N_{design}, gyrations	50	50	50	50	50	50	50	50
NMAS, mm	12.5	12.5	9.5	9.5	12.5	12.5	9.5	9.5
Asphalt Content, %	5.99	6.11	6.08	6.14	5.64	5.78	5.77	6.18
Rice SG (G_{mm})	2.568	2.513	2.553	2.677	2.500	2.472	2.423	2.537
VTM, %	3.4	5.0	4.1	2.6	5.1	4.2	3.5	4.3
VMA, %	17.8	19.2	18.6	18.1	18.0	17.0	16.4	18.7
VFA, %	80.9	73.8	77.8	85.8	71.8	75	79	77
FA Ratio	1.11	1.06	1.16	1.15	0.82	0.7	0.8	0.9
Mixture Bulk SG (G_{mb})	2.481	2.387	2.448	2.608	2.373	2.367	2.337	2.428
Effective SG (G_{se})	2.838	2.773	2.822	2.989	2.734	2.704	2.642	2.808
Performance Properties at Optimum Asphalt Binder Content								
Cantabro Mass Loss, %	7.9	10.8	7.5	6.6	8.0	--	--	--
CT index at 25°C	119	51	82	82	83	--	--	--
Strength at 54.4°C, kPa	106	211	172	174	114	--	--	--
Gradation / Sieve Size	**% Passing**							
3/4 inch (19.0 mm)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1/2 inch (12.5 mm)	94.8	98.5	98.5	100.0	97.9	98.0	100.0	100.0
3/8 inch (9.5 mm)	85.6	90.4	94.7	96.1	88.8	89.0	94.0	94.0
No. 4 (4.75 mm)	59.7	59.3	66.6	66.8	57.4	63.0	57.0	64.0
No. 8 (2.36 mm)	38.5	38.3	44.8	45.6	38.1	46.0	34.0	41.0
No. 16 (1.18 mm)	26.5	26.8	28.4	32.8	28.7	32.0	26.0	30.0
No. 30 (600 µm)	19.4	19.1	18.7	23.7	19.9	22.0	20.0	21.0
No. 50 (300 µm)	15.1	13.5	13.0	16.6	11.7	15.0	14.0	15.0
No. 100 (150 µm)	10.9	9.4	9.5	11.0	6.9	7.0	9.0	9.0
No. 200 (75 µm)	6.6	6.5	7.1	7.0	4.6	3.9	4.8	5.2

-- = not available; BMD = balanced mix design; CT = cracking tolerance; E = extremely heavy traffic; FA = fines to aggregate; G_{se} = aggregate effective specific gravity; NMAS = nominal maximum aggregate size; PG = performance grade; RA = recycling agent; RAP = reclaimed asphalt pavement; S = standard traffic; SG = specific gravity; VFA = voids filled with asphalt; VMA = voids in mineral aggregate; VTM = voids in total mixture; WMA = warm mix additives.

Ruggedness Analysis and Results

Ruggedness Levels

The ruggedness analysis, as outlined in ASTM C1067, requires assigning two levels to each prescribed factor—quantitative or qualitative (ASTM International, 2020). The seven factors identified in the preceding section were evaluated in this part of the study. Table 2 presents the identified factors along with the levels selected for each. The determination of these levels was informed by existing practices in relevant standards (e.g., AASHTO). Where standards were silent, the research team relied on practices potentially employed in the field and their theoretical and practical knowledge.

Table 2. Factors and Levels for the Ruggedness Analysis

Factor ID	Factor Description	Plus (+) Level	Minus (–) Level
F1	Homogenization of mixture sample containers	Two boxes (60–70 lb each)	Single box (60–70 lb)
F2	Mixture sample splitting method	AASHTO R 47—Quartering with template	Scooping
F4	Mixture specimen thickness in the container	1 inch thick (AASHTO R 30compliant)	3 inches thick
F8	Mixture sample heating	6.5–7.0 hours at compaction temperature	3.0–3.5 hours at compaction temperature
F9	Mixture specimen heating	1.5–2.5 hours at compaction temperature	Time required for specimen to reach compaction temperature
F15	Reheating mixture specimens multiple times	Reheat and compact specimen pans the day after splitting	Compact specimen pans immediately after sample splitting
F16	Mixture specimen heating process	Stagger pans in oven (each 15 minutes apart)	Pans in oven simultaneously

AASHTO = American Association of State Highway and Transportation Officials.

Two levels for each factor were designated using plus (+) and minus (–) signs, as in Table 2 shows. Following the ASTM C1067 procedure (ASTM International, 2020), combinations of factor levels for each determination, defined as the numerical value of a test specimen characteristic measured in accordance with the specified test method, were assigned per the layout in Table 3. This process resulted in eight determinations (designated D1 through D8). Using this framework, test specimens from Mix 1 through Mix 5 were prepared for each determination, and IDT-CT and the IDT-HT test were conducted in accordance with VDOT test methods. It must be noted that determination number 4 (D4) closely aligns with VTRC's current specimen preparation practices.

Table 3. Pattern of Assigning Levels to Seven Factors in Accordance with ASTM C1067 (ASTM International, 2020)

Factor ID	Determination Number							
	D1	D2	D3	D4	D5	D6	D7	D8
F1	-[a]	-	-	-	+	+	+	+
F2	-	-	+	+	-	-	+	+
F4	-	+	-	+	-	+	-	+

Factor ID	Determination Number							
	D1	D2	D3	D4	D5	D6	D7	D8
F8	+	+	-	-	-	-	+	+
F9	+	-	+	-	-	+	-	+
F15	+	-	-	+	+	-	-	+
F16	-	+	+	-	+	-	-	+

ASTM = American Society for Testing and Materials. [a] Plus (+) and minus (–) signs indicate the factor levels defined in Table 2.

Ruggedness Analysis Results

Figure 6 presents the CT index results for each determination across the evaluated mixtures. The results demonstrate that CT index magnitudes are sensitive to variations in specimen preparation practices. Although the statistically significant impacts of these variations are discussed in following sections, it is important to highlight that differences in preparation procedures can also lead to meaningful practical implications.

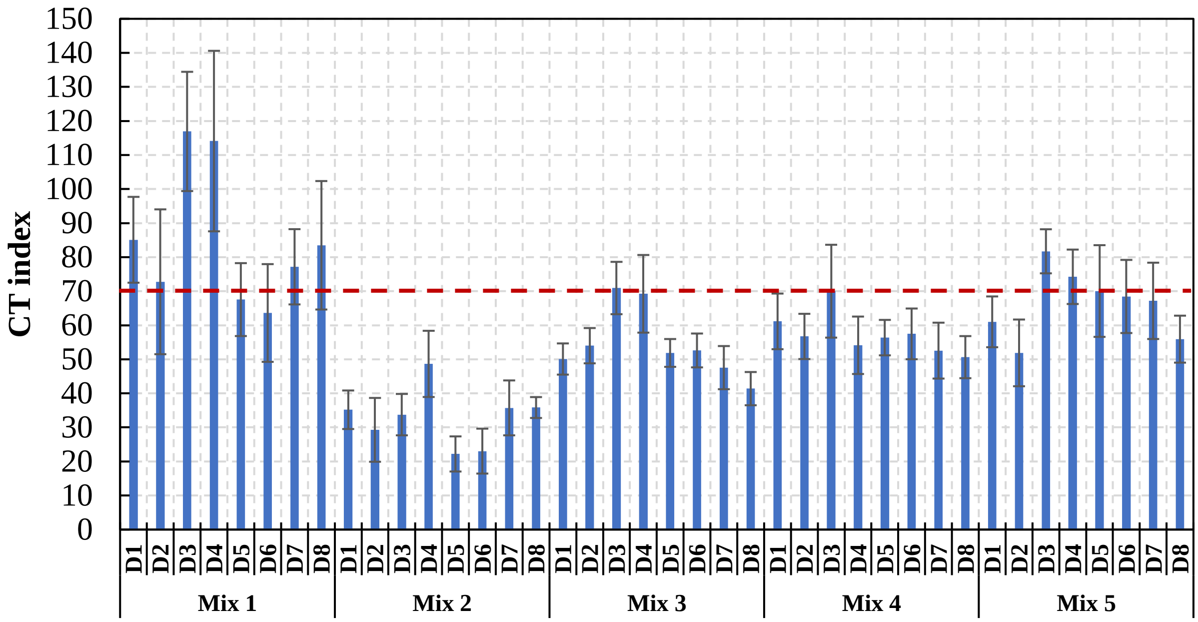

Figure 6. CT Index Results for Each Determination in Accordance with ASTM C1067 Procedure across Evaluated Mixtures (ASTM International, 2020). Red dashed line indicates VDOT's current CT index threshold of 70 for reheated mixtures. Error bars indicate plus or minus 1 standard deviation. ASTM = American Society for Testing and Materials; CT = cracking tolerance.

Using D4 as a reference point, the mixture's acceptance or compliance status—based on the established CT index threshold of 70—may vary depending on the preparation procedure employed. For Mix 1, using D4 as the reference, the mixture is deemed acceptable for cracking resistance because the CT index exceeds 70. However, under determinations D5 and D6, the CT index falls below 70, suggesting the mixture would be classified as unacceptable, thus yielding a false negative outcome. Similarly, Mix 5 also demonstrates false negative outcomes under certain preparation procedures.

Conversely, for Mix 3 and Mix 4, false positive outcomes could be observed. Under D4, both mixtures fall below the threshold and are regarded as unacceptable. However, if the mixtures were prepared following the procedure represented by D3, they would meet the acceptance criterion. An additional noteworthy observation from Figure 6 is the magnitude of relative range (expressed as a percentage), calculated as the difference between the maximum and minimum values (range), divided by the average (mean), and multiplied by 100. The relative ranges for Mix 1 through Mix 5 are 62.7%, 80.3%, 54.0%, 33.8%, and 45.0%, respectively, indicating substantial variability across different specimen preparation procedures.

The test for equal variances, conducted at a 95% confidence level on each mixture set, indicated that different specimen preparation procedures did not have statistically significant impacts on the variability of the CT index, as evidenced by p-values greater than 0.05. In other words, each specimen preparation procedure resulted in statistically similar variances across the five mixtures evaluated.

Figure 7 presents the strength results for each determination across the evaluated mixtures. Similar to the CT index results, strength magnitudes are sensitive to variations in specimen preparation practices. However, within this dataset, no practical impacts—such as false positive or false negative outcomes—were observed. The relative range for strength values for Mix 1 through Mix 5 is 32.1%, 40.2%, 23.9%, 12.8%, and 59.7%, respectively. This range suggests that specimen preparation variability notably influences strength measurements. On the other hand, the test for equal variances, conducted at a 95% confidence level on each mixture set, indicated that different specimen preparation procedures did not significantly affect the variability of the strength results, as evidenced by p-values greater than 0.05.

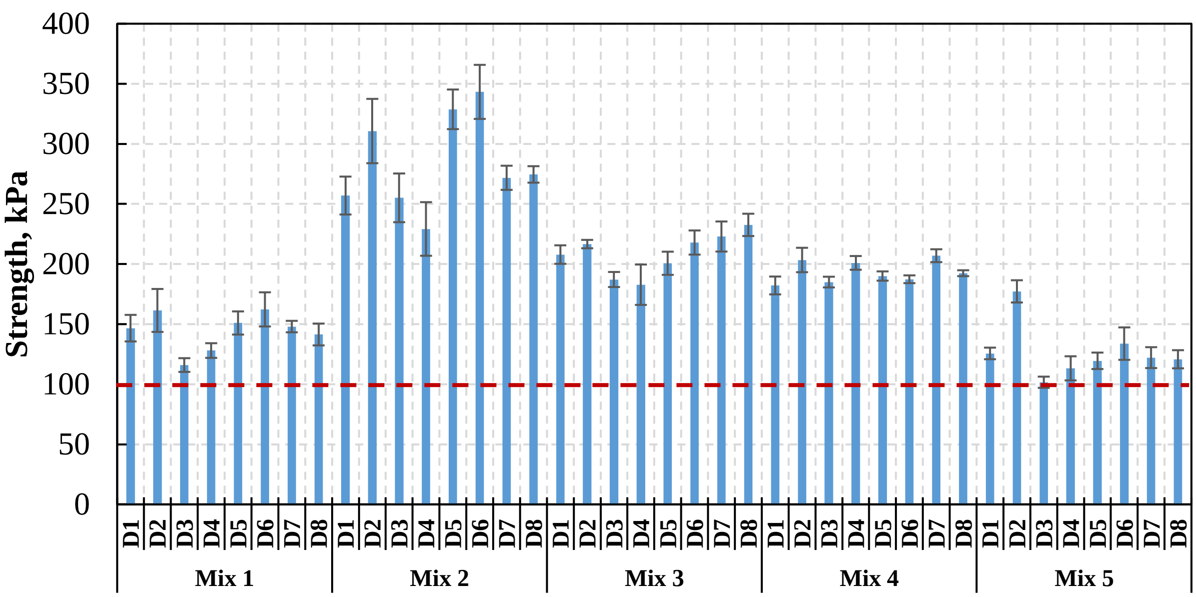

Figure 7. Strength Results for Each Determination in Accordance with ASTM C1067 Procedure across Evaluated Mixtures (ASTM International, 2020). Red dashed line indicates VDOT's current strength threshold of 100 kPa for nonreheat and reheated mixtures conditioned in water without sealing. Error bars indicate plus or minus 1 standard deviation. ASTM = American Society for Testing and Materials.

Overall, the findings emphasize the critical importance of implementing unified specimen preparation practices to improve the reliability and consistency of the test results for both IDT-CT and the IDT-HT test.

The ruggedness analyses for the CT index and strength, summarized in Tables 4 and 5, respectively, were conducted in accordance with the statistical procedures outlined in ASTM C1067 at a 95% confidence level (ASTM International, 2020). In these analyses, computed F values were compared with a defined significance threshold of 5.32, in which an F value exceeding 5.32 indicates that the corresponding factor has a statistically significant effect on the test results. The ruggedness analyses provide critical insights into the influence of preparation-related factors on the repeatability and reliability of tests. Across both test methods, several common observations emerged.

Table 4. Ruggedness Analysis Results for Cracking Tolerance Index Measurements in Accordance with ASTM C1067 (ASTM International, 2020)

Factor ID	Factor Description	F Value*				
		Mix 1	Mix 2	Mix 3	Mix 4	Mix 5
F1	Homogenization of mixture sample containers	40.61	27.14	20.90	9.62	0.77
F2	Mixture sample splitting method	45.40	58.82	3.41	0.30	11.14
F4	Mixture specimen thickness in the container	0.70	2.96	0.08	6.89	12.64
F8	Mixture sample heating	8.20	2.12	21.53	4.42	49.85
F9	Mixture specimen heating	1.33	1.95	0.47	5.87	0.20
F15	Reheating mixture specimens multiple times	1.68	12.09	1.27	3.24	0.94
F16	Mixture specimen heating process	0.00	13.92	0.01	1.10	1.86

ASTM = American Society for Testing and Materials. *The computed F values at a 95% confidence interval were evaluated for their statistical significance across five asphalt mixtures, with an F value exceeding 5.32 indicating significance. Cells highlighted in yellow correspond to factors that exhibited an F value exceeding 5.32 for a given mixture.

Table 5. Ruggedness Analysis Results for Strength Measurements in Accordance with ASTM C1067 (ASTM International, 2020)

Factor ID	Factor Description	F Value*				
		Mix 1	Mix 2	Mix 3	Mix 4	Mix 5
F1	Homogenization of mixture sample containers	24.79	60.41	59.96	0.54	6.95
F2	Mixture sample splitting method	74.99	99.99	2.93	10.11	144.32
F4	Mixture specimen thickness in the container	9.65	5.25	9.29	8.00	86.58
F8	Mixture sample heating	15.67	3.41	79.28	9.41	89.98
F9	Mixture specimen heating	4.74	0.36	4.80	59.84	37.73
F15	Reheating mixture specimens multiple times	4.11	19.58	4.03	5.92	46.37
F16	Mixture specimen heating process	2.25	11.36	0.28	0.88	8.87

ASTM = American Society for Testing and Materials. *The computed F values at a 95% confidence interval were evaluated for their statistical significance across five asphalt mixtures, with an F value exceeding 5.32 indicating significance. Cells highlighted in yellow correspond to factors that exhibited an F value exceeding 5.32 for a given mixture.

The homogenization of mixture sample containers (F1) consistently demonstrated significant effects, appearing as a key contributor to variability (in terms of parameter magnitude) in four of five mixtures for both CT index and strength. This activity highlights the essential need for proper homogenization prior to splitting and specimen preparation to reduce variability introduced by material inconsistencies. In addition, the main effect analysis, as conducted in accordance with ASTM E1169 at a 95% confidence level (ASTM International, 2021), indicated that the minus level (a single box) resulted in higher CT index values or lower strength values on average compared with the plus level (homogenizing). This effect is likely related to the longer oven conditioning times required when homogenizing larger sample masses, as in the case of combining multiple boxes (two in the case of this study). These observations indicate that the minimum sample size required for each test method should be obtained in accordance with AASHTO R 97. Specifically, results for a given test should be obtained from a single, unified sample, not from multiple separate samples. When smaller containers are used during sampling, these containers should be thoroughly combined before splitting to ensure uniformity and consistency across replicate specimens.

The mixture sample splitting method (F2) also emerged as a consistently significant factor, influencing four out of five mixtures in strength and three out of five mixtures in CT index. In addition, the main effect analysis, as conducted in accordance with ASTM E1169 at a 95% confidence level (ASTM International, 2021), indicated that the plus level (quartering in accordance with AASHTO R 47) resulted in higher CT index values or lower strength values on average compared with the minus level (scooping). These observations highlight the importance of standardized and carefully controlled splitting practices to ensure uniformity across replicate specimens. Nonstandard practices, such as scooping, introduce significant variability in test results. Therefore, strict adherence to the procedures outlined in AASHTO R 47 for obtaining representative specimens is critical to maintaining the reliability and consistency of test outcomes.

The specimen thickness in the container (F4) was another significant factor, affecting four of five mixtures for strength and three of five mixtures for CT index. Therefore, maintaining consistent specimen thickness during preparation is essential to minimize variability in test outcomes. In addition, the main effect analysis, as conducted in accordance with ASTM E1169 at a 95% confidence level (ASTM International, 2021), indicated that the minus level (3-inch-thick nonstandard container) resulted in higher CT index values or lower strength values on average compared with the plus level (1-inch-thick AASHTO R 30-compliant pan). It should be noted that although AASHTO R 30 is designed for short-term oven aging of laboratory-produced mixtures to match the aging of plant-produced mixtures, the research team viewed the pan requirements in AASHTO R 30 as a good baseline equipment requirement to help achieve more consistent results between laboratories. Although specimen thickness in the container is a defined factor under AASHTO R 30, anecdotal evidence suggests that some laboratories use containers in practice such as pails (rather than pans), with specimens covered by foil. This factor was further explored and analyzed to better understand the effect of this practice on test results, with the findings presented in the fine-tuning section of this study.

Like F1, the mixture sample heating (F8) consistently demonstrated significant effects, emerging as a key contributor to variability in four out of five mixtures for both CT index and strength measurements. In addition, the main effect analysis, as conducted in accordance with ASTM E1169 at a 95% confidence level (ASTM International, 2021), indicated that the minus level (sample heating for 3.5 hours) resulted in higher CT index values or lower strength values on average compared with the plus level (sample heating for 7 hours). These observations underscore the essential need for standardized and carefully controlled heating practices to minimize variability introduced during sample conditioning. In this ruggedness phase of the study, F8 was explored using specific heating times applicable to the particular container type employed—namely, 10 × 10 × 10-inch cardboard boxes. However, it is important to note that field practices often involve using different container types and sizes, such as smaller cardboard boxes or canvas bags, when sampling asphalt mixtures. Given the mass differences associated with these varying container types, the time required to achieve proper sample heating would differ accordingly. Therefore, this factor was further explored and analyzed, with the detailed findings presented in the fine-tuning section of this study.

The specimen heating (F9) showed more selective significance, influencing the variability in two out of five mixtures for strength and only one mixture for CT index. However, a review of the specimen tracking records indicated that most determinations maintained statistically equivalent temperatures during compaction. This uniformity may have masked the potential effect of this factor on the observed test results. Consequently, this factor was further explored and analyzed, with the detailed findings presented in the fine-tuning section of this study.

Reheating the specimens multiple times (F15) showed selective significance for the CT index, influencing only one mixture. However, it emerged as a meaningful factor for strength, affecting three out of five mixtures. These findings highlight the need to specify a standard practice for this factor—such as either splitting the mixture into pans and compacting the following day after reheating, or alternatively, reheating and splitting the mixture on the same day before compaction—to ensure greater uniformity and consistency in laboratory procedures.

The specimen heating process (F16) also showed selective significance, influencing one out of five mixtures for the CT index and two out of five mixtures for strength. Although this factor was not ranked among the most influential, caution should be exercised regarding its potential effect on test outcomes, particularly when handling large volumes of test specimens, for which heating consistency among replicates may become more difficult to maintain.

Fine-Tuning Analysis—Evaluation of Significant Factors

F4—Specimen Thickness in the Container

The influence of loose mixture specimen thickness in containers (F4) was further evaluated in this part of the study through a targeted investigation comparing the effects of different container types and the practice of covering containers with aluminum foil during oven conditioning. Specimens prepared from Mix 1 through Mix 3 under these varied conditions were subsequently fabricated and tested using both IDT-CT and the IDT-HT test to assess the potential effects on test outcomes.

Two container types commonly used in practice were selected for this evaluation. The first was a shallow pan, resulting in a loose mixture thickness of approximately 1 inch (25 millimeters), representing the lower end of the AASHTO R 30 allowable range. The second was a pail—specifically, a 5-quart bucket Virginia contractors frequently use—which, under typical use, produced loose mixture thicknesses near the upper end of the AASHTO R 30 allowable range (approximately 50 millimeters or 2 inches). Figure 8 illustrates the two container types used in this study. For each container type, two conditions were evaluated: (1) mixtures covered with aluminum foil during oven conditioning and (2) mixtures left uncovered. In each scenario, a temperature probe was placed in the middle of the loose mixture to monitor internal temperatures and the time required to reach the target compaction temperature. For each pan type in this study, a separate metal funnel was used to transfer the mix to the gyratory mold in one lift per the requirements in AASHTO T 312.

(a) (b)

Figure 8. Container Types Used for Evaluating Loose Mixture Specimen Thickness in Containers: (a) Shallow Pan; (b) 5-Quart Pail

Figure 9 shows the average specimen time in the oven for Mix 1 through Mix 3 under four conditioning combinations—pail with no foil, pail with foil, pan with no foil, and pan with foil. The results clearly show that specimens conditioned in pails consistently required longer oven times to reach the compaction temperature compared with those specimens in pans, regardless of whether they were covered with foil. Furthermore, the use of aluminum foil further extended the oven time for both container types. The longest oven times were observed for the pail with foil configuration, and the shortest times were consistently associated with the pan without foil.

Figure 10 presents the CT index results for Mix 1 through Mix 3 under four different conditioning combinations—pail with no foil, pail with foil, pan with no foil, and pan with foil. Although container type and cover condition produced some variation in CT index values, the overall magnitudes within each mixture remained relatively consistent. Notably, the lowest CT index values were generally observed for mixtures conditioned in a pan with a 1-inch thickness and without foil covering compared with those mixtures with foil covering or thicker layer thicknesses. Figure 10 shows that although this observation was not statistically different from the other conditions, it can be attributed to the potential adverse effects of oxidation and thermal gradients during conditioning, which may lead to less favorable CT index outcomes.

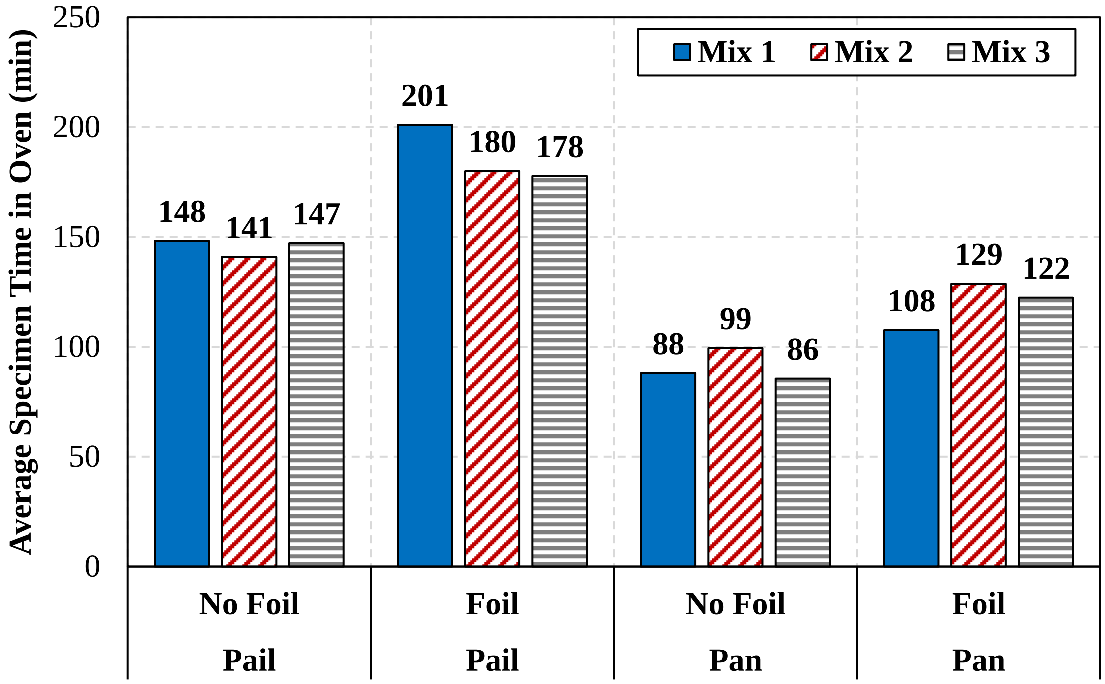

Figure 9. Average Specimen Time in Oven under Different Conditioning Combinations, Measured as the Time Required to Reach the Target Compaction Temperature

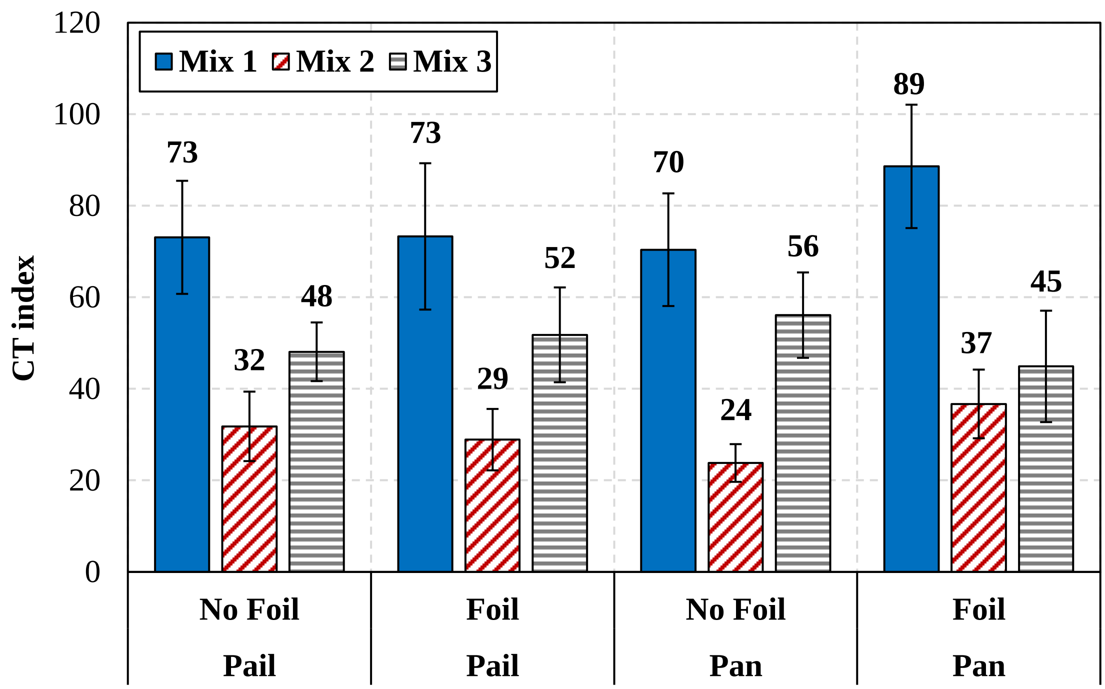

Figure 10. CT Index Results under Different Conditioning Combinations. Error bars indicate plus or minus 1 standard deviation. CT = cracking tolerance.

Table 6 presents the results of Tukey pairwise analysis, conducted at a 95% confidence level, which evaluated the statistical similarity of the different container and foil covering combinations for each mixture. Combinations sharing the same letter are statistically similar, meaning no significant difference was detected between them. As shown, the effect of container type and covering condition did not result in statistically significant differences in CT index values for all mixture combinations evaluated, except in the case of Mix 2 for the pan with foil covering condition. This isolated observation is likely attributable to the inherent variability of the CT index. In fact, when considering the single-operator precision estimate for the CT index (18.3%), this case still represents statistically similar results compared with the other combinations within Mix 1.

Table 6. Tukey Pairwise Analysis at a 95% Confidence Level for Cracking Tolerance Index Results

Pan Type	Foil Covering	Mix 1	Mix 2	Mix 3
		Statistical Grouping*		
Pan	Yes	A	A	A
	No	A	B	A
Pail	Yes	A	A B	A
	No	A	A B	A

*Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

A similar analysis conducted for F4 (loose mixture specimen thickness in containers) for the CT index was also performed for the strength parameter obtained from the IDT-HT test. Figure 11 presents the strength results for Mix 1 through Mix 3 under four different conditioning combinations—pail with no foil, pail with foil, pan with no foil, and pan with foil. Table 7 presents the results of the Tukey pairwise analysis, conducted at a 95% confidence level.

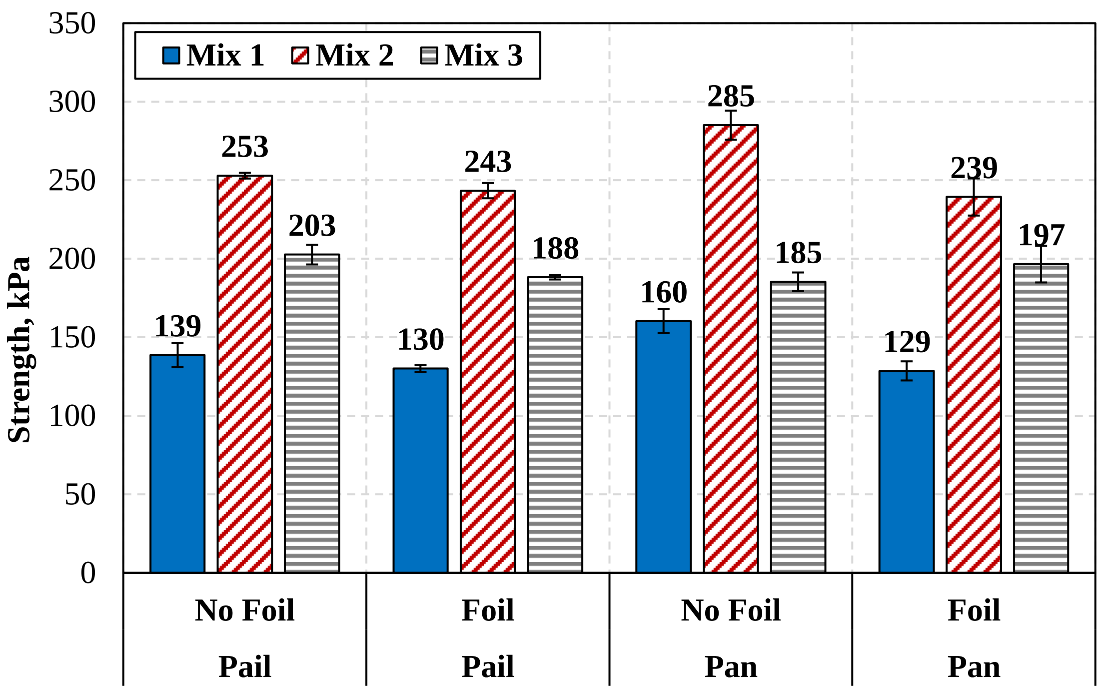

Figure 11. Strength Results under Different Conditioning Combinations. Error bars indicate plus or minus 1 standard deviation.

Table 7. Tukey Pairwise Analysis at a 95% Confidence Level for Strength Results

Pan Type	Foil Covering	Mix 1	Mix 2	Mix 3
		Statistical Grouping*		
Pan	Yes	B	B	A B
	No	A	A	B
Pail	Yes	B	B	A B
	No	B	B	A

*Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

The results indicate similar trends to those observed for the CT index. However, the pan with no foil condition significantly affected the strength results statistically for all three mixtures evaluated. This observation aligns with the findings from the CT index, where conditioning in thinner specimen thicknesses without foil covering resulted in increased strength values. This trend is expected to correspond with a decrease in CT index values, as was previously discussed.

Notably, because the IDT-HT test exhibits greater repeatability compared with IDT-CT, the effect of this factor was more pronounced for the strength parameter. Nevertheless, when considering the single-operator variability for the strength parameter (10.8%), the observed differences would not be considered statistically significant because they fall within the expected precision range of the test method.

Evaluation of this factor (F4) indicated that using containers compliant with AASHTO R 30, which specifies that the loose mixture specimen thickness be maintained between 1 and 2 inches (25 and 50 millimeters) during conditioning, may help minimize variability and contribute to more consistent outcomes, supporting its use as a guideline for specimen preparation procedures. The analysis also showed that using foil coverings during oven conditioning did not significantly affect test results statistically. However, it was observed to introduce unnecessary additional costs and, more importantly, to increase the time required to reach the target compaction temperature.

F8—Mixture Sample Heating

This factor demonstrated a statistically significant influence on test results during the ruggedness evaluation, particularly regarding the specific heating times applicable to the container type employed—namely, the 10 × 10 × 10-inch cardboard boxes that VTRC typically uses to sample its mixtures for testing and evaluation. As previously noted, field practices often involve using different container types and sizes, such as smaller cardboard boxes or canvas bags, when sampling asphalt mixtures. These varying container types inherently require different heating times to achieve proper and uniform sample conditioning.

To address this variance, the factor was further explored and analyzed as part of this study phase using additionally sampled mixtures (Mix 6 through Mix 8) placed in containers of varying sizes. As Figure 12 shows, three container types were used—the standard 10 × 10 × 10-inch cardboard box that VTRC uses (designated as the large box), a smaller 10 × 10 × 6-inch box that some VDOT districts and contractors commonly use (designated as the small box), and canvas bags also frequently used in VDOT field practices. Mixtures sampled in each container were subjected to three different heating durations of 2.5 ± 0.5 hours, 4.0 ± 0.5 hours, and 5.5 ± 0.5 hours. In each scenario, a temperature probe was inserted into the center of the container to monitor internal temperatures. Following conditioning and processing, the specimens were tested using IDT-CT and the IDT-HT test to evaluate the effect of this factor on test results.

Figure 12. Containers Evaluated for Heating Time Effects (left to right): Virginia Transportation Research Council's 10 × 10 × 10-Inch Box, 10 × 10 × 6-Inch Box, and Canvas Bags Used in VDOT Field Sampling.

Figure 13 presents the measured sample temperatures at the center of the mixture for Mix 6 through Mix 8 across three container types—canvas bags, large boxes, and small boxes under three heating durations: 2.5 ± 0.5 hours, 4.0 ± 0.5 hours, and 5.5 ± 0.5 hours. As expected, the results show that longer heating durations led to higher internal sample temperatures overall across all container types. However, it is important to note that none of the evaluated conditions, except for Mix 6 in the canvas bags at 4.0 ± 0.5 hours and 5.5 ± 0.5 hours heating durations, achieved the target compaction temperatures at the center of the mixture. Despite this observation, the mixtures were still workable and could be processed and compacted even at the shortest heating duration of 2.5 ± 0.5 hours.

Table 8 presents the average CT index values and results from a Tukey pairwise analysis conducted at a 95% confidence level for Mix 6 through Mix 8 across different container types and heating durations. For Mix 6, no statistically significant differences were observed across most conditions, with all combinations grouped primarily under A or A/B. For Mix 7, greater statistical separation was observed, particularly under longer heating durations in the large box and canvas bag configurations, with groupings ranging from A to D. For Mix 8, all combinations remained statistically similar under grouping A, indicating no significant differences despite variations in container type or heating time.

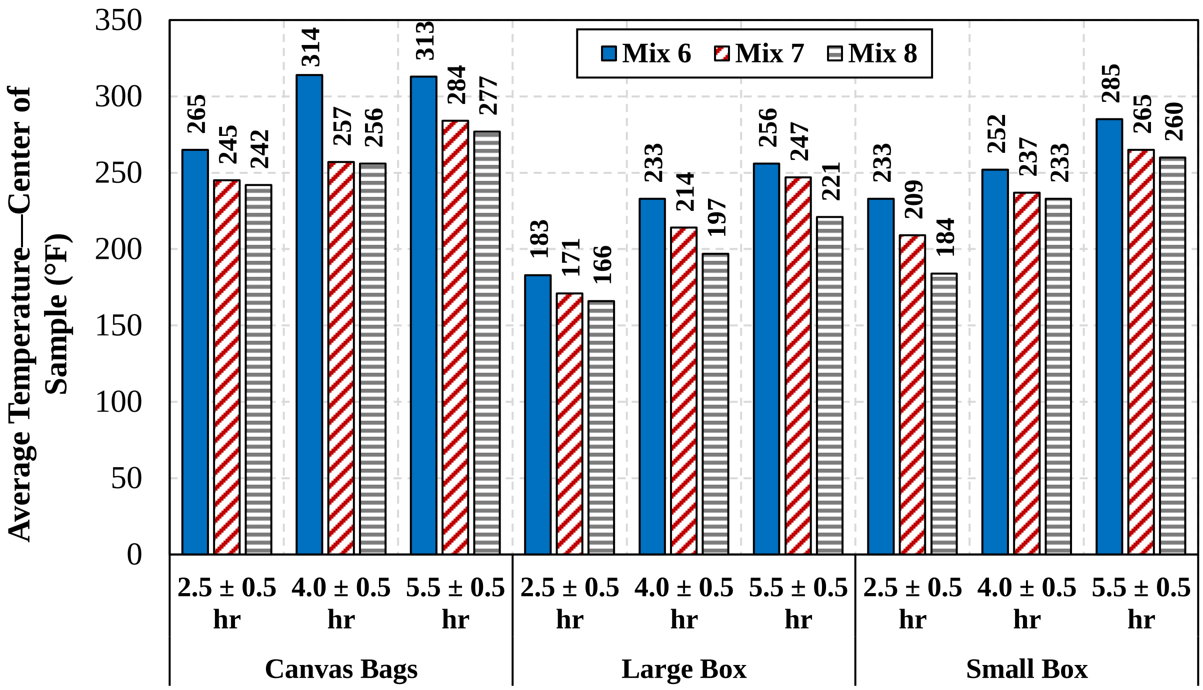

Figure 13. Average Temperatures Measured at the Center of Samples for Different Container Types and Three Heating Durations. hr = hour.

Table 8. Average CT Index Values and Tukey Pairwise Analysis at a 95% Confidence Level for Mix 6 through Mix 8 across Container Type and Heating Durations

Container Type	Container Heating Time (hours)	Mix 6		Mix 7		Mix 8	
		CT Index	Grouping*	CT Index	Grouping*	CT Index	Grouping*
Small Box	2.5 ± 0.5	64.0	A B	53.3	A	104.8	A
	4.0 ± 0.5	66.5	A B	50.7	A B	94.5	A
	5.5 ± 0.5	63.3	A B	44.2	A B C	90.0	A
Large Box	2.5 ± 0.5	76.9	A	47.5	A B	89.8	A
	4.0 ± 0.5	72.1	A B	38.3	B C D	92.1	A
	5.5 ± 0.5	64.3	A B	32.9	C D	95.7	A
Canvas Bags	2.5 ± 0.5	72.2	A B	51.4	A B	97.7	A
	4.0 ± 0.5	56.7	B C	49.0	A B	91.0	A
	5.5 ± 0.5	41.9	C	27.7	D	82.4	A

CT = cracking tolerance. *Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

Importantly, the data show a general trend in which increasing the heating duration from 2.5 to 5.5 hours resulted in decreased CT index values across all container types. This pattern was most pronounced in Mix 7, for which longer heating times were associated with statistically lower CT index groupings. However, a similar downward trend was observed in Mix 6 and Mix 8, even when not statistically significant. Although some minor variations between container types were also noted, the most consistent and notable trend across the dataset was the negative relationship between heating time and CT index. Although CT index magnitudes showed some variability overall, the statistical analysis suggests that these differences largely fall within the expected variability range, especially when considering the single-operator variability of the test method.

Table 9 presents the average strength values (in kPa) and results from a Tukey pairwise analysis conducted at a 95% confidence level for Mix 6 through Mix 8 across different container types and heating durations. For Mix 6, strength values increased with longer heating durations, with the highest strength recorded in the canvas bags at 5.5 hours (261.4 kPa), showing clear statistical separation compared with shorter durations. For Mix 7, a similar trend was observed in which strength generally increased with heating time. The highest strength was again recorded in the canvas bags at 5.5 hours (179.7 kPa), and shorter durations in the small box showed significantly lower values. For Mix 8, strength values also increased with heating duration, with the highest values again in the canvas bags at 5.5 hours (155.2 kPa), contrasting with lower strength values under shorter durations.

Table 9. Average Strength Values and Tukey Pairwise Analysis at a 95% Confidence Level for Mix 6 through Mix 8 Across Container Type and Heating Durations

Container Type	Container Heating Time (hours)	Mix 6		Mix 7		Mix 8	
		Strength, kPa	Grouping*	Strength, kPa	Grouping*	Strength, kPa	Grouping*
Small Box	2.5 ± 0.5	163.3	D	111.0	E	126.9	C D
	4.0 ± 0.5	204.0	B C	124.5	E	136.0	B C D
	5.5 ± 0.5	-	-	142.4	B C	139.7	B C
Large Box	2.5 ± 0.5	182.5	C D	129.3	C D	128.8	C D
	4.0 ± 0.5	182.9	C D	137.8	B C D	134.5	C D
	5.5 ± 0.5	205.6	B	151.5	B	131.2	C D
Canvas Bags	2.5 ± 0.5	178.0	D	131.6	C D	122.6	D
	4.0 ± 0.5	221.0	B	132.9	C D	149.0	A B
	5.5 ± 0.5	261.4	A	179.7	A	155.2	A

- = data not available (combination not tested). *Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

Based on the analysis of CT index and strength results across different container types and heating durations, limiting sample heating time to a maximum of 4 hours, with a suggested minimum heating time of 2.5 hours, will help produce more consistent test results. Although container type did not have a statistically significant effect overall, the data showed that heating samples in canvas bags may lead to more pronounced changes in both CT index and strength values compared with heating samples in boxes.

F9—Mixture Specimen Heating

Building on the findings from the ruggedness phase, in which the specimen heating factor (F9) showed selective but limited statistical significance—influencing variability in only a few mixtures and potentially masked by uniform compaction temperatures—this factor was further investigated in the fine-tuning phase of the study. The objective of this phase was to more closely examine the effects of specimen heating time to better understand how this factor influences the test outcomes.

To that end, loose mixture specimens from Mix 1 through Mix 3 were prepared and placed in an oven set at the higher end of the compaction temperature range specified in each mixture's job mix formula. This evaluation utilized the AASHTO R 30-compliant (9 × 13 × 1-inch) pans discussed previously for all specimens. The specimens were subjected to three designated oven durations of 45, 90, and 135 minutes, each with a tolerance of ±15 minutes. As a note, the same forced convection oven was used to reheat every sample and specimen in this study for consistency. A temperature probe was inserted into the center of each specimen to monitor internal temperatures throughout conditioning. Once the designated time was reached, the specimens were processed for compaction and subsequent testing to investigate the effect of specimen heating time on the IDT-CT and IDT-HT test results.

Figure 14 presents the measured average specimen temperatures after each heating duration for Mix 1 through Mix 3. As expected, longer oven times led to progressively higher internal specimen temperatures across all mixes. For the 45 ± 15-minute condition, average temperatures remained less than 270°F, indicating that the specimens had not yet reached the target compaction temperature but were still workable and able to be compacted. In contrast, the 90 ± 15-minute and 135 ± 15-minute conditions yielded average temperatures near or greater than 285°F, confirming that specimens were at or above the compaction temperature at these longer heating durations.

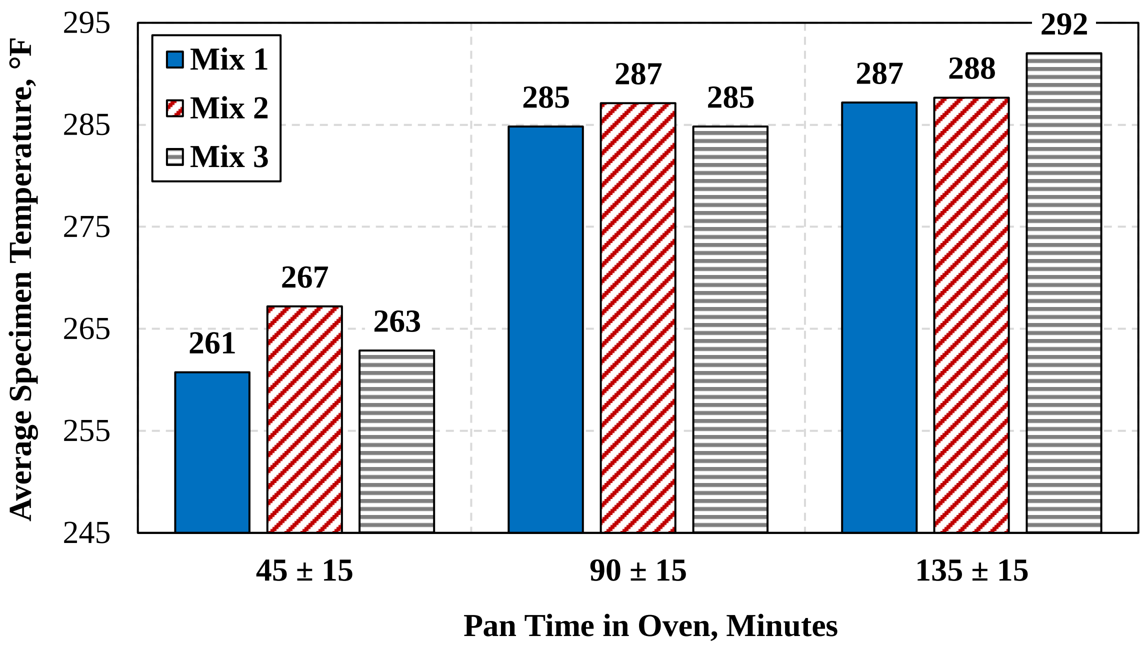

Figure 14. Average Internal Specimen Temperatures at the Completion of Oven Heating

Table 10 presents the average CT index values and Tukey pairwise analysis results at a 95% confidence level for Mix 1 through Mix 3 across three specimen heating durations. For Mix 1, the CT index decreased notably as oven time increased, with values dropping from 91.6 at 45 minutes (group A) to 63.4 at 135 minutes (group B), showing a statistically significant difference between short and long heating durations. For Mix 2, although a slight decrease in CT index from 30.8 to 25.1 occurred across the heating durations, all conditions remained within the same statistical grouping (A), indicating no statistically significant difference. For Mix 3, a similar pattern was observed, with CT index values decreasing slightly from 53.8 to 44.6 as oven time increased. However, all results fell within the same statistical group (A), again indicating no statistically significant differences across heating durations.

Table 10. Average CT Index Values and Tukey Pairwise Analysis at a 95% Confidence Level for Mix 1 through Mix 3 across Specimen Heating Durations

Oven Time (minutes)	Mix 1		Mix 2		Mix 3	
	CT Index	Grouping*	CT Index	Grouping*	CT Index	Grouping*
45 ± 15	91.6	A	30.8	A	53.8	A
90 ± 15	74.0	B	26.8	A	46.0	A
135 ± 0.5	63.4	B	25.1	A	44.6	A

CT = cracking tolerance. *Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

Table 11 presents the average strength values and the results of a Tukey pairwise analysis at a 95% confidence level for Mix 1 through Mix 3 across the three specimen heating durations. For Mix 1, strength increased progressively with longer heating durations, rising from 113.5 kPa at 45 minutes (group C) to 160.5 kPa at 135 minutes (group A), reflecting a statistically significant change in strength with extended heating. For Mix 2, a similar trend was observed, with strength increasing from 248.1 kPa at 45 minutes (group B) to 288.6 kPa at 135 minutes (group A), again showing a statistically significant difference between short and long heating times. For Mix 3, strength also increased with longer heating, moving from 194.7 kPa at 45 minutes (group B) to 209.5 kPa at 135 minutes (group A), indicating statistically significant changes at longer durations.

Table 11. Average Strength Values and Tukey Pairwise Analysis at a 95% Confidence Level for Mix 1 through Mix 3 across Specimen Heating Durations

Oven Time (minutes)	Mix 1		Mix 2		Mix 3	
	Strength, kPa	Grouping*	Strength, kPa	Grouping*	Strength, kPa	Grouping*
45 ± 15	113.5	C	248.1	B	194.7	B
90 ± 15	136.5	B	263.0	B	180.1	C
135 ± 0.5	160.5	A	288.6	A	209.5	A

*Combinations sharing the same letter are statistically similar, indicating no significant difference between them.

Analysis of the data across different specimen heating durations indicated that allowing specimens to remain in the oven for more than 45 minutes after reaching the target compaction temperature may lead to less consistent results. To help ensure consistency, staggering the placement of pans in the oven was evaluated during the ruggedness phase. It was found not to have a statistically significant effect on test outcomes. One approach that may help optimize workflow while minimizing unnecessary heating time involves placing the three pans intended for IDT-HT specimens into the oven approximately 30 minutes after the five pans for IDT-CT specimens. Although not required, this practice was found to support efficient specimen handling and maintain consistent conditioning.

Development of Best Practices and Guidelines

Based on the results and analyses conducted throughout this study, a draft VTM was developed to incorporate the recommended specimen preparation guidelines and procedures. This draft test method is provided in Appendix C. It outlines the standardized practices necessary to ensure consistent, reliable preparation of asphalt mixture specimens for IDT-CT and IDT-HT testing. The proposed method draws directly from the data-driven findings of this study and is intended to support improved repeatability, accuracy, and alignment with VDOT's performance testing objectives.

CONCLUSIONS

- *Stakeholders' input revealed a clear recognition that subtle differences in handling and heating protocols can significantly influence test outcomes, reinforcing the value of collaborative efforts to define and control these factors.*

- *Multiple preparation-related factors can meaningfully influence both CT index and strength outcomes.* Even when statistical significance was not always affected, practical differences were evident. Inconsistent preparation practices can lead to false positive or false negative assessments of mixture performance, undermining confidence in acceptance decisions, especially based on the CT index.

- *Unified, well-defined specimen preparation procedures are critical for improving the reliability, repeatability, and interpretability of both IDT-CT and IDT-HT test results.* By implementing consistent and standardized practices, laboratories can minimize procedural variability and ensure that test results reflect the actual performance characteristics of asphalt mixtures rather than unintended variability introduced by procedural inconsistencies.

- *This study developed a draft VTM, provided in Appendix C, to standardize specimen preparation for IDT-CT and IDT-HT testing.* Key factors such as mixture homogenization, splitting method, specimen thickness, and heating duration were found to influence test outcomes significantly.

RECOMMENDATIONS

1. *VDOT's Materials Division should adapt the VTM developed in this study as part of the standard specimen preparation procedures for VDOT district laboratories and asphalt contractors to use within the BMD process.* This implementation effort should involve coordination with stakeholders to refine the procedures proposed in this study and ensure consistency, clarity, and practicality across all users.

2. *VDOT's Materials Division should consider benchmarking the variability of IDT-CT and IDT-HT test results (i.e., CT index and strength) following the adoption of the VTM developed in this study to assess improvements in consistency and reliability*.

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 process 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

Regarding Recommendation 1, VDOT's Materials Division will collaborate with key stakeholders, including VDOT district laboratories, asphalt contractors, and VTRC, to review and refine the procedures developed in this study. The finalized guidance may be implemented either as a standalone VTM or integrated into the existing VTMs, such as those for IDT-CT (i.e., VTM-143) and the IDT-HT test (i.e., VTM-145), to avoid the need for an additional separate method. This collaborative effort may involve forming a dedicated task force or working group to gather input, address stakeholder feedback, and ensure the resulting procedures are practical, clearly understood, and aligned with both field and laboratory needs. As part of this process, further discussions will help clarify definitions related to reheat and nonreheat mixtures, homogenization, sample sizes, and container types. This work is expected to be completed no later than September 2026.

Regarding Recommendation 2, VTRC, with support from VDOT's Materials Division, will submit a research needs statement to initiate a benchmarking effort to evaluate the variability of IDT-CT and IDT-HT test results. A research needs statement will be drafted and submitted to the VTRC Pavement Research Advisory Subcommittee B by fall 2027.

Benefits

Implementing the recommendations from this study offers several important benefits to VDOT's performance testing program. By adopting the proposed VTM for specimen preparation, VDOT can significantly improve the reliability, repeatability, and consistency of IDT-CT and IDT-HT test results across district laboratories and contractor operations. Improved consistency will help minimize unnecessary testing variability and, critically, reduce the likelihood of false positive or false negative outcomes in mixture acceptance decisions. This consistency ensures that asphalt mixtures are evaluated more accurately against performance criteria, strengthening confidence in decisions made within the BMD process.

As outlined in the implementation plan, the collaborative review and finalization of the VTM with stakeholders will further ensure that the adopted procedures are practical, clearly communicated, and aligned with field realities. Additionally, the planned benchmarking effort will help quantify improvements in test consistency, demonstrating tangible benefits compared with prior practices. Over time, these enhancements are expected to improve VDOT's overall quality assurance framework, support the selection of better-performing asphalt mixtures, and contribute to longer lasting pavements and more efficient use of maintenance resources.

ACKNOWLEDGMENTS

The authors are grateful to the following individuals who served on the technical review panel for this study: Angela Beyke, Project Champion and Assistant State Materials Engineer, VDOT Materials Division; Bryan Smith, Senior Asphalt Pavement Field Engineer, VDOT Materials Division; Candice Entwisle, Assistant State Asphalt Program Manager, VDOT Materials Division; Donald Clyde Landreth, Materials Pavement Engineer, VDOT Salem District; Erin Robartes, Research Scientist, VTRC; Harold (Don) French, District Materials Engineer, VDOT Lynchburg District; Kwame Adu-Gyamfi, District Materials Engineer, VDOT Richmond District; Michael Dudley, Director, Virginia Asphalt Association; Thomas (Ben) Carter, Engineering Technician III, VDOT Materials Division; and Mourad Bouhajja formerly of VDOT Materials Division.

The authors thank Andrew Barbour, Troy Deeds, Derek Lister, and Jennifer Samuels of VTRC and Tina Taylor, Jason Moore, Nathan Cantrell, Cullen Blanton, and Sam Dunlop of NCAT for their assistance with this study.

REFERENCES

American Association of State Highway and Transportation Officials. *Standard Specifications for Transportation Materials and Methods of Sampling and Testing, and AASHTO Provisional Standards*. Washington, DC, 2019.

American Association of State Highway and Transportation Officials. *Standard Specifications for Transportation Materials and Methods of Sampling and Testing, and AASHTO Provisional Standards.* Washington, DC, 2020.

ASTM International. *Standard Practice for Conducting a Ruggedness Evaluation or Screening Program for Test Methods for Construction Materials*. ASTM C1067-20. West Conshohocken, PA, 2020.

ASTM International. *Standard Practice for Conducting Ruggedness Tests*. ASTM E1169-21. West Conshohocken, PA, 2021.

Bennert, T., Hass, E., Wass, E., and Berger, B. Indirect Tensile Testing for Balanced Mixture Design and Quality Control Performance Testing. *Journal of Asphalt Paving Technology*, Vol. 90, September 2020, pp. 363–390.

Boz, I., Diefenderfer, S.D., Habbouche, J., and Seitllari, A. *Interlaboratory Study for the Indirect Tensile at High Temperature Test and Ideal Rutting Test*. VTRC 25-R14. Virginia Transportation Research Council, Charlottesville, VA, 2025.

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*, Vol. 2676, No. 5, May 2022, pp. 225–241.

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

Diefenderfer, S.D., and Bowers, B.F. Initial Approach to Performance (Balanced) Mix Design: The Virginia Experience. *Transportation Research Record*, Vol. 2673, No. 2, February 2019, pp. 335–345.

Diefenderfer, S.D., 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, VA, 2021.

Habbouche, J., Boz, I., and Diefenderfer, S.D. Validation of Performance-Based Specifications for Surface Asphalt Mixtures in Virginia. *Transportation Research Record*, Vol. 2676, No. 5, May 2022a, pp. 277–296.

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

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

Moore, N., and Taylor, A. *Guide on Asphalt Mixture Specimen Fabrication for BMD Performance Testing*. Publication IS-145. National Asphalt Pavement Association, Greenbelt, MD, 2023.

Sias, J.E., Dave, E.V., and Myers, L.A. *Practices for Fabricating Asphalt Specimens for Performance Testing in Laboratories*. NCHRP Synthesis 552. Transportation Research Board, Washington, DC, 2020.

Taylor, A.J., Moore, J.R., and Moore, N. *NCAT Performance Testing Round Robin*. NCAT Report 22-01. National Center for Asphalt Technology, Auburn, AL, 2022.

Virginia Department of Transportation. Determination of Asphalt Content from Asphalt Paving Mixtures by the Ignition Method—(Asphalt Lab). VTM-102. Richmond, VA, 2013.

Virginia Department of Transportation. *Road and Bridge Specifications*. Richmond, VA, 2020.

Virginia Department of Transportation. Determination of Cracking Tolerance Index for Asphalt Mixture Using the Indirect Tensile Cracking Test (IDT-CT) at Intermediate Temperature. VTM-143. Richmond, VA, 2023a.

Virginia Department of Transportation. Method of Test for Determining Rutting Susceptibility of Asphalt Mixtures Using the Indirect Tensile at High Temperature (IDT-HT) Test. VTM-145. Richmond, VA, 2023b.

West, R., Willis, J.R., and Marasteanu, M. *Improved Mix Design, Evaluation, and Materials Management Practices for Hot Mix Asphalt with High Reclaimed Asphalt Pavement Content*. NCHRP Report 752. Transportation Research Board, Washington DC, 2013.

APPENDIX A: SURVEY FOR ASPHALT CONTRACTORS IN VIRGINIA

Ruggedness Study of Specimen Preparation and Fine-Tuning of Test Methods for IDT-CT and IDT-HT

This survey is designed to collect key information related to specimen preparation processes that would affect the test results for the indirect tensile cracking test (IDT-CT) and indirect tensile at high temperature (IDT-HT) test during both the design and production stages, including specimens subject to reheating. This survey is part of an ongoing collaborative research effort with the National Center for Asphalt Technology (NCAT) aimed at conducting a ruggedness study on key factors that necessitate "standardization" and more comprehensive guidelines during the specimen preparation of mixtures for the IDT-CT and IDT-HT test. For additional information, please visit https://vtrc.virginia.gov/projects/all-projects/124496/.

Note: A similar survey is being sent to VDOT district laboratories and research partner laboratories.

Disclaimer: The survey responses will be kept confidential and will be analyzed *anonymously*.

Thank you in advance for your participation.

Definitions:

The following definitions are used to clarify the scope in this questionnaire:

- **Mix Design:** During the mix design stage, specimens are fabricated in the laboratory from loose mixtures created by blending raw materials such as aggregates, asphalt binders, reclaimed asphalt pavement, additives, etc.
- **Production Non-Reheats:** During the production phase, specimens are compacted on site without reheating the loose mixture sampled at the plant.
- **Production Reheats:** Option 1: The plant sampled mixture is let to cool down to room temperature in a sample container (boxes, bags, or large pans) and is then reheated until the mixture reaches a workable temperature, then separated to specimen size, brought to compaction temperature, and compacted. Option 2: When sampled, the material is immediately separated to specimen size, quickly cooled down to room temperature, then reheated and compacted.

Section 1: Contact Information

Q1—Please provide the following:

Name	
Position / Title	
Agency / Company	
Address (City, State, and Zip Code)	
Phone	
Email	

May we contact you for more information?

☐Yes.

☐Please contact this person instead (please provide name, phone number, and/or email address).

☐No.

<u>Section 2</u>: Assessment of Key Factors

<u>Q2</u>—Please provide a rating between 1 and 10 for each of the following factors (please rate them <u>independently!!</u>), with "1" indicating no importance and "10" indicating high importance in terms of achieving consistent results and minimizing variability for the test results. In case the question or factor does not cover a practice that you do, please select "N/A" for "Not Applicable" as your response. Each factor will have three categories for ratings related to specimen fabrication: during the design phase, during the production phase for nonreheats, and during the production phase for reheats (Option 1 or 2). In addition, please elaborate on your ratings by providing comments and insights to support and clarify your answers.

1. **Homogenization of mixtures and sample containers:**
   This factor involves determining whether it is necessary to thoroughly blend materials sampled from various spots of the asphalt mixture pile (either from the mixture dumped on the ground during production or the mixture sampled from the back of a truck) using multiple buckets or containers before or during the splitting process.

2. **Mixture sample splitting method:**
   This factor involves, for example, choosing between the quartering method, a mechanical splitter, or scooping from a container.

3. **Container types where mixture samples are conditioned:**
   This factor involves choosing whether to heat samples in pans or buckets, for example.

4. **Mixture sample thickness in the container:**
   This could involve selecting either a standard thickness or a thickness based on the nominal maximum aggregate size.

5. **Covering of mixture samples:**
   This practice may vary regarding whether to cover samples when heating or to leave the pans uncovered in the oven.

6. **Location of mixture sample when placed in the oven.**

7. **Stirring versus not stirring during the heating and conditioning of the mixture sample.**

8. **Mixture sample heating to a specific duration:**

This factor addresses variations in heating among loose mixture samples within the same batch during specimen fabrication and determines whether mixture sample should be heated for a specific duration, regardless of the temperature.

9. **Mixture sample heating to a specific temperature:**
   This factor addresses variations in heating among loose mixture samples within the same batch during specimen fabrication and determines whether mixture sample should be heated to a specific temperature, irrespective of duration.

10. **Oven type:**
    This factor involves the use of various types of ovens, such as a forced-draft oven and a nonforced-draft oven.

11. **Types of material transfer device used to move mixture sample from the pans to the compaction molds (e.g., funnel, scoop, others).**

12. **The act of transferring the mixture sample from the pans to the compaction molds:**
    This factor involves using a funnel or another device to transfer the mixture sample from the pans to the molds, scooping, or transferring the mixture sample directly from the conditioning pans into the compaction mold (with no transfer device).

13. **Temperature of material transfer device (e.g., funnel, scoop, others).**

14. **Temperature of compaction molds.**

15. **Reheating mixture samples multiple times:**
    This factor involves allowing mixture samples to be reheated multiple times.

16. **Mixture sample heating process:**
    This factor involves placing all mixture samples in the oven simultaneously (at the same time) or in stages.

17. **Time lapsed until compaction (lag time):**
    Lag time is the time between when mixture is sampled and then compacted.

18. **Time lapsed until testing (dwell time):**
    Dwell time is the time between when specimens are compacted and tested.

19. **Equipment used by laboratories to fabricate performance test specimens:**
    This includes the type of equipment used to sample, split, and compact IDT-CT and IDT-HT specimens, such as mix splitter, gyratory compactor, compaction mold, core drying device, conditioning ovens, vacuum-bag sealing device, and others.

20. **Storage time of loose mixture samples:**
    This involves determining the maximum allowable storage time for loose asphalt mixtures.

21. **Cooling mechanism of asphalt mixture specimen post-compaction.**
    This factor may involve various practices, including leaving the specimens in front of a fan, using an accelerated cooling systems such air conditioner, and others.

22. **Number of technicians involved in the fabrication of specimens collected from the same batch of mixtures.**

Q3—Please list any additional key factors related to specimen preparation processes that you believe could contribute to an increase in the variability of both the CT index (from the IDT-CT) and strength (from the IDT-HT test). If your list includes more than 5 factors, please list the top 5 with the highest importance according to your experience.

Acknowledgment
The research team thanks you for your time, effort, and information.
This completes the survey.

APPENDIX B: SUMMARY OF SURVEY RESPONSES

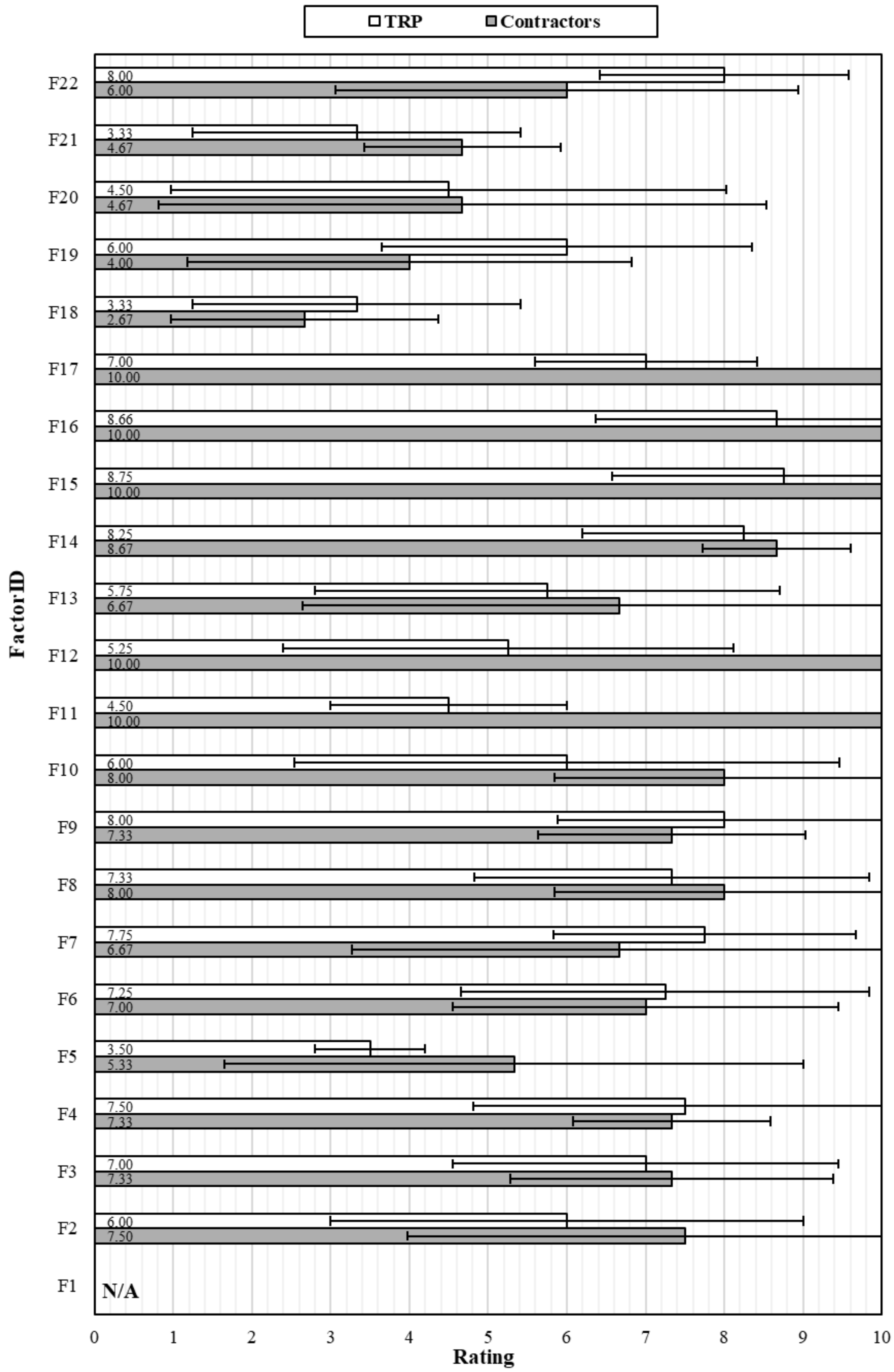

Figure B1. Factors Rated by TRP Members and Contractors during the Mix Design Stage. Error bars indicate plus or minus 1 standard deviation. N/A = not applicable; TRP = technical review panel.

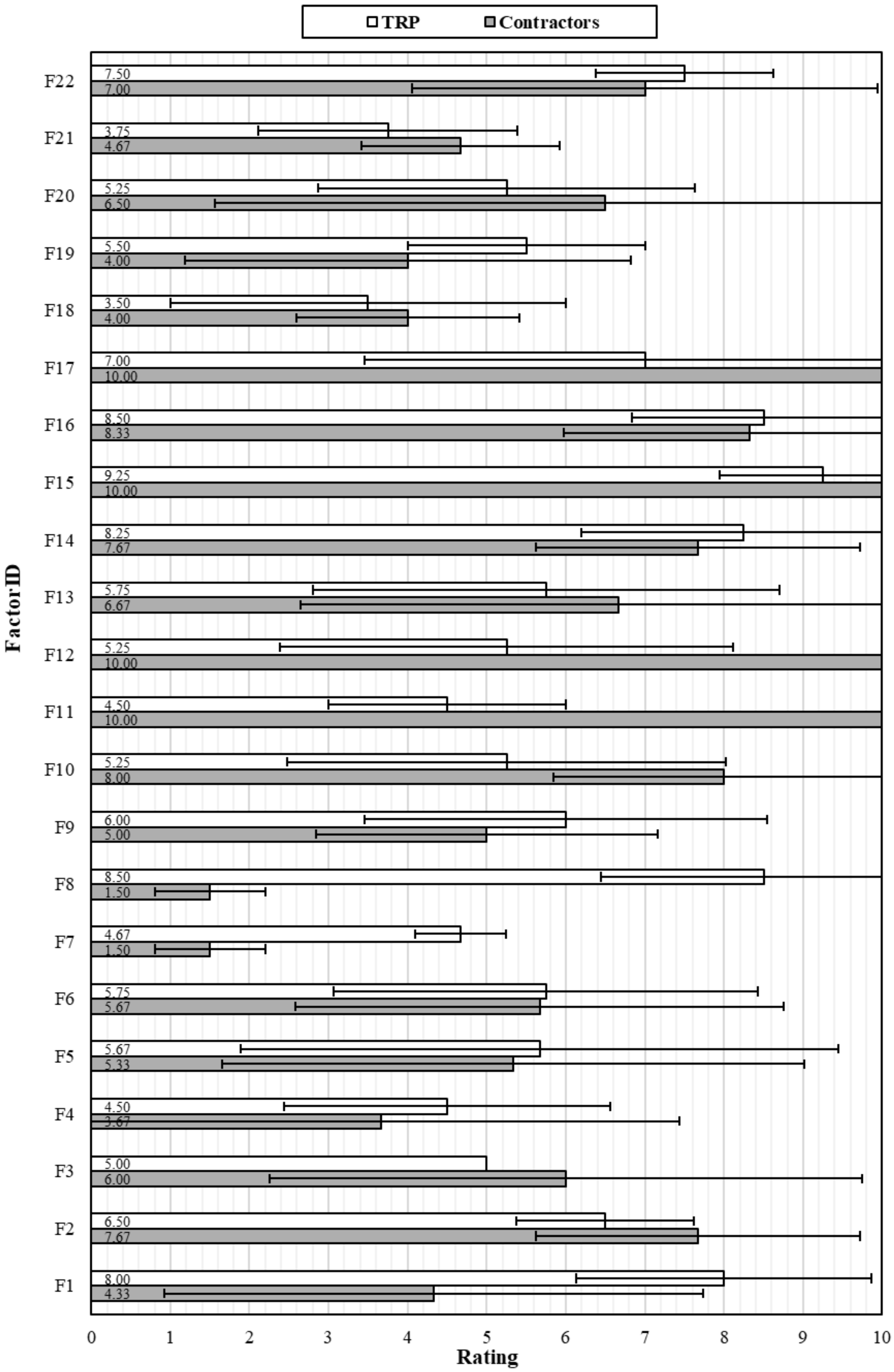

Figure B2. Factors Rated by TRP Members and Contractors during Production for Nonreheated Specimens. Error bars indicate plus or minus 1 standard deviation. TRP = technical review panel.

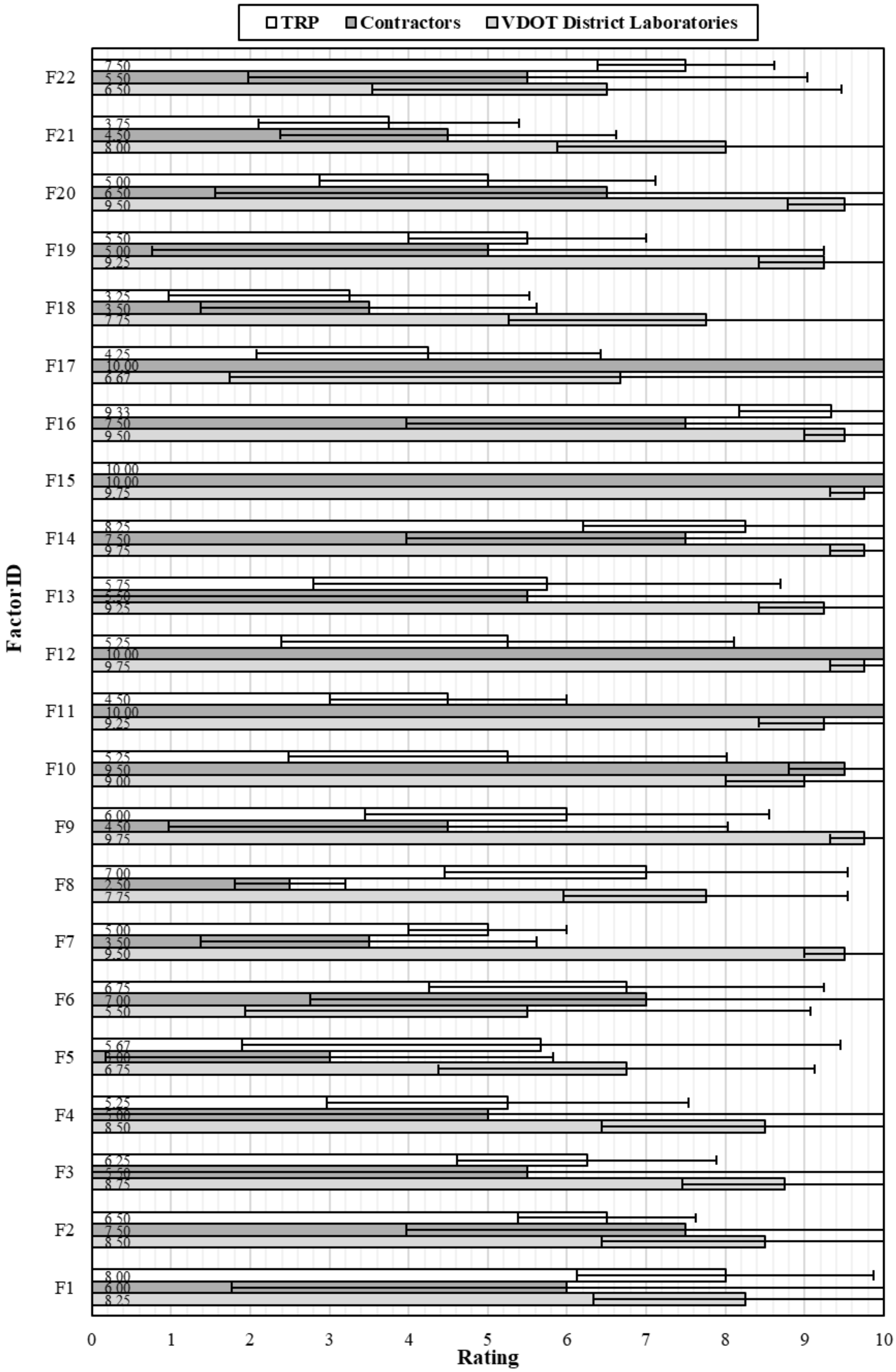

Figure B3. Factors Rated by TRP Members, Contractors, and VDOT District Laboratories during Production for Reheated Specimens. Error bars indicate plus or minus 1 standard deviation. TRP = technical review panel.

Table B1. Summary of Survey Responses by TRP, Contractors, and VDOT District Laboratories

ID	Details	Mix Design				Production for Non-Reheated Specimens				Production for Reheated Specimens					
		TRP		Contractors		TRP		Contractors		TRP		Contractors		VDOT District Laboratories	
		Avg	Stdev	Avg	Stdev	Avg	Stdev	Avg	Stdev	Avg	Stdev	Avg	Stdev	Avg	Stdev
F1	Mix homogenization of mixture sample containers	0.00	0.00	0.00	0.00	8.00	1.87	4.33	3.40	8.00	1.87	6.00	4.24	8.25	1.92
F2	Mixture sample splitting method	6.00	3.00	7.50	3.53	6.50	1.12	7.67	2.05	6.50	1.12	7.50	3.53	8.50	2.06
F3	Container types where mixture samples are conditioned	7.00	2.45	7.33	2.05	5.00	0.00	6.00	3.74	6.25	1.64	5.50	6.36	8.75	1.30
F4	Mixture sample thickness in the container	7.50	2.69	7.33	1.25	4.50	2.06	3.67	3.77	5.25	2.28	5.00	5.65	8.50	2.06
F5	Covering of mixture samples	3.50	0.70	5.33	3.68	5.67	3.78	5.33	3.68	5.67	3.78	3.00	2.82	6.75	2.38
F6	Location of mixture sample when placed in the oven	7.25	2.59	7.00	2.45	5.75	2.68	5.67	3.09	6.75	2.49	7.00	4.24	5.50	3.57
F7	Stirring vs not stirring during the heating/conditioning of the mixture sample	7.75	1.92	6.67	3.40	4.67	0.57	1.50	0.70	5.00	1.00	3.50	2.12	9.50	0.50
F8	Mixture sample heating to a specific duration	7.33	2.51	8.00	2.16	8.50	2.06	1.50	0.70	7.00	2.55	2.50	0.70	7.75	1.79
F9	Mixture sample heating to a specific temperature	8.00	2.12	7.33	1.70	6.00	2.55	5.00	2.16	6.00	2.55	4.50	3.53	9.75	0.43
F10	Oven type	6.00	3.46	8.00	2.16	5.25	2.77	8.00	2.16	5.25	2.77	9.50	0.70	9.00	1.00
F11	Types of material transfer device being used to move mixture sample from the pans to the compaction molds (e.g., funnel, scoop, others)	4.50	1.50	10.00	0.00	4.50	1.50	10.00	0.00	4.50	1.50	10.00	0.00	9.25	0.83
F12	The act of transferring the mixture sample from the pans to the compaction molds	5.25	2.86	10.00	0.00	5.25	2.86	10.00	0.00	5.25	2.86	10.00	0.00	9.75	0.43
F13	Temperature of material transfer device (e.g., funnel, scoop, others)	5.75	2.95	6.67	4.03	5.75	2.95	6.67	4.03	5.75	2.95	5.50	6.36	9.25	0.83
F14	Temperature of compaction molds	8.25	2.05	8.67	0.94	8.25	2.05	7.67	2.05	8.25	2.05	7.50	3.53	9.75	0.43
F15	Reheating mixture specimen pan multiple times	8.75	2.17	10.00	0.00	9.25	1.30	10.00	0.00	10.00	0.00	10.00	0.00	9.75	0.43
F16	Mixture sample heating process	8.66	2.30	10.00	0.00	8.50	1.66	8.33	2.36	9.33	1.15	7.50	3.53	9.50	0.50
F17	Time lapsed until compaction (lag time)	7.00	1.41	10.00	0.00	7.00	3.54	10.00	0.00	4.25	2.17	10.00	0.00	6.67	4.93
F18	Time lapsed until testing (dwell time)	3.33	2.08	2.67	1.70	3.50	2.50	4.00	1.41	3.25	2.28	3.50	2.12	7.75	2.49
F19	Equipment used by laboratories to fabricate performance test specimens	6.00	2.35	4.00	2.82	5.50	1.50	4.00	2.82	5.50	1.50	5.00	4.24	9.25	0.83
F20	Storage time of loose mixture samples	4.50	3.53	4.67	3.86	5.25	2.38	6.50	4.94	5.00	2.12	6.50	4.94	9.50	0.71
F21	Cooling mechanism of asphalt mixture specimen post-compaction	3.33	2.08	4.67	1.25	3.75	1.64	4.67	1.25	3.75	1.64	4.50	2.12	8.00	2.12
F22	Number of technicians being involved in the fabrication of specimens collected from the same batch of mixtures	8.00	1.58	6.00	2.94	7.50	1.12	7.00	2.94	7.50	1.12	5.50	3.53	6.50	2.96

TRP = technical review panel.

APPENDIX C: VIRIGINA TEST METHOD

Virginia Test Method—XXX

Specimen Preparation Procedures for Balanced Mix Design Testing During Production

Month XX, 202Y

1. Scope

1.1. This test method covers the procedures for sampling, reheating, splitting, and compacting asphalt mixture specimens during production for balanced mix design (BMD) testing. This method applies only to BMD practices during production, pertaining to both production reheated and production non-reheated specimens. Procedures related to mix design development are beyond the scope of this test method.

2. Referenced Documents

- AASHTO R 47 Reducing Sample of Asphalt Mixtures to Testing Size.
- AASHTO R 97 Sampling Asphalt Mixtures.
- AASHTO T 312 Preparing and Determining the Density of Asphalt Mixture Specimens by Means of the Superpave Gyratory Compactor.

3. Definitions

3.1. Production Non-Reheat: Specimens compacted in the laboratory without reheating the loose mixture sampled at the plant.

3.2. Production Reheat: Specimens compacted and evaluated in the laboratory by reheating the loose mixture sampled at the plant.

4. Significance and Use

4.1. This test method establishes standardized procedures for the preparation of asphalt mixture specimens during production for BMD testing. It ensures that specimen integrity is maintained throughout the reheating, splitting, and compaction processes, thereby providing consistent and reliable samples for performance evaluation.

5. Hazards

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

6. Procedure

6.1. Sampling Asphalt Mixture

6.1.1. Determine the amount of asphalt mixture required to produce a set of BMD test specimens from a production sample. During sampling, collect at least 20% more asphalt mixture than is necessary for BMD performance testing in accordance with AASHTO R 97.

Note 1—The quantity obtained should be adequate to produce the necessary number of Superpave gyratory compactor specimens, including additional material for trial compaction and for determining the theoretical maximum specific gravity (G_{mm}) and asphalt binder content (AC%).

Note 2—Sample an additional 20% asphalt mixture to ensure sufficient material for (1) producing the necessary testing specimens and (2) preventing segregation during the splitting process, especially for the specimens near the end of the splitting process.

Note 3—Excessive reheating durations in canvas bags have been observed to adversely affect the cracking tolerance index results. Rigid containers, such as cardboard boxes or metal buckets, have been observed to be less susceptible to significant changes in mixture properties during reheating.

Note 4—When sampling into multiple containers, randomly select containers for BMD performance testing and volumetric property testing to ensure unbiased representation of the asphalt mixture.

6.1.2. If the asphalt mixture is to be tested as a production nonreheat sample, proceed directly to splitting, per section 6.3.

6.2. Reheating of Mixture Samples

6.2.1. Set the oven to the upper limit of the compaction temperature range specified in the job mix formula. It is permissible for the sample to be in the oven as it is heating to the set point.

6.2.2. Reheat the full sample in its original sampling container at the specified compaction temperature for a duration not less than 2.5 hours and not exceeding 4 hours.

Note 5—The mixture does not need to reach full compaction temperature before splitting.

6.2.3. Proceed to splitting the mixture, per section 6.3, once sufficient workability is achieved.

Note 6—For samples cooled to room temperature in cardboard or canvas containers, reheating for approximately 2.5 hours is adequate to achieve workability.

6.3. Splitting Mixture Samples into Test Specimens

6.3.1. Homogenize the reheated asphalt mixture sample in accordance with AASHTO R 47. If multiple containers were used for sampling, combine and blend thoroughly before splitting. Use blending methods outlined in AASHTO R 47, Section 7.

Note 7—A "Mechanical Splitter—Type A" from AASHTO R 47 has proven effective for homogenizing multiple field sample containers.

6.3.2. Split the homogenized mixture into individual test specimens using procedures specified in AASHTO R 47.

Note 8—If the sample is too large to split all at once, place approximately one-half of the blended sample in a separate container and keep in the oven while the first portion is being split. Split the remaining portion immediately after the first portion.

6.3.3. Transfer each specimen to a pan, spreading the mixture evenly to a thickness of 1.0–2.0 inches (25–50 mm).

6.3.4. Place specimens in the oven to reheat for compaction within 24 hours of splitting.

6.4. Heating of Mixture Specimens Prior to Compaction

6.4.1. Heat specimen containers uncovered (no lids, foil, or paper) to the compaction temperature specified in the job mix formula.

6.4.2. Compact specimens when they reach the low temperature limit of the job mix formula-specified compaction temperature range.

6.4.3. Compact specimens within 45 minutes of reaching the compaction temperature.

6.4.4. Do not hold specimens in the oven for more than 2.5 hours prior to compaction.

Note 9—For production reheat specimens, stagger the timing of placing the pans in the oven to ensure equivalent heating times for compaction across different test types. For example, place indirect tensile at high temperature pans in the oven 30 minutes after indirect tensile cracking test pans to manage compaction timing.

6.5. Compaction of Mixture Specimens

6.5.1. Compact specimens in accordance with AASHTO T 312.