VTRC
Virginia Transportation Research Council

We Bring Innovation to Transportation

Simple and Practical Tests for Rutting Evaluation of Asphalt Mixtures in the Balanced Mix Design Process

http://www.virginiadot.org/vtrc/main/online_reports/pdf/23-r11.pdf

ILKER BOZ, Ph.D.
Research Scientist, Virginia Transportation Research Council

JHONY HABBOUCHE, Ph.D., P.E.
Senior Research Scientist, Virginia Transportation Research Council

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

GRIFFIN P. COFFEY
Former Graduate Research Assistant, University of Virginia

OSMAN E. OZBULUT, Ph.D.
Associate Professor, University of Virginia

AKSEL SEITLLARI, Ph.D.
Assistant Professor, State University of New York at Canton

Final Report VTRC 23-R11

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

Standard Title Page - Report on Federally Funded Project

<table>
<tr><td colspan="2">1. Report No.:
FHWA/VTRC 23-R11</td><td colspan="2">2. Government Accession No.:</td><td colspan="2">3. Recipient's Catalog No.:</td></tr>
<tr><td colspan="4" rowspan="2">4. Title and Subtitle:
Simple and Practical Tests for Rutting Evaluation of Asphalt Mixtures in the Balanced Mix Design Process</td><td colspan="2">5. Report Date:
January 2023</td></tr>
<tr><td colspan="2">6. Performing Organization Code:</td></tr>
<tr><td colspan="4">7. Author(s):
Ilker Boz, Ph.D., Jhony Habbouche, Ph.D., P.E., Stacey D. Diefenderfer, Ph.D., P.E., Griffin P. Coffey, Osman E. Ozbulut, Ph.D., and Aksel Seitllari, Ph.D.</td><td colspan="2">8. Performing Organization Report No.:
VTRC 23-R11</td></tr>
<tr><td colspan="4" rowspan="2">9. Performing Organization and Address:
Virginia Transportation Research Council
530 Edgemont Road
Charlottesville, VA 22903</td><td colspan="2">10. Work Unit No. (TRAIS):</td></tr>
<tr><td colspan="2">11. Contract or Grant No.:
117336</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agencies' Name and Address:
Virginia Department of Transportation
1401 E. Broad Street
Richmond, VA 23219</td><td colspan="2" rowspan="2">
Federal Highway Administration
400 North 8th Street, Room 750
Richmond, VA 23219-4825</td><td colspan="2">13. Type of Report and Period Covered:
Final</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code:</td></tr>
<tr><td colspan="6">15. Supplementary Notes:
This is an SPR-B report.</td></tr>
</table>

16. Abstract:

The Virginia Department of Transportation (VDOT) is currently using the asphalt pavement analyzer (APA) as a testing tool to screen the rutting potential of asphalt mixtures as part of its balanced mix design (BMD) method. However, the cost and availability of APA equipment in VDOT and contractor laboratories and the speed of testing are main barriers to the routine use of APA in the BMD process, especially during the production of asphalt mixtures. The existence of alternative tests that are simple, practical, and performance indicative can help facilitate the implementation of BMD into practice. The monotonic loading tests (hereinafter "monotonic tests") have been proposed to address the need for simpler tests.

This study was undertaken to assess the feasibility of using monotonic tests as a screening tool to evaluate the rutting potential of dense-graded asphalt surface mixtures as part of the BMD initiative in Virginia. In this effort, three monotonic tests, the indirect tensile at high temperature (IDT-HT) test, rapid rutting (RR) test, and Marshall stability and flow (MS) test, identified from the literature, were evaluated using 17 plant-produced asphalt mixtures with "A" and "D" designations. The results were used to assess the monotonic tests and the APA test relative to each other through several performance metrics and correlations with fundamental rutting tests and mechanistic-empirical–based simulations. The results were also used to develop performance threshold criteria for the considered tests.

The results showed that the IDT-HT and RR tests can be used to screen the rutting potential of asphalt mixtures meeting VDOT mixture volumetric and gradation requirements for the A and D mixtures. The initial performance criteria for the IDT-HT and RR tests were established for these mixtures. Based on the results, the IDT-HT test was found to be the most suitable alternative test to the APA test.

The study recommends using the IDT-HT test as part of the BMD initiative with a corresponding minimum strength of 133 kPa as a performance criterion based on the testing conditions used in this study. Implementation of traffic-based performance criteria for the A and D mixtures should be explored, as the current practice does not differentiate the performance between these mixtures. The results of the IDT-HT, RR, and the APA tests should be compared and correlated with those from fundamental rutting tests and with performance predictions obtained from mechanistic-empirical pavement design simulations using a larger number of mixtures to verify the findings of this study. In addition, the results should be compared to field performance results for full assurance that the implemented tests and associated threshold values are appropriate. Fine-tuning the testing process for the IDT-HT, RR, and APA tests and establishing precision estimates and statements are necessary for proper quality assurance practices. The study further recommends hands-on training and demonstration of the tests being considered by VDOT as part of the BMD implementation.

<table>
<tr><td colspan="3">17 Key Words:
Rutting, asphalt mixtures, balanced mix design, performance testing, indirect tensile rutting at high temperature test, IDEAL RT, rapid rutting test, rutting tolerance index, asphalt pavement analyzer, performance criteria.</td><td colspan="3">18. Distribution Statement:
No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.</td></tr>
<tr><td colspan="2">19. Security Classif. (of this report):
Unclassified</td><td colspan="2">20. Security Classif. (of this page):
Unclassified</td><td>21. No. of Pages:
56</td><td>22. Price:</td></tr>
</table>

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

FINAL REPORT

SIMPLE AND PRACTICAL TESTS FOR RUTTING EVALUATION OF ASPHALT MIXTURES IN THE BALANCED MIX DESIGN PROCESS

Ilker Boz, Ph.D.
Research Scientist
Virginia Transportation Research Council

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

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

Griffin P. Coffey
Former Graduate Research Assistant
University of Virginia

Osman E. Ozbulut, Ph.D.
Associate Professor
University of Virginia

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

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

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

Charlottesville, Virginia

January 2023
VTRC 23-R11

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 (VDOT) is currently using the asphalt pavement analyzer (APA) as a testing tool to screen the rutting potential of asphalt mixtures as part of its balanced mix design (BMD) method. However, the cost and availability of APA equipment in VDOT and contractor laboratories and the speed of testing are main barriers to the routine use of APA in the BMD process, especially during the production of asphalt mixtures. The existence of alternative tests that are simple, practical, and performance indicative can help facilitate the implementation of BMD into practice. The monotonic loading tests (hereinafter "monotonic tests") have been proposed to address the need for simpler tests.

This study was undertaken to assess the feasibility of using monotonic tests as a screening tool to evaluate the rutting potential of dense-graded asphalt surface mixtures as part of the BMD initiative in Virginia. In this effort, three monotonic tests, the indirect tensile at high temperature (IDT-HT) test, rapid rutting (RR) test, and Marshall stability and flow (MS) test, identified from the literature, were evaluated using 17 plant-produced asphalt mixtures with "A" and "D" designations. The results were used to assess the monotonic tests and the APA test relative to each other through several performance metrics and correlations with fundamental rutting tests and mechanistic-empirical–based simulations. The results were also used to develop performance threshold criteria for the considered tests.

The results showed that the IDT-HT and RR tests can be used to screen the rutting potential of asphalt mixtures meeting VDOT mixture volumetric and gradation requirements for the A and D mixtures. The initial performance criteria for the IDT-HT and RR tests were established for these mixtures. Based on the results, the IDT-HT test was found to be the most suitable alternative test to the APA test.

The study recommends using the IDT-HT test as part of the BMD initiative with a corresponding minimum strength of 133 kPa as a performance criterion based on the testing conditions used in this study. Implementation of traffic-based performance criteria for the A and D mixtures should be explored, as the current practice does not differentiate the performance between these mixtures. The results of the IDT-HT, RR, and the APA tests should be compared and correlated with those from fundamental rutting tests and with performance predictions obtained from mechanistic-empirical pavement design simulations using a larger number of mixtures to verify the findings of this study. In addition, the results should be compared to field performance results for full assurance that the implemented tests and associated threshold values are appropriate. Fine-tuning the testing process for the IDT-HT, RR, and APA tests and establishing precision estimates and statements are necessary for proper quality assurance practices. The study further recommends hands-on training and demonstration of the tests being considered by VDOT as part of the BMD implementation.

FINAL REPORT

SIMPLE AND PRACTICAL TESTS FOR RUTTING EVALUATION OF ASPHALT MIXTURES IN THE BALANCED MIX DESIGN PROCESS

Ilker Boz, Ph.D.
Research Scientist
Virginia Transportation Research Council

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

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

Griffin P. Coffey
Former Graduate Research Assistant
University of Virginia

Osman E. Ozbulut, Ph.D.
Associate Professor
University of Virginia

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

INTRODUCTION

The Virginia Department of Transportation (VDOT) is currently in the process of implementing the balanced mix design (BMD) concept to increase asphalt mixture durability and to allow responsible incorporation of innovative materials in asphalt mix designs (Diefenderfer et al., 2021a). As part of this effort, the asphalt pavement analyzer (APA) is being used as a testing tool to screen asphalt mixtures that may be susceptible to permanent deformation, i.e., rutting. However, there are several impediments to contractors and agencies adopting the APA test. The APA device is costly to purchase and maintain and is not readily available for use by many VDOT district and contractor laboratories. Moreover, the time needed to test asphalt mixtures using the APA is lengthy; at least 8.5 hours are usually required to evaluate an asphalt mixture from conditioning the specimen to conducting the test. This makes the use of the APA test challenging during the mix design stage and impractical for the purposes of quality assurance during plant production. Thus, VDOT is interested in evaluating alternative tests that can overcome these impediments.

The researchers hypothesized that rutting could be evaluated through the use of monotonic loading tests (hereinafter "monotonic tests") at appropriate high temperatures. These tests have several advantages for implementation such as the availability and cost of the equipment required to run the test; simplicity of the test; ease and time of specimen preparation; speed of testing; test repeatability and reproducibility; and, most important, correlation with in-service performance. The rutting evaluation of asphalt mixtures with these tests can be performed with equipment that either is available in many asphalt laboratories or can be readily purchased. For example, many laboratories have loading frames that are already used for other tests such as indirect tensile–based cracking tests, semi-circular bending–based cracking tests, and moisture susceptibility tests and that can be readily used for evaluating the rutting potential of asphalt mixtures. In addition, the time required to condition test specimens and perform a rutting test would be significantly shorter for the monotonic tests as compared to the APA test.

PURPOSE AND SCOPE

The purpose of this study was to assess the feasibility of using monotonic tests as screening tools to evaluate the rutting potential of dense-graded asphalt surface mixtures in Virginia. A total of 17 plant-produced mixtures with "A" and "D" designations were evaluated with the APA test and three selected monotonic tests. A and D mixtures are designated for traffic loads of 0 to 3 million equivalent single axle loads and 3 to 10 million equivalent single axle loads, respectively. Several analyses were conducted to identify monotonic tests for BMD use and to develop corresponding threshold criteria for the tests. In addition, the study included the use of a digital image correlation (DIC) method for forensic investigation of the monotonic tests.

METHODS

Eight tasks were performed to achieve the study objectives:

1. A literature review was undertaken to summarize the state of the art and practice regarding the use of monotonic tests to evaluate the rutting potential of asphalt mixtures. This task was also aimed at identifying the appropriate test methods for further evaluation.

2. Plant-produced mixtures were sampled, and mixture volumetric properties, extracted binder properties, and aggregate gradations were determined.

3. A series of tests were conducted on a subset of the sampled mixtures to determine an appropriate rutting test temperature for the selected monotonic tests.

4. Laboratory testing was conducted on the mixtures using the selected monotonic tests alongside the APA test. Comparative analyses among the tests were performed to select the most suitable monotonic tests for further evaluation. The results were analyzed in terms of repeatability, performance discrimination potential, ranking

among the asphalt mixtures, sensitivity of the tests to changes in volumetric properties, relationship to the binder non-recoverable creep compliance, and correlation among tests.

5. A forensic analysis of the selected monotonic tests was performed with the DIC method. Specimen full-field deformation and strain characteristics captured during the tests were examined to understand the mechanism by which asphalt mixtures fail under each monotonic test and to determine which test configuration might be the best indicator of the rutting susceptibility of asphalt mixtures. The details of this task are presented in Appendix A.

6. A suite of fundamental rutting tests, i.e., the dynamic modulus, confined flow number (FN), and stress sweep rutting (SSR) tests, was performed on six selected mixtures in addition to the monotonic tests and APA test. Comparative analyses were performed to assess further the ability of the monotonic tests and APA test to capture the rutting susceptibility of asphalt mixtures as compared to the fundamental rutting tests.

7. Mechanistic-empirical (ME)–based simulations using AASHTOWare Pavement ME Design software (hereinafter "Pavement ME Design") were carried out to predict the long-term performance of the selected mixtures placed on typical pavement structures. Both material level and project level analyses were conducted. The results obtained through ME analyses were compared to the results of the selected monotonic tests and APA test.

8. Four different approaches were undertaken to develop traffic-based performance criteria for the selected monotonic tests for A and D mixtures. An exercise was also carried out to verify the existing performance criterion for the APA test and to develop traffic-based performance criteria.

Literature Review

The state of the art and practice information related to the objective of this study was summarized through a comprehensive literature review performed by searching various transportation engineering–related databases and search engines such as the Transport Research International Documentation bibliographic database, the Catalog of Worldwide Libraries, and Google Scholar.

Laboratory Testing and Evaluation

Volumetric Properties and Aggregate Gradations

Seventeen plant-produced dense-graded asphalt surface mixtures, designated A through Q and having a diverse range of mixture components, were sampled from various plants in Virginia and further evaluated in the laboratory. The volumetric properties and gradations of the mixtures were determined. The data collected included asphalt content and gradation; bulk and

Rice mixture specific gravities (G_{mb} and G_{mm}); air voids (voids in total mix [VTM]), voids in mineral aggregate [VMA], and voids filled with asphalt [VFA]); bulk and effective aggregate specific gravities (G_{sb} and G_{se}); fines/asphalt ratio; percent binder absorbed (P_{ba}); and effective binder content (P_{be}).

Asphalt Binder Testing

Extraction of asphalt binder from the mixtures was conducted in accordance with AASHTO T 164, Quantitative Extraction of Asphalt Binder From Hot Mix Asphalt (HMA), Method A, using *n*-propyl bromide as the solvent. The asphalt binder was then recovered from the solvent using the Rotavap recovery procedure specified in AASHTO T 319, Quantitative Extraction and Recovery of Asphalt Binder From Asphalt Mixtures. The multiple stress creep recovery test was performed on the extracted and recovered asphalt binders in accordance with AASHTO M 332, Standard Specification for Performance-Graded Asphalt Binder Using Multiple Stress Creep Recovery (MSCR) Test, to determine the non-recoverable creep compliance, J_{nr}.

Monotonic Tests

Indirect Tensile at High Temperature (IDT-HT) Test

The indirect tensile at high temperature (IDT-HT) test is conducted by applying a constant rate of axial displacement on the diametrical plane of a test specimen. The test can be conducted at a wide range of loading rates and temperatures on cylindrical specimens of various heights. In this study, a loading rate of 50 mm/min was selected for several reasons including practicality and consistency. It was practical because the selection of this loading rate will result in a very short testing duration, resulting in no need to use a testing chamber during the testing itself. In addition, this loading rate provides consistency as it is the same as that for other tests such as the cracking and moisture damage tests, which will prevent confusion and a frequent change of machine setups. This loading rate also allows the Marshall press testing equipment to be used to perform the rutting test.

The test temperature was determined to be 54.4°C. The rationale for selecting this temperature is provided later. Three replicate 150-mm-diameter by 62-mm-tall specimens were fabricated at 7% ± 0.5% target air voids to match the specimen dimensions and target air void content for the indirect tensile cracking test at intermediate temperature currently being used as part of the BMD effort in Virginia. The specimens were conditioned for 2 hours at the test temperature before being tested, as recommended by Bennert et al. (2018). Once the IDT-HT test was complete, the indirect tensile strength test was then determined using Equation 1. A higher strength value indicates a better resistance to rutting.

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

where

S_t= indirect tensile strength, kPa

P_{max} = maximum load, N
t = specimen thickness, mm
D = specimen diameter, mm.

Rapid Rutting (RR) Test

Researchers at the Texas Transportation Institute recently proposed a monotonic test, called the rapid rutting (RR) test and also known as the IDEAL-RT (Zhou et al., 2019). The RR test was conducted in a manner similar to that of the IDT-HT test except that a shear fixture was used in lieu of an IDT-HT fixture. Testing was conducted on three replicate specimens having the same size and target air void content as the IDT-HT test. The specimens were conditioned for 2 hours at the test temperature of 54.4°C prior to testing. A test specimen was placed on a U-shaped shear fixture, and a loading rate of 50 mm/min was applied along the diametrical plane of the specimen.

The rutting susceptibility of asphalt mixtures from the RR test is quantified through the rutting tolerance (RT) index, as shown in Equation 2. A higher RT index indicates a greater resistance to rutting.

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

where

P_{max} = maximum load, N
t = specimen thickness, m
w = width of upper loading strip, 0.0191 m.

Marshall Stability and Flow (MS) Test

The Marshall stability and flow (MS) test is used as part of the Marshall mix design procedure and evaluation and as a quality assurance tool for monitoring the production process for asphalt mixtures. The test measures the resistance to plastic deformation of asphalt mixtures. During the MS test, cylindrical specimens are placed between two cylindrical segments or test heads having a certain inside radius of curvature dependent on the diameter of the test specimen and loaded with a uniform vertical movement of 50.8 mm/min in a manner similar to that of the IDT-HT and RR tests. The test is typically performed at 60°C for hot mix asphalt.

The MS tests were conducted at a loading rate of 50 mm/min in accordance with ASTM D5581, Standard Test Method for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus (6 in. Diameter Specimen). Three replicate specimens were conditioned in an oven at 54.4°C for 2 hours prior to testing. The specimen dimensions and target air void content were the same as for the IDT-HT and RR tests. The peak load is used as the rutting performance criterion for this test.

Conventional Tests

Asphalt Pavement Analyzer (APA) Test

The APA test is currently used as part of VDOT's BMD specification and is conducted in accordance with AASHTO T 340, Standard Method of Test for Determining Rutting Susceptibility of Hot Mix Asphalt (HMA) Using the Asphalt Pavement Analyzer (APA). The APA test is performed on 150-mm-diameter and 75-mm-tall specimens compacted to an air void level of 7% ± 0.5%. VDOT's provisional BMD specification allows a maximum APA rut depth of 8 mm at 64°C and 8,000 loading cycles for A and D mixtures.

Fundamental Rutting Tests

Dynamic Modulus Test

The dynamic modulus tests were performed using the Asphalt Mixture Performance Tester in accordance with AASHTO T 378, Standard Method of Test for Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Asphalt Mixture Performance Tester (AMPT). Three replicate 100-mm-diameter by 150-mm-tall specimens having air void contents of 7.0 ± 0.5% were tested. Three test temperatures (4, 21, and 40°C) and frequencies (0.1, 1, and 10 Hz) were used; an additional testing frequency of 0.01 Hz was used at 40°C. The dynamic modulus tests were performed in the uniaxial mode without confinement, with testing performed from the lowest to the highest temperature. At each temperature, the tests were conducted from the highest to the lowest frequency. An applied strain range of 75 to 125 microstrains was maintained throughout the test.

Stress Sweep Rutting (SSR) Test

The SSR test was also used to evaluate the rutting potential of asphalt mixtures. The SSR tests were performed using the Asphalt Mixture Performance Tester in accordance with AASHTO TP 134, Standard Method of Test for Stress Sweep Rutting (SSR) Test Using the Asphalt Mixture Performance Tester (AMPT). The SSR tests were conducted at two test temperatures (26 and 55°C) under a confining pressure of 69 kPa with three 200-cycle loading blocks at three deviatoric stress levels. The tests were performed on three replicate 100-mm-diameter by 150-mm-tall specimens fabricated at an air void level of 7 ± 0.5%. As outlined in AASHTO TP 134, the rutting strain index (RSI) was computed from the SSR test to differentiate the rutting performance of the mixtures.

Repeated Load Permanent Deformation (RLPD) Test

The repeated load permanent deformation (RLPD) test is used as an indicator of the rutting performance for asphalt mixtures and has been adopted for the calibration of the Pavement ME Design rutting model. The RLPD tests were conducted in general accordance with the protocol in AASHTO T 378, Standard Method of Test for Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Asphalt Mixture Performance Tester (AMPT). The protocol does not include any practice on testing conditions (confined or unconfined) or the deviator stress level. In this study, a deviator stress of 482.6 kPa and a confinement pressure of 68.9 kPa were used as recommended in NCHRP Report 719 (Von Quintus et al., 2012). The tests were performed at three different temperatures (30, 40, and 50°C) using a minimum of two replicate specimens at each temperature. The RLPD tests were performed on 100-mm-diameter by 150-mm-tall specimens fabricated with air void contents of 7 ± 0.5%. The FN index as defined by Zhang et al. (2013) was used to differentiate the rutting performance of the mixtures.

Test Temperature Selection for Monotonic Tests

Tests used to evaluate the rutting potential of asphalt mixtures are performed at a relatively high temperature that is determined as a function of local climate conditions and pavement layer depth. In this study, four temperatures were identified for evaluation in determining the critical high testing temperature for the monotonic tests considered. The first temperature identified was the high temperature performance grade (64°C) for asphalt binders typically used for A and D mixtures, which is also the test temperature for the APA test. A test temperature of 54.4°C was identified as it represents the 7-day average maximum high temperature at 50% reliability in Virginia through the pavement depth computed using LTPPBind (Apeagyei et al., 2011). A test temperature of 49°C was identified based on the approach proposed by Christensen and Bonaquist (2007) for the IDT-HT test, which was determined as 10°C below the critical pavement temperature at 20 mm below the pavement surface at 50% reliability. A temperature of 40°C was identified in accordance with the recommended procedure in NCHRP Report 704 (Moulthrop and Witczak, 2011). The lowest (40°C) and highest (64°C) temperatures were discounted from consideration. On the one hand, based on engineering judgment, a temperature of 40°C is expected to be too low to discriminate the mixture rutting susceptibility. On the other hand, a temperature of 64°C is expected to make test specimens vulnerable to creep and susceptible to damage during handling. Thus, a series of tests were performed on five mixtures at 49°C and 54.4°C to investigate any potential differences between the two test temperatures.

Mechanistic-Empirical Rutting Simulations

Pavement ME Design was introduced as an improved pavement design methodology. The design package is an ME method that relies on fundamental material properties and calculates the accumulated damages, such as rutting, as a function of climate and traffic levels. The mechanistic part consists of calculations of fundamental pavement responses (stresses, strains, and deflection). It uses transfer functions to relate these to potential field performance, which constitutes the empirical part of the package. Pavement ME Design uses hierarchical design levels: Level 1 assembles a direct measure of material and traffic parameters; Level 2 uses regression equations to estimate the corresponding Level 1 parameters; and Level 3 contains general “global” values (Mallela et al., 2009).

VDOT has adopted Pavement ME Design for routine pavement design for new construction for interstate and primary routes. As part of its implementation strategy, VDOT performed local calibration of the design software, primarily focusing on fatigue cracking and rutting distress modes. The local calibration factors were developed from a pool of asphalt mixtures representing the commonly designed and produced mixtures. Concerns had arisen, however, that the developed calibration factors could be insensitive when mixture types that were not part of the calibration pool were used. Given these concerns and to improve the evaluation of the tests used in this study, Pavement ME Design analyses for the selected mixtures were performed using (1) Level 1 inputs (dynamic modulus |E*|, mixture volumetric properties [i.e., asphalt content, air voids, and unit weight]) and VDOT's recommended local calibration coefficients, and (2) material-specific calibration coefficients. For the latter, RLPD data were used to calibrate the asphalt concrete layer rutting coefficient β_i shown in Equation 3.

$$\varepsilon_p = \varepsilon_r k_z \beta_{r1} 10^{k_1} T^{k_2 \beta_{r2}} N^{k_3 \beta_{r3}} \quad \text{[Eq. 3]}$$

where

ε_p = measured permanent strain
ε_r= resilient strain
T = temperature (°F)
N = number of load repetitions
k_z= depth correction function; during the calibration process, k_z was set equal to 1 as the laboratory specimens were subjected to uniform confinement pressure
$\beta_{r1}, \beta_{r2}, \beta_{r3}$= local calibration coefficients
k_1, k_2, k_3= global field calibration coefficients (k_1 = −3.35412, k_2 = 1.5606, k_3 = 0.4791), Pavement ME Design, Version 2.2.6.

A typical Virginia pavement structure was considered for the analysis. The thickness of the surface material was varied (30 mm to 64 mm) to take into account the minimum and maximum nominal maximum aggregate size (NMAS) of 9.5 mm and 12.5 mm, respectively, for surface mixtures. Analyses for varying surface thicknesses were performed for the six selected mixtures. The input values for the rest of the structure were based on recommended VDOT values for pavement design (Diefenderfer, 2011). Similarly, vehicle class distribution sets were obtained from the Virginia Statewide Traffic database. Three average annual daily truck traffic (AADTT) values were used: 299, 699, and 999. The selected values were based on VDOT traffic volume estimates and were designated for the A and D (i.e., A = ≤299 AADTT, D = 300 to 999 AADTT) asphalt mixtures. One climatic station (Charlottesville) was selected for the analysis.

Performance Criteria Development

One of the tasks involved within the BMD framework is to adopt practical test methods. Establishing performance criteria is vital to optimize the performance of asphalt mixtures within a BMD framework. Four approaches were undertaken in this study to establish performance criteria for the tests considered.

1. *Approach I.* This approach assumed that the asphalt mixtures tested would have adequate rutting performance in the field because they were designed and produced

meeting VDOT's volumetric specifications. The results of the tests were evaluated based on the minimum (in the case of the monotonic tests) or maximum (in the case of the APA test) value to establish the performance criterion.

2. *Approach II.* This approach used the correlation between the J_{nr} parameter and the results of the considered tests to establish the performance criterion.

3. *Approach III.* This approach applied the relationship between the monotonic test results and the APA test results to establish the performance criterion.

4. *Approach IV.* This approach compared the monotonic test results to the results of the Pavement ME Design analysis to establish the performance criterion.

RESULTS AND DISCUSSION

Literature Review

The literature review was conducted to document the state of the art and practice regarding the monotonic tests that are used to evaluate the rutting potential of asphalt mixtures. The review also aimed to identify the most suitable tests for further evaluation. The review revealed that there are a number of monotonic tests available for the purpose, and the versatility of such tests in screening the rutting potential of asphalt mixtures is evidenced through numerous laboratory and field studies. The existence of such simple, practical, performance-indicating tests is especially important as these tests can be used in a timely manner during the production and acceptance processes for asphalt mixtures, in addition to their use during mix design.

Based on the findings in the literature, three monotonic tests were selected for evaluation: the IDT-HT test, the RR test, and the MS test. These tests were selected because they addressed logistical and practical needs and the findings regarding these tests in the literature were encouraging. Many laboratories already have loading frames that are used to perform the moisture damage test (known as the tensile strength ratio test) in accordance with AASHTO T 283, Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage, or other similar tests. Those loading frames can also be used for assessing the rutting potential of asphalt mixtures with the selected monotonic tests. Thus, the cost and issues of equipment availability for these tests would be minimized. The MS test was selected because the results of this test are used to measure the resistance to plastic deformation of asphalt mixtures. The test is performed in accordance with AASHTO T 245, Standard Method of Test for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus; ASTM D6927, Standard Test Method for Marshall Stability and Flow of Asphalt Mixtures; or ASTM D5581, Standard Test Method for Resistance to Plastic Flow of Bituminous Mixtures Using Marshall Apparatus (6 in. Diameter Specimen), depending on the diameter of the specimen used for testing.

The study also included a review of the practices of other states with regard to the APA test along with threshold failure criteria for screening out rut-susceptible asphalt mixtures. A brief summary of the review is provided here for each selected monotonic test and the APA test. Further information on the monotonic tests and how they relate to the rutting of asphalt mixtures can be found elsewhere (Boz et al., 2022).

Indirect Tensile at High Temperature (IDT-HT) Test

The monotonic indirect tensile test has been used as a simple and practical test to characterize several performance aspects of asphalt mixtures for several decades (Boz et al., 2021a), in particular the rutting susceptibility, since the early 2000s (Boz et al., 2022). For instance, the indirect tensile test, as performed in accordance with AASHTO T 322, Standard Method of Test for Determining the Creep Compliance and Strength of Hot Mix Asphalt (HMA) Using the Indirect Tensile Test Device, can be used along with several other tests to assess low-temperature properties of asphalt mixtures. It can also be used to determine long-term moisture damage susceptibility in accordance with AASHTO T 283, Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage, and to estimate the fatigue cracking potential at intermediate temperatures in accordance with ASTM D8225, Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature.

The review indicated that the IDT-HT test performed at a relatively high temperature can be used as a screening tool for evaluating the rutting potential of asphalt mixtures during mix design and for quality control purposes. The IDT-HT test has been reported to be a simple, quick, repeatable, and reproducible test with correlations to other well-established simulative and fundamental rutting tests and field performance. The IDT-HT test showed a good repeatability with a single-operator coefficient of variation (COV) of 8.2% and a multiple-laboratory COV of 11.8%, specific to the testing conditions reported in the respective study (Bennert et al., 2021). The IDT-HT test was correlated with wheel tracker tests, such as the APA test and the Hamburg wheel test (Bennert et al., 2021; Yin et al., 2020), and to fundamental rutting tests such as the repeated shear at constant height test (Christensen et al., 2000), the RLPD test (Wen and Bhusal, 2013), and the SSR test (Meroni et al., 2021). The results of the IDT-HT test were also strongly correlated with rut depths resulting from full-scale accelerated loading wheel passes (Christensen et al., 2004). Finally, the IDT-HT test also indicated a good agreement with in-service performance (Christensen and Bonaquist, 2007). The promising results led to the development of performance criteria from the IDT-HT results to assess rutting resistance. Table 1 summarizes the IDT-HT performance criteria from different studies as a function of the design traffic level and temperature at a loading rate of 50 mm/min. Several agencies have already implemented or are in the process of implementing the IDT-HT test for the rutting evaluation of asphalt mixtures, such as the Alabama Department of Transportation (DOT) and the New Jersey DOT.

Rapid Rutting (RR) Test

The RR test has been reported to be a simple, quick, repeatable, and reproducible test having correlations with other well-established simulative and fundamental rutting tests and field performance (Zhou et al., 2019; Zhou et al., 2020; Zhou et al., 2021). The RR test was shown to be sensitive to key components of asphalt mixtures such as binder content, binder type, air voids, aggregate shape, reclaimed asphalt pavement, recycled asphalt shingles, and aging conditions.

The repeatability of the test, as quantified by the COV, was 6.7%, lower than that of other rutting tests including simulative tests (e.g., the APA test) and fundamental rutting tests (e.g., the FN test). In addition, the RR test ranked the performance of the mixtures in the same manner as other tests such as the APA test, the Hamburg wheel test, and the FN test. Further, the RR test showed good correlations with the performance of Texas FM468 test sections, MnROAD 2008 test sections, and WesTrack test sections. Recently, the following criteria were proposed for the RR test performed at 50°C using a loading rate of 50 mm/min on mixtures conditioned at a compaction temperature for 2 hours (Zhou, 2021):

- *For mixtures with PG64-XX (or lower) with 95% confidence:* RT index ≥ 60.
- *For mixtures with PG70-XX with 95% confidence:* RT index ≥ 65.
- *For mixtures with PG76-XX (or higher) with 98% confidence:* RT index ≥ 75.

Table 1. Examples of Rutting Resistance Limits for the IDT-HT Test

Reference	Traffic/Mix Type	Temperature, °C	Strength, kPa	Notes
Christensen and Bonaquist, 2007	3M ESALs	49	107.9	4% V_a, 115-mm-tall specimens
Christensen and Bonaquist, 2007	10M ESALs	49	173.1	4% V_a, 115 mm tall specimens
Bennert et al., 2021	High-RAP surface course (PG64E-22)	44	324.1	6.5% V_a, 95-mm-tall specimens
Bennert et al., 2021	High-RAP surface course (PG64S-22)	44	158.6	6.5% V_a, 95-mm-tall specimens
Yin and West, 2021	10M ESALs	50	137.9	7% V_a, 62-mm-tall specimens

IDT-HT = indirect tensile at high temperature; ESAL = equivalent single axle load; V_a = air void content; RAP = reclaimed asphalt pavement; PG = performance grade.

Marshall Stability and Flow (MS) Test

There is ample literature with regard to the MS test and its sensitivity to changes in asphalt mixture properties and its ability to indicate the resistance of asphalt mixtures to rutting. The test is reported to be sensitive to changes in aggregate type, gradation, shape, maximum size, and asphalt content (Roberts et al., 2002; Rushing, 2009). Although it was used by many DOTs and other agencies in the United States in the 1980s (Kandhal and Koehler, 1985; Roberts et al., 2002), and is still used by some U.S. agencies and other agencies around the world, the test's ability to capture the rutting performance of asphalt mixtures is debatable. Kandhal and Koehler (1985) reported the state of the practice of the MS test in the United States. They found that the most common ranges for the stability were 7,117 N to 8,000 N and for the flow measures were 8 to 16 and 8 to 18 flow units to minimize the rutting in asphalt pavements by heavy truck traffic. On the other hand, it was also reported that the MS test is a poor indicator of an asphalt mixture's resistance to permanent deformation in the field (Brown et al., 2001; Roberts et al., 2002).

Asphalt Pavement Analyzer (APA) Test

There is also ample literature with regard to the APA test and its sensitivity to changes in asphalt mixture properties and its ability to indicate the resistance of asphalt mixtures to rutting. As a result of such studies, many state agencies in the United States have adopted the APA test to evaluate the rutting susceptibility of their asphalt mixtures. A review of current APA threshold failure criteria used by U.S. state DOTs as of May 1, 2019, was compiled in AASHTO MP 46-2020, Standard Practice for Balanced Mix Design. The findings of this review are presented in Table 2. Table 2 shows that the APA test rutting criteria differ among and within agencies as a function of traffic type or mixture type. Further, the test temperature and specimen air void contents varied among the agencies. The majority of the agencies used a hose pressure of 689 kPa (100 psi) and a wheel load of 445N (100 lbf) for the APA test configuration.

Mixture Volumetric Properties and Gradations

A total of 17 plant-produced mixtures with A and D designations were evaluated in this study. The volumetric properties and aggregate gradations were determined at the Virginia Transportation Research Council (VTRC) laboratory and are summarized in Table 3. For some mixtures, specific volumetric property and gradation values were not within the allowable ranges in the VDOT specifications (VDOT, 2020). The majority of these mixtures were borderline with respect to the allowable specification ranges with the exception of Mixture G. However, it must be noted that the results were obtained from testing two replicate samples of mixtures collected from a randomly selected production lot at the asphalt plants. Mixture G was produced for evaluation at VDOT's accelerated pavement testing facility and was purposely designed not to meet VDOT specifications.

Asphalt Binder Properties

Table 3 also presents the J_{nr} parameter of the extracted and recovered binders at a stress level of 3.2 kPa and test temperature of 64°C. This parameter is used to characterize the resistance of asphalt binders to rutting and defines the high temperature performance grade of asphalt binders. AASHTO M 332 specifies a maximum J_{nr} requirement for standard (S), heavy (H), very heavy (V), and extremely heavy (E) traffic of 4.5, 2.0, 1.0, and 0.5 kPa^{-1}, respectively. VDOT specifications require of the use of PG64S-16 and PG64H-16 asphalt binders for A and D dense-graded surface mixtures, respectively.

Table 3 indicates that the binders from three mixtures (Mixtures F, G, and H) fell under the standard traffic (S) category and met the VDOT specification criterion for the high temperature PG for A mixtures. The J_{nr} parameter for the binder from Mixture O was slightly higher than the J_{nr} requirement for the standard traffic loading; Mixture O was not used for VDOT work but was included in the experimental plan because of its potentially high susceptibility to rutting. The binders from Mixtures H and I were adequate for heavy traffic (H) loading. Although Mixture H was specified as an A mixture, it exceeded the VDOT specification criterion. Mixture I was specified as a D mixture and met the VDOT specification criterion. All other mixtures, regardless of their designations, were adequate for either very heavy (V) or extremely heavy (E) traffic and exceeded VDOT specification criteria from the standpoint of high temperature binder grade. The intermediate and low-temperature properties of the binders are presented in Appendix B.

Table 2. APA Performance Criteria Used by U.S. State Agencies

State	Binder and Mixture Types	Criteria at 8,000 Cycles	Notes
Alabama	Mixtures subjected to 10 to 30 million ESALs	4.5 mm at 67°C	V_a at design number of gyrations
Alaska	All	3.0 mm at 40°C	6 ± 1% V_a
Arkansas	75 and 115 gyrations 160 and 205 gyrations	8.0 mm at 64°C 5.0 mm at 64°C	7 ± 1% V_a
Georgia	19 mm and 25 mm NMAS 9.5 mm and 12.5 mm NMAS	5.0 mm at 49°C 5.0 mm at 64°C	V_a at the design number of gyrations
Idaho	75 and 100 gyrations	5.0 mm at binder PG-HT	Compaction to 7% ± 1% V_a
North Carolina	9.5 mm NMAS, <0.3 million ESALs 9.5 mm NMAS, 0.3 to 3 million ESALs 9.5 mm NMAS, 3 to 30 million ESALs 9.5 mm NMAS, >30 million ESALs 12.5 mm NMAS, 3 to 30 million ESALs 12.5 mm NMAS, >30 million ESALs	11.5 mm at binder PG-HT 9.5 mm at binder PG-HT 6.5 mm at binder PG-HT 4.5 mm at binder PG-HT 6.5 mm at binder PG-HT 4.5 mm at binder PG-HT	4 ± 0.5% V_a
New Jersey	High performance thin overlay Bituminous rich intermediate course Bridge deck waterproof surface course Bituminous rich base course High recycled asphalt pavement mix, PG64-22 High recycled asphalt pavement mix, PG76-22	4.0 mm at 64°C (mix design) 5.0 mm at 64°C (production) 6.0 mm at 64°C (mix design) 7.0 mm at 64°C (production) 3.0 mm at 64°C 5.0 mm at 64°C 7.0 mm at 64°C 4.0 mm at 64°C	Field air void level
Ohio	Non-polymer mix Heavy surface and high stress mix Bridge deck waterproofing mix	5.0 mm at 48.9°C 3.0 mm at 54.4°C 4.0 mm at 64°C	7 ± 1% Va; hose pressure of 100 psi and wheel load of 115 lbf
Oregon	80 gyrations, PG58-xx 80 gyrations, PG64-xx 80 gyrations, PG70-xx 100 gyrations, PG64-xx 100 gyrations, PG70-xx 100 gyrations, PG76-xx	6.0 mm at 64°C 5.0 mm at 64°C 4.0 mm 64°C	V_a at design number of gyrations
South Carolina	PG76-22 PG64-22	3.0 mm at 64°C 5.0 mm at 64°C	4 ± 1% V_a
South Dakota	Truck ADT <75 Truck ADT 76 to 250 Truck ADT 251 to 650 Truck ADT >651	8.0 mm at binder PG-HT 7.0 mm at binder PG-HT 6.0 mm at binder PG-HT 5.0 mm at binder PG-HT	V_a at design number of gyrations

APA = asphalt pavement analyzer; ESAL = equivalent single axle load; V_a = air void content; NMAS = nominal maximum aggregate size; PG = performance grade; PG-HT = PG at high temperature; ADT = annual daily traffic.

Table 3. Volumetric and Binder Properties and Gradations for All Mixtures

Mixture	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Mixture Type	9.5A	12.5A	12.5D	9.5D	9.5A	9.5A	9.5A	9.5A	9.5D	9.5D	9.5A	9.5D	9.5A	9.5A	12.5A	12.5A	9.5A
RAP, %	30	30	26	26	26	40	45	40	26	30	45	30	45	60	0	15	30
Volumetric Property																	
NMAS, mm	9.5	12.5	12.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	12.5	9.5	9.5	12.5	12.5	9.5
AC, %	5.64	5.04	5.39	6.19	5.47	5.61	6.96	5.45	6.03	6.09	6.34	5.76	6.16	5.88	5.26	5.80	5.42
VTM, %	3.3	3.1	3.0	2.2	4.5	2.7	0.7	4.2	2.3	5.9	2.6	4.0	2.6	2.4	4.4	2.3	3.6
VMA, %	16.1	15.3	15.8	16.3	17.5	16.7	16.9	17.7	16.5	19.5	17.3	17.2	17.0	15.8	16.2	15.5	17.0
VFA, %	79.5	79.6	81.3	86.2	74.2	83.9	96.1	76.0	86.3	69.8	84.7	77.0	84.8	84.6	72.9	85.0	78.7
FA Ratio	0.87	1.28	1.36	1.03	1.25	1.13	1.25	1.23	1.10	1.01	1.23	1.03	1.20	1.33	1.02	1.42	1.27
G_{se}	2.651	2.916	2.890	2.691	2.834	2.921	2.820	2.931	2.878	2.792	2.821	2.787	2.801	2.803	2.655	2.834	2.924
G_{sb}	2.648	2.901	2.872	2.680	2.830	2.920	2.799	2.930	2.858	2.776	2.801	2.773	2.789	2.777	2.649	2.800	2.921
P_{ba}, %	0.04	0.18	0.22	0.16	0.05	0.01	0.27	0.01	0.25	0.21	0.26	0.19	0.16	0.34	0.09	0.44	0.04
P_{be}, %	5.59	4.87	5.18	6.04	5.42	5.60	6.70	5.43	5.79	5.89	6.10	5.58	6.01	5.56	5.17	5.38	5.39
Binder Property																	
J_{nr}, kPa^{-1}	1.85	0.43	0.50	1.00	0.28	2.39	3.90	1.02	1.13	0.57	0.65	0.66	0.37	0.24	4.54	3.69	0.92
Gradation, percent passing																	
¾ in (19.0 mm)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
½ in (12.5 mm)	98.9	97.3	99.3	100.0	99.4	100.0	99.2	100.0	98.9	99.5	99.0	100.0	99.3	98.5	91.8	94.7	100.0
3/8 in (9.5 mm)	93.3	86.6	92.0	91.9	94.7	95.5	93.8	96.6	93.4	94.2	95.0	96.3	93.0	93.2	70.6	87.5	96.9
No. 4 (4.75 mm)	60.7	55.6	58.5	56.7	65.9	63.7	62.5	63.9	63.0	65.1	68.0	67.3	62.2	65.4	57.5	56.3	66.9
No. 8 (2.36 mm)	43.9	39.2	39.6	40.6	47.1	42.3	41.0	41.5	42.3	40.1	43.3	41.5	41.5	43.0	39.7	35.9	42.2
No. 16 (1.18 mm)	35.0	30.3	28.4	33.3	34.9	30.5	27.3	29.6	28.9	24.1	27.3	30.3	27.5	29.0	26.0	25.7	29.9
No. 30 (600 µm)	25.8	23.5	20.1	25.0	24.5	22.1	19.4	21.7	19.8	15.8	18.6	24.1	19.2	20.8	16.4	20.0	21.5
No. 50 (300 µm)	17.0	16.8	13.7	13.1	15.4	15.1	14.1	15.2	13.0	10.5	12.9	14.1	12.9	14.3	9.9	16.4	15.2
No. 100 (150 µm)	8.4	10.3	9.8	8.0	9.5	9.6	10.6	10.0	8.8	7.6	9.5	8.1	9.3	9.9	7.0	12.4	10.2
No. 200 (75 µm)	4.9	6.2	7.1	6.2	6.7	6.3	8.4	6.7	6.4	5.9	7.5	5.7	7.2	7.4	5.3	7.6	6.9

Red text indicates values that did not meet the VDOT production criteria for gradations and volumetric properties, respectively (VDOT, 2020).

RAP = reclaimed asphalt pavement; NMAS = nominal maximum aggregate size; AC = asphalt content; VTM = voids in total mixture; VMA = voids in mineral aggregate; VFA = voids filled with asphalt; FA ratio = fines/asphalt ratio; G_{se} = aggregate effective specific gravity; G_{sb} = aggregate bulk specific gravity; P_{ba} = absorbed asphalt content; P_{be} = effective asphalt content; J_{nr} = non-recoverable creep compliance at 3.2 kPa and 64°C.

Selection of Test Temperature for Monotonic Tests

Two temperatures, 49.0°C and 54.4°C, were selected for further evaluation to determine the critical testing temperature for the monotonic tests. A series of tests were performed on five mixtures (Mixtures A through E) at 49.0°C and 54.4°C to investigate any potential differences between the results at the two temperatures. Figure 1 shows the correlations between the test results at the selected temperatures. The extent of the correlation between the results was evaluated in terms of a coefficient of determination (R^2) with respect to the ranking criteria as follows: excellent (≥90%), good (70%-80%), fair (40%-70%), poor (20%-39%), and very poor (≤19%) (Tran and Hall, 2005). As seen, the correlation was excellent for the IDT-HT test, good for the RR test, and poor for the MS test. It can be inferred from the position of the data with respect to the equality line in the figures that the tests were able to capture the expected trend with respect to temperature, the increase in measured property with a decrease in temperature, except for one mixture under the MS testing. In addition, the repeatability of the tests was improved at 54.4°C for two tests (the RR and MS tests), although the repeatability, as quantified by the COV, for these mixtures was under 15% at both temperatures. Moreover, the 54.4°C tests resulted in a higher performance discrimination potential, especially for the RR and MS tests. Thus, 54.4°C was selected as the testing temperature for the monotonic tests.

Comparative Evaluation of the Monotonic Tests and the APA Test

The monotonic tests were first evaluated relative to the APA test from several perspectives. As indicated previously, the results of the tests were analyzed in terms of repeatability, performance discrimination potential, ranking among the asphalt mixtures, sensitivity of the tests to mixture properties, relationship to J_{nr}, and correlation with each other. This evaluation included 16 mixtures (Mixtures A through P). Mixture Q was included later in the experimental program of the study due to material shortages of the other mixtures.

Test Results

Following the selection of the test temperature, the remainder of the mixtures were tested at 54.4°C with the three monotonic tests, and all mixtures were subjected to the APA test at 64°C in accordance with AASHTO T 340. The average results were obtained from three replicate measurements from the monotonic tests, and the APA results were obtained from two replicate tests with each replicate being composed of two specimens. The results of all of the tests are shown in Figure 2. It is expected that a higher strength, RT index, or peak load from the monotonic tests and a lower rut depth from the APA test will correspond to a lower rutting potential in the field. The wide range of test results indicated a wide potential applicability of the outcomes of this study.

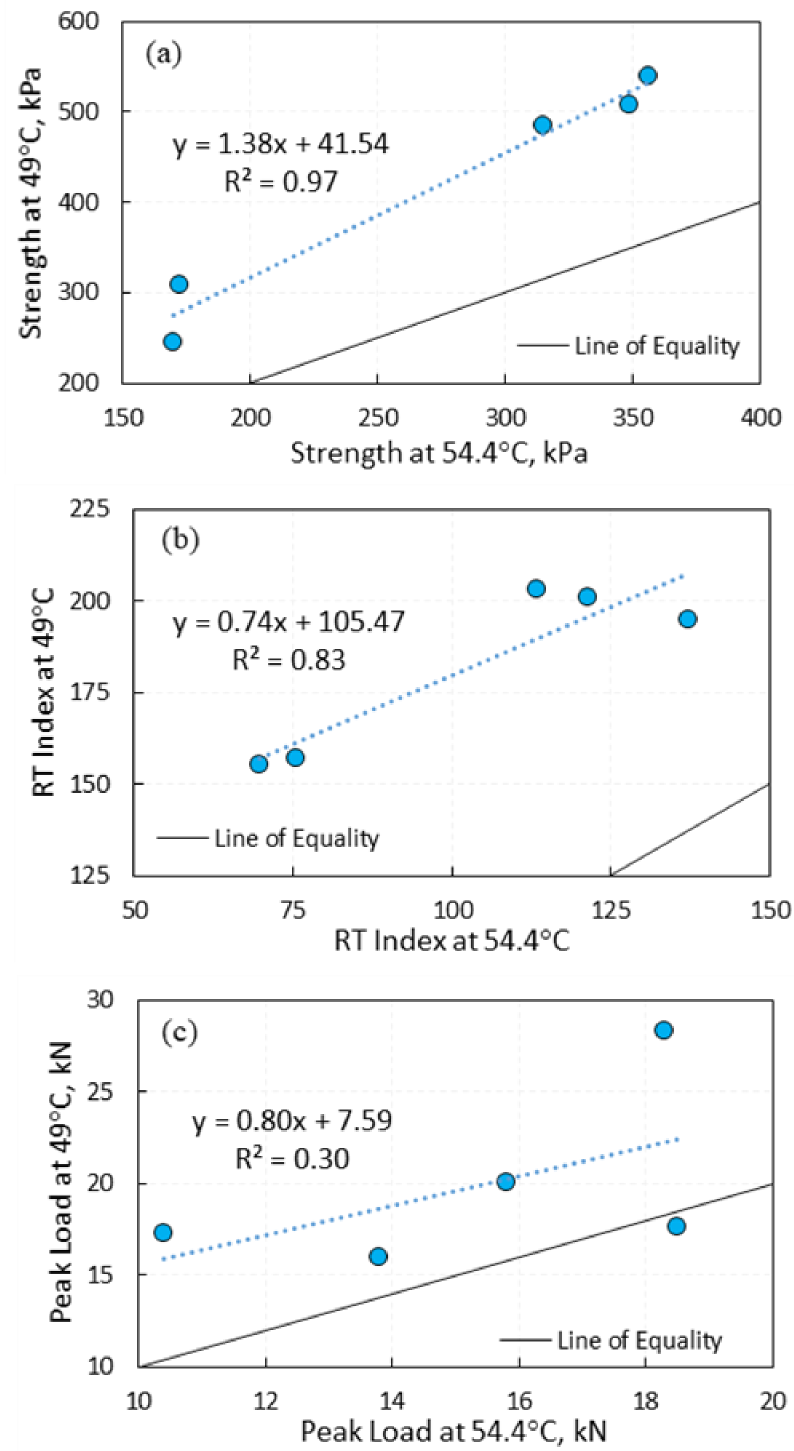


Figure 1. Correlation Between Results of Tests: (a) IDT-HT test; (b) RR test; (c) MS test at 49°C and 54.4°C. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow.

Repeatability

The repeatability (single-operator precision) and reproducibility (multi-laboratory precision) of a test method are quantitative expressions of variability that represent the two extremes of test method precision, and they must be established if a test method is to be incorporated in any quality assurance specifications (Hand et al., 2021).

The knowledge of variability provides a rational basis for establishing tolerance limits and acceptance practice. For example, a test method with low variability provides a higher precision, which requires fewer replicate specimens for testing, and is more likely to differentiate asphalt mixtures with a range of performance levels compared to a test method with high variability (Zhou et al., 2017). In addition, a test method with low variability can be considered more advantageous during the quality measurement process because statistical differences between two sets of a given mixture (i.e., multi-laboratory evaluation) can be confidently determined without the masking effect from a test method with high variability (Boz et al., 2021b; Habbouche et al., 2021a). Moreover, the variability parameters (whether single-operator or multi-laboratory) can be incorporated into the process of establishing a performance threshold criterion for a given test method (Bennert et al., 2018; Diefenderfer et al., 2021b; Habbouche et al., 2021b).

The precision estimates of the considered tests based on the mixture types and operating conditions used in this study such as test temperature and specimen dimensions are not yet available. However, the repeatability of the tests can be evaluated by the measure of intermediate precision. For this study, the intermediate precision refers to the precision calculated from the results of testing of the mixtures conducted under the same conditions by the same operator in the same laboratory over the course of the presented work. Since the magnitude of each test output differed significantly, the COV was used as the estimate of intermediate precision.

Figure 3 presents the COV for each mixture for each test. It can be seen that the spread and overall magnitude of the COV for the IDT-HT and RR tests are narrower and smaller than those for the MS and APA tests. The COV for the IDT-HT test ranged from 2.2% to 21.6%, with an average of 13.1%, whereas it ranged from 0.6% to 26.4%, with an average of 12.2% for the RR test. On the other hand, the average COVs for the MS and APA tests were 16.6% and 16.4%, respectively, with a COV ranging from about 3.5% to 38% for both tests. The results indicated that the IDT-HT and RR tests have similar repeatability characteristics and provide a better measure of repeatability than the MS and APA tests.

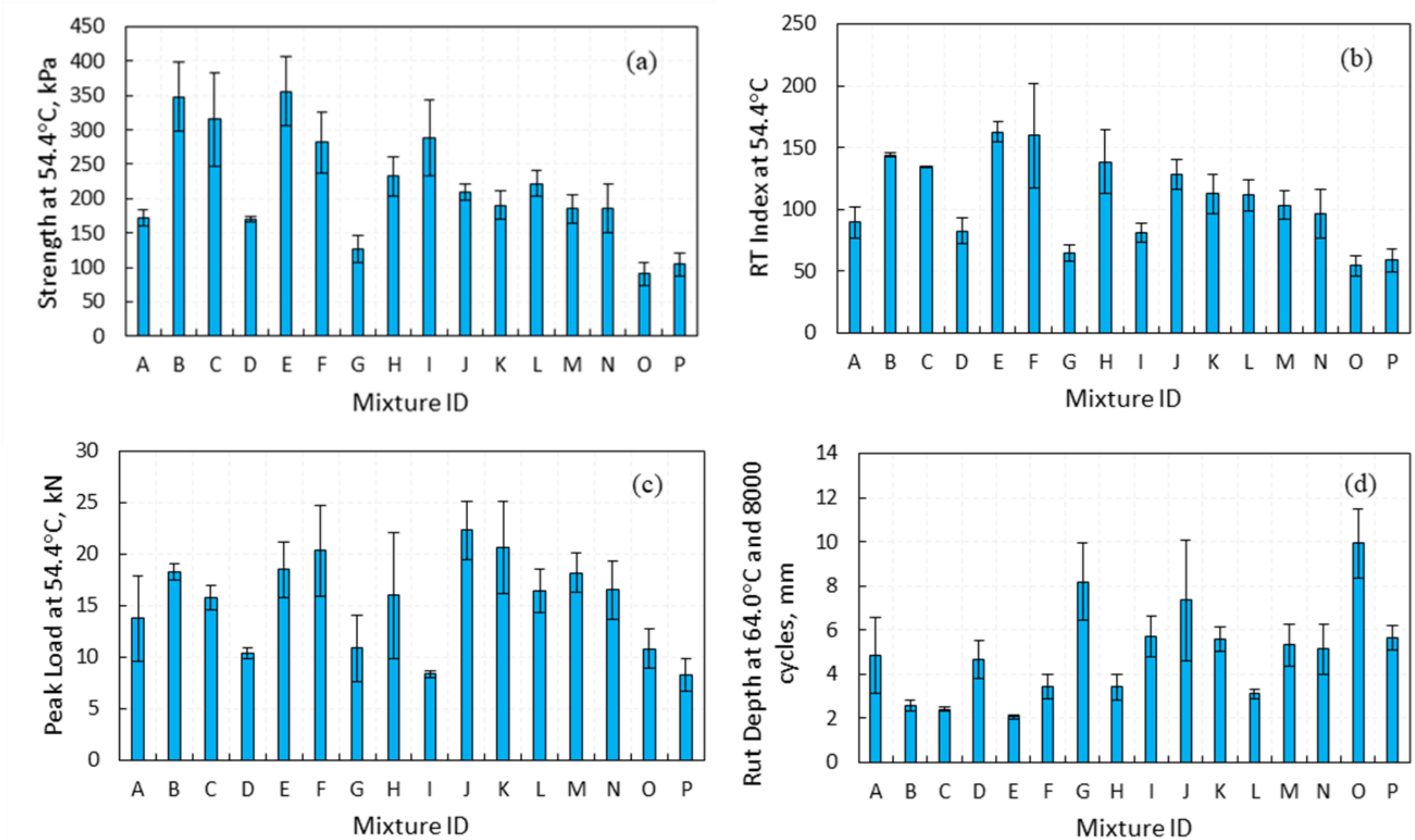


Figure 2. Test Results: (a) IDT-HT test; (b) RR test; (c) MS test; (d) APA test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer.

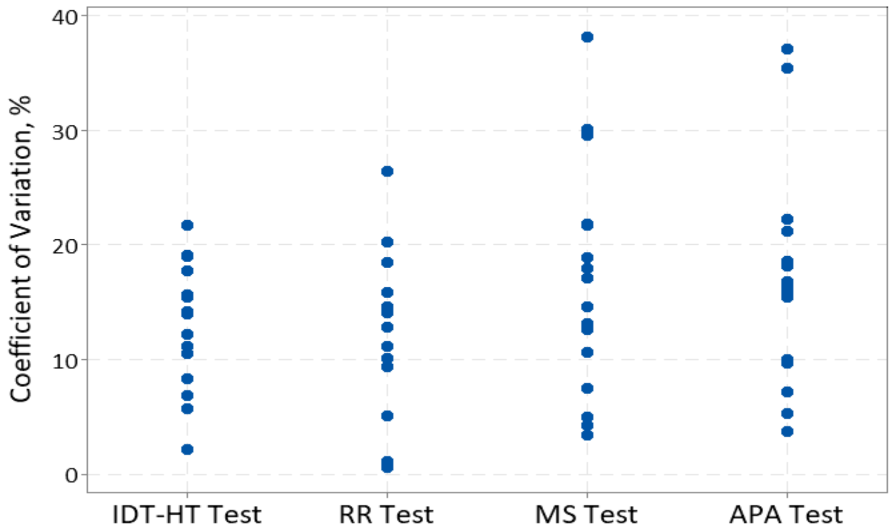


Figure 3. Spread of the Coefficient of Variation for Each Test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer.

Performance Discrimination Potential

The ability of the test methods to discriminate or differentiate the rutting performance of the mixtures is assessed through statistical analysis. The one-way analysis of variance (ANOVA) with Tukey's multiple comparison method at a 95% confidence interval was used for that purpose. Prior to the ANOVA, the data for each test were checked for the assumptions of normality and equal variances at a 95% confidence interval. The results confirmed that the data from all tests were normally distributed and had equal variances. The ANOVA results showed statically significant differences for the test results among the mixtures, and the Tukey test was performed to identify the differences or similarities between specific pairs of the mixture groups for a given test. The results of the Tukey test are presented in Table 4. The mean values for mixtures sharing the same letter in the table are statistically similar. As seen, the IDT-HT and APA tests classified the mixtures into six statistically distinct groups, making these two tests the most discriminating regarding the performance of the mixtures. The RR test resulted in five groups, whereas the MS test categorized the mixtures into only three statistical groups, indicating that the MS test was the least sensitive test among the tests considered.

Performance Ranking Order

An analysis was conducted to compare the tests in terms of their potential to rank the rutting performance of the mixtures. The mixtures were sorted from the most rut-resistant to the least rut-resistant to determine their order of ranking. The results are presented in Table 5. As shown, Mixture E was consistently identified as the most rut-resistant mixture, and Mixture O was identified as the most rut-susceptible mixture by all tests except the MS test. A close inspection of the table indicates that there were differences in the overall ranking of the individual mixtures among the tests. When these differences were assessed with respect to the performance discrimination analysis in the preceding section and the correlation analysis presented later in this report, the order of the ranking became statistically insignificant.

However, this does not negate the fact that these tests identified the most and least rut-resistant mixtures. For the MS test, the ranking results showed a more scattered trend compared to those of the other three tests, potentially because of the high variability in the test results.

Table 4. Tukey Pairwise Statistical Comparisons of the Tests

Mix ID	IDT-HT Test	RR Test	MS Test	APA Test
A	d/e/f	c/d/e	a/b/c	c/d/e/f
B	a	a/b/c	a/b/c	e/f
C	a/b	a/b/c/d	a/b/c	e/f
D	d/e/f	d/e	b/c	c/d/e/f
E	a	a	a/b	f
F	a/b/c/d	a/b	a	d/e/f
G	e/f	e	b/c	a/b
H	b/c/d	a/b/c	a/b/c	d/e/f
I	a/b/c	d/e	c	b/c/d
J	c/d/e	a/b/c/d	a	a/b/c
K	c/d/e/f	b/c/d	a	b/c/d
L	b/c/d/e	b/c/d	a/b/c	d/e/f
M	d/e/f	c/d/e	a/b	b/c/d/e
N	d/e/f	c/d/e	a/b/c	c/d/e
O	f	e	b/c	a
P	f	e	c	b/c/d

The mean values for mixtures sharing the same letter in the table were statistically similar.
IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer.

Table 5. Mix Rutting Performance Ranking From the Most Resistant to the Most Susceptible

IDT-HT Test	RR Test	MS Test	APA Test
E	E	J	E
B	F	K	C
C	B	F	B
I	H	E	L
F	C	B	H
H	J	M	F
L	K	N	D
J	L	L	A
K	M	H	N
N	N	C	M
M	A	A	K
A	D	G	P
D	I	O	I
G	G	D	J
P	P	I	G
O	O	P	O

IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer. The letters indicate the Mix ID.

Sensitivity to Mixture Properties

Statistical analysis techniques were also used to investigate the sensitivity of the test results to the mixture volumetric parameters, gradation, and J_{nr}. An ANOVA at a 95% confidence interval was conducted to determine the statistically significant factors for each test.

The following volumetric properties were calculated at the 7% air void content at which the tests were performed and included in this analysis: fines to asphalt ratio (F/A), VMA, and bulk specific gravity of the aggregate blends (G_{sb}). The effective volume of asphalt ($\%V_{be}$) and the VFA factors were not included in the analysis as they are highly correlated with VMA at the same air void content. Each gradation sieve and the J_{nr} parameter were included in the analysis. The statistically significant factors identified from the analysis are presented in Table 6. As shown, the IDT-HT test has six statistically significant factors followed by the APA test with four statistically significant factors affecting the test results. The RR test results were statistically sensitive to changes in four factors, whereas the number of statistically significant factors was two for the MS test. It is worth noting that the majority of the tests identified G_{sb} and J_{nr} as significant factors influencing the rutting potential of the mixtures.

Table 6. Statistically Significant Factors Affecting the Results of the Tests

Test	Factor	
	Sieve Size	**Volumetric/Binder Property**
IDT-HT	9.5 mm, No. 16, No. 30, and No. 200	G_{sb} and J_{nr}
RR	No. 8 and No. 200	G_{sb} and J_{nr}
MS	No. 4	J_{nr}
APA	No. 30	G_{sb}, VMA, and J_{nr}

IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer; G_{sb} = bulk specific gravity of the aggregate blend; J_{nr} = non-recoverable creep compliance at 3.2 kPa and 64°C; VMA = voids in mineral aggregate.

Correlations With the J_{nr} Parameter

Figure 4 shows the correlations between the test results and the J_{nr} parameter. As seen in Figure 4(a), the correlation of the IDT-HT test with the J_{nr} rutting parameter was fair using a power fit resulting in an R^2 of 54%, and the correlation trend was in the right direction, where the higher the strength, the lower the J_{nr} parameter. The magnitude of the correlation increased to an R^2 of 70% when the data point encircled in the figure was excluded from the analysis, indicating a good correlation of the J_{nr} parameter with the IDT-HT test results. Similar observations were made with the RR test, as shown in Figure 4(b). There was a fair correlation between the RT index and the J_{nr} parameter using a power fit resulting in an R^2 of 53%, which increased to an R^2 of 80% after the same data point was removed, showing a good relationship with the J_{nr} parameter. The correlations between the J_{nr} parameter and the MS and APA tests were more scattered compared to the IDT-HT and RR tests. As shown in Figure 4(c), the correlation between the J_{nr} parameter and the MS test was fair using a power fit with an R^2 of 56%. Likewise, the correlation between the J_{nr} parameter and the APA test was fair as quantified through a linear fit with an R^2 of 54%, as shown in Figure 4(d). The results indicated that, overall, the tests were able to capture the effect of binder on the mixture performance response, especially the IDT-HT and RR tests.

Correlation Among the Tests

The existence of potential correlations among the results of the monotonic tests was investigated through graphical comparisons with fitted functions. Figure 5 presents the correlations.

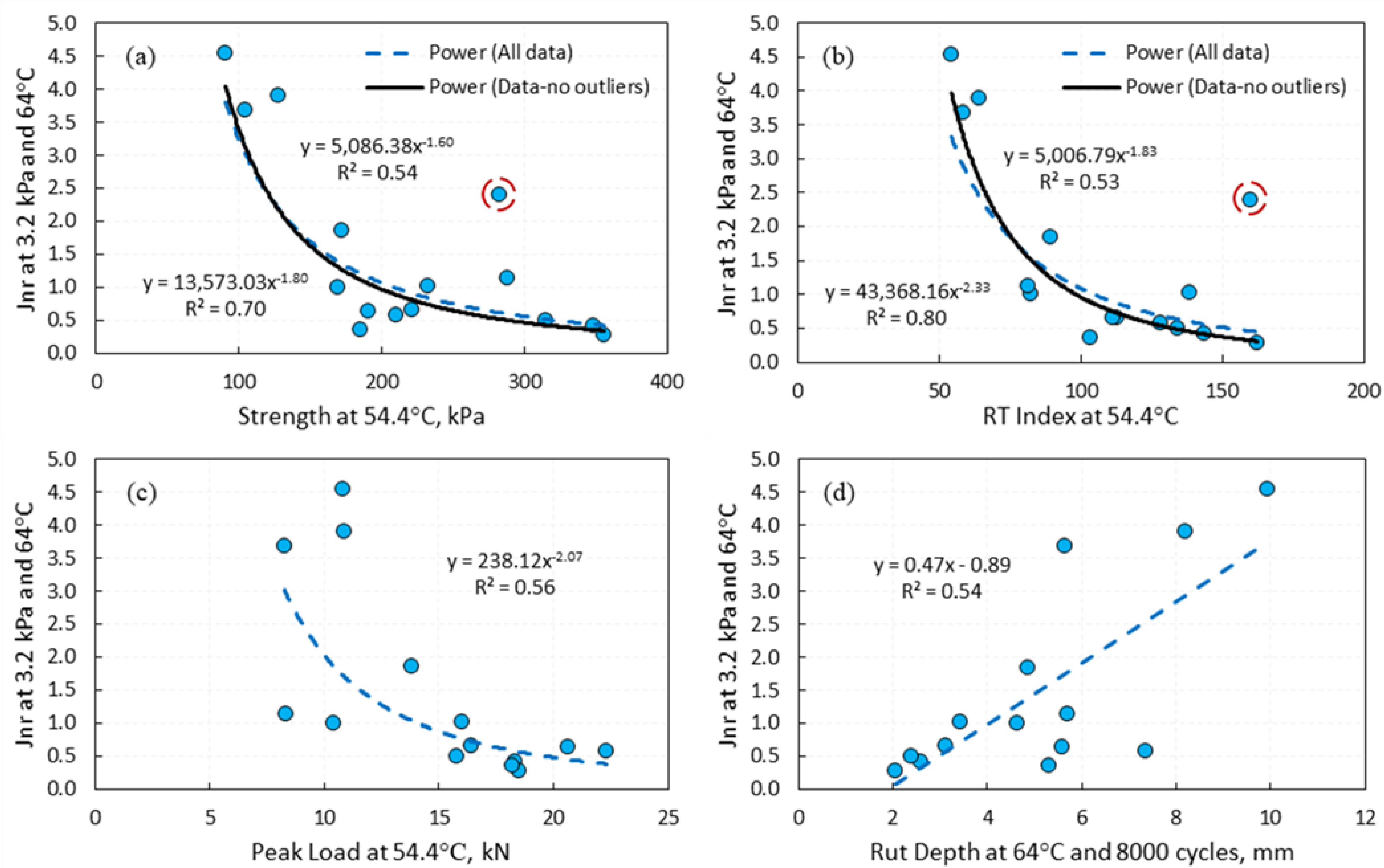


Figure 4. Correlation Between the Non-Recoverable Creep Compliance (J_{nr}) and Other Test Results: (a) IDT-HT; (b) RR; (c) MS; (d) APA. The circled data points appear to be an outlier. R^2 improves when the circled points are removed from the analysis. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer.

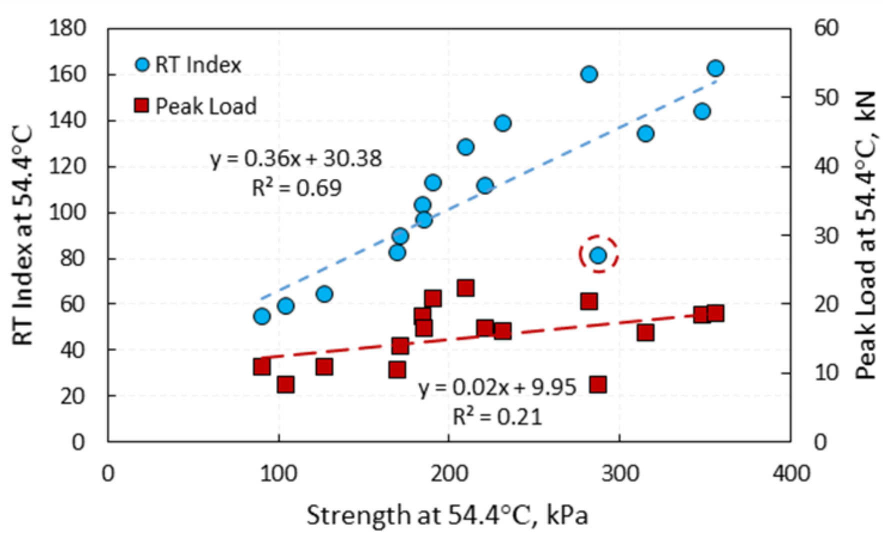


Figure 5. Correlation Among the Results of the Monotonic Tests. The circled data point appears to be an outlier. The R^2 shown is for all data points including the encircled potential outlier.

As shown, the correlation between the IDT-HT strength and RT index was fair through a linear fit with an R^2 of 69%. When the potential outlier data point encircled in the figure was excluded, the goodness of the linear fit increased to 85% (although not shown in the figure), indicating a good correlation between the results of the two tests. Although the trend of the correlation is in the right direction, there was no correlation between the results of the MS test and the IDT-HT test (R^2 of 21%) or the RR test (not shown).

The correlations between the results of the monotonic tests and the APA test were also evaluated. Figure 6(a) indicates a fair correlation between the results of the IDT-HT test and the APA test established through a power fit with an R^2 of 65%. The figure also shows the confidence interval at a 95% level for the test results, and it is seen that the majority of the data points fall within the expected variation of the test results. If the three data points outside the confidence interval (potential outliers) are removed from the analysis, the goodness of the correlation between the results of the two tests increases to excellent with an R^2 of 91% (the fit is not shown in the figure). Such a level of correlation between the results of the two tests was also reported by others (Bennert et al., 2018; Bennert et al., 2021; Zaniewski and Srinivasan, 2004). Thus, the results here indicated the robustness of the correlation between the results of the two tests.

Figure 6(b) shows that a fair correlation exists between the results of the RR and APA tests through a power fit with an R^2 of 60%. Similar to the correlation between the results of the IDT-HT test and the APA test, most of the data points for the RR tests also fall within the confidence interval limits at 95%. Removing three data points outside the confidence interval resulted in a good correlation with an R^2 of 85%. Zhou et al. (2020) reported a higher level of correlation between the tests with an R^2 of 92% through an exponential fit. As shown in Figure 6(c), there was no correlation between the results of the MS and APA tests (R^2 of 13%), although the trend of the results was in the right direction. The results indicated that the results of the IDT-HT and RR tests are correlated with those of the APA test, with the IDT-HT test having a slightly better correlation.

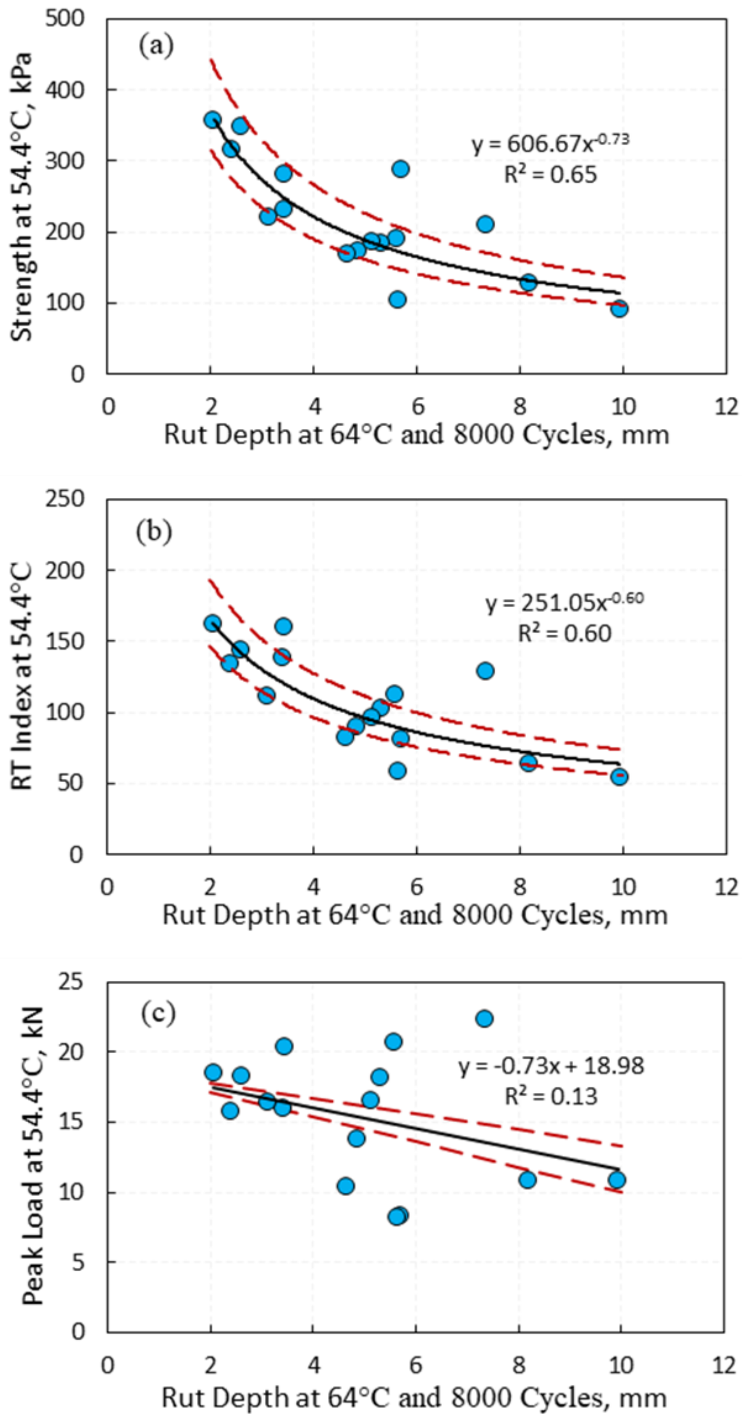


Figure 6. Correlation Between Results of APA Test and Other Tests: (a) IDT-HT test; (b) RR test; (c) MS test. Red dashed lines indicate the 95% confidence interval for the APA test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow.

The forensic analysis through the DIC method, as presented in Appendix A, indicated that the IDT-HT and RR tests each induce a different state of strain within the body of specimens. The magnitudes of horizontal strain and maximum principal strain on the surface of the specimen at the peak load are also different. There is also no correlation between the strain values at the peak load, although this lack of correlation may be due to the limited number of mixtures tested under this task. Nevertheless, further analysis is needed with a larger number of mixtures to investigate the correlation between the two tests and how the test (or other) parameters obtained from the tests provide an indication of the rutting potential of asphalt mixtures.

The overall evaluation of the tests based on the analysis is summarized in Table 7. The tests were ranked based only on the relative comparisons of the results obtained in this task. The test meeting the desired characteristics for each evaluation parameter is identified in bold italic text. It seen that the IDT-HT test ranked first and the RR test ranked second with regard to meeting the desired characteristics. These were followed by the APA test and the MS test as the third and fourth ranked tests, respectively.

Table 7. Overall Evaluation of the Tests

Parameter	Test			
	IDT-HT	**RR**	**MS**	**APA**
Repeatability	Good	***Very Good***	Poor	Fair
Discrimination Potential	***Very Good***	Good	Fair	***Very Good***
Performance Ranking	Similar	Similar	Scattered	Similar
Sensitivity to Mixture Properties	***Very Good***	Fair	Poor	Good
Correlation with the J_{nr} Parameter	Good	***Very Good***	Fair	Fair

Bold italic text indicates that the test met the desired characteristic.
IDT-HT = indirect tensile at high temperature; RR = rapid rutting; MS = Marshall stability and flow; APA = asphalt pavement analyzer; J_{nr} = non-recoverable creep compliance.

Comparison to Fundamental Rutting Tests

The results of the IDT-HT, RR, and the APA tests were compared to the results of the fundamental rutting tests to evaluate further the ability of the tests to capture the rutting potential of asphalt mixtures. The MS test was not included in this part of the study because of the findings in the preceding sections. The fundamental rutting tests conducted in this study were the dynamic modulus test, SSR test, and repeated load permanent deformation (RLPD) test. Six mixtures (Mixtures G, K, L, O, P, and Q) were used for testing and evaluation. These mixtures were selected to provide a wide range of rutting performance relative to the results of the monotonic tests and the APA test.

Dynamic Modulus Test

Figure 7 plots the dynamic modulus (|E*|) master curves for the six mixtures. The generalized logistic model alongside the polynomial shift factor was used to construct the |E*| master curves at a reference temperature of 21°C (Boz et al., 2017). In general, a higher |E*| value at a given temperature and frequency (loading) indicates a more rut-resistant mixture. Several index parameters from the |E*| master curve can be obtained to quantify mixture rutting performance. The |E*| values at 38°C and 0.1 Hz can be used for that purpose (Apeagyei et al., 2011). Through the time-temperature superposition principle, the |E*| values at 38°C and 0.1 Hz were first determined from the curves and then compared to the results of the IDT-HT, RR, and APA tests.

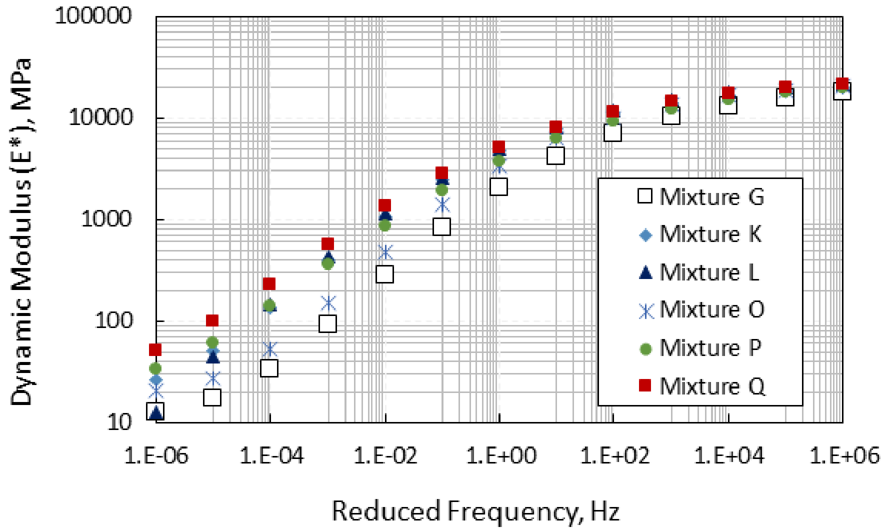


Figure 7. Dynamic Modulus Mastercurves for the Mixtures at a Reference Temperature of 21°C

Figure 8 presents the correlations between the results of the tests and the |E*| rutting parameter. As shown, the correlations with the |E*| rutting parameter were established through a linear fit having an R^2 of 74%, 71%, and 75% for the IDT-HT, RR, and APA tests, respectively. The correlation for the tests was described as good. The highest correlation with the |E*| rutting parameter at 38°C and 0.1 Hz was with the APA test, followed closely by the IDT-HT test and then the RR test. The trends were in the expected direction such that as the APA rut depth increased, the E* value decreased, and as the IDT-HT strength and RT index increased, the E* value increased.

Stress Sweep Rutting (SSR) Test

The RSI results were compared to the results of the IDT-HT, RR, and APA tests, as shown in Figure 9. The relationship between the IDT-HT strength and the RSI parameter was good through a power fit with an R^2 of 76%, whereas it was excellent through a power fit with an R^2 of 91% for the RR test. The APA rut depth was fairly correlated to the RSI parameter through a power fit with an R^2 of 65%. The results showed that the results of the tests followed the expected trends with respect to the RSI parameters, as a lower RSI value indicates a more rut-resistant mixture.

Repeated Load Permanent Deformation (RLPD) Test

The RLPD tests were performed at three temperatures, 30°C, 40°C, and 50°C, for each mixture, but only the results at 50°C are presented here for the sake of brevity. The results of this test at the three temperatures were used as input in Pavement ME Design for the ME analysis. Figure 10 presents comparisons between the FN index from the RLPD test and the results of the IDT-HT, RR, and APA tests for the five mixtures.

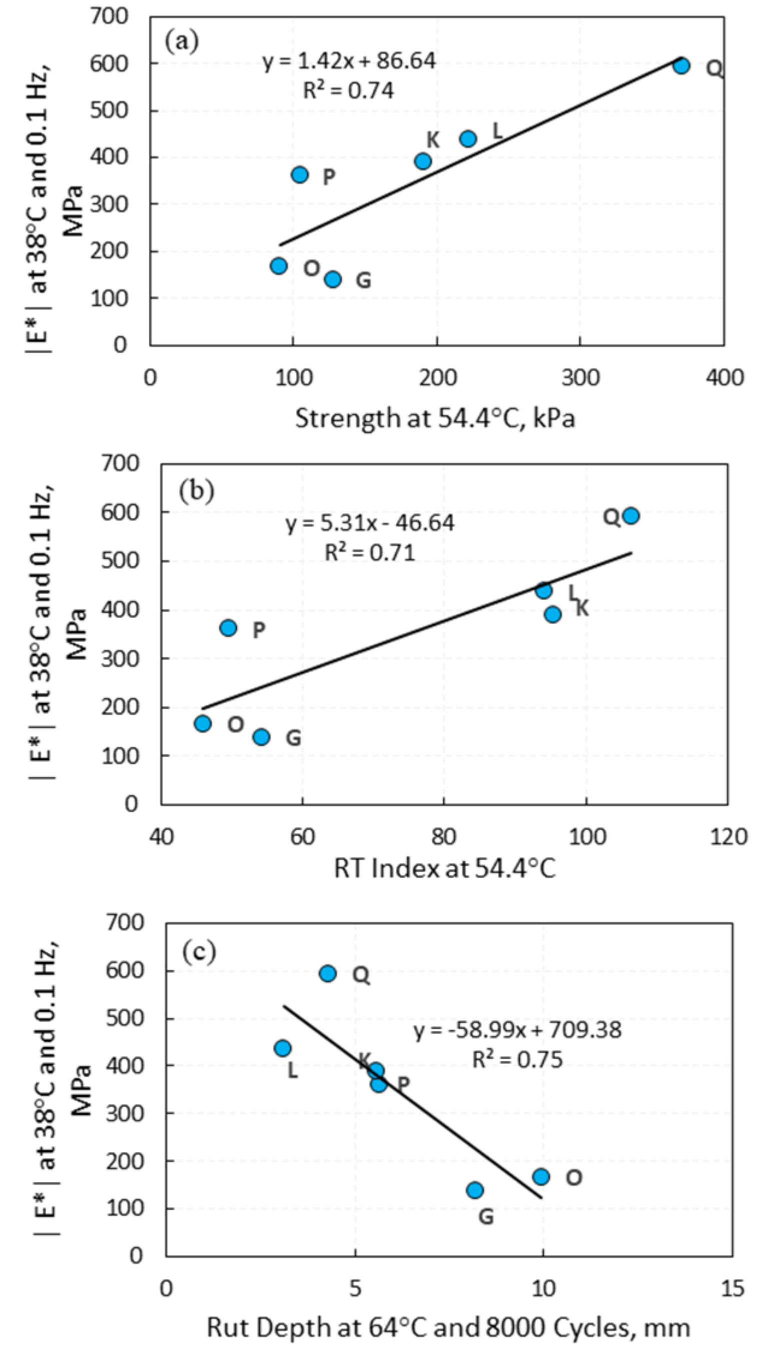


Figure 8. Correlation Between Results of |E*| at 38°C and 0.1 Hz and Other Tests: (a) IDT-HT test; (b) RR test; (c) APA tests. |E*| = dynamic modulus; IDT-HT = indirect tensile at high temperature; RR = rapid rutting; APA = asphalt pavement analyzer.

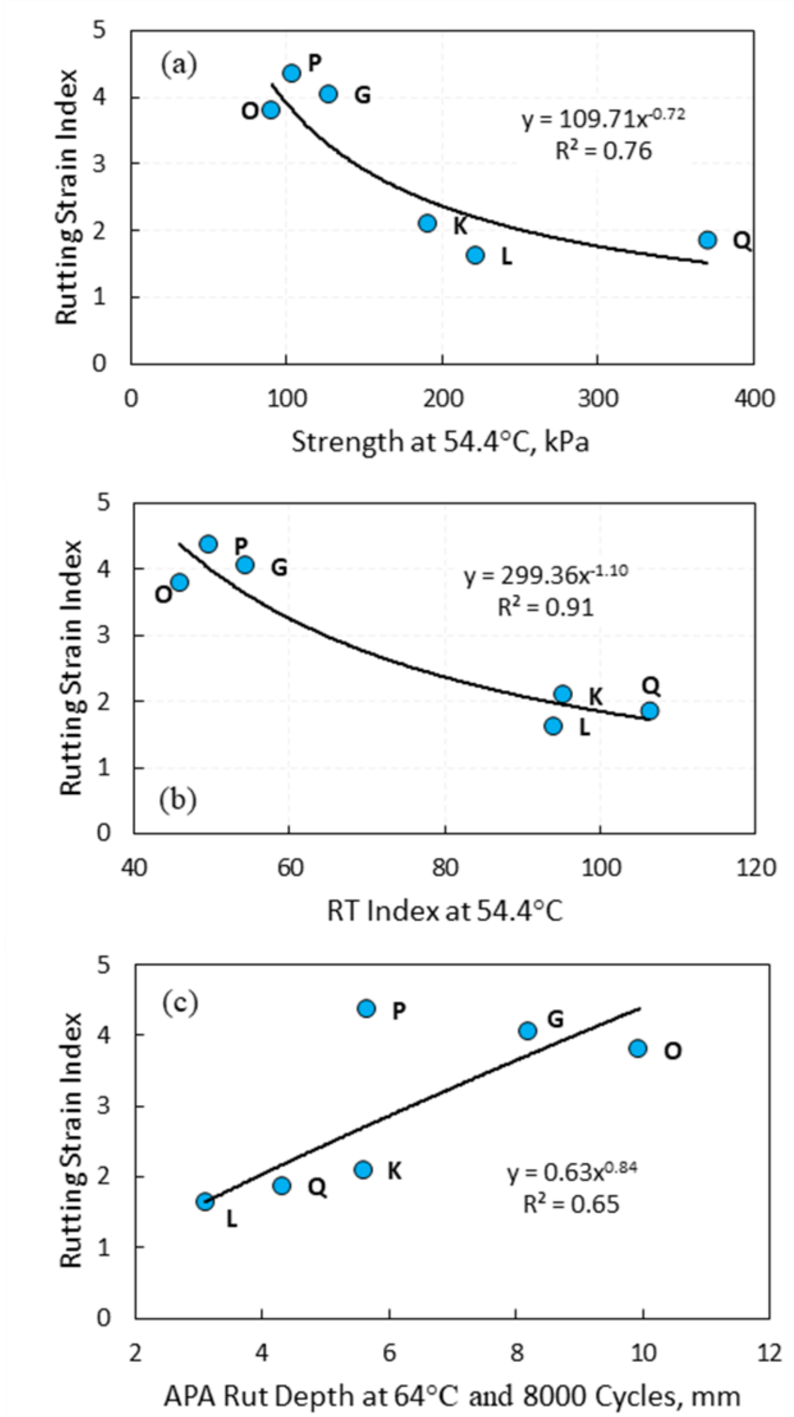


Figure 9. Correlation Between Results of Rutting Strain Index and Other Tests: (a) IDT-HT test; (b) RR test; (c) APA tests. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; APA = asphalt pavement analyzer.

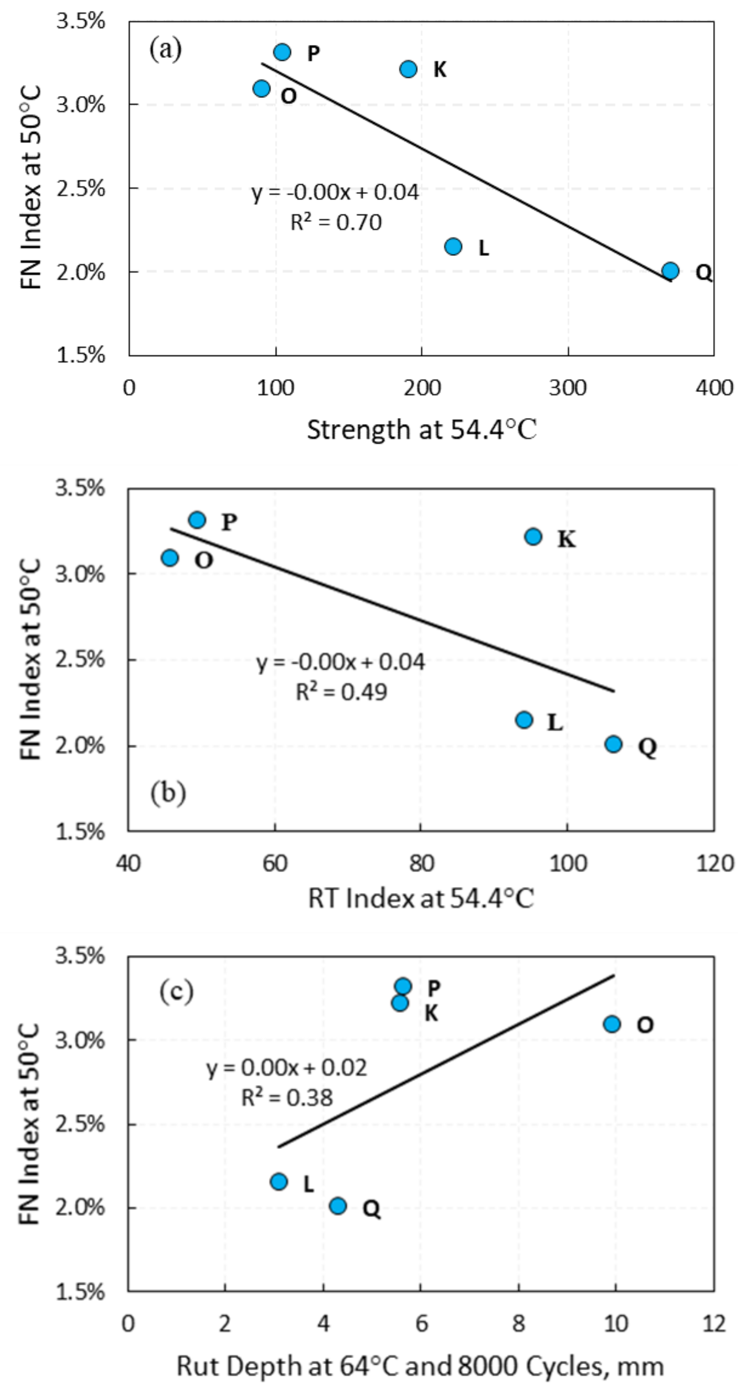


Figure 10. Correlation Between Results of FN Index and Other Tests: (a) IDT-HT test; (b) RR test; (c) APA test. FN = flow number; IDT-HT = indirect tensile at high temperature; RR = rapid rutting; APA = asphalt pavement analyzer.

Mixture G was excluded from this analysis, as it failed drastically under the RLPD test; the inclusion of Mixture G with the other five mixtures resulted in no correlations shown between the FN index and the results of any of the other three tests. Mixture G was ranked as one of mixtures having the lowest resistance to rutting by all other tests used in this study, including the fundamental rutting tests. When variations in the test results were considered, all tests ranked Mixture G in the lowest rutting resistance category.

However, none of these tests was able to differentiate significantly Mixture G from other mixtures with the lowest resistance to rutting to the extent that the FN index did. Mixture G drastically failed under accelerated full-scale testing. This observation warrants further investigations on test methods and their ability to capture and discriminate the rutting susceptibility of asphalt mixtures, especially for those not meeting asphalt mixture volumetric properties such as Mixture G. Nevertheless, as shown in Figure 10, the correlations between the FN index and the results of the tests were established through a linear fit. The IDT-HT showed a good correlation with the FN index with an R^2 of 70%, whereas the RR and APA had a fair correlation with the FN index with an R^2 of 50% and 38%, respectively. The results indicated that the results of the tests followed the expected trends with respect to the FN index, as a lower FN index value indicates a more rut-resistant mixture.

The overall evaluation of the tests based on the correlation analysis is summarized in Table 8. Ranking of the tests in the table was exclusively established based on relative comparisons of the results obtained in this part of the study. The test meeting the desired characteristics for each evaluation parameter is identified in bold italic text. As shown in Table 8, each test met the desired characteristic for once. Further inspection of the table indicated that the IDT-HT test overall had the highest degree of correlation to the fundamental rutting tests, followed by the RR and APA tests.

Table 8. Agreement Between IDT-HT, RR, and APA and Fundamental Rutting Test Parameters

Parameter	IDT-HT	RR	APA
E* at 38°C and 0.1 Hz	Medium	Low	***High***
Rutting strain index	Medium	***High***	Low
FN index at 50°C	***High***	Medium	Low

Bold italic text indicates that the test meets the desired characteristic.
E* = dynamic modulus; FN = flow number; IDT-HT = indirect tensile at high temperature; RR = rapid rutting; APA = asphalt pavement analyzer.

Development of Performance Thresholds

Based on the analysis presented thus far, the IDT-HT test and the RR test can be used to evaluate the rutting potential of asphalt mixtures. This prompted the need to establish performance threshold criteria for both tests. Four different approaches were undertaken to establish rutting thresholds for the two tests. This study also provided an opportunity to verify the existing rutting threshold for the APA test. Moreover, some of these approaches provided an opportunity to establish preliminary traffic-based rutting thresholds.

Approach I

In this approach, a performance threshold for each test considered (i.e., IDT-HT, RR, and APA) was determined for each mixture category, A and D, using the results from the 16 mixtures tested in this study. The mixture designations were identified based on the J_{nr} parameter and not on their designations in the job-mix formula submittals. As indicated previously, three of the 17 mixtures met the VDOT specification criterion for an A designation. The remainder of the mixtures met or exceeded the VDOT specification criterion for a D designation. Two mixtures (Mixtures G and O) did not meet the VDOT specification criterion and thus were excluded from this analysis.

With Approach I, the minimum strength and RT index values were determined from the mixtures for each mixture category, and then these values were selected as the rutting thresholds. Similarly, the maximum APA rut depth from the mixtures was determined for each mixture category, and then the values were selected as the rutting thresholds. This approach assumes that these plant-produced mixtures will not have a notable rutting problem in the field. It must be noted that Approach I was used to establish and verify the existing APA criterion (maximum rut depth of 8 mm) established by the research team in previous studies; however, that effort was undertaken without differentiating the mixture categories because the majority of the mixtures included in the studies were D mixtures (Bowers et al., 2022; Diefenderfer et al., 2021b). This was also the case for this study; nevertheless, the results of this analysis are shown in Table 9, where it can be seen that the APA rut depth criteria do not conform to practical expectations of a higher rutting criterion for the lower traffic designation. This observation from the APA test was considered in the development of preliminary criteria for the mixtures.

Table 9. Performance Criteria Based on Approach I

Test	Parameter	Mix Designation	Performance Criterion
IDT-HT	Strength	A	≥105 kPa
		D	≥170 kPa
RR	RT Index	A	≥59
		D	≥82
APA	Rut Depth	A	≤5.6 mm
		D	≤7.4 mm

IDT-HT = indirect tensile at high temperature; RR = rapid rutting; RT = rutting tolerance; APA = asphalt pavement analyzer.

Approach II

As indicated previously, AASHTO M 332 specifies a maximum J_{nr} limit at 3.2 kPa for different traffic levels. The J_{nr} requirements for standard (S), heavy (H), very heavy (V), and extremely heavy (E) traffic are 4.5, 2.0, 1.0, and 0.5 kPa^{-1}, respectively. These proposed specification limits from AASHTO M 332 along with the correlations established earlier between the J_{nr} and the test parameters provided an opportunity to establish a performance threshold for each test based on the proposed binder specifications or evaluate the recommended or established performance threshold (in the case of the APA test). This exercise was carried out with respect to the VDOT specification requirements for A and D surface mixtures, which call for a minimum of PG64S-16 and PG64H-16 asphalt binders, respectively.

The correlations between the tests and the J_{nr} parameter shown in Figure 4 were used to establish the performance thresholds. The improved equations (i.e., the equations resulting from removing one data point) obtained for the IDT-HT test and the RR tests as shown in Figure 4 were used for this analysis. Table 10 presents the established thresholds for each test.

Table 10. Performance Criteria Based on Approach II

Test	Parameter	Mix Designation	Performance Criterion
IDT-HT	Strength	A	≥110 kPa
		D	≥150 kPa
RR	RT Index	A	≥59
		D	≥78
APA	Rut Depth	A	≤11.5 mm
		D	≤6.1 mm

IDT-HT = indirect tensile at high temperature; RR = rapid rutting; RT = rutting tolerance; APA = asphalt pavement analyzer.

Approach III

VDOT's provisional BMD specification requires that A and D surface mixtures have a maximum APA rut depth of 8 mm at 64°C and 8,000 loading cycles. By use of the rutting threshold of 8 mm along with the correlations in Figure 6, rutting thresholds were developed for the IDT-HT and the RR tests. The minimum threshold for the IDT-HT test is a strength of 133 kPa, and the minimum threshold for the RR test is an RT index of 72. These thresholds are most likely appropriate for D mixtures because the 8 mm APA threshold was established primarily from the evaluation of D mixtures.

Approach IV

Approach IV used the results of pavement design simulations using Pavement ME Design. The rutting performance predictions of six mixtures (Mixtures G, K, L, O, P, and Q) were evaluated with respect to the IDT-HT, RR and APA test results. As indicated previously, two analyses were performed: one with project level calibration coefficients, and the other with material level calibration coefficients. The project level analysis included mixture dynamic modulus and binder shear modulus values as input, whereas the material level analysis included the project level data and RLPD data as input. Prior to the analyses, the impact of asphalt layer thickness and truck traffic volume on the rutting potential of asphalt mixtures were investigated.

The impact of surface asphalt layer thickness on rutting performance prediction was evaluated by varying the surface layer thickness of simulations while keeping all other structural characteristics the same. Four surface layer thicknesses, 30 mm, 38 mm, 50 mm, and 64 mm, were selected based on the nominal maximum aggregate sizes of the mixtures. The results, not provided in this report for the sake of brevity, indicated that the surface layer thickness has a negligible effect on the predicted deformation regardless of whether project level or material level calibration coefficients were used.

The impact of truck traffic volume on rutting performance prediction was evaluated by varying the traffic volume. The following three traffic volume categories were considered: Category I with a maximum of 299 AADTT, Category II with a maximum of 999 AADTT, and Category III with a maximum of 699 AADTT. The first two categories correspond to the traffic volumes for A and D mixtures as per VDOT specifications, and the third was selected as a mid-value for the analysis. As expected, the increase in truck traffic volume resulted in an increased accumulated deformation. It was also observed that the effect of truck traffic volume on rutting appears less significant when material level calibration is used.

Based on these observations, the project level and material level analyses were performed on a selected layer thickness of 38 mm and a traffic level of 999 AADTT. This combination is the most conservative combination to consider for analysis because the surface layer thickness is the thinnest allowable and the traffic volume is the highest allowable for the mixture types used in this study.

The IDT-HT test results versus rutting predicted by Pavement ME Design for project and material level calibrations are shown in Figure 11. The predicted rut depths in this figure (and for Figures 12 and 13) are based on the rut depth of the entire pavement structure. It is evident from the figure that the predicted rut depths for the analysis level used were similar for the mixtures with the exception of a single mixture (Mixture G) for the material level analysis. The rut depth for Mixture G was much higher than 10 mm but was set to that number to provide clarity for other data points in the figure. As is apparent from the figure, Mixture G was not predicted to have a rutting problem at the project level; however, it failed in the material level analysis. This highlighted the differences among the project and material levels of rutting prediction. For this particular case, the observed rutting depth for Mixture G in the material level analysis was mostly influenced by the RLPD results. As indicated earlier, Mixture G failed drastically under the RLPD test and the rutting level observed for Mixture G was significantly different than the rutting levels observed from the other five mixtures evaluated. On the other hand, although Mixture G was one of the least rutting-resistant mixtures per the |E*| test, it was not that different from the other mixtures of the same category when the variation in test results was considered. Since the |E*| test is the performance test used for the project level analysis in Pavement ME Design, the predicted rutting depths were very similar for the mixtures in the project level analysis. Mixture G did not meet the volumetric design criteria as it contained too much asphalt, resulting in excessively low voids and high VFA, and drastically failed under accelerated full-scale testing.

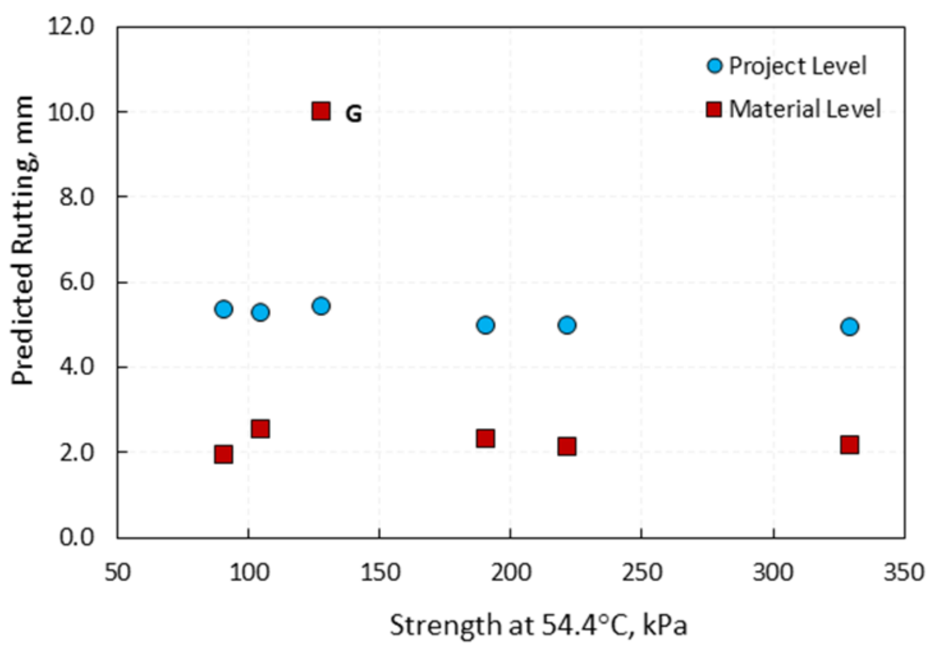


Figure 11. Predicted Rutting and IDT-HT Test Results. The rut depth for Mixture G was set to 10 mm to provide clarity for other data points. IDT-HT = indirect tensile at high temperature.

These observations indicated that caution should be exercised when implementing the tests for mixtures beyond the volumetric criteria, as this is one of the approaches in BMD design. Mixture G was excluded from the threshold analysis due to the contradiction among the test results.

The IDT-HT test was better able to differentiate the performance of the mixtures compared to the predicted rut depths. This was also the case for other tests considered in this study including the fundamental rutting tests (i.e., dynamic modulus and SSR tests). Obviously, such an observation is expected as the analysis involved in Pavement ME Design is non-linear and the combined effects of input parameters from performance tests and volumetric properties are dictating the results, as opposed to the performance tests alone in which a single magnitude of the test output is used to quantify the rutting potential. This indicates that the performance tests during the design and production of asphalt mixtures might be used as "go or no-go" tests unless an effort is made to establish performance prediction models for these tests. It is also apparent from the figure that the results indicate higher accumulated deformations for project level calibration when compared to the accumulated deformations measured when material level calibration coefficients are used.

Nevertheless, a rutting threshold for the IDT-HT test was developed with respect to the rut depths obtained from Pavement ME Design considering the material level calibration data only. Inferential statistics were adopted to explore the results obtained from each test. Confidence interval analyses with a 5% error and the COV for each test were used. The predicted rutting results indicated no failure for the mixtures per the specified maximum total pavement deformation of 6.6 mm. Since no failure was observed, the confidence interval analysis was performed on the mixture with the highest predicted rut value, which corresponds to a strength of 104 kPa with an average COV of 13.7%. From the analysis, the strength ranged from 88 kPa to 120 kPa. The lower level of the interval (88 kPa) was excluded from the analysis as no mixture was tested at that strength. A strength value of 104 kPa should be an appropriate criterion for mixtures with an A designation, whereas for mixtures with a D designation, a conservative approach was taken and a strength value of 120 kPa was deemed appropriate as a threshold criterion for the IDT-HT test.

The observations noted for the IDT-HT tests were also seen for the RR and APA tests, as can be inferred from Figures 12 and 13. The performance criteria for the RR and APA tests were established following the framework used for the IDT-HT test. From Figure 12, the mixture with an RT index of 50 and an average COV of 12.4% for the RT index in this study was used to calculate the confidence interval for the purpose. The RT index range for this mixture was 43 to 57. The lower bound (RT index = 43) was excluded from the analysis, and an RT index of 50 was established as a threshold criterion for mixtures with an A designation. The upper bound (RT index = 57) was established as a criterion for mixtures with a D designation. From Figure 13, one mixture failed with respect to the 8 mm APA threshold. However, the predicted rutting results indicated no failure. Thus, the mixture with the highest APA rut depth of 9.8 mm was used for development of the performance threshold as it performed satisfactorily according to the ME analysis.

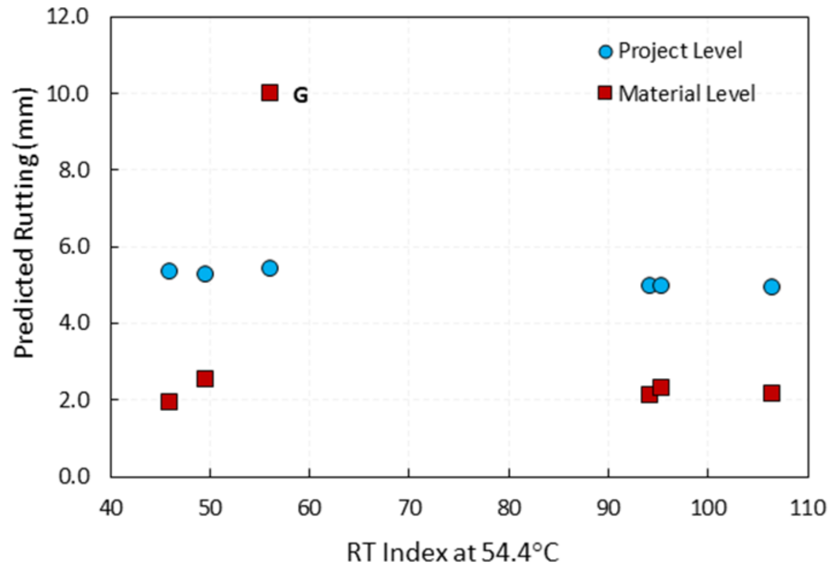


Figure 12. Predicted Rutting and RR Test Results. The rut depth for Mixture G was set to 10 mm to provide clarity for other data points. RT = rutting tolerance.

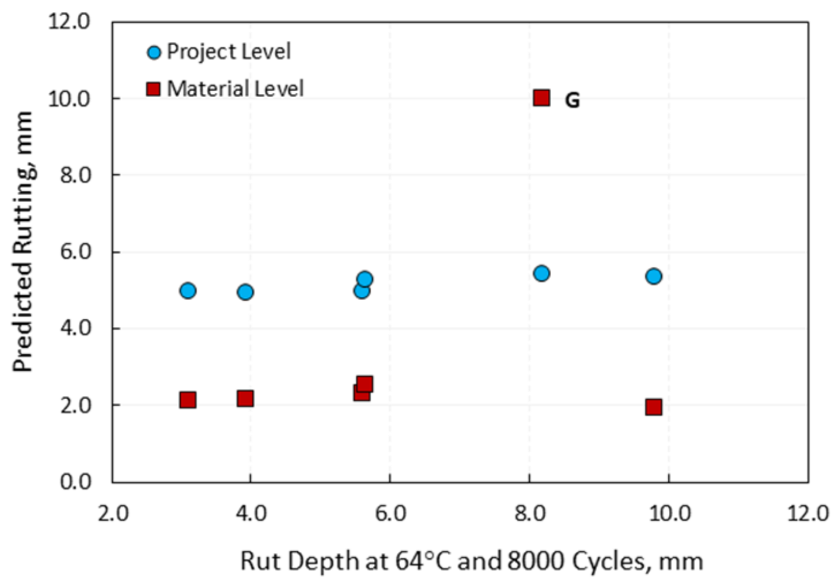


Figure 13. Predicted Rutting and APA Test Results. The rut depth for Mixture G was set to 10 mm to provide clarity for other data points. APA = asphalt pavement analyzer.

With an average COV of 16.1% for the APA test in this study, the confidence interval for this mixture was 8.3 mm to 11.3 mm. The upper value of 11.3 mm was excluded from further consideration as no ME analyses were performed on any mixtures that had rut depths greater than 9.8 mm. Thus, the APA rut depths of 9.8 mm and 8.3 mm were found to be appropriate thresholds for A and D mixtures, respectively. The summary of the performance criteria established using Approach IV is presented in Table 11.

Table 11. Performance Criteria Based on Approach IV

Test	Parameter	Mix Designation	Performance Criterion
IDT-HT	Strength	A	≥104 kPa
		D	≥120 kPa
RR	RT Index	A	≥50
		D	≥57
APA	Rut Depth	A	≤9.8 mm
		D	≤8.3 mm

IDT-HT = indirect tensile test at high temperature; RR = rapid rutting; RT = rutting tolerance; APA = asphalt pavement analyzer.

Table 12 shows the summary of the performance threshold criteria determined from the four approaches for each mixture type per each test.

Table 12. Summary of the Performance Criteria From the Different Approaches

Approach	Test	Parameter	Mix Designation	Performance Criterion
I	IDT-HT	Strength	A	≥105 kPa
			D	≥170 kPa
II			A	≥110 kPa
			D	≥150 kPa
III			A	N/A
			D	≥133 kPa
IV			A	≥104 kPa
			D	≥120 kPa
I	RR	RT Index	A	≥59
			D	≥82
II			A	≥59
			D	≥78
III			A	N/A
			D	≥72
IV			A	≥50
			D	≥57
I	APA	Rut Depth	A	≤5.6 mm
			D	≤7.4 mm
II			A	≤11.5 mm
			D	≤6.1 mm
III			A	N/A
			D	N/A
IV			A	≤9.8 mm
			D	≤8.3 mm

N/A = not applicable. IDT-HT = indirect tensile at high temperature; RR = rapid rutting; RT = rutting tolerance; APA = asphalt pavement analyzer.

Selection of Performance Criteria for the Tests

The VDOT BMD specification currently requires the same threshold criteria for its dense-graded A and D mixtures regardless of the modes of distress considered. This results in overdesigned or very conservatively designed A mixtures, especially from the rutting performance perspective. Based on the previous analysis, a preliminary suggestion for a performance criterion for each rutting test considered was developed for A and D mixtures separately.

Performance Thresholds for A Mixtures

The number of A mixtures evaluated in this study was limited, and thus the developed thresholds should be considered preliminary and to require further evaluation and validation. This concern especially manifests itself for the APA performance criterion for A mixtures established using Approach I where it was lower than the criterion for D mixtures. Thus, Approaches II and IV were evaluated comparatively to establish the thresholds for each test. Approach III was not applicable for A mixtures because there were no data.

As can be seen from Table 12, for the IDT-HT test, the strength values of 110 kPa and 104 kPa were obtained as performance criteria for A mixtures from Approach II and Approach IV analyses, respectively. Both values are within the typical variability observed in this study for this test and within the confidence interval of the relationship established for performance criteria development as part of Approach II. To be conservative, a value of 110 kPa was selected as a suggested preliminary performance criterion for A mixtures from the IDT-HT test.

For the RR test, the RT index values of 59 and 50 were obtained as performance criteria for A mixtures from Approach II and Approach IV analyses, respectively. Similarly, these values were within the test variability and the confidence interval of the Approach II criterion. Thus, an RT index of 59 was selected as a suggested preliminary performance criterion for A mixtures for this test.

The analyses resulted in having APA rut depths of 11.3 mm and 9.8 mm for A mixtures using Approach II and Approach IV, respectively. The previous variability and confidence interval discussions are also applicable for the APA test. The rut depth of 9.8 mm was selected as a suggested preliminary performance criterion for A mixtures; however, this number was rounded to 9.5 mm.

Performance Thresholds for D Mixtures

As shown in Table 12, the criteria for the IDT-HT and RR tests from Approach I for D mixtures are always higher than the criteria for other approaches in terms of rutting potential. This indicates that the mixtures evaluated in this study were not approaching a point where they might have a rutting issue, particularly compared to the criteria for the other approaches. In other words, the mixtures may be overly stiff when designed using the criteria from Approach I. The performance criteria from Approach III for both tests fell between the criteria from Approaches II and IV and were statistically similar to both ends individually when the variability of the tests and the confidence intervals for the established performance criteria for each test from each approach were considered. Approach II values may be too high to consider as the criteria because there were mixtures produced with lower values and no notable rutting issues have been observed so far for these mixtures. Approach IV values were developed based on Pavement ME Design simulations and represent the low end of the test results in this study. The values from Approach III provide balance between the two other approaches. Thus, the strength of 133 kPa from the IDT-HT test and the RT index of 72 from the RR test were selected as minimum performance values from Approach III for D mixtures. The APA criterion for D mixtures from Approach III was 6.1 mm; however, the correlation coefficient for the relationship between the J_{nr} parameter and the APA test results used to develop the criterion was relatively low. The APA performance criteria for the same mixtures were 7.4 mm and 8.3 mm based on Approach I and Approach IV, respectively. These values were within the variability of the existing APA threshold of 8 mm. Thus, it is suggested that the APA performance criterion for D mixtures remain at 8 mm rut depth, considering the results of the initial performance development and verification studies. Table 13 presents the suggested preliminary performance thresholds for A and D mixtures from each test.

Table 13. Summary of Preliminary Mixture Type Performance Criteria

Test	Parameter	Mix Designation	Performance Criterion
IDT-HT	Strength	A	≥110 kPa
		D	≥133 kPa
RR	RT Index	A	≥59
		D	≥72
APA	Rut Depth	A	≤9.5 mm
		D	≤8 mm

IDT-HT = indirect tensile at high temperature; RR = rapid rutting; RT = rutting tolerance; APA = asphalt pavement analyzer.

CONCLUSIONS

- *Based on the mixtures tested in this study and applicable only for dense-graded asphalt surface mixtures with A and D designations:*

  — *A test temperature of 54.4°C is appropriate to evaluate the rutting potential of asphalt mixtures using the monotonic tests considered in this study*.

  — *The comparative analysis of the test results between the monotonic tests and APA test indicates that the IDT-HT and RR tests provide either similar or better performance evaluation characteristics of asphalt mixtures than the APA test.* The IDT-HT and RR tests have similar repeatability characteristics and provide a better measure of repeatability compared to the APA and MS tests. The IDT-HT and APA tests have a similar performance discrimination potential, followed by the RR test. The MS test is the least sensitive test among the tests considered in this study. The IDT-HT, RR, and APA tests provide a similar performance ranking for the rutting potential of asphalt mixtures. The IDT-HT and APA tests are the tests most sensitive to changes in asphalt mixture composition.

  — *The correlations between tests indicated that the overall rutting potential of asphalt mixtures identified by the IDT-HT, RR, and APA tests is in agreement with that of fundamental rutting tests.*

  — *Caution should be exercised when evaluating, through the tests considered in this study including the fundamental rutting tests, the rutting potential of asphalt mixtures that do not meet the requirements for certain volumetric properties (e.g., VMA and VFA).*

— *Although the performance tests including fundamental rutting tests provide a range of rutting potential, the rutting predictions through mechanistic-based simulations are similar regardless of whether the project level or material level analysis is used.* Performance predictions determined using ME simulations at the project level provide lower magnitudes of rutting levels compared to the rutting levels predicted at the material level.

— *The IDT-HT test is the most suitable alternative test to the APA test for use in Virginia's BMD framework.* Preliminary performance criteria requiring minimum strength values of 110 kPa and 133 kPa are deemed suitable for asphalt mixtures with A and D designations, respectively.

— *The RR test is also a suitable alternative test to the APA test for use in BMD process.* Preliminary performance criteria requiring minimum RT index values of 59 and 72 are deemed suitable for asphalt mixtures with A and D designations, respectively.

— *The current performance criterion limiting APA rut depth to a maximum of 8.0 mm is reasonable for D mixtures.* A preliminary performance criterion requiring a maximum APA rut depth of 9.5 mm is deemed suitable for asphalt mixtures with an A designation.

RECOMMENDATIONS

1. *VDOT's Materials Division should adopt the IDT-HT test for BMD implementation with an initial minimum strength criterion of 133 kPa for A and D mixtures based on the testing conditions used in this study.* Further assessment of the relationships and correlations among the results of the IDT-HT, RR and APA tests, the results of fundamental rutting tests, ME simulations and analyses, and field performance is needed to ensure that the most appropriate test(s) and corresponding threshold criteria are considered as part of BMD implementation. Adjustments to the implemented tests and/or the initial criteria may be necessary as additional data from a more diverse range of mixtures become available.

2. *VTRC should conduct a study to fine-tune the testing protocols and determine the acceptable variability for the IDT-HT, RR, and APA tests.*

3. *VTRC should further explore the development of traffic-based performance criteria for A and D mixtures.* This should include evaluating and benchmarking low volume traffic mixtures and using ME evaluation methods to assess the necessary performance properties of an A mixture.

4. *VTRC and VDOT's Materials Division should continue to organize hands-on training and demonstrations of the IDT-HT and RR tests being considered by VDOT as part of the BMD implementation.*

IMPLEMENTATION AND BENEFITS

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

Implementation

With regard to Recommendation 1, VDOT's Materials Division has already included the IDT-HT test and associated criterion into the 2023 BMD specification for mix design. VTRC is working closely with VDOT's Materials Division to include the IDT-HT test and its performance criterion into the 2024 BMD specification for both mix design and production. Moreover, the ongoing VTRC Project 120747, Mechanistic-Based Evaluation of Performance Thresholds for Engineered Surface Asphalt Mixtures, is evaluating the relationship and correlations between the performance test results and ME analysis simulations using a diverse range of asphalt mixtures. However, a research needs statement on evaluating the rutting tests with respect to the field performance will be drafted and submitted to the respective VTRC Pavement Research Advisory Subcommittee by no later than Fiscal Year 2024.

With regard to Recommendation 2, VTRC Project 121388, Inter-Laboratory Study for the Indirect Tensile Test at High Temperature and Rapid Rutting Test, is ongoing. One of many objectives of this project is to address the impact on the test results from several factors such as device type, loading rate, conditioning environment, reheating, and storage life. Another objective is to develop precision estimates and statements for the IDT-HT and RR tests.

With regard to Recommendation 3, VTRC Project No. 120747, Mechanistic-Based Evaluation of Performance Thresholds for Engineered Surface Asphalt Mixtures, is ongoing. The overarching objective of this effort is to evaluate the feasibility of existing performance cracking and rutting thresholds with a major focus on D mixtures. Moreover, VTRC will draft a research needs statement to evaluate the same concept for A mixtures and will submit it to the appropriate VTRC Pavement Research Advisory Subcommittee by no later than Fiscal Year 2024.

With regard to Recommendation 4, VTRC and VDOT's Materials Division, with the help of the Virginia Asphalt Association and Germanna Community College through the Virginia Education Center for Asphalt Technology (VECAT), organized and hosted five in-person BMD training classes in 2022. In addition to the other topics related to VDOT's BMD initiative, the first half of the curriculum covered course work and presentations (study guides) on BMD tests and associated performance thresholds, which also included the IDT-HT and RR tests. The second half of the curriculum included hands-on training to perform BMD tests, including IDT-HT and RR tests. In addition, a proficiency testing program for both the IDT-HT and RR tests is planned as part of VTRC Project 121388, Inter-Laboratory Study for the Indirect Tensile Test at High Temperature and Rapid Rutting Test. Moreover, the IDT-HT test will be covered as part of the Virginia Education Center for Asphalt Technology Asphalt Mix Design and Plant Technician Training Classes in 2023.

Benefits

Performance tests are central to the BMD method, and the use of simple, practical, and performance-indicating tests is vital for the success of BMD implementation. This study showed the viability of using simple and practical test methods to evaluate the rutting performance of asphalt mixtures as part of VDOT's BMD initiative. The proposed test methods make use of a single equipment/loading frame that is already available in the majority of asphalt laboratories to evaluate several performance aspects of asphalt mixtures, requiring minimal equipment expenditures for VDOT and contractors. Asphalt mixture specimens for the proposed tests are compacted to the same dimensions and air void content as the specified BMD cracking test and the testing protocol is very similar to that of the cracking test, thereby requiring minimal training needs and providing convenience during the specimen preparation process. Further, the proposed tests have a quick turn-around time that is much needed for timely decisions, especially during mixture production.

ACKNOWLEDGMENTS

The authors are grateful to the following individuals who served on the technical review panel for this study: Robert Crandol (Project Champion and Assistant State Maintenance Engineer, VDOT Maintenance Division); Sungho Kim (formerly of VDOT); Jason Provines (Senior Research Scientist, VTRC); Todd Rorrer (Asphalt Pavement Field Engineer, VDOT Materials Division); David Shields (District Materials Engineer, VDOT Northern Virginia District); and Bryan Smith (Asphalt Pavement Field Engineer, VDOT Materials Division). The authors thank Troy Deeds, Danny Martinez, and Jennifer Samuels of VTRC for their outstanding efforts in sample collection and testing. Appreciation is also extended to Linda Evans of VTRC for her editorial support.

REFERENCES

Apeagyei, A.K., Diefenderfer, B.K., and Diefenderfer, S.D. Rutting Resistance of Asphalt Concrete Mixtures That Contain Recycled Asphalt Pavement. *Transportation Research Record: Journal of Transportation Research Board*, 2208(1), 2011, pp. 9-16.

Bennert, T., Haas, E., and Wass, E. Indirect Tensile Test (IDT) to Determine Asphalt Mixture Performance Indicators During Quality Control Testing in New Jersey. *Transportation Research Record: Journal of the Transportation Research Board*, 2672(28), 2018, pp. 394-403.

Bennert, T., Haas, E., Wass, E., and Berger, B. Indirect Tensile Testing for Balanced Mixture Design and Quality Control Performance Testing. *Journal of the Association of Asphalt Paving Technologists*, Vol. 89, 2021, pp. 363-389.

Bowers, B.F., Diefenderfer, S.D., Moore, N., and Lynn, T. Balanced Mix Design and Benchmarking: A Case Study in Establishing Performance Test Thresholds. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(12), 2022.

Boz, I., Tavassoti-Kheiry, P., and Solaimanian, M. The Advantages of Using Impact Resonance Test in Dynamic Modulus Master Curve Construction Through the Abbreviated Test Protocol. *Materials and Structures*, Vol. 50, No. 3, 2017, pp. 1-12.

Boz, I., Habbouche, J., and Diefenderfer, S.D. The Use of Indirect Tensile Test to Evaluate the Resistance of Asphalt Mixtures to Cracking and Moisture-Induced Damage. In *Proceedings of the International Airfield and Highway Pavements Conference*. American Society of Civil Engineers, Reston, VA, 2021a, pp. 104-114.

Boz, I., Habbouche, J., Diefenderfer, S., and Bilgic, Y. Precision Estimates and Statements for Performance Indices From the Indirect Tensile Cracking Test at Intermediate Temperature. *Transportation Research Record: Journal of Transportation Research Board*, 2676(5), 2021b, pp. 225-241.

Boz, I., Coffey, G.P., Habbouche, J., Diefenderfer, S.D., and Ozbulut, O.E. A Critical Review of Monotonic Loading Tests to Evaluate Rutting Potential of Asphalt Mixtures. *Construction and Building Materials*, Vol. 335, 127484, 2022.

Brown, E.R., Kandhal, P.S., and Zhang, J. *Performance Testing for Hot Mix Asphalt*. NCAT Report -05. National Center for Asphalt Technology, Auburn, AL, 2001.

Christensen, D., and Bonaquist, R. Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing. In *Transportation Research Circular E-C124*: *Using the Indirect Tensile Test to Evaluate Rut Resistance in Developing Hot-Mix Asphalt Designs.* Transportation Research Board, Washington, DC, 2007, pp. 62-77.

Christensen, D.W., Bonaquist, R.F., and Jack, D.P. *Evaluation of Triaxial Strength as a Simple Test for Asphalt Concrete Rut Resistance*. FHWA-PA-2000-010-97-04. Pennsylvania Department of Transportation, University Park, 2000.

Christensen, D.W., Bonaquist, R.F., Anderson, D.A., and Gokhale, S. Performance Tests for Asphalt Mixes. In *Transportation Research Circular E-C068: Indirect Tension Strength as a Simple Performance Test.* Transportation Research Board, Washington, DC, 2004, pp. 44-57.

Diefenderfer, S.D. Evaluation of the MEPDG Using Asphalt Material Inputs Obtained From Plant Mix. *Journal of the Association of Asphalt Paving Technologists*, Vol. 80, 2011, pp. 101-134.

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

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, 2021b.

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, 2021a.

Habbouche, J., Boz, I., and Diefenderfer, S.D. Validation of Performance-Based Specifications for Surface Asphalt Mixtures in Virginia. *Transportation Research Record: Journal of the Transportation Research Board*, 2676(5), 2021b, pp. 277-296.

Hand, A.J.T., Boz, I., Piratheepan, M., Hierholzer, F., Diefenderfer, B.K., Jones, D., Habbouche, J., and Louw, S. Ruggedness Evaluation and Precision Estimates for Newly Developed Test Methods for Asphalt-Treated Cold Recycled Pavements. *Transportation Research Record: Journal of Transportation Research Board*, 2675(11), 2021, pp. 348-358.

Kandhal, P.S., and Koehler, W.S. Marshall Mix Design Practice: Current Practices. *Journal of the Association of Asphalt Paving Technologists*, Vol. 54, 1985, pp. 284-314.

Mallela, J., Glover, L.T., Darter, M.I., Von Quintus, H., Gotlif, A., Stanley, M., and Sadasivam, S. *Guidelines for Implementing NCHRP 1-37A ME Design Procedures in Ohio: Volume 1: Summary of Findings, Implementation Plan, and Next Steps*. FHWA/OH-2009/9A. Applied Research Associates, Inc., Champaign, IL, 2009.

Meroni, F., Flintsch, G.W., Habbouche, J., Diefenderfer, B.K., and Giustozzi, F. Three-Level Performance Evaluation of High RAP Asphalt Surface Mixes. *Construction and Building Materials*, Vol. 309, 125164, 2021.

Moulthrop, J., and Witczak, M. *NCHRP Report 704: A Performance-Related Specification for Hot-Mix Asphalt*. Transportation Research Board, Washington, DC, 2011.

Roberts, F.L., Mohammad, L.N., and Wang, L.B. History of Hot Mix Asphalt Mixture Design in the United States. *Journal of Materials in Civil Engineering*, Vol. 14, No. 4, 2002, pp. 279-293.

Rushing, J.F. *Development of Ndesign Criteria for Using the Superpave Gyratory Compactor to Design Asphalt Pavement Mixtures for Airfields*. ERDC/GSL TR; 09-39. U.S. Army Corps of Engineers, Washington, DC, 2009.

Tran, N., and Hall, K. Evaluating the Predictive Equation in Determining Dynamic Moduli of Typical Asphalt Mixtures Used in Arkansas. *Journal of the Association of Asphalt Paving Technologists*, Vol. 74, 2005, pp. 1-17.

Von Quintus, H.L., Mallela, J., Bonaquist, R., Schwartz, C.W., and Carvalho, R.L. *NCHRP Report 719: Calibration of Rutting Models for Structural and Mix Design*. Transportation Research Board, Washington, DC, 2012.

Wen, H., and Bhusal, S.A. Laboratory Study to Predict the Rutting and Fatigue Behavior of Asphalt Concrete Using the Indirect Tensile Test. *Journal of Testing and Evaluation*, Vol. 41, No. 2, 2013, pp. 299-304.

Yin, F., and West, R.C. *Balanced Mix Design Resource Guide*. No. IS-143. National Asphalt Pavement Association, Greenbelt, MD, 2021.

Yin, F., Taylor, A.J., and Tran, N. *Performance Testing for Quality Control and Acceptance of Balanced Mix Design*. NCAT Report 20-02. National Center for Asphalt Technology, Auburn, AL, 2020.

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

Zaniewski, J.P., and Srinivasan, G. *Evaluation of Indirect Tensile Strength to Identify Asphalt Concrete Rutting Potential*. Asphalt Technology Program, Department of Civil and Environmental Engineering, West Virginia University, 2004.

Zhang, J., Alvarez, A.E., Lee, S.I., Torres, A., and Walubita, L.F. Comparison of Flow Number, Dynamic Modulus, and Repeated Load Tests for Evaluation of HMA Permanent Deformation. *Construction and Building Material*, Vol. 44, 2013, pp. 391-398.

Zhou, F. *NCHRP Project 20-44(16): Implementation of the IDEAL Cracking Test for Asphalt Mix Design and QC/QA*. Transportation Research Board, Washington, DC, 2021.

Zhou, F., Im, S., Sun, L., and Scullion, T. Development of an IDEAL Cracking Test for Asphalt Mix Design and QC/QA. *Road Materials and Pavement Design*, Vol. 18, No. 4, 2017.

Zhou, F., Crockford, B., Zhang, J., Hu, S., Epps, J., and Sun, L. Development and Validation of an Ideal Shear Rutting Test for Asphalt Mix Design and QC/QA. *Journal of the Association of Asphalt Paving Technologists*, 2019.

Zhou, F., Hu, S., and Newcomb, D. Development of a Performance-Related Framework for Production Quality Control With Ideal Cracking and Rutting Tests. *Construction and Building Materials*, Vol. 261, 120549, 2020.

Zhou, F., Steger, R., and Mogawer, W. Development of a Coherent Framework for Balanced Mix Design and Production Quality Control and Quality Acceptance. *Construction and Building Materials*, Vol. 287, 123020, 2021.

APPENDIX A

EVALUATING THE IDT-HT AND RR TESTS WITH FULL-FIELD DISPLACEMENT MEASUREMENTS USING THE DIGITAL IMAGE CORRELATION METHOD

An optical DIC measurement system was employed to obtain the full-field deformation (strain) fields during the IDT-HT and RR tests. DIC is a real-time, full-field, and non-contact optical measurement system that uses a series of sequential images captured during loading to track and correlate patterns within a subset space, which can in turn be used to describe deformation. This task was undertaken to gain better insight into how asphalt mixtures fail under these tests and to determine which test configuration might be the best indicator of the rutting potential of asphalt mixtures. To this extent, DIC measurements were taken on four mixtures (Mixtures K, O, P, and Q) while being subjected to the IDT-HT and RR tests at 54.4°C. For image acquisition and the subsequent application of the DIC method, a surface treatment of the specimen was needed. The surface pattern on the specimen should be non-periodic, isotropic, and of high contrast to guarantee accurate measurements by the DIC system with the lowest possible noise. To prepare the specimens for DIC measurements, a paint roller was used to apply flat exterior ultra-white paint onto the flat side of each specimen. Two layers of paint were necessary to create a completely white surface. Once the white paint was dry, the black speckle pattern was applied to each specimen.

The DIC testing setup consisted of a load frame, a customized test fixture, a high-quality camera and tripod, a lighting fixture, a DIC calibration card, and a computer with Vic-Snap and Vic-2D software from Correlated Solutions. A Point Grey camera and Schneider Kreuznach lens were used for image capture. A tripod was used to align the camera with the specimen, making sure the optical axis of the lens was perpendicular to the specimen face. A lighting fixture was used to illuminate the face of the specimen during testing. Using Vic-Snap, the focus was set properly on the speckle pattern prior to each test. The camera resolution was 2448 x 2048 pixels with a charge-coupled device sensor that has a format of 2/3 in. An image acquisition interval of 100 ms was used for the DIC tests in this study. The entire DIC test setup is shown in Figure A1. The three replicate specimens were conditioned in an environmental chamber at 54.4°C for 2 hours and tested within 2 minutes of removal from the chamber. Vic-2D software was used to analyze the image data after testing and obtain full-field axial and shear strain data throughout the test.

Full-Field Strain Maps

Figure A2 shows the full-field evolution of horizontal strain distribution under the IDT-HT and RR tests on Mixture O. The horizontal strain fields ε_{xx} on the overall surface of the specimen for eight loading stages, including at peak load, are shown in the figure. For the given color legend, red indicates high strain values and purple indicates low strain values. For the IDT-HT test, as the loading progresses, the magnitude of horizontal strains increases over the central area of the specimen where high tensile stress is concentrated.

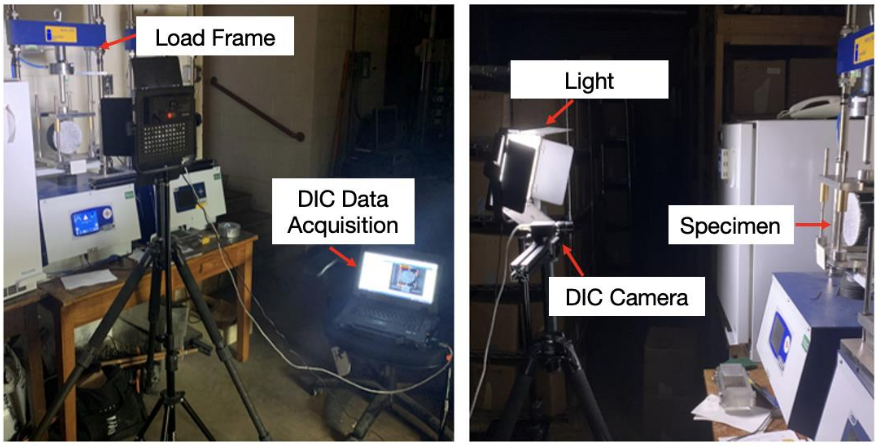
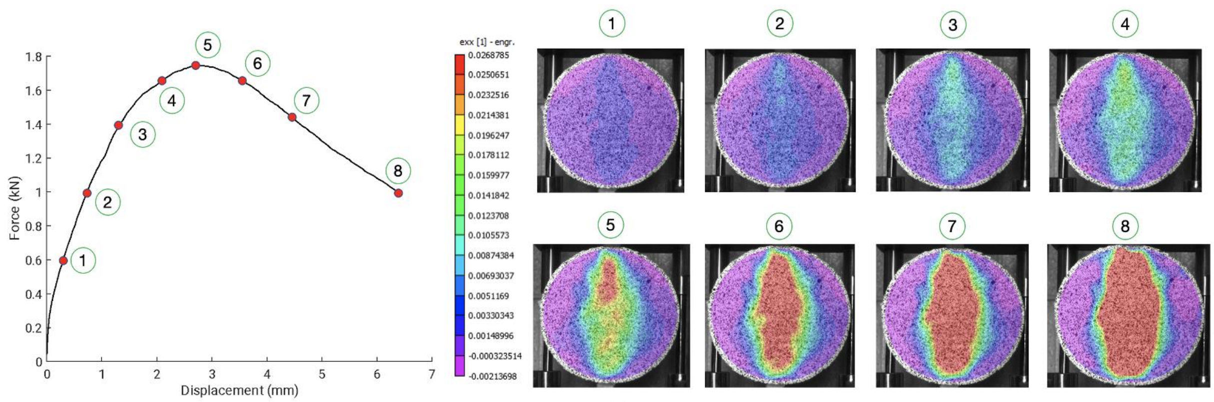
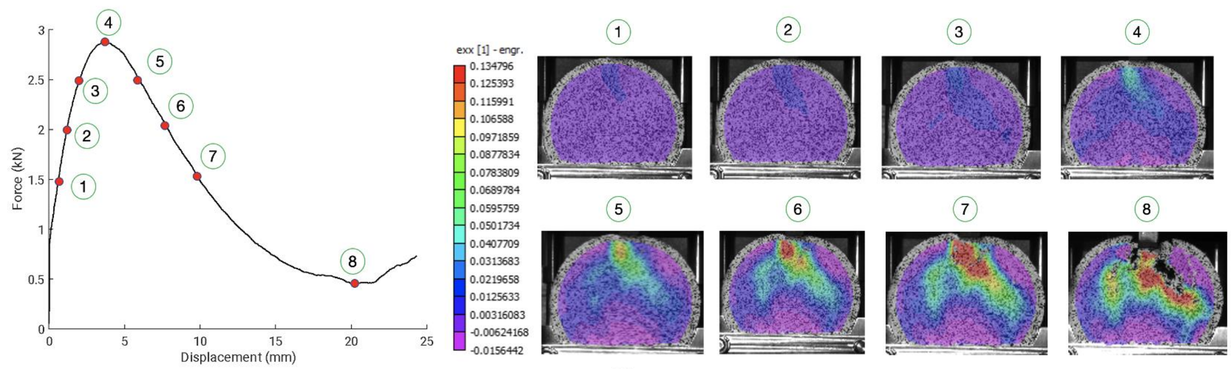


Figure A1. Test Setup for DIC Measurements. DIC = digital image correlation.



(a)

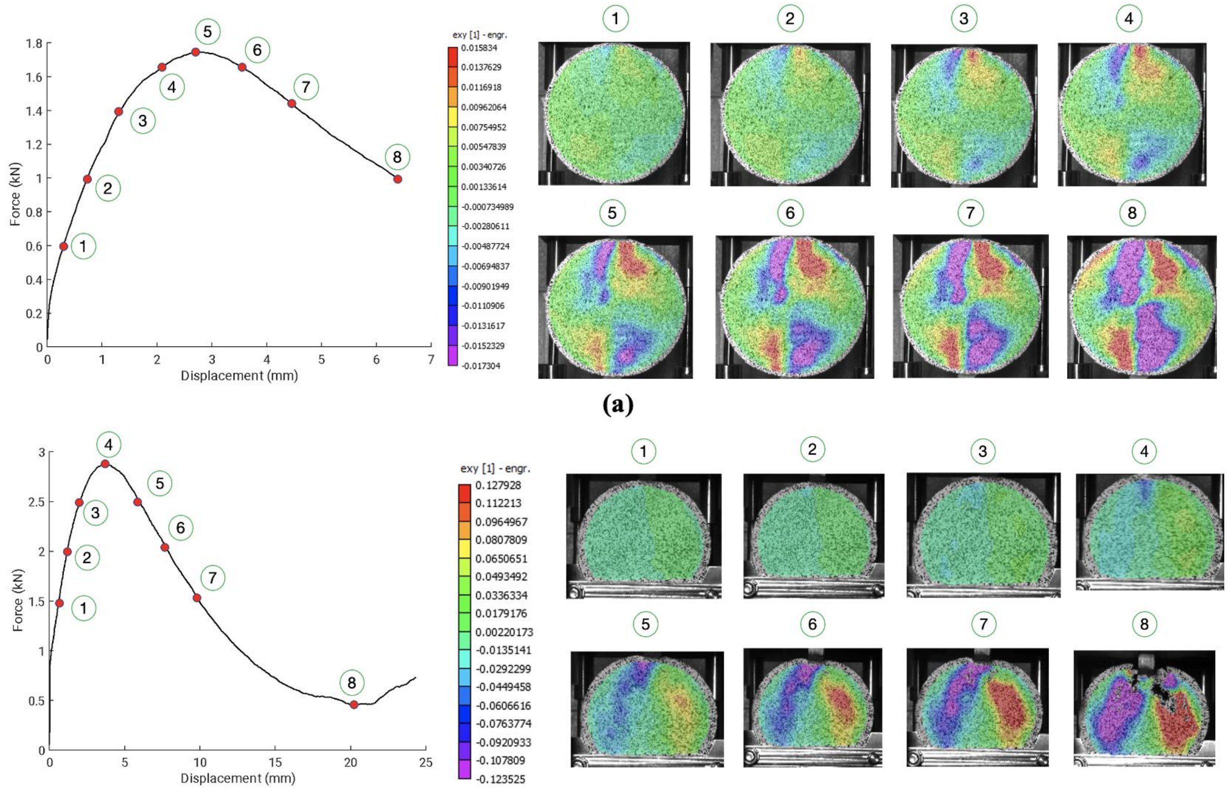


(b)

Figure A2. Evolution of Full-field Horizontal Strain Distribution: (a) IDT-HT test; (b) RR test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting. For the given color legend, red indicates high strain values and purple indicates low strain values.

As shown in Image 5 of Figure A2(a), high horizontal strains are developed under the maximum load at the top central portion of the specimen, and such high strains propagate throughout the complete central region after the maximum. For the RR test, high horizontal strains were developed only in a restricted zone located at the top of the specimen near the load contact point and after the maximum load had been applied.

Figure A3 shows the full-field evolution of shear strain distribution under the IDT-HT and RR tests on the same mixture. The shear strain fields ε_{xy} on the overall surface of the specimen for eight loading stages, including the peak load, are shown. Shear strains attain high positive values (indicated in red) and high negative values (indicated in purple); green indicates low shear strains. For the IDT-HT test, high shear strains occur close to the top and bottom loading points at the central part of the specimen. For the RR test that includes one top loading point but two bottom loading points, the shear strains propagate from top loading point to two bottom supports. The experimental strain maps obtained using the DIC technique agree well with numerical results in the literature (Zhou et al., 2019).



(b)

Figure A3. Evolution of Full-field Shear Strain Distribution: (a) IDT-HT test; (b) RR test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting. For the given color legend, red indicates high strain values and purple indicates low strain values. Green indicates low shear strains.

Peak Strains

The peak values of horizontal strain ε_{xx} and maximum principal strain ε_1 on the surface of the specimen at the maximum load point are obtained from the DIC measurements for each specimen of the mixtures. Figures A4 and A5 show the mean of the peak values of ε_{xx} and ε_1 for each mixture together with the error bars under the IDT-HT and RR tests. For most of the mixtures, the peak values of ε_{xx} and ε_1 at maximum load in the RR test are somewhat higher than those in the IDT-HT test.

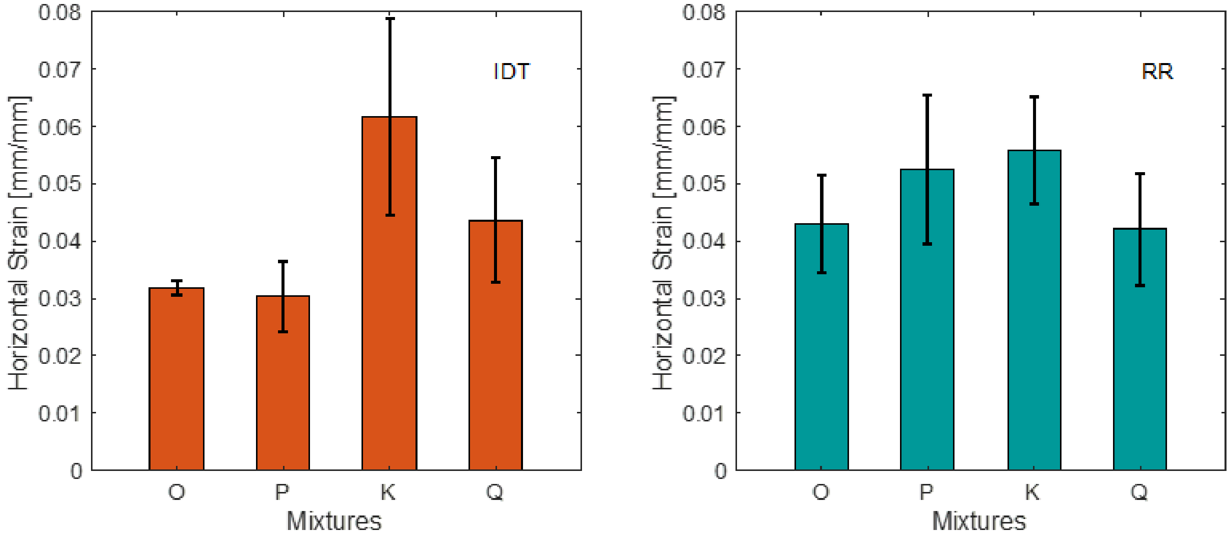


Figure A4. Peak Values of Horizontal Strain ε_{xx} at Maximum Load: (a) IDT-HT test; (b) RR test. ID-HT = indirect tensile at high temperature; RR = rapid rutting.

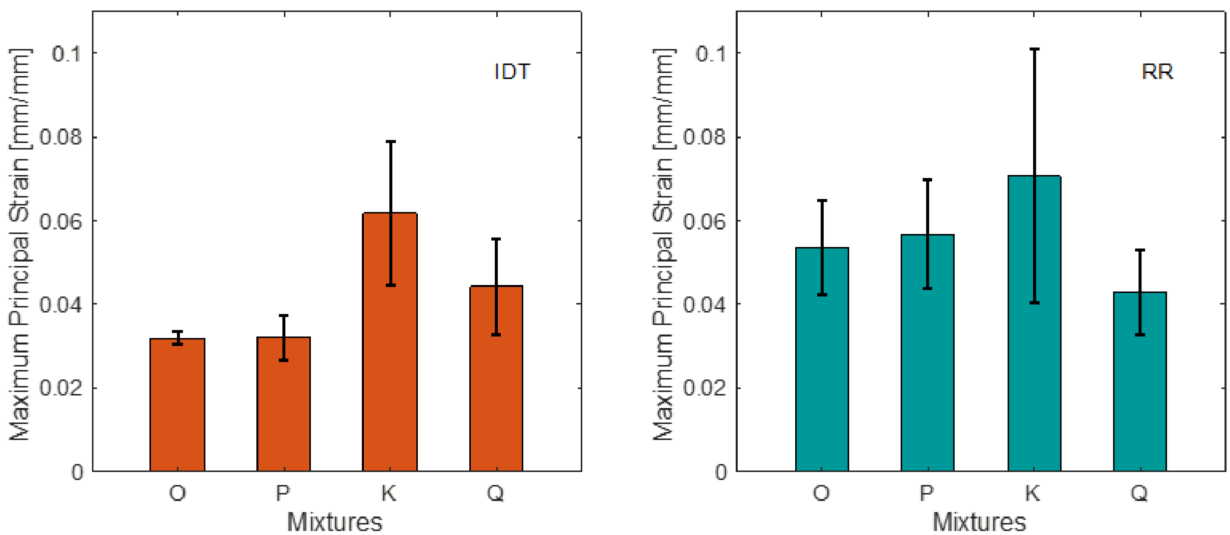


Figure A5. Peak Values of Maximum Principal Strain ε_1 at Maximum Load: (a) IDT-HT test; (b) RR test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting.

Similarly, the peak values of shear strain γ_{xy} and maximum shear strain γ_{max} on the surface of the specimen at the maximum load point are obtained from the DIC measurements for each specimen of the mixtures. Figures A6 and A7 show the mean of the peak values of γ_{xy} and γ_{max} for each mixture together with the error bars under the IDT-HT and RR tests. For Mixtures O and P, the shear strain γ_{xy} attains higher values in the RR test than in the IDT-HT test. For Mixture K, the shear strain γ_{xy} is somewhat higher in the IDT-HT test, and a similar peak value of shear strain γ_{xy} was observed in the IDT-HT and RR tests for Mixture Q. On the other hand, the maximum shear strain γ_{max} attains a peak value of about 1.57 for all the tested specimens for both tests.

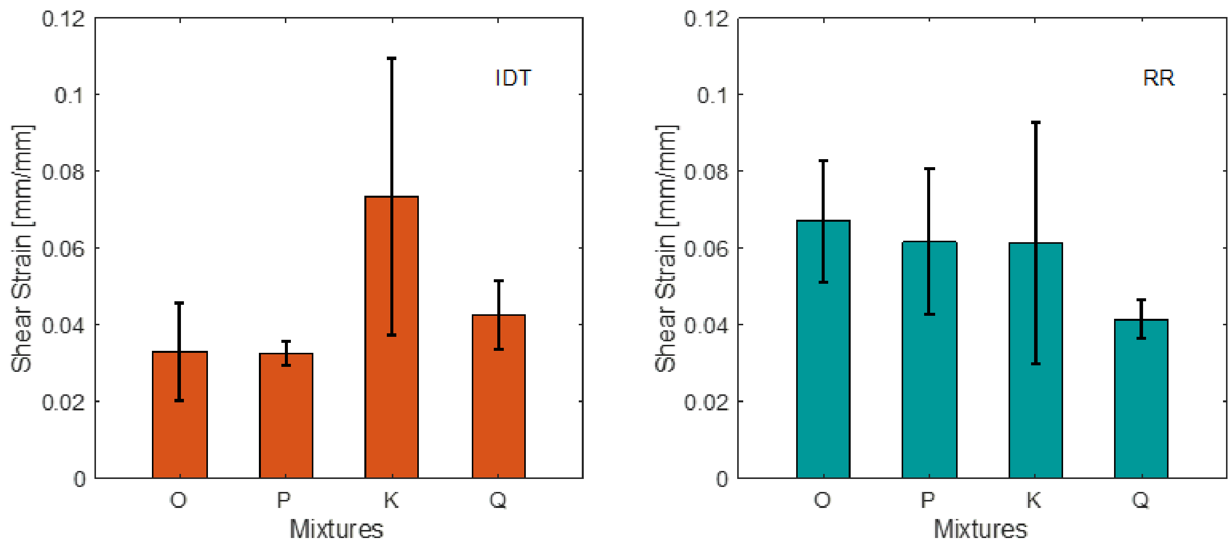


Figure A6. Peak Values of Shear Strain γ_{xy} at Maximum Load: (a) IDT-HT test; (b) RR test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting.

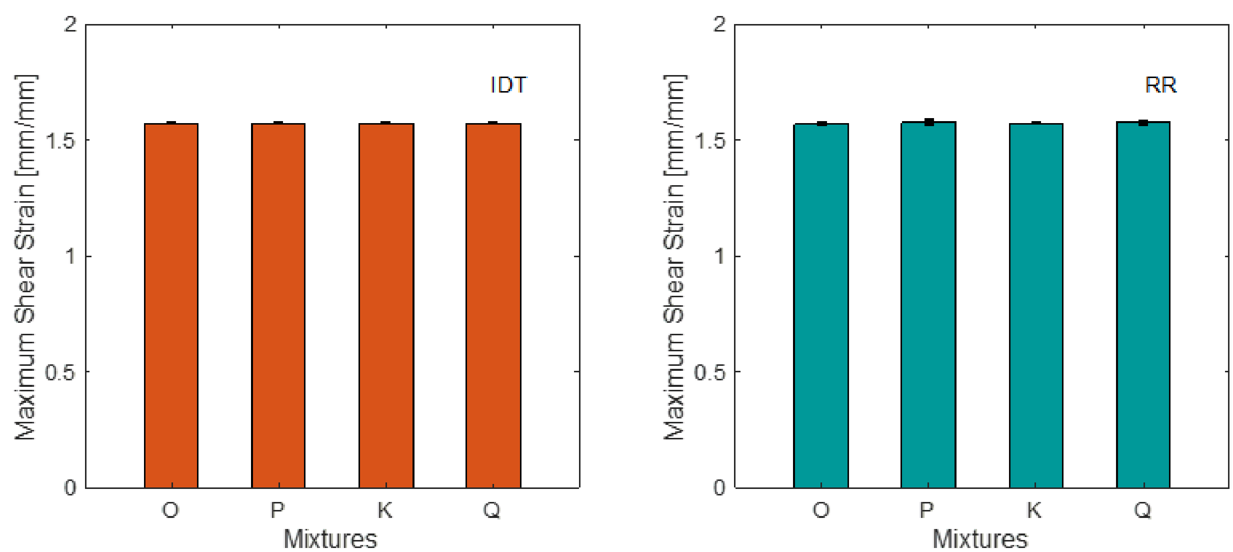


Figure A7. Peak Values of Maximum Shear Strain γ_{max} at Maximum Load: (a) IDT-HT test; (b) RR test. IDT-HT = indirect tensile at high temperature; RR = rapid rutting.

Overall, the results indicate that peak shear strains induced in the specimens during the IDT-HT test were somewhat lower for two of the mixtures and either higher or similar for the other two mixtures compared to those during the RR test. Further, the peak value of the maximum shear strain at the peak load was the same for both tests.

APPENDIX B

INTERMEDIATE AND LOW TEMPERATURE BINDER PROPERTIES

Table B1. Part I: Intermediate and Low Temperature Binder Properties

Property		Mixture ID								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Dynamic Shear, 10 rad/sec, specification: $\|G^*\| \cdot \sin \delta < 5000$ kPa										
PAV $\|G^*\| \cdot \sin \delta$, kPa	19°C	-	-	-	-	-	-	6,436	-	-
	22°C	5,418	-	-	5,989	-	7,133	4,613	5,910	5,840
	25°C	3,861	6,109	5,166	4,337	-	4,959	-	4,250	4,120
	28°C	-	4,500	3,894	-	5,035	-	-	-	-
	31°C	-	-	-	-	3,832	-	-	-	-
PAV Failure Temperature, °C		**22.7**	**27.0**	**25.4**	**23.7**	**28.1**	**24.9**	**24.3**	**23.5**	**23.3**
Creep Stiffness, 60 sec, specification: Stiffness (S) < 300 MPa and m-value > 0.300										
Stiffness (S), MPa	-6ºC	-	126	106	101	126	-	84	-	-
	-12ºC	157	206	214	189	226	217	148	163	170
	-18ºC	340	-	-	-	-	381	-	319	332
	-24ºC	-	-	-	-	-	-	-	-	-
m-value	-6ºC	-	0.318	0.325	0.338	0.301	-	0.347	-	-
	-12ºC	0.313	0.275	0.290	0.297	0.273	0.305	0.284	0.313	0.315
	-18ºC	0.265	-	-	-	-	0.252	-	0.275	0.268
	-24ºC	-	-	-	-	-	-	-	-	-

PAV = pressure aging vessel; - = no data collected.

Table B2. Part II: Intermediate and Low Temperature Binder Properties

Property		Mixture ID							
		J	**K**	**L**	**M**	**N**	**O**	**P**	**Q**
Dynamic Shear, 10 rad/sec, specification: |G*|·sin δ < 5000 kPa									
PAV |G*|·sin δ, kPa	19°C	-	-	-	-	-	6,988	6,744	-
	22°C	-	-	-	-	6,536	4,692	4,601	-
	25°C	5,235	5,492	-	5,158	4,708	-	-	6,720
	28°C	3,969	4,119	-	3,877	-	-	-	4,620
	31°C	-	-	-	-	-	-	-	-
PAV Failure Temperature, °C		**25.5**	**26.0**	**-**	**25.3**	**24.5**	**21.5**	**21.3**	**27.2**
Creep Stiffness, 60 sec, specification: Stiffness (S) < 300 MPa and m-value > 0.300									
Stiffness (S), MPa	-6ºC	96	131	-	96	95	-	-	-
	-12ºC	188	228	-	191	188	-	-	230
	-18ºC	-	-	-	-	-	-	-	434
	-24ºC	-	-	-	-	-	-	-	-
m-value	-6ºC	0.319	0.348	-	0.314	0.327	-	-	-
	-12ºC	0.294	0.294	-	0.284	0.292	-	-	0.304
	-18ºC	-	-	-	-	-	-	-	0.258
	-24ºC	-	-	-	-	-	-	-	-

PAV = pressure aging vessel; - = no data collected.