Performance Evaluation of In-situ and Laboratory Prepared Asphalt Materials: Practical Considerations for Sample Preparation and Testing Methodology

Pejoohan Tavassoti, Ph.D.
Assistant Professor, Jr. Norman W. McLeod Chair in Sustainable Pavement Engineering
Department of Civil and Environmental Engineering, University of Waterloo

Amma Wakefield, Ph.D., P.Eng.
Canadian & Research Engineer
Asphalt Institute

Xiomara Sanchez, Ph.D., P.Eng.
Assistant Professor and D.C. Campbell Chair in Highway Construction and Pavement Research
Department of Civil Engineering, University of New Brunswick

Sina Varamini, Ph.D., P.Eng.
Director, Pavements and Materials Group
Engtec Consulting Inc.

Paper prepared for presentation at the
Soils and Materials Session

2022 Annual Conference and Exhibition of the Transportation Association of Canada (TAC)
Edmonton, AB

Abstract

Research and implementation of Performance Engineered Pavements (PEP) have been gaining momentum, especially during the past two decades. Several factors including but not limited to the sample preparation methods, specimen dimensions, compaction level, performance testing method, variability of results, minimum and maximum thresholds, and sensitivity to materials sources can be named among other major contributors in this area. This paper provides a summary of major procedures and protocols used by national and international transportation agencies toward performance testing. Furthermore, this paper presents a synthesis of the critical steps that should be considered by highway agencies in utilizing performance testing into design, performance verification, and/or quality assurance part of contracts. Successful implementation of this approach requires understanding the rationale behind development of the commonly used methods. Therefore, this paper also provides relevant aspects of the research work behind development of most specified test methods, and how several factors could influence the accuracy and variability of the test results. To this end, factors such as: air voids, thickness variation, cutting and coring methods, and compaction using gyratory or slab compactors, are considered and discussed in this paper.

Introduction

Consideration of performance testing is becoming more important for asphalt mix from two perspectives, namely: (1) understanding of mix performance against extreme traffic and temperature events, and (2) complementing the volumetric parameters and by adding a mechanical dimension to mix design procedures. These two perspectives combined can help designers or decision makers understand performance (i.e., benchmark the performance), and better change certain aspects of the mix toward certain property (i.e., better cracking resistant). The latter has been a focal point of mix design for the last few years in Canada, where Balanced Mix Design (BMD) is gaining more attraction to become a design choice. As of now, in Canada, performance testing is mainly used as part of product development and academic research programs. A level of performance testing is not yet adopted as part of mix design, and designs are further based only on volumetric properties.

Performance testing of asphalt mixes is a broad topic, and it is outside the scope of this paper to explain all the performance testing available to the asphalt industry. Rather the focus of this paper is those methods that are being proposed as Quality Assurance (QA) friendly types and the methods with capability of being adopted as part of mix designs. Therefore, this paper is only focused on the methods listed below:

- Hamburg Wheel Tracking (HWT) test used in evaluating permanent deformation (i.e., rutting susceptibility);
- Semi-Circular Bend (SCB) test to evaluate cracking resistance at intermediate temperature;
- Disk-shaped Compacted Tension (DCT) test to evaluate low temperature cracking; and
- IDEAL type of test in evaluating rutting at higher and cracking at intermediate temperatures (IDEAL-RT and IDEAL-CT, respectively).

A brief description of each test is provided, while the sensitivity of test to air voids, fabrication method, and type of mix has been literature searched and summarized in this paper.

Hamburg Wheel Tracking (HWT) Test

Wheel tracking devices such as Asphalt Pavement Analyzer (APA) and Hamburg Wheel Tracking (HWT) equipment were originally developed to assess the resistance of asphalt concrete mixes against permanent deformation [1]. Use of HWT devices has been researched extensively over the course of the past 20 years. Several State DOTs have either adopted or are in the process of adopting HWT test as a part of their specifications, where some are using it as a part of the implementation of balanced mix design framework, and some just for QA and QC related purposes [2]. Although, HWT test is known as a very promising tool toward assessing rutting and moisture susceptibility of mixes, it should be noted that several details may vary from one equipment to another and from one specification to another, which can ultimately contribute to the variability of the test results or impose challenges in directly comparing the results from different labs. This section provides a summary of the major aspects that should be considered to properly incorporate the HWT testing as a part of the performance engineered pavements framework (PEP).

Like any other mix performance test, sample preparation practices can affect the results of HWT testing as well. The test can be performed on both slabs and cylindrical-shaped briquettes prepared under controlled laboratory conditions or retrieved from production facility. Field cores with in-place density can be also tested. However, this test methods have been predominately used for laboratory-produced laboratory-compacted (referred to as "LPLC") specimens, and other sample types less commonly used. It should be mentioned that typically Plant-Produced Field-Compacted (PPFC) and field cores are mostly used for quality assurance purposes or forensic of premature failures. Both mix/specimen types can be still considered for HWT testing but may require extra considerations as explained in this paper. As for the specification to follow, AASHTO T 324 is the method followed predominantly in Canada and the United States (U.S.). This specification allows for the air void content of the SGC specimens to be in the range of 7.0±0.5 which is intended to replicate the in-situ density of the compacted mix (minimum 93.0% of Gmm) [3]. For the slabs the tolerances are slightly higher at 7±1.0 instead due to difficulty of achieving 0.5% variation after using slab compactors. In terms of conditioning the specimen prior to testing, generally the specifications require a short-term aging of the loose mix for the laboratory produced specimens prior to the laboratory compaction according to AASHTO R30 or the equivalent specifications for the jurisdiction. Often AASHTO R35 is used for warm mix type of mixes. Other aging practices such as those on compacted specimens and long-term aging simulation of disk-shaped specimens has not been addressed in the standards, mainly given that rutting is generally believed to be an early-age distress, which may typically be counteracted with the hardening effect of longer-term aging in the materials. While this may be true in the case of rutting resistance of mixes, this phenomenon might have a different implication with respect to moisture induced damage given that binder aging could decrease the adhesion between asphalt and aggregates.

The specimen height can generally vary between 38 to 100 mm in different mounting configuration. Nevertheless, it is required to maintain a minimum specimen height of twice the nominal maximum aggregate size of the mix. A thickness of 62 ± 2.0 mm is commonly used in many setups when testing LPLC specimens or those field cores with sufficient thickness that can make up 62 mm thickness. However, in case the specimen thickness is shorter than this range, shimming and using a quick setting gypsum plaster (mostly referred to as "Plaster of Paris (P.O.P)") can be used to compensate for the difference in the specimen height. It should be noted that POP can be only used for a limited thickness increase (mostly up to 5 mm), where shimming is more recommended due to higher reliability. The material to be used for shimming is recommended to be aluminum. Figure 1 provides an example of using aluminum shimming for a field core with thickness of less than 62 mm. if nominal core diameter is less than 150 mm, it is then recommended to use POP to consistently fill any gaps between the edge of testing mold and outside perimeter of the specimen. This is to ensure the water during the testing is allowed to penetrate top-down in simulating moisture damage, as opposed to side-way ingression. Figure 1, shows use of POP in closing any gaps between outside diameter of a field core and testing mold. While different techniques are recommended to accommodate for such geometry difference, the effect of the specimen thickness on the test results has not been fully investigated in the past and should be further studied in the future work. This becomes also relevant as in many cases the cores and slabs obtained from the field compacted pavements would be somewhere in the range of 38 mm and above [4].

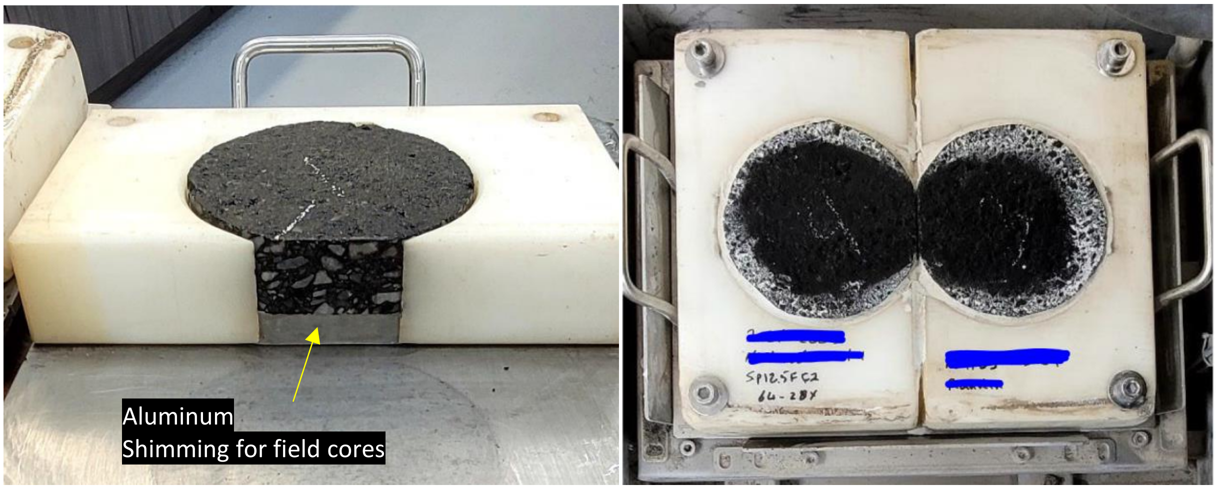

Figure 1. Use of Aluminum Shimming to Raise Field Core Sample to the Thickness of 62mm (Right), Use of P.O.P in Consistent and Tight Filling of any Gaps Between Specimen Outside Perimeter and the Mold (Photo Courtesy of Engtec Consulting Inc.)

Furthermore, the AASHTO T 324 indicates the test can be performed on both field cores and lab prepared specimens, it should be noted that even the AASHTO T 324 standard states that the field compacted samples have not shown to be sufficiently controlled for inclusion in a precision and bias statement [3]. This will emphasize the need for better reflecting the effect of compaction degree on the test results and the ultimate interpretation of the results as compared to the corresponding in-service performance of mixes. In case of adjusting the height of the specimens it is crucial to verify that the load applied by the steel wheels remain within the 705 ±4.5 N range.

Unlike many other mechanical testing standards that specify the allowable delay between sample preparation and testing to minimize the discrepancies rooted in testing specimens at different times, there is not an explicitly stated requirement with this respect. However, the standard specifies the duration of preconditioning in order to reach the test temperature that will be discussed in the next section.

The HWT testing temperature is not explicitly indicated in the AASHTO T 324 testing protocol. However, different agencies have a modification of the original AASHTO standard, where they indicate either a single testing temperature or a value corresponding to the High-performance grade (H-PG) temperature of the binder used in preparing the specimens [1, 5]. The testing temperature currently ranges from 40°C up to 56°C. In case of specifying the testing temperature based on the PG of the virgin binder, one of the gray areas that also needs to be better investigated is the appropriate testing temperature when a high amount of Reclaimed Asphalt Pavement (RAP) is present in the mix, where it can affect the H-PG of the binder blend.

The two main schools of thought are testing asphalt mixes at a constant temperature which is representative of the climatic condition of the region versus performing the test at a temperature depending on the H-PG of the binder used in the mix. Reviewing the literature indicates that even the latter was originally meant to represent different regions in the same state/province where the minimum binder grade requirements are dominated by the climatic conditions. However, this can be mistakenly overlooked if the specification does not explicitly state the minimum performance grade requirement as the criterion for selecting the testing temperature. In other words, when looking at the test as a tool for performance assessment, as long as the mix can withstand the damaging effect of loading at the expected high temperature end, a higher-grade binder which is typically costlier should increase the reliability of the performance at a given climatic condition. In some cases, where the goal is to contrast the ability of different mixes in terms of their ultimate durability, using higher testing temperatures becomes the only way to distinguish the performance of different mixes within a reasonable number of wheel passes. An example of this can be the study performed under the Wisconsin Highway Research Program (WHRP), where the researchers investigated the HWTT results at 40, 45, 50°C, where the research found that testing at both 45 and 50°C could result in similar assessment, with the latter being able to demonstrate the effect of aggregate source as well. Testing at 40^{o}C on the other hand was not found to be a promising option [6].

In addition to the testing temperature, a survey conducted by the Louisiana State University indicates that the temperature control tolerance of different available HWT devices in the states ranges from ±0.3^{o}C up to ±1.0^{o}C [7]. The accuracy of the temperature measurements can also vary based on the number of the probes and their locations within the tank. The AASHTO standard requires that the average of four measurements in the tank be within ±1.0^{o}C of the specified test temperature. The standard also recommends a 45-minute of preconditioning time to reach the temperature equilibrium but indicates that this time should not exceed 60±5 minutes. Although such considerations may seem minutiae at the first glance, the ability to maintain a uniform water bath temperature, and hence a temperature equilibrium of the specimen, within the desired tolerance would be crucial in producing reliable results through HWT devices. Regardless of the exact values to be set in the specification, the very first key step would be to provide clear and consistent procedural details to minimize any discrepancies rooted in the potential inherent subjectivity of users' interpretations of the testing protocol.

The minimum requirements for the impression measurements in the test are at the midpoint of the tracks ± 12.5 mm. These measurements should be taken at least every 400 passes. As for the track length, almost all the commercially available devices use a track length of about 228 mm. The minimum range for the LVDTs used for impression measurements is specified as 20 mm and an accuracy level of 0.15 mm. Some studies evaluated the use of additional measurement points along the wheel path to better understand the permanent deformation evolution with number of passes, while this has not been adopted by any transportation agencies as of the date of this publication. To this end, using a minimum of number of data collection points across the specimen should be considered as an important aspect. This parameter seems to be adjustable to as few of 5 points or up to 227 points in different HWT devices [8]. Nevertheless, evaluating the effect of HWT devices and their capabilities across different vendors can be recognized as an important step toward performance based and BMD implementation of the HWT testing based on previous experiences [8, 9].

Once a consistent sample preparation and testing protocol is established, the results collected from HWT devices can be post-processed in a relatively standard way to derive index-based indicators of how the mix performs either in terms of its resistance to permanent deformation or moisture-induced damaged. Among 14 State DOTs most of them use HWT test results to evaluate both moisture and rutting resistance, while only three agencies solely rely on the test as an indication of the rutting performance. As illustrated by Figure 2, a typical test would result in three distinct phases, namely post-compaction, creep stage, and stripping stage. The characteristic features of each stage can then be used to calculate indices to better understand the performance of mixes, which is typically carried out using: (1) number of passes to rutting failure, (2) Stripping Inflection Point (SIP), (3) creep slope, and (4) stripping slope. Based on the research results from a study conducted by the Iowa DOT, some agencies have started considering the ratio between the stripping and creep slopes into account to determine whether SIP is a valid indicator for a given mix or not [10]. To this end, a ratio greater than 2.0 can validate the SIP while values less that 1.0 indicate that SIP cannot be used as a reliable indicator of the stripping behavior.

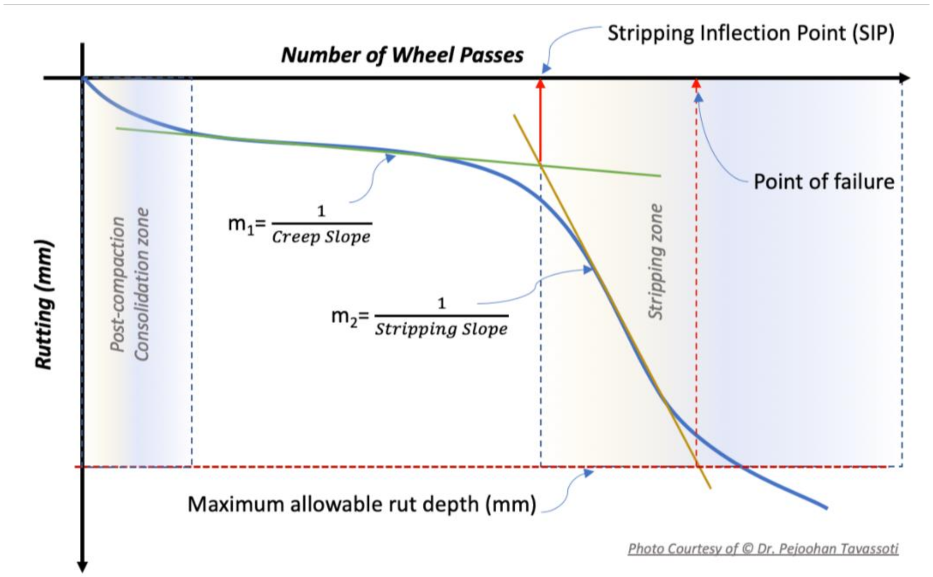

Photo Courtesy of © Dr. Pejoohan Tavassoti

Figure 2. Typical stages of HWT test and the standard parameters.

Furthermore, review of the existing specifications indicate that the failure criteria is defined differently across the north American transportation agencies. Among those considering the different High temperature PG of the binders, several agencies such as in California, Illinois, Oklahoma, Texas, and Wisconsin require meeting a given maximum rut depth (e.g., 12.5mm) but at different number of wheel passes for each testing temperature. Whereas some agencies such as Utah, Montana, and Colorado specify a certain maximum rut depth at a fixed number of wheel passes. It should be mentioned that some agencies select test temperatures based on the nominal Critical Stripping Temperature (CST) established for AASHTO M320 asphalt binder performance grades by the Utah Department of Transportation [11] to evaluate the moisture susceptibility of the asphalt concrete mixtures. These temperatures can be predicted based on the temperature at which the G*/sinδ parameter in Dynamic Shear Rheometer (DSR) testing is equal to 2.2 kPa-1 using Equation 1. This provides a consistent reference point based on the viscosity of the binder as related to the performance grade.

$$\boldsymbol{CST}\ (^{\circ}\mathrm{C}) = \mathbf{0.42}\left(\boldsymbol{T_{G^*/sin\delta=2.2\ kPa}}\right) + \mathbf{20.5} \qquad \textbf{Eq. 1}$$

where

CST is the Critical Stripping Temperature in °C;
TG*/sinδ = 2.2kPa-1 is the temperature at which DSR G*/sinδ = 2.2kPa-1 in °C

Regardless, of the specification details used, the published report indicate that different agencies have been able to achieve a relatively acceptable level of success with using the HWT test for rutting and moisture assessment. It is noteworthy, however, that the process is still being revisited for potential improvements by these agencies, and especially toward fine-tuning the BMD implementation efforts.

Aside from the standard parameters mentioned above, researchers have been looking into non-standard parameters such as Corrected Rut Depth (CRD) approach by Yin et al. [12], Rutting Resistance Index (RRI) by Wen et al. [13] and other parameters derived from polynomial fitting of the data as opposed to the traditional bi-linear approach. Although these efforts are important to better understand the correlation between the field and lab performance of the mixes and distinguishing the creep contribution from that of stripping, further investigation is still needed and hence none of the north American state DOTs have adopted these non-standard parameters yet.

Semi-Circular Bend (SCB) Test

The Semi-Circular Bend (SCB) Test was developed as a fracture test to characterize the low-temperature cracking resistance of asphalt mixtures to differentiate mixtures whose service life might be compromised by cracking. The SCB test method is generally valid for specimens that are tested at temperatures of 10 °C or below [14]. The University of Illinois modified the SCB procedure, called the Illinois Flexibility Index Test (I-FIT), based on a thorough investigation of test temperatures, loading rates and sample geometry to quantify the cracking potential of asphalt mixtures at intermediate temperatures [15].

The SCB I-FIT is a monotonic bending test developed to predict the crack propagation properties of asphalt mixes at intermediate temperatures. The provisional standard test method, AASHTO TP-124, "Determining the Fracture Potential of Asphalt Mixtures Using the Semi-Circular Geometry at Intermediate Temperature," calls for a 150-mm diameter by 50-mm thick semi-circular specimen with a 15-mm notch to be simply supported by two bars on the flat surface, as pictured in Figure 3. Prior to testing, the test specimen is conditioned for two hours at 25°C, the standard test temperature. The load is applied to the curved surface above the notch at a vertical rate of 50 mm/min. Load and vertical displacement are recorded until the load drops below 0.1 kN. The fracture energy (G_f) is calculated from the area beneath the load displacement curve to 0.1 kN (Eq.2). The post-peak slope at inflection point of the load displacement curve, |m|, is an indicator of the brittle to ductile failure. The Flexibility Index (FI) parameter is calculated by multiplying the fracture energy by a scaling factor constant and dividing by the slope (Eq.3) [16]. The FI is an indication of the resistance of the mix to crack propagation at intermediate temperature, where higher values of FI are favorable.

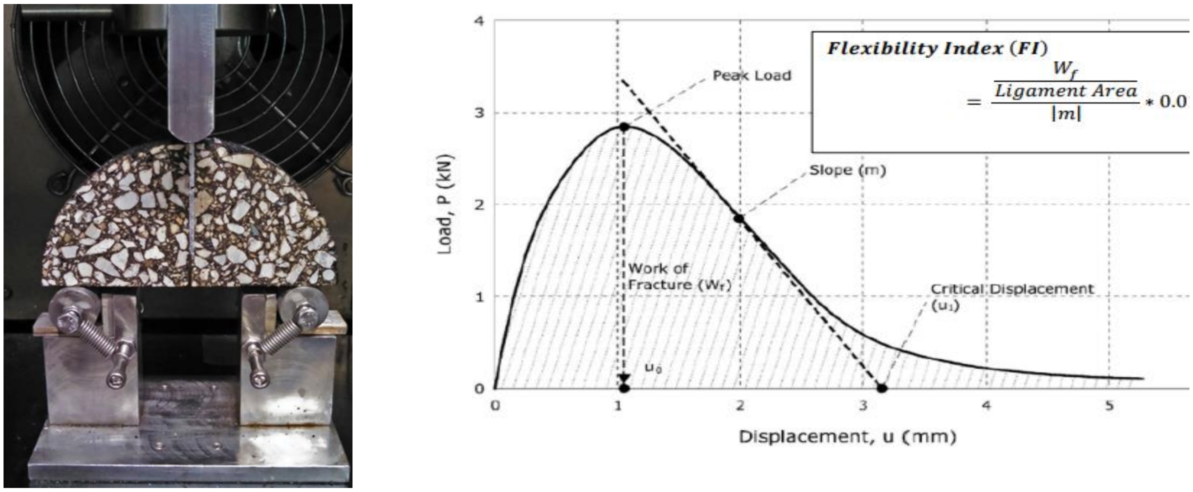

Figure 3. I-FIT Test Setup and Example Plot of Results [17]Click or tap here to enter text.

$$\boldsymbol{G_f = \frac{W_f}{A_{lig.}} \times 10^6} \qquad \textbf{Eq. 2}$$

where

G_f is the fracture energy (J/m^2)
W_f is the work of fracture (J)
$A_{lig.}$ Is the ligament area (mm^2)

$$\boldsymbol{FI = \frac{G_f}{|m|} \times 0.01} \qquad \textbf{Eq. 3}$$

where

G_f is the fracture energy (J/m^2)
$|m|$ is the post-peak slope at the inflection point

Researchers have found that the SCB I-FIT test shows consistent and repeatable trends for changes in asphalt mix design properties, and the FI parameter is shown to provide a greater distinction between mixtures' fracture properties relative to the total fracture energy parameter alone. As such, the I-FIT test is considered as a reliable test to identify the cracking potential of asphalt mixtures at intermediate temperatures and ready for implementation in Wisconsin [17-19] Some researchers have expressed concerns about the SCB I-FIT. Difficulties have been documented with asphalt mixtures containing elevated levels of recycled materials that were reportedly caused by the relatively high loading rate (50 mm/min) applied during the test [20, 21].

To address these difficulties, Kasser et al. developed an alternative to FI, the Cracking Resistance Index (CRI), which implements the peak load from the same test instead of the post peak value slope. They reported that CRI values provide greater discrimination to differentiate mixtures with distinct characteristics. Also, CRI values provided less variability than FI values because the peak load did not show significant variability from test to test, as compared to the post peak slope. Thus, both indices, the FI and

CRI, were observed to be sensitive to different mixture characteristics, such as the binder performance grade (PG) and recycled material content. However, both also had difficulty distinguishing asphalt content variations [22].

The preparation of the test specimens for SCB-IFIT is somewhat complicated and requires cutting and notching, which could cause unexpected testing variance. Such variance could be worsened if the created notch ends in the middle of a large aggregate as shown in Figure 4. This has been raised as a concern by some researchers and has not been yet documented if it could add any variability to the test results. But it is understood that SCB is capturing flexibility of the mastic portion of the specimen, and as such the crack needs to start in the mastic and further propagate through the mastic. Figure 4 shows an example of a notch started in mastic portion of the specimen.

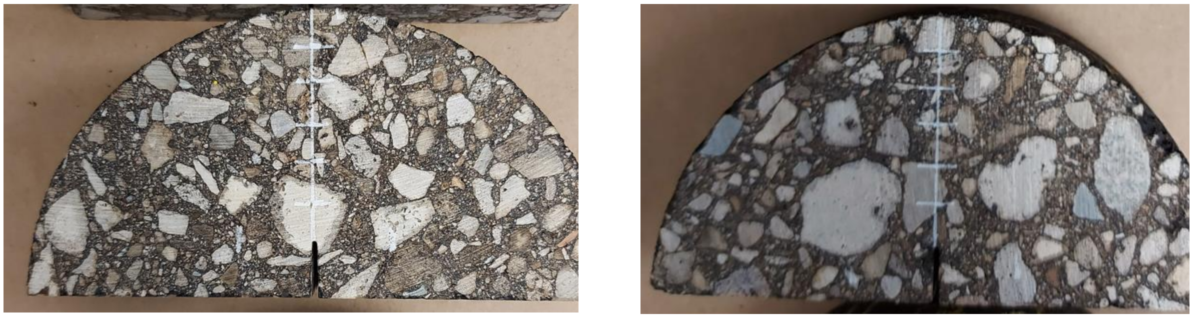

Figure 4. Comparison of Two SCB Specimens from the Same Mix: notch cut through largest aggregate size (Right), and notch cut through Mastic Portion of the mix (Left)

The conditioning of specimens prior to testing is two hours at the testing temperature (25°C). One important factor to consider is the condition method: (1) environmental chamber (external to test setup), (2) environmental chamber (same chamber housing the test setup), and (3) water bath. The first and last option of conditioning methods often used for those setups that are not housed inside a chamber and are stand-alone type (i.e., Marshall press types). It is important to understand any variability between these setups if an agency intends on implementing SCB FI for QA or contractual acceptance purposes. When it comes to accuracy, it is preferable to test for SCB when the setup and specimen conditions are both in one chamber.

In addition, the recently developed indirect tensile asphalt cracking test (IDEAL-CT), which requires no cutting or notching of specimens, is also becoming widely considered as a possible practical test. More explanation on this test is provided separately in a following section of this paper.

The effects of specimen geometry and air void content has also been studied for SCB-IFIT. Al-Qadi et al. used I-FIT to investigate the relationship between the FI and specimen thickness, with thicknesses that varied from 25 to 62.5 mm. They observed a linear reduction in the FI value with increasing specimen thickness, which can be explained by the effect of specimen thickness on the post peak value. They also observed a variation in fracture energy values that was not statistically significant enough to establish a correlation. The researchers recommended a simple linear correction factor for the FI that is based upon a standard 50-mm thick specimen. [23] The recommended correction factors for air void and thickness limitations have shown mixed results. [18]

With respect to air void content of test specimens for SCB-IFIT, Kaseer et al. documented that FI values are heavily dependent on specimen air void content and may require different correction factors to adjust the FI values to standard conditions. [22] Additionally, work done by Barry showed a consistent decrease in peak load with increased air void content in test specimens. [20] Although less significant, there was also an impact observed for the fracture energy values whereby they decreased with an increase in air void content. The peak load and air void content were more strongly correlated than the fracture energy and air void content [20]. Other research has noted significant effects of air void content and specimen thickness on FI values, where higher FI values are observed with an increase in air void content and decrease in specimen thickness. [18] The importance of attaining higher density in the asphalt mat is promoted to improve the durability of asphalt pavements. Therefore, further investigations should focus on understanding the effects and correlations between influential factors and cracking parameters.

The effects of specimen geometry and air void content on the I-FIT parameters becomes even more significant when field compacted specimens are tested and analyzed. Laboratory-compacted specimens can be more carefully prepared to meet established standard test requirements, while field-compacted specimens present inherit variabilities that can affect the test results. This is important as agencies explore using the I-FIT as a construction QC and/or QA test.

Disc-Shaped Compact Tension (DCT) Test

The Disc-Shaped Compact Tension Test was developed at the University of Illinois at Urbana–Champaign. It is used to determine the fracture energy of asphalt-aggregate mixture. The fracture energy is a fundamental property of a material and is defined as the energy required to open a unit of area the crack surface; it is measured in J/m². The higher the fracture energy, the more stress the specimen can withstand, and since the test is conducted at low temperature, this means the material has better resistance to thermal cracking.

Fracture energy (G_f) is calculated using the following equation:

$$\boldsymbol{G_f = \frac{AREA}{B \cdot (W - a)}} \quad \textbf{Eq. 4}$$

where

$AREA$: area under the load vs CMOD curve (N·m)
B: Specimen thickness (m)
$(W - a)$: Initial ligament length (m)

The disk-shaped compact tension test is conducted in accordance with ASTM D7313 Standard. In this test, a notch and two loading holes are placed on a 150mm diameter and 50mm thick circular sample that is subjected to tension, and the crack mouth opening displacement (CMOD) is measured with a clip-on gage. The test setup and geometry of the sample are shown in Figure 5.

This test is conducted at a test temperature of 10°C above the low-temperature performance grade of the asphalt binder, as suggested by the specification. However, it is not yet confirmed if this temperature should be changed to better reflect the low PG grade of the asphalt binder.

One of the variables that can affect the fracture energy is the thickness of the specimen. Wagoner et al. conducted a brief study which concluded that fracture energy increases as a function of the specimen thickness. They also concluded that fracture energy also increases with the increase of temperature. It was demonstrated that fracture energy is sensitive to changes in the properties of the mix, particularly the asphalt binder grade, and softer binders tend to have higher fracture energy. [24]

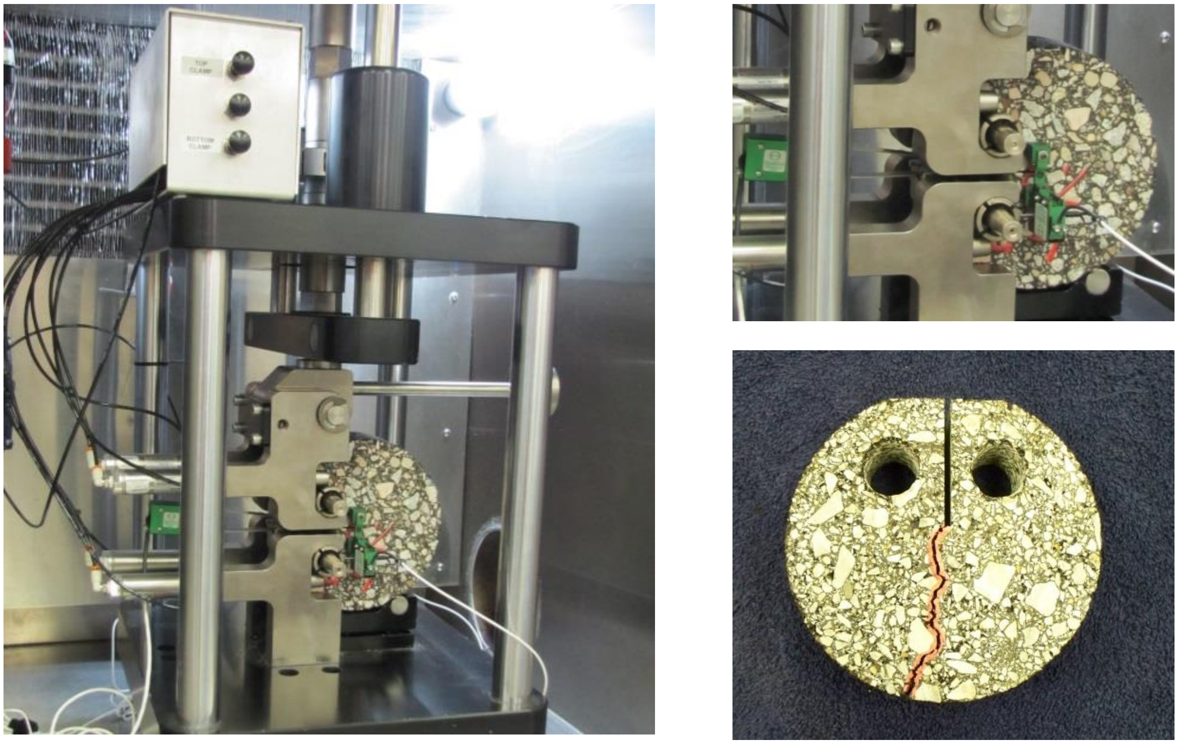

Figure 5. Disk-Shaped Compact Tension Test Setup (Photo Courtesy of Asphalt Institute)

IDEAL-CT and IDEAL-RT Test Methods

The IDEAL-CT was developed by the Texas A&M Transportation Institute, which later became standardized by the American Society of Testing Materials in the ASTM D8225-19. The test is conducted at the performance grade intermediate temperature; however, a common test temperature is 25°C. Specimens should be conditioned for at least two hours at the test temperature, and testing on a minimum of three specimens is recommended. This test requires only gyratory compacted specimens with 150 mm diameter, 62 mm height and 7± 0.5 percent air voids. There is no need for cutting or notching, and it uses the same breaking head used for the IDT test on 6-inch diameter specimens (Figure 6).

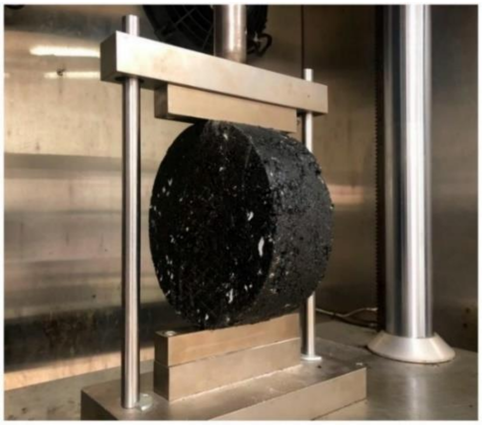

Figure 6. IDEAL-CT Test Setup

In general, a higher CT_{index} and FI suggest asphalt mixture's ability in exhibiting more ductile behavior, which may translate to a better cracking resistance at the testing temperatures. In contrary, lower CT_{index} and FI indicate a mixture’s tendency to exhibit more brittle failure which could translate to higher cracking susceptibility. IDEAL-CT and SCB example load versus displacement diagrams are presented in Figure 7.

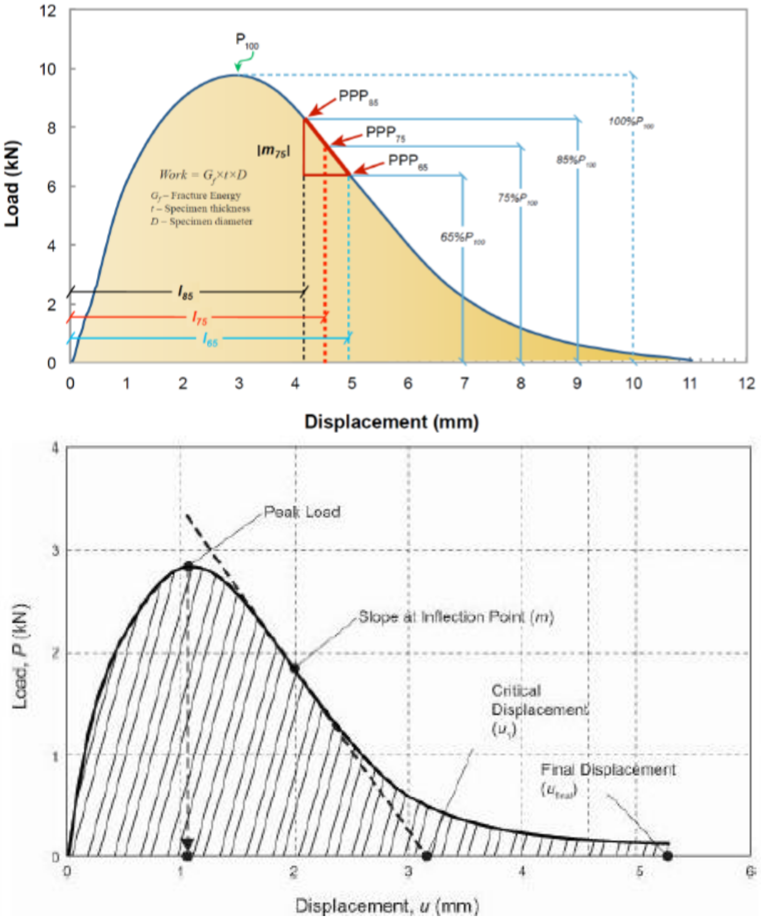

Figure 7. IDEAL-CT (Top) [25] Versus and SCB I-FIT (Bottom) [26] Loading and Analysis

In the test, a monotonic load at a rate of 50mm/min is applied, then load and displacement data are recorded. The output of the test is the CT index, which is a function of the fracture energy and the slope of the post-peak curve at 75 percent of the peak load, using the following equation:

$$\boldsymbol{CT_{Index}} = \frac{\boldsymbol{G_f}}{|\boldsymbol{m_{75}}|} \times \frac{\boldsymbol{l_{75}}}{\boldsymbol{D}} \qquad \textbf{Eq. 5}$$

where

G_f: Failure energy (area under the load vs displacement curve divided by cross-sectional area $D \times t$)
$|m_{75}|$: the absolute value of the post-peak slope at 75 percent of the peak load
l_{75}: Displacement at 75 percent of the peak load
D: specimen diameter

A correction factor ($t/62$) should be applied for field cores with thickness different from 62 mm.

Yan et al. conducted research to compare the SCB-IFIT with and without notch and the IDEAL-CT. The comparison focused on the specimen preparation, testing set-up, testing variance, and the capability of discriminating between different mixtures. The flexibility index (FI) results from the SCB with and without notch show a strong linear correlation (R^2 of 0.95). The results showed that the notched SCB-IFIT specimens had a larger variance in their results than un-notched specimens, indicating that the notch can be eliminated to reduce the variance of results. The cracking tolerance index (CT_{index}) results from IDEAL-CT were found to correlate very well with the FI from SCB-IFIT, and CT_{index} results show less variation than the FI results (average COV of 5.3% versus 23.0%). Since the IDEAL-CT requires no specimen cutting or notching, Yan et al. concluded that it could be a promising alternative to evaluate cracking resistance of asphalt mixtures [21].

The CT index is unitless. A higher CT index represents better resistance to cracking. Given the ease of testing, IDEAL-CT has been considered for several U.S. Departments of Transportation for quality assurance purposes. Agencies that have adopted the test include Alabama, Idaho, Kentucky, Missouri, Oklahoma, Tennessee, and Virginia [27]. However, each agency is responsible for defining the threshold for the CT index.

As part of the NCHRP 195, several analyses demonstrated that the IDEAL-CT is sensitive to RAP and RAS content, asphalt binder type, asphalt binder content, air voids and ageing conditions. The maximum coefficient of variability reported was 23.5 percent, with most of the results below 20 percent. IDEAL-CT also demonstrated a good correlation with field cracking performance.

A factor that significantly influences the IDEAL-CT results is air voids. The test is considered rugged for test temperature, specimen thickness, and loading rate [25]. Also, a study was conducted to examine the variability due to the device used, which did not exceed the repeatability level of 20 percent established for the test [28].

The IDEAL-RT, also developed at the Texas A&M Transportation Institute, requires the same specimen geometry and compaction as the IDEAL-CT previously described. Rutting is caused mainly by shearing; hence the test uses the same loading frame and loading rate as IDEAL-CT but a different test fixture (Figure 8) which has a cradle assembled on the base plate to create shearing stress instead of tension stress.

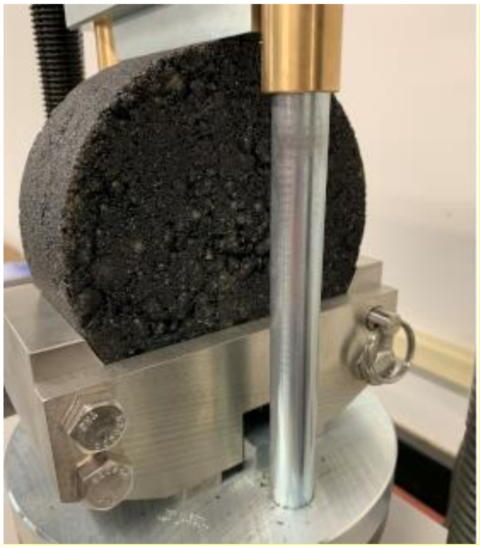

Figure 8. IDEAL-RT Test Setup

The standard is under development by the American Society of Testing Materials ASTM WK71466. The test is performed at a high temperature which may be selected based on local climate data or using LTPP Bind. However, it is commonly conducted at 50°C to match the temperature used by several agencies for the Hamburg wheel tracking test. Specimens should be conditioned for at least two hours at the test temperature. Testing on a minimum of three specimens is recommended.

The output of the test is the RT index, which is the shear strength determined from the following equation:

$$\boldsymbol{RT_{Index} = 0.356\frac{P_{max}}{t \times w}} \qquad \textbf{Eq. 6}$$

where

P_{max}: maximum load in N
t: specimen thickness in mm
w: width of the loading strip (19.05 mm)

A correction factor ($t/62$) should be applied for field cores with thickness different from 62 mm.

The units of the RT index are MPa. A higher RT index represents better resistance to rutting. The test has not yet been adopted by any agency for quality assurance purposes and is still undergoing further investigation by researchers.

IDEAL RT is sensitive to RAP and RAS content, asphalt binder type, asphalt binder content, aggregate type, air voids and ageing conditions. The coefficient of variability is below 10 percent, and it also has a good correlation with existing rutting tests, such as Asphalt Pavement Analyzer and Hamburg Wheel Tracking Test; and with field rutting performance [25]. A ruggedness test has not been conducted for this test yet.

Summary and Recommendations

This paper identified and described common asphalt mixture performance evaluation methods that various road agencies in North America are considering for quality control and quality assurance purposes. The tests investigated included HWT, SCB, DCT and IDEAL CT and RT.

Regardless of the test selected, developing a robust testing database using results of a wide range of mixes commonly used in the jurisdiction can be identified as one of the necessary steps for balanced mix design implementation purposes. Such database should preferably cover different Nominal Maximum Aggregate Sizes (NMAS), binder grades, traffic levels, aggregate sources, and Recycled Materials sources and contents. To develop such database, the partner laboratories should be able to well control the details of sample preparation/collection and testing configurations. Collecting the relevant information about different testing devices can also help with establishing precision and bias statements once sufficient level of data is collected.

The review of the current practices in using performance testing indicate that the following areas still require further investigations to properly realize the potential of these tests, especially toward implementation of a balanced mix design framework:

- Further study is needed to investigate the effect of sample preparation practices. This will help eliminate potential conflicts of using different specimen heights as well as differences in the details of compacting the original cylinders using Superpave gyratory compactor in the lab.
- A provincial level database to collectively gather the test results needs to be established to better correlate the field performance with the test results for the range of asphalt concrete mixes produced using the domestically available materials. The database will be crucial in establishing proper pass/fail criteria, distinguishing poor- from well-performing mixes for QA/QC purposes and defining precision and bias statements.
- Harmonizing the sample preparation, testing protocol details, including testing conditioning and testing temperature and devices requirements is deemed crucial for a reliable interpretation of the collected test results and minimizing any discrepancies in the mix assessment process. Consistent sample collection and preparation to obtain a representative sample continues to be important, which includes mixture conditioning time of plant-produced samples prior to briquette preparation. For examples, using a standard box size for QA programs to collect field samples would ensure that collected material amount remains consistent for split-samples for the different testing laboratories. It is important to understand how these variables affect the results and its impact on the data quality of any database.

References

[1] Tavassoti, P. and Baaj, H., 2020. Moisture Damage in Asphalt Concrete Mixtures: State of the Art and Critical Review of the Test Methods. In Transportation Association of Canada 2020 Conference and Exhibition-The Journey to Safer Roads.

[2] West, R., Yin, F., Rodezno, C. and Taylor, A., 2021. Balanced Mixture Design Implementation Support. Wisconsin Highway Research Program.

[3] American Association of State Highway and Transportation Officials (AASHTO). 2019. T 324-19, Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures, Standard Specifications for Transportation Materials and Methods of Sampling and Testing.

[4] Tavassoti-Kheiry, P., Solaimanian, M., Qiu, T. and Boz, I., 2017. High-Frequency Shear and Young's Modulus of Asphalt Concrete Through Resonant Column Testing of Small-Scale Specimens (No. 17-06822).

[5] Yin, F., Chen, C., West, R., Martin, A.E. and Arambula-Mercado, E., 2020. Determining the relationship among hamburg wheel-tracking test parameters and correlation to field performance of asphalt pavements. Transportation Research Record, 2674(4), pp.281-291.

[6] Swiertz, D., Ling, C., Teymourpour, P. and Bahia, H., 2017. Use of the Hamburg wheel-tracking test to characterize asphalt mixtures in cool weather regions. Transportation Research Record, 2633(1), pp.9-15.

[7] Mohammad, L.N., Elseifi, M., Cao, W., Raghavendra, A. and Ye, M., 2017. Evaluation of various Hamburg wheel-tracking devices and AASHTO T 324 specification for rutting testing of asphalt mixtures. Road Materials and Pavement Design, 18(sup4), pp.128-143.

[8] Hajj, E.Y., Aschenbrener, T. and Nener-Plante, D., 2022. Examples of Successful Practices with State Implementation of Balanced Design of Asphalt Mixtures. Transportation Research Record, p.03611981221084696.

[9] Hajj, E.Y., Aschenbrener, T.B. and Nener-Plante, D., 2021. Case Studies on the Implementation of Balanced Mix Design and Performance Tests for Asphalt Mixtures: Virginia Department of Transportation (VDOT). UNR Pavement Engineering & Science Program.

[10] Schram, S. and Williams, R. C., 2012. Ranking of HMA moisture sensitivity tests in Iowa, Report No. RB00-012, Federal Highway Administration (FHWA).

[11] Sias, J.E., Dave, E.V. and McCarthy, L.M., 2020. Practices for fabricating asphalt specimens for performance testing in laboratories (No. Project 20-05, Topic 50-14).

[12] Yin, F., Arambula, E., Lytton, R., Martin, A.E. and Cucalon, L.G., 2014. Novel method for moisture susceptibility and rutting evaluation using Hamburg wheel tracking test. Transportation Research Record, 2446(1), pp.1-7.

[13] Wen, H., Wu, S., Mohammad, L.N., Zhang, W., Shen, S. and Faheem, A., 2016. Long-term field rutting and moisture susceptibility performance of warm-mix asphalt pavement. Transportation Research Record, 2575(1), pp.103-112.

[14] American Society for Testing and Materials (ASTM). 2020. D7313-20, Standard Test Method for Determining Fracture Energy of Asphalt Mixtures Using the Disk-Shaped Compact Tension Geometry. ASTM International.

[15] Albritton, G.E., Barstis, W.F. and Crawley, A.B., 1999. Polymer modified hot mix asphalt field trial (No. FHWA/MS-DOT-RD-99-111). Mississippi. Dept. of Transportation.

[16] West, R., Rodezno, C., Leiva, F. and Yin, F., 2018. Development of a framework for balanced mix design. Project NCHRP, pp.20-07.

[17] Illinois-DOT. 2016. Determining the Fracture Potential of Asphalt Mixtures Using the Illinois Flexibility Index Test (I-FIT).

[18] Batioja-Alvarez, D., Lee, J. and Haddock, J.E., 2019. Understanding the Illinois flexibility index test (I-FIT) using Indiana asphalt mixtures. Transportation Research Record, 2673(6), pp.337-346.

[19] Ling, C., Swiertz, D., Mandal, T., Teymourpour, P. and Bahia, H., 2017. Sensitivity of the Illinois flexibility index test to mixture design factors. Transportation Research Record, 2631(1), pp.153-159.

[20] Barry, M.K., 2016. An analysis of impact factors on the Illinois flexibility index test.

[21] Yan, C., Zhang, Y. and Bahia, H.U., 2020. Comparison between SCB-IFIT, un-notched SCB-IFIT and IDEAL-CT for measuring cracking resistance of asphalt mixtures. Construction and Building Materials, 252, p.119060.

[22] Kaseer, F., Yin, F., Arámbula-Mercado, E., Martin, A.E., Daniel, J.S. and Salari, S., 2018. Development of an index to evaluate the cracking potential of asphalt mixtures using the semi-circular bending test. Construction and Building Materials, 167, pp.286-298.

[23] Al-Qadi, I., Ozer, H., Lambros, J., Lippert, D., El Khatib, A., Khan, T., Singh, P. and Rivera-Perez, J.J., 2015. Testing protocols to ensure mix performance w/high RAP and RAS. Illinois Center for Transportation, 1, p.209.

[24] Wagoner, M.P., Buttlar, W.G., Paulino, G.H. and Blankenship, P., 2005. Investigation of the fracture resistance of hot-mix asphalt concrete using a disk-shaped compact tension test. Transportation Research Record, 1929(1), pp.183-192.

[25] Zhou, F., 2019. Development of an IDEAL cracking test for asphalt mix design, quality control and quality assurance. NCHRP-IDEA Program Project Final Report, (195).

[26] American Association of State Highway and Transportation Officials (AASHTO). 2018. TP-124-20, Standard Method of Test for Determining the Fracture Potential of Asphalt Mixtures Using the Flexibility Index Test (FIT).

[27] Yin, F. and West, R.C., 2021. Balanced mix design resource guide (No. IS-143).

[28] Moore, N., Steger, R., Bowers, B. and Taylor, A., 2022. Investigation of IDEAL-CT Device Equivalence: Are All Devices Equal?. Transportation Research Record, p.03611981221091551.