Airport Asphalt Pavement Technology Program

Balanced Mix Design: Rutting Performance Tests

Final Report

August 2025

Credit: NAPA

Airport Asphalt Pavement Technology Program

PREPARED BY:

University of Nevada, Reno
Elie Y. Hajj
Nicole G. Elias (currently California State Polytechnic University, Pomona)
Bipin Khanal
Ashraf Alrajhi
Adam Hand

Rutgers University
Thomas Bennert

Texas A&M Transportation Institute
Fujie Zhou
Jon Epps
Kin Ming Chan

RDM International, Inc.
Christopher S. Decker
Harsh Patel

Duval Engineering LLC
John Duval

NAPA NATIONAL ASPHALT PAVEMENT ASSOCIATION

FEDERAL AVIATION ADMINISTRATION

The Airport Asphalt Pavement Technology Program (AAPTP) is a cooperative agreement effort between the **National Asphalt Pavement Association** (NAPA) and the **Federal Aviation Administration** (FAA) to advance asphalt pavements and pavement materials. The AAPTP advances solutions for asphalt pavement design, construction, and materials deemed important to airfield reliability, efficiency, and safety. The program leverages NAPA's unique technology implementation capabilities with assistance from the FAA and industry to advance deployment and adoption of innovative asphalt material technologies.

Acknowledgments

The authors would like to acknowledge Dr. Richard Willis, NAPA Vice President for Engineering, Research, & Technology, and project panel members, including Brett Williams, Navneet Garg, Sadie Casillas, Chuck Mills, Brandon Brever, Alfredo Castro, Tim Peffer, and Kelly Senger for overseeing the project and reviewing project deliverables. The authors appreciate the effort of Dario Batioja-Alvarez (ARA, Inc./FAA Technical Center) for his continuous help in sampling materials and testing core samples, as well as his participation in the Interlaboratory Study (ILS) along with Wade Collins (Pavement Technology Inc.). The authors would also like to acknowledge Edwin Haas, Shelby Maigis, Edward Wass Jr., and Drew Tulanowski for specimen preparation and laboratory experimental testing, as well as the following personnel who helped coordinate the sampling process, provided mix design and acceptance data, and collected materials from several current airfield projects: Howard Moseley (Florida Department of Transportation), Robbie Robinson (Associated Asphalt Partners), Edgard Hitti and Dylan Stutters (Granite Construction), Jennifer Lombardo, Mahear Abou Eid, Andrew Pavey, and Mitchel Miller (Alaska Department of Transportation & Public Facilities), Frank Rancadore (Granite Rock), Sean Papathakis (Sacramento International Airport), Karen Sepulveda (Burbank Airport), Mark Puttock (Metropolitan Washington Airports Authority), and Chris Gardner (DiDonato Associates).

Credit: NAPA

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AASHTO	American Association of State Highway and Transportation Officials
AC	Asphalt concrete
APA	Asphalt Pavement Analyzer
AV	Air voids
BMD	Balanced Mix Design
CAIT	Center for Advanced Infrastructure and Transportation
CBR	California Bearing Radio
CI	Confidence interval
COV	Coefficient of variation
CRD	Corrected rut depth
DOT	Department of Transportation
EWR	Newark Liberty International Airport
FAA	Federal Aviation Administration
FAARFIELD	FAA Rigid and Flexible Iterative Elastic Layered Design
FMFC	Field-mixed field-compacted
FMLC	Field-mixed laboratory-compacted
FOS	Factor of safety
GAW	Gross aircraft weight
HT-IDT	High temperature indirect tensile test
HWTT	Hamburg wheel-tracking test
IRT	Ideal rutting test
ILS	Interlaboratory study
JMF	Job mix formula
LMLC	Laboratory-mixed laboratory-compacted
LTPP	Long-Term Pavement Performance
LVDT	Linear variable differential transformer
MEPDG	Mechanistic-Empirical Pavement Design Guide
NAPA	National Asphalt Pavement Association
NAPTF	National Airport Pavement Test Facility
NMAS	Nominal maximum aggregate size
PANYNJ	Port Authority of New York and New Jersey
PG	Performance grade
PHL	Philadelphia International Airport
RPMLC	Reheated plant-mixed laboratory-compacted
QA	Quality assurance
QC	Quality control
RLT	Repeated load triaxial
RT_{Index}	Rutting tolerance index
RTS	Reno Stead Airport

SFO	San Francisco International Airport
SMF	Sacramento International Airport
SN	Stripping number
TPA	Tampa International Airport
TTI	Texas A&M Transportation Institute
TEB	Teterboro Airport
WRSC	Western Regional Superpave Center

Executive Summary

This study aimed to establish representative rutting test protocols and criteria tailored to airfield asphalt mixtures, supporting the Federal Aviation Administration's (FAA's) balanced mix design (BMD) efforts at both the mix design and production stages. Four rutting test methods were evaluated, with an emphasis on laboratory protocols that best simulate field conditions by accounting for specimen preparation, air void (AV) levels, aging state, conditioning method, and test temperatures.

Experimental results revealed strong correlations between the Asphalt Pavement Analyzer (APA) at both 100 psi/100 lb and 250 psi/250 lb settings, the high temperature indirect tensile strength test, and the ideal rutting test. Improved correlations were observed when using Hamburg wheel-tracking test rut depths at 5,000 passes rather than 20,000 passes. An AV level of 7±0.5 percent was determined to be representative of in-place AVs in airfield pavements and was recommended for all rutting tests to ensure consistent specimen preparation.

A mechanistic-empirical approach was applied to refine the FAA's APA 250 psi/250 lb rutting test criterion by accounting for aircraft speed and load. The framework used the 3D-Move Analysis software tool to model pavement responses under varying temperatures, speeds, and loads, producing stress states representative of field conditions. These stress conditions were then applied in the repeated load triaxial test to develop laboratory-based rutting performance models for selected airfield mixtures. By combining mechanistic pavement responses with these performance models, the study enabled rutting prediction and quantified the sensitivity of airfield mixtures to operational conditions. This approach led to revised test criteria for both slow/stationary aircraft and general airfield pavements.

Laboratory verification of the recommended criteria was conducted using field cores from airfield pavement sections with known performance histories. Revisions to FAA's P-401/P-403 asphalt mixtures specifications are proposed. To expand BMD implementation into production, pilot projects are recommended to validate the proposed protocols and identify practical challenges. Long-term monitoring of sampled pavement sections will further refine the correlations between laboratory criteria and in-service performance of airfield asphalt pavements.

Chapter 1. Introduction

Asphalt mix design methods have been progressively subjected to several improvements, targeting a superior performance for flexible pavements. The latest state-of-the-art methodology for designing asphalt mixtures, called Balanced Mix Design (BMD), involves the implementation of new specifications for asphalt concrete (AC) pavements that represent better engineering insight on actual mixture performance in the field. Accordingly, BMD has been widely implemented by several highway agencies (Hajj, Aschenbrener, & Nener-Plante, 2022b; Elias, et al., 2022; TRB, 2022). Concurrently, the Federal Aviation Administration (FAA) is considering implementing a BMD framework in its subsequent specifications update for Item P-401/P-403, Asphalt Mix Pavement, in advisory circular 150/5370-10H (FAA, 2018).

The prospective BMD framework targets improved airfield pavement performance by incorporating rutting and cracking performance criteria based on laboratory mechanical testing. Rutting, or permanent deformation, is one of the major distresses in AC pavements. Rutting is mostly generated by slow-moving or standing heavy aircraft traffic, for example during stacking on taxiways, coupled with high pavement temperatures. The rutting mechanism is further exacerbated on airfield pavements under aircraft with high wheel loads and tire pressure. Nowadays, with the new larger and heavier generation of aircraft, manufacturers tend to increase tire pressure in order to increase payload or to add more wheels to maintain the load limit on each wheel (Wang, Li, Garg, & Zhao, 2020; White, 2016; Rushing & Garg, 2017). The gross aircraft weight (GAW) and the relative gear configuration dictate the load distribution per wheel that exceeds, in most cases, the wheel load of a highway truck. The pavement-tire interaction and resulting state of stresses depend mostly on the tire pressure. For large commercial aircraft, wheel loads typically range from 14,000 to 77,000 lb, with tire pressures between 150 and 240 lb per square inch (psi), compared to wheel loads of 4,500 lb and tire pressures ranging from 85 to 110 psi for truck trailers (FAA, 2022; Song & Garg, 2010; Christensen, 2008).

Rutting is commonly evaluated through laboratory mechanical testing that can investigate the engineering properties and mechanical behavior of asphalt mixtures. Several highway agencies have successfully implemented mechanical tests as part of their asphalt mix design and/or quality assurance (QA) procedures. However, for the FAA BMD framework, the laboratory mechanical testing and criteria need to be thoroughly tailored to airfield conditions in terms of test temperature, wheel load, specimen air void (AV) level, tire pressure, etc. Following an extensive research effort initiated in 2012, FAA successfully incorporated laboratory mechanical rutting tests into its specifications for asphalt mixtures, designated as Item P-401 and Item P-403 (Rushing & Garg, 2017; Rushing, Little, & Garg, 2012; Rushing, Little, & Garg, 2014). These specifications include three key laboratory mechanical tests: the Asphalt Pavement Analyzer (APA) under 250 psi hose pressure/250 lb wheel load test; the APA under 100 psi hose pressure/100 lb wheel load test; and the Hamburg wheel-tracking test (HWTT) (AASHTO, 2023a; AASHTO, 2023b).

While the implementation of rutting tests for mix design represents a major advancement, the current tests are conducted at a single temperature with a uniform criterion, regardless of an airfield's geographical location or load carrying capacity (FAA, 2018). To that end, the FAA recently initiated a research effort to re-evaluate current rutting test criteria to consider key parameters such as different aircraft load levels, aircraft speed, and airfield environmental conditions. Accordingly, the refinement of current FAA rutting specifications requires a thorough mechanistic analysis and modeling of actual airfield pavements under representative aircraft loading conditions. The mechanistic analysis performed in this project accounted for key differences between highway and airfield pavements, including variations in load levels, tire pressure, axle configuration, and pavement structure.

As part of an inclusive BMD framework, the FAA also plans to investigate new surrogate rutting mechanical tests that could be easily adopted during production.

In summary, the new recommendations for rutting test specifications aim to improve or address the current limitations in terms of the following:

- Testing asphalt mixtures at a single test temperature regardless of the geographical location of the project (i.e., climatic conditions) or the location of the asphalt mixture within the pavement structure.
- Using a single rutting test criterion regardless of the aircraft traffic mix and volume.
- Allowing agencies to test compacted samples for rutting either at mix design AV or at 7 percent AV, while evaluating them against the same test criterion. Previous airfield research studies that led to current FAA rutting test criteria conducted the APA at mix design AV (FAA, 2018; Rushing, Little, & Garg, 2012). On the other hand, current standard test methods mentioned in the FAA advisory circular (i.e., AASHTO T 340 and AASHTO T 324) require preparing the samples at 7±0.5 percent AV (AASHTO, 2023a; AASHTO, 2023b).
- Recommending the current HWTT maximum rut depth criterion without enough research on the correlation of the HWTT set threshold with actual airfield pavement conditions.
- Considering only repeated load rutting tests, while lacking surrogate rutting tests that could be conducted efficiently and reliably during production at a representative testing frequency.

Objective and Scope

Considering the main differences between highway and airfield pavement conditions, and the potentially significant impact that test protocols can have on the resulting data, this research study thoroughly emphasized setting representative test protocols that best simulate actual airfield conditions. To make BMD implementation as efficient as possible, the following key factors were considered in this project when evaluating test methods and criteria for mix design, control strip, and QA:

- Specimen geometry.
- Specimen target AV level.
- Sample preparation (cutting, gluing, etc.).
- Conditioning/aging temperature and time.
- Test temperature.
- Test loading conditions.

The project's Phase I and Phase II tasks, with corresponding inputs and outputs to illustrate the relationship and dependence of information from each, are shown in Table 1. The flowcharts in Figure 1 and Figure 2 summarize the main approach followed in this study to achieve its objectives. The tasks and findings from Phase I were previously documented in an interim report (Hajj, et al., 2025b).

The primary deliverables of this study included the following:

- **Preliminary rutting test criteria.** Preliminary test criteria were established and presented in an interim project report (Hajj, et al., 2025b). These criteria were not based on experimental testing conducted during this project but were primarily derived from the literature review and limited mechanistic analysis.
- **Representative specimen conditions for rutting test methods.** This effort was heavily based on the review and analysis of actual airfield data for in-place densities. The findings were presented in Technical Memo 1 (Hajj, et al., 2025a) and Technical Memo 2 (Hajj, et al., 2025c).
- **Refined rutting test criteria.** The established criteria were based on comprehensive evaluations of laboratory-mixed laboratory-compacted (LMLC) and reheated plant-mixed laboratory-compacted (RPMLC) airfield asphalt mixtures complemented with mechanistic pavement analyses.
- **Verification of the recommended rutting test criteria.** The rutting test criteria were verified by testing field-mixed field-compacted (FMFC) samples (i.e., field cores) from airfield pavement sections that had experienced different levels of field rutting.
- **Interlaboratory study (ILS).** A simplified ILS was conducted to investigate the precision and bias of the candidate rutting test methods at the established testing conditions.

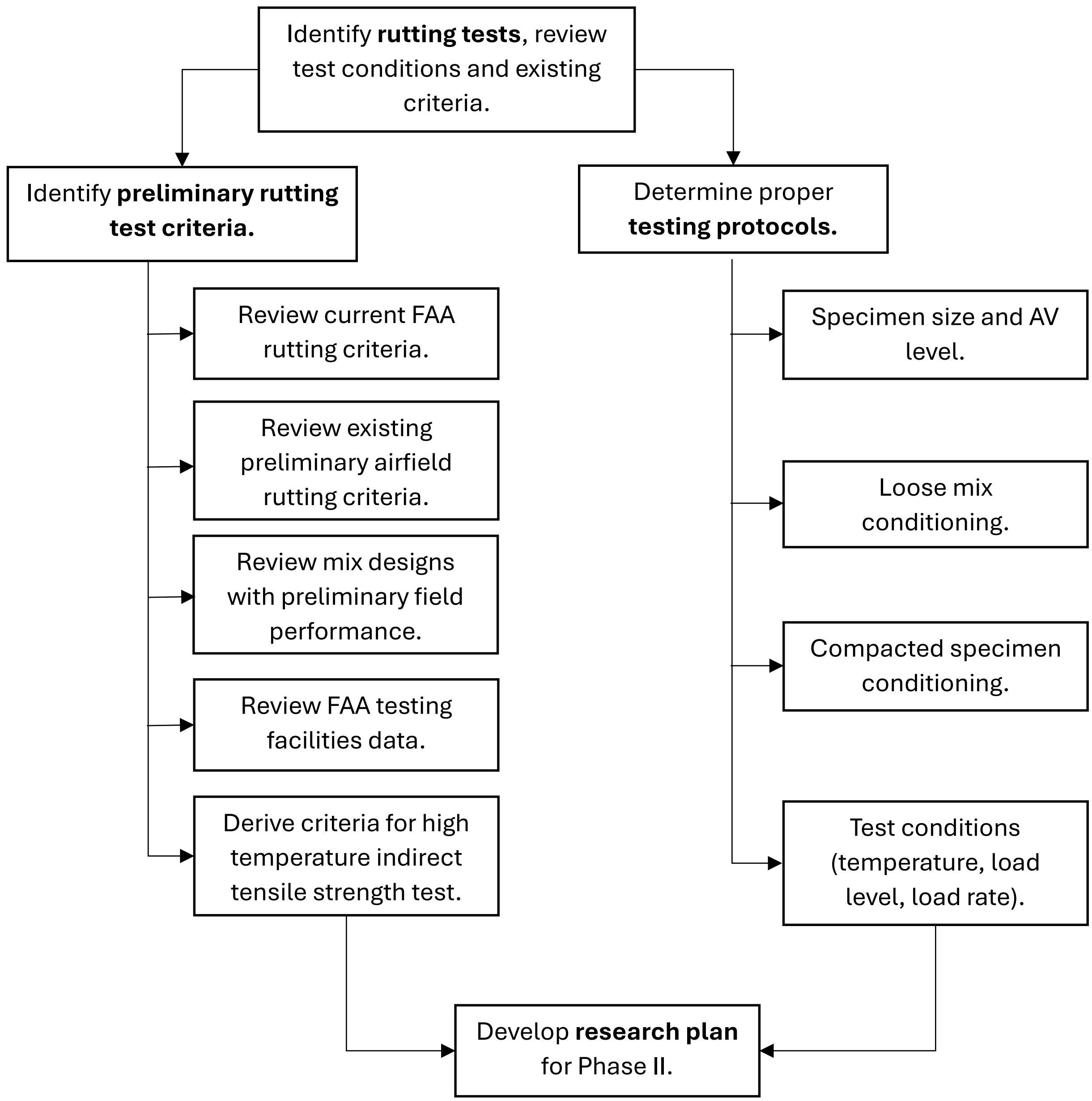
Table 1. Research Plan Breakdown Summarizing Associated Inputs and Outputs by Task

Phase1

Inputs	Task	Outputs
1. Airfield and highway data literature review. 2. FAA studies on airfield pavement performance and laboratory rutting tests review. 3. Asphalt mixtures performance review.	**Task 1** Gather Information	1. Available rutting test methods and specifications/laboratory criteria. 2. Handling, conditioning, aging, and testing protocols. 3. Validation of rutting tests. 4. Correlation between rutting tests. 5. Rutting performance of FAA mixtures.
1. Task 1 inputs and outputs. 2. Candidate rutting tests. 3. Candidate airfield pavements identification. 4. Field performance data review.	**Task 2** Verify and Establish Preliminary Test Criteria	1. Selection, planning, and sampling of materials from airfield projects. 2. Collection of mix designs and QA data. 3. Proposed preliminary test criteria.
1. Task 1 and 2 inputs and outputs. 2. Identified refinements to project team research approach. 3. Proposed rutting test methods. 4. Proposed handling, conditioning, aging, and testing protocols. 5. Proposed asphalt mixtures for laboratory evaluation.	**Task 3** Develop Phase II Research Plan	1. Updated Phase II research plan based on Task 1 and 2 findings with activities to execution experimental design detailed. 2. Proposed laboratory testing details and data analysis. 3. Updated schedule for delivering final deliverables.
1. Tasks 1–3 inputs and outputs. 2. Regular team meetings. 3. Project reporting and meeting requirements.	**Task 4** Prepare Interim Report	1. Interim report documenting Phase I, Tasks 1–3, for project panel review and consideration. 2. Project panel meeting.

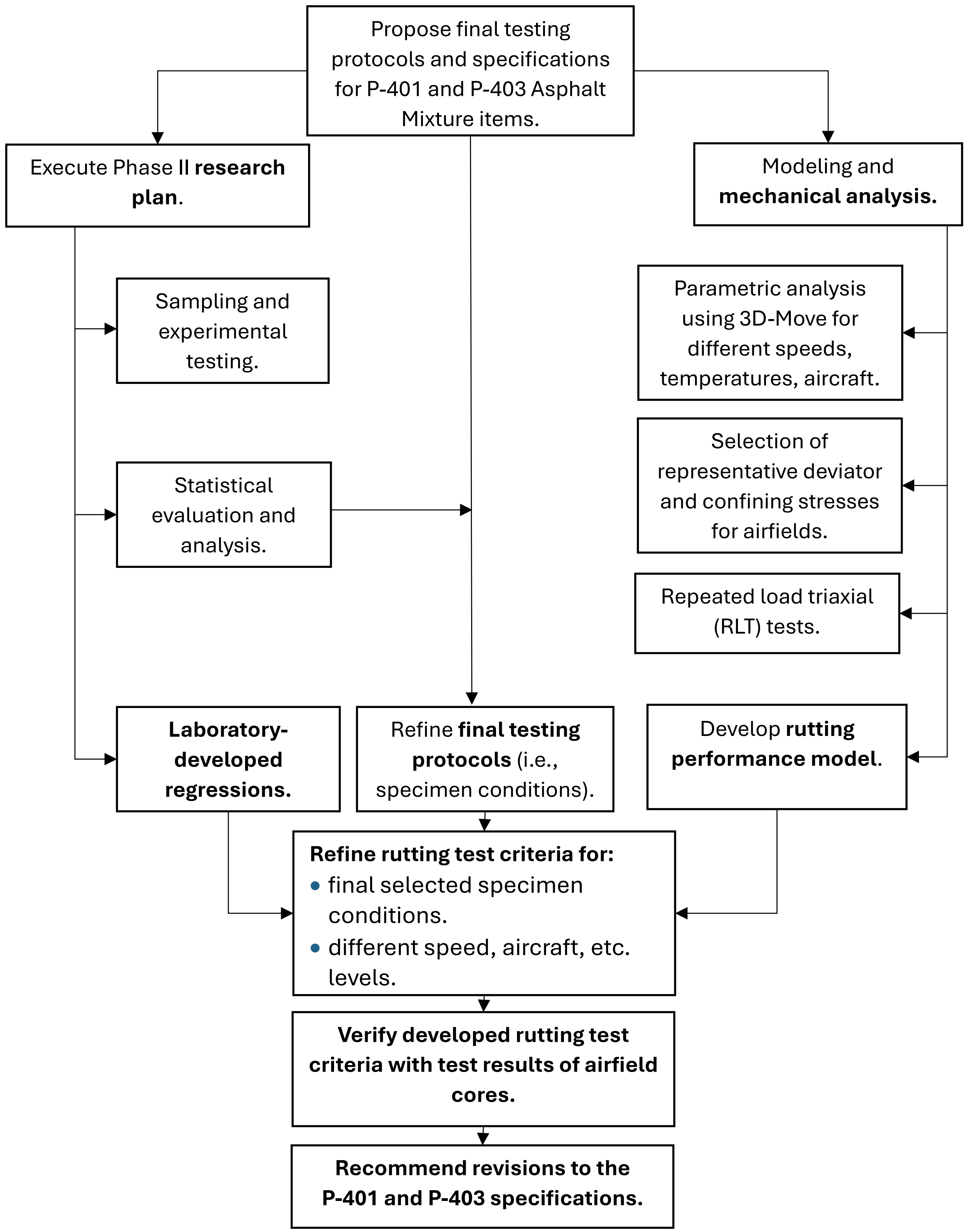
Phase II

Inputs	Task	Outputs
1. Raw materials sampling from recent airfield projects. 2. Plant/field loose asphalt mixtures and field cores sampling from recent airfield projects. 3. Rutting test protocols.	**Task 5** Execute Research Plan	1. Rutting test data for varying airfield asphalt mixtures evaluated under different test conditions. 2. Regressions between different rutting test parameters.
1. Analysis and statistical evaluation of the data generated from Task 5 outputs. 2. Refine testing protocols and select final testing parameters. 3. Mechanistic analysis and modeling to refine test criteria for selected testing protocols.	**Task 6** Recommend Revisions to FAA P-401 and P-403 Specifications	1. Selected final testing protocols for ease of BMD implementation during mix design as well as during production. 2. Proposed final test criteria and specifications for P-401 and P-403 airfield asphalt mixtures.
1. Task 6 outputs. 2. ILS to evaluate the variability of the recommended rutting test criteria.	**Task 7** Develop an Implementation Plan	1. Implementation plan targeting the applicability of the recommended test methods and criteria based on common field practices and industry feedback.
1. Tasks 5–7 inputs and outputs. 2. Regular team meetings. 3. Project reporting and meeting requirements.	**Task 8** Prepare Final Report and Deliverables	1. Final report documenting Tasks 5–7 for project panel review and consideration. 2. Project panel meeting. 3. Webinar summarizing details of the research approach, findings, and recommendations.



Source: University of Nevada, Reno

Figure 1. Phase I Overall Scope of Work



Source: University of Nevada, Reno

Figure 2. Phase II Overall Scope of Work

Literature Summary

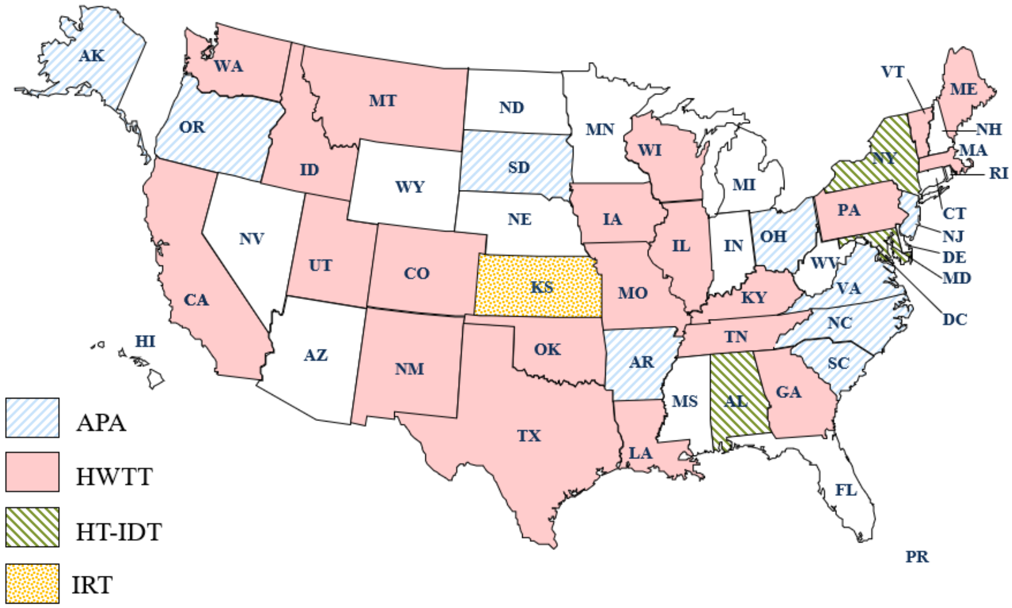
The literature review section elaborates on the different rutting test methods that were selected in this project for future specifications. Table 2 presents the four main rutting mechanical tests under two different modes of testing (i.e., monotonic and repeated loading) identified in this project to be candidates for mix design and/or acceptance during production of airfield AC pavements. Factors such as efficiency, practicality, common availability and affordable cost, repeatability, sensitivity to mixture components, simulation of rutting mechanism, and correlation with field performance were considered (Rushing, Little, & Garg, 2014; Zhou, et al., 2019; Christensen & Bonaquist, 2007; West, Rodezno, Leiva, & Yin, 2018; Hajj, Hand, Chkaiban, & Aschenbrener, 2019; Hajj, Aschenbrener, & Nener-Plante, 2022a).

Current implementation of rutting mechanical tests varies among State departments of transportation (DOTs) (Figure 3) (NCAT, 2023; NAPA, 2023; AASHTO, 2022d). Twenty-two State DOTs have implemented the HWTT, and nine have implemented APA, as per the National Asphalt Pavement Association (NAPA) BMD Guide published in June 2024 (NAPA, 2023; AASHTO, 2022d). Four State DOTs use the high temperature indirect tensile strength test (HT-IDT), while several others are currently investigating its use as part of BMD (e.g., Maine DOT). The ideal rutting test (IRT), for which a standard test method was recently published (ASTM D8360-22), is being implemented by two State DOTs and is currently being considered along with other rutting mechanical tests by several others (ASTM, 2022).

A detailed review of the literature on the application of the four candidate rutting tests for airfield pavements, along with relative test conditions, criteria, and findings, can be found in Appendix B (Hajj, et al., 2025b). Appendix B also presents the main findings on the repeatability of each rutting test and its sensitivity to several asphalt mixture characteristics.

Table 2. Candidate Rutting Mechanical Tests

Test	Standard Test Method	Mode of Testing	Outcome
APA	AASHTO T 340-23 (AASHTO, 2023b)	Repeated loading.	Rut depth.
HWTT	AASHTO T 324-23 (AASHTO, 2023a)	Repeated loading.	Rut depth, number of passes to failure, rutting resistance index, corrected rut depth.
HT-IDT	ASTM D6931-17 (ASTM, 2017) ALDOT-458 (Alabama DOT, 2022)	Monotonic.	Indirect tensile strength.
IRT	ASTM D8360-22 (ASTM, 2022)	Monotonic.	Rutting tolerance index (RT_{Index}).



Alabama implemented both HWTT and HT-IDT
Missouri implemented HWTT and IRT
Virginia implemented both APA and HT-IDT

Source: National Asphalt Pavement Association

Figure 3. Rutting Tests Implemented by State DOTs as Current State of Practice (NAPA, 2023)

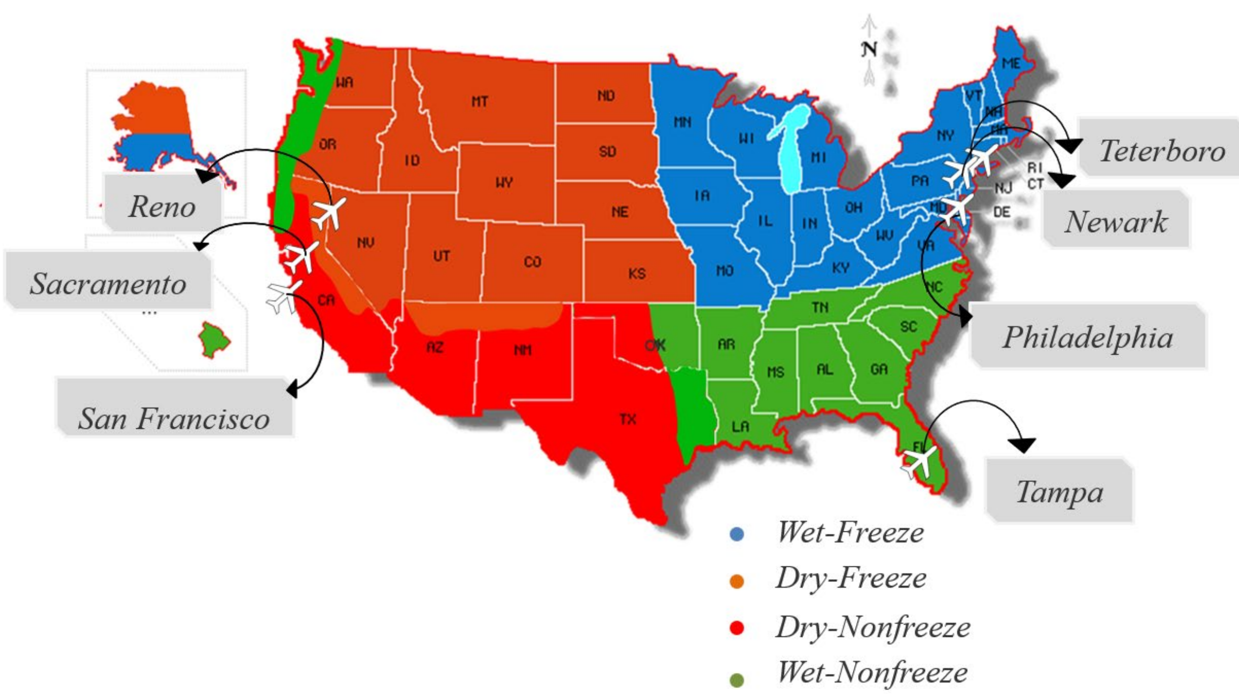
Several airfield research studies using APA, HWTT, and HT-IDT were documented in the literature. For the IRT, which was recently developed, the literature detailed some highway studies using this test. As expected, based on the substantial research studies, the APA test was the most common rutting mechanical test to evaluate the susceptibility of airfield asphalt mixtures to rutting. These research studies were used to establish the current rutting test criteria implemented in the FAA advisory circular (FAA, 2023). The highest within-laboratory COV reported for the monotonic tests was equivalent to 21.6 and 26.4 percent for the HT-IDT and IRT, respectively, compared to a 38 percent maximum within-laboratory COV for the APA test and 51.7 percent for the HWTT. As detailed in Appendix B, the four candidate rutting tests were shown to be sensitive to the main asphalt mixture components, such as asphalt binder content, asphalt binder grade, AV level, and gradation (Hajj, et al., 2025b).

Chapter 2. Research Plan

The primary goal when developing the research plan was to increase the variety of materials included in the testing matrix to encompass a wide range of factors. These factors included asphalt binder grade, modification type, mix design method, nominal maximum aggregate size (NMAS), and aggregate lithology. The following section elaborates on the materials sampled from several airfield projects, the testing protocols, the experimental testing plan, and the mechanistic pavement analysis approach adopted in this study.

Materials

Raw materials and plant-produced asphalt mixtures from various airfield projects in different Long-Term Pavement Performance (LTPP) climatic zones were collected to ensure a diverse range of materials for testing purposes (Schwartz, et al., 2015). The locations of these airfield projects are highlighted on the map in Figure 4. A summary of the airfield projects included in the study can be found in Table 3 along with their airport code, construction date, category, and hub size (according to the FAA classification), maximum GAW, and LTPP climatic zone. This information was gathered to provide a complete understanding of the conditions and factors affecting pavement performance in various locations across the United States.



Source: University of Nevada, Reno

Figure 4. Airports Identified for Sampling Materials

Table 3. Airport Locations and Mixture Characteristics of Sampled Projects (FAA, 2018; FAA, 2024b; FAA, 2021b; FAA, 2024d)

Airport Characteristics

Airport Code	EWR	PHL	RTS	SFO	SMF	TEB	TPA
Airport Name	Newark Liberty International Airport	Philadelphia International Airport	Reno Stead Airport	San Francisco International Airport	Sacramento International Airport	Teterboro Airport	Tampa International Airport
Construction Date	Aug.–Sept. 2022	May 2018	Oct. 2022	Spring 2023	Sept. 2022	July–Aug. 2022	Oct. 2022
Classification/ Hub	Primary/ Large	Primary/ Large	Reliever/ NA	Primary/ Large	Primary/ Medium	General aviation/ NA	Primary/ Large
GAW (lb)	>100,000	>100,000	≤100,000	>100,000	>100,000	≤100,000	>100,000
LTPP Climatic Zone	Wet-Freeze	Wet-Freeze	Dry-Freeze	Dry-Nonfreeze	Dry-Nonfreeze	Wet-Freeze	Wet-Nonfreeze

Asphalt Mixture Characteristics

Airport Code	EWR	PHL	RTS	SFO	SMF	TEB	TPA
Mixture Type	Modified P-401 Surface	P-401 Surface	P-401 Surface (bottom lift)	P-401 Surface	P-401 Surface	Modified P-401 Surface	P-404 Surface
PG	82-22	82-22	64-28NV	76-22M	76-22M	64-22[1]	82-22 Fuel-resistant
Gradation	Mix 2 (PANYNJ Specification Section 321218)	Grad 1 (401-3.3)	Grad 2 (401-3.3)	Grad 1 (401-3.3)	Grad 2 (401-3.3)	FAA Mix 3 (PANYNJ Specification Section 321218)	Grad 3 (401-3.3)
Aggregate Lithology	Gneiss, Braen Stone Industries, Sparta, NJ	Trap Rock, Dyer Quarry, Birdsboro, PA	Andesite, Lockwood, NV	A.R. Wilson Quarry, Aromas, CA	Alluvial sand and gravel, Western Aggregates, Inc.,	Gneiss, Tilcon New York, Inc., Mt. Hope, NJ	Granite, Nova Scotia
NMAS, mm	19	19	12.5	19	12.5	19	9.5

[1]Unmodified asphalt binder.

NA = not applicable; M = modified; NV = polymer-modified in accordance with Nevada DOT standard specifications; PANYNJ = Port Authority of New York and New Jersey.

Table 3 also summarizes the asphalt mixture characteristics of sampled airfield projects, including the corresponding asphalt mixture type, binder performance grade (PG), gradation, and NMAS. While none of the sampled projects included reclaimed asphalt pavement, all projects were identified as P-401 except for the following slight modifications (FAA, 2018):

- The Newark Liberty International Airport (EWR) and Teterboro Airport (TEB) mixtures are designed per the Port Authority of New York and New Jersey (PANYNJ) Specification Section 321218, which includes the requirements of FAA advisory circular 150/5370 Item P-401 with FAA-approved modifications.
- The Tampa International Airport (TPA) airfield mixture is designed as a fuel-resistant P-404 mixture targeting low permeability (finer gradation, lower compaction effort, lower design AV, etc.) and using a highly modified binder grade.

Testing Protocols

Incorporating a new test method into the specifications involves establishing proper test criteria as well as defining proper laboratory rutting test protocols covering the entire testing procedure. Testing protocols include several key parameters—such as laboratory compaction method, AV level, specimen size and preparation method (cutting, coring, etc.), loose mixture aging temperature and time, compacted specimen conditioning, test temperature, and load level—reflecting actual flexible airfield pavement conditions (Elias, 2024). Accordingly, the rutting test protocols were classified into the following four main categories:

- Specimen characteristics including specimen AV level, size, and preparation method.
- Conditioning of loose mixture prior to compaction.
- Conditioning of compacted specimen at relative test temperature.
- Test conditions including test temperature, load level, and load rate.

Changing the parameters of a test protocol can influence the final laboratory test results and lead to improper rutting characterization of asphalt mixtures. Therefore, representative rutting test protocols should be carefully defined to best simulate actual field conditions of airfield pavements. The two AV levels of 5±0.5 and 7±0.5 percent considered in the experimental plan were based on an analysis of in-place density data for an array of airfield projects (Hajj, et al., 2025a; Elias, 2024). The 5 percent AV matches the 75th percentile of mat core AV data, whereas the 7±0.5 percent AV range was further suggested based on in-place joint density (Hajj, et al., 2025a; Elias, 2024).

Moreover, the limitations of laboratory practices and sample preparation are one of the main considerations for future proposed specifications. While field conditions serve as a key reference for identifying appropriate laboratory test protocols, it is essential to investigate the practicality and effectiveness of any proposed test condition. As an example, the laboratory compaction effort needed to reach a representative AV level under a certain specimen height is one of the key parameters assessed in this study (Elias, 2024). This was done to avoid excessive compaction effort in the laboratory that may cause aggregate breakdown or damage to the asphalt mixture skeleton during laboratory compaction.

Accordingly, the experimental plan involved cutting the 62-mm specimens targeting 5±0.5 percent AV from thicker samples at a 165-mm height. On the other hand, the APA samples compacted to a 75-mm height reached the 5 and 7 percent target AV within a reasonable number of gyrations due to the thicker specimen geometry. Therefore, the cutting technique was solely examined for specimens targeting 5±0.5 percent AV at a 62-mm height (i.e., HWTT, HT-IDT and IRT) (Hajj, et al., 2025a; Elias, 2024). Further investigations of specimen preparation methods may evaluate the need for cutting to maintain a reasonable number of gyrations, based on the ratio of the mix NMAS to the specimen thickness at a certain AV level.

Table 4 summarizes the testing parameters selected in this study for the rutting mechanical tests of flexible airfield pavements and provides a brief justification for each. The justifications presented were based on a combination of the previous literature review, common test methods, and an analysis of actual airfield section data (Hajj, et al., 2025b).

Table 4. Selected Rutting Test Protocols for Airfield Asphalt Mixtures

Category	Factor	Project Recommendation	Justification
Specimen Characteristics	AV level.	5±0.5% (cut from 165-mm specimens). 7±0.5% (directly molded).	In-place density for airfield mat cores (cut to maintain reasonable gyration number) (Elias, 2024). In-place density for airfield joint cores (Elias, 2024).
	Specimen size.	APA: 150 mm by 75±2 mm. HWTT, HT-IDT, IRT: 150 mm by 62±1 mm (unified height for ease of implementation).	AASHTO T 340 (AASHTO, 2023b). AASHTO T 324, ASTM D6931, and ASTM D8360 (AASHTO, 2023a; ASTM, 2022; ASTM, 2020).
Loose Mixture Conditioning	Laboratory-prepared, loose mixture short-term oven aging.	2 hr at compaction temperature.	AASHTO R 30 (AASHTO, 2022a).
	Plant-mixed, loose mixture reheating.	5-gal bucket at compaction temperature for 90 min, then transferred to large pans and reheated for 60 min, followed by splitting per AASHTO R 47 (AASHTO, 2022b).	Protocol developed to minimize aging and conditioning time as much as possible (Elias, 2024).
	Lag time, laboratory-prepared, loose mixture compaction.	Sample mixing and compaction same day.	Set by the study to minimize asphalt mixture aging.
Compacted Specimen Conditioning	Dwell time.	Maximum 7 calendar days.	Set by the study to minimize aging while allowing enough time for cutting and drying.
	Pre-test conditioning to bring the sample to test temperature.[1]	APA: 6 hr ±10 min in the temperature-calibrated test chamber. HWTT: 45 min in water. HT-IDT/IRT: 60±5 min in temperature-controlled water bath.	AASHTO T 340 (AASHTO, 2023b). AASHTO T 324 (AASHTO, 2023a). ASTM D6931/D8360 (ASTM, 2022; ASTM, 2020).

Category	Factor	Project Recommendation	Justification
Test Conditions	Test temperature.	LTPPBind Online environmental PG (no grade bumping), 12.5 mm rut depth, 50% reliability, at surface (FHWA, 2024).	NCHRP 9-33 (Advanced Asphalt Technologies, LLC, 2011). Equipment limitations.
	Test load level.	APA: 250 psi (250 lb) and 100 psi (100 lb). HWTT: 158 lb.	AASHTO T 340 (AASHTO, 2023b). AASHTO T 324 (AASHTO, 2023a).
	Test load rate.	HT-IDT and IRT: 50 mm/min.	ASTM D6931/D8360 (ASTM, 2022; ASTM, 2020).

[1]All conditioning times were monitored through a sample probe to reach target test temperature.
LTPP = Long-Term Pavement Performance.

The detailed loose mixture conditioning protocols for laboratory and plant mixtures developed in this study, along with the dwell and lag time, were set in a manner to minimize any additional aging or stiffening in the asphalt mixtures throughout the testing (Elias, 2024; Hajj, et al., 2025b).

Lag time refers to the duration between asphalt mixture sampling and sample compaction. Considering that different lag times between laboratories may create discrepancy in test results, the prospective FAA rutting specifications would need to define and identify lag time. Another parameter that will be included in the future FAA rutting specifications is the dwell time, which refers to the duration between asphalt mixture compaction and mechanical testing.

It is worth noting that while the experimental plan of this study investigated testing at two different AV levels (as shown in Table 4), a single AV level was ultimately further considered for the final specifications.

Experimental Testing Plan

A detailed laboratory experimental plan was developed to ensure high consistency with minimal variability in testing results between the different laboratories of the research team. The first step in the experimental plan was to verify that the plant-produced asphalt mixtures as well as the raw materials conform to the job-mix formula (JMF) within the acceptable production tolerance prescribed by FAA specifications (FAA, 2018). The mix design verification for the plant-produced mixtures included measurements of the asphalt binder content, aggregate gradation, and volumetric properties.

Accordingly, the theoretical maximum specific gravity, bulk specific gravity, and asphalt binder content (by centrifuge extraction) of the sampled plant-produced mixtures were evaluated against the JMF or production data (i.e., quality control [QC] and acceptance). The extracted aggregate gradations were assessed based on the control chart limits set in the FAA advisory circular (FAA, 2018). The detailed mix design verification data and experimental test results can be found in Appendix D.

The rutting tests were conducted under various predefined conditions that are summarized in Table 5. As previously noted, to ensure consistency during testing across the involved laboratories, the raw and extracted aggregate gradations were verified with the JMF control chart limits (i.e., limits for individual measurements and range-based limits). The experimental matrix incorporated five rutting mechanical tests performed by the following American Association of State Highway and Transportation Officials (AASHTO)-accredited laboratories:

- Western Regional Superpave Center (WRSC) laboratory at the University of Nevada, Reno.
- Center for Advanced Infrastructure and Transportation (CAIT) at Rutgers University.
- Texas A&M Transportation Institute (TTI) at Texas A&M University.

The rutting tests were defined at the set AV levels, along with the test temperature, specimen size, specimen preparation method, and number of replicates for each condition.

Table 5. Experimental Matrix for Rutting Tests

Test Temperature	LTPPBind Online environmental PG (no grade bumping), 12.5-mm rut depth, 50% reliability, at surface
Loose Mix Conditioning (Hajj, et al., 2025b)	Based on developed LMLC loose mix conditioning protocol (2 hr at compaction temperature)

Test	APA (100 psi/100 lb)		APA (250 psi/250 lb)		HWTT		HT-IDT		IRT	
AV Level, %	5±0.5	7±0.5	5±0.5	7±0.5	5±0.5	7±0.5	5±0.5	7±0.5	5±0.5	7±0.5
Specimen Diameter, mm	150	150	150	150	150	150	150	150	150	150
Specimen Height, mm	75	75	75	75	165	62	165	62	165	62
Cutting (Yes or No)	No	No	No	No	Yes	No	Yes	No	Yes	No
Final Specimen Height, mm	75	75	75	75	62	62	62	62	62	62
Number of Samples per Mixture per Combination	4 or 6	4 or 6	4 or 6	4 or 6	4	4	3	3	3	3
Total Number of Samples per Mixture	12		12		8		6		6	

Mechanistic Pavement Analysis

Following the experimental testing, a thorough mechanistic analysis and modeling of actual airfield pavements under representative aircraft loading conditions was conducted. The mechanistic analysis accounted for key differences between highway and airfield pavements, including variations in load levels, tire pressure, axle configuration, and pavement structure.

The project aimed to establish a mechanistic framework to refine the current FAA APA 250 psi/250 lb test criterion for different speeds and aircraft. The criteria adjustments were founded on airfield pavement mechanistic responses coupled with rutting performance models developed from the RLT test. The first step in the framework involved analytical modeling using 3D-Move Analysis software to evaluate airfield pavement responses under different loading conditions. The 3D-Move results were used for the following two main purposes:

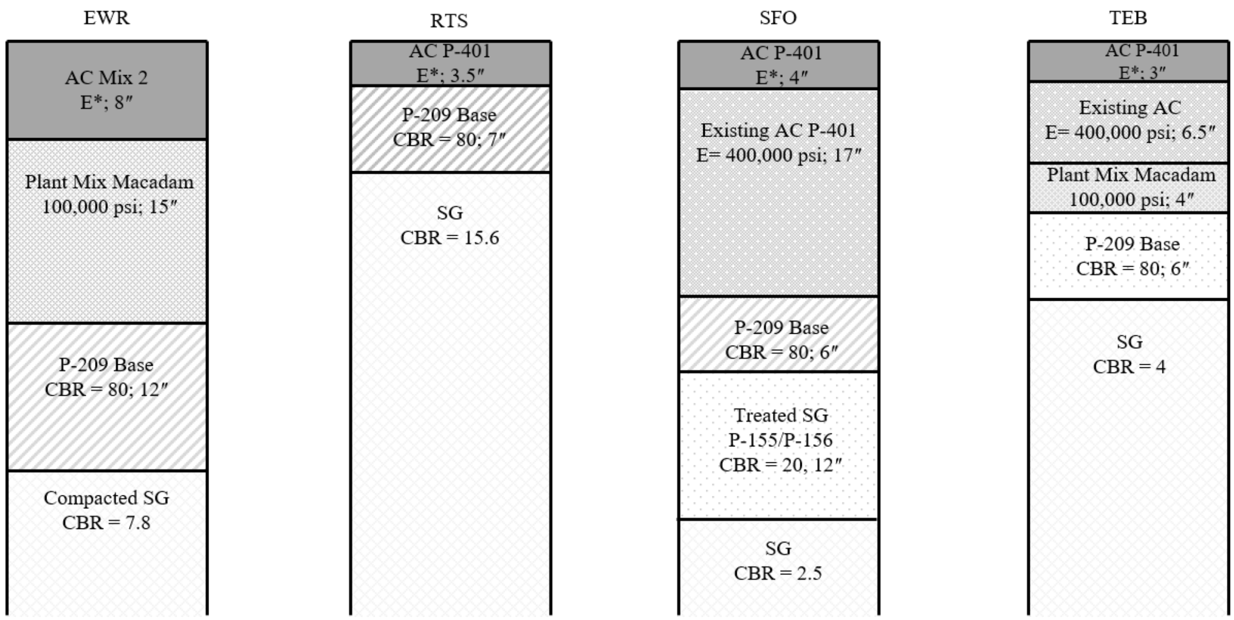
- To calculate confining and deviatoric stress conditions that are representative of actual airfield pavements for use in the RLT test.
- To determine the load-induced resilient strains in the AC layers for use in the rutting performance models generated from the RLT test.

In summary, the mechanistic-empirical framework developed in this study comprises three main components: 3D-Move modeling, experimental testing (i.e., dynamic modulus [E*] and RLT), and a sensitivity analysis of rut depth to key parameters based on the developed rutting models. This approach was applied to four airfield pavement projects representing different LTPP climatic zones: Newark Liberty International Airport (EWR), Reno Stead Airport (RTS), San Francisco International Airport (SFO), and Teterboro Airport (TEB). These projects consisted of asphalt mixtures with unmodified and polymer-modified asphalt binders of different grades. The pavement structures for these airfields and relevant airport characteristics are summarized in Table 5 and Table 6, respectively.

Table 6. Characteristics of Modeled Airfields for Mechanistic Analysis

Airport Code	Airport	Construction Date	Classification/Hub	GAW (lb)	LTPP Climatic Zone
EWR	Newark Liberty International Airport	Aug.–Sept. 2022	Primary/Large	>100,000	Wet-Freeze
RTS	Reno Stead Airport	Oct. 2022	Reliever/NA	≤100,000	Dry-Freeze
SFO	San Francisco International Airport	Spring 2023	Primary/Large	>100,000	Dry-Nonfreeze
TEB	Teterboro Airport	July–Aug. 2022	General aviation/NA	≤100,000	Wet-Freeze

NA = not applicable.



CBR = California Bearing Ratio; E = Modulus of Elasticity; E* = Dynamic modulus for viscoelastic materials; SG = subgrade.

Source: University of Nevada, Reno

Figure 5. Pavement Structures of Modeled Airfield Projects

To refine the rutting test criteria for airfield pavements, mechanistic modeling using 3D-Move included a parametric analysis focusing on three key parameters: temperature, speed, and aircraft type. Given the viscoelastic behavior of asphalt mixtures, three temperatures and three speeds were selected for the analysis, as shown in Table 7. The selected temperatures included the following:

- The environmental testing temperature based on the LTPPBind Online environmental PG at surface with 50 percent reliability and 12.5 mm rut depth.
- The 64 °C temperature specified in the current FAA APA 250 psi/250 lb test (AASHTO, 2023b; FHWA, 2024).

The environmental testing temperature in this study was set to reflect actual environmental conditions at each airfield project (Hajj, et al., 2025b). For speed parameter, 5 mph and 15 mph were selected to represent the critical minimum speeds typically observed on taxiways, while 45 mph was selected to simulate loading due to a slow-moving aircraft on a runway (Christensen, 2008; FAA, 2024a).

Moreover, five aircraft types were modeled for each evaluated airfield project. These aircraft were selected after evaluating the air traffic mix observed at each airport in 2023, as illustrated in the SFO example in Figure 6. Each modeled aircraft falls within one of the four GAW categories, as shown in Table 7, except the last category (>100,000 lb), which includes two aircraft: Boeing 737 with single axle dual tire configuration and Boeing 777 with a tridem axle dual tire configuration (FAA, 2021c).

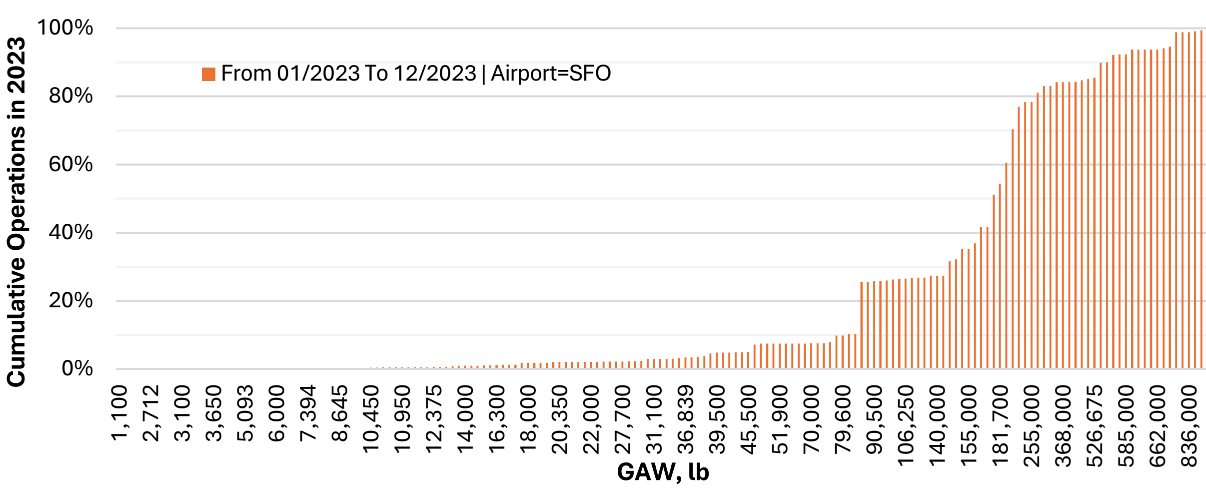
Table 7. Input of Parametric Analysis

Parameter	EWR	RTS	SFO	TEB
Temperature, °C	52, 58[1], and 64	52[1], 58, and 64	40[1], 52, and 64	52[1], 58, and 64

Parameter	Category	Value
Speed, mph (Christensen, 2008; FAA, 2024a)	Taxiway	5 and 15
	Runway	45
Traffic (FAA, 2021c)	Maximum GAW, lb	Aircraft
	≤12,500	Beechcraft King Air B200 (MTOW = 12,590 lb)
	<60,000	CRJ2–Bombardier CRJ-200 (MTOW = 47,700 lb)
	≤100,000	GA6C–G-7 Gulfstream G600 (MTOW = 94,600 lb)
	>100,000	B38M–Boeing 737 MAX 8 (MTOW = 181,700 lb)
		B772–Boeing 777-200 (MTOW = 547,000 lb)

[1]Environmental baseline testing temperature, LTPPBind Online final PG at surface with 50% reliability, 12.5 mm target rut depth, without grade bumping.

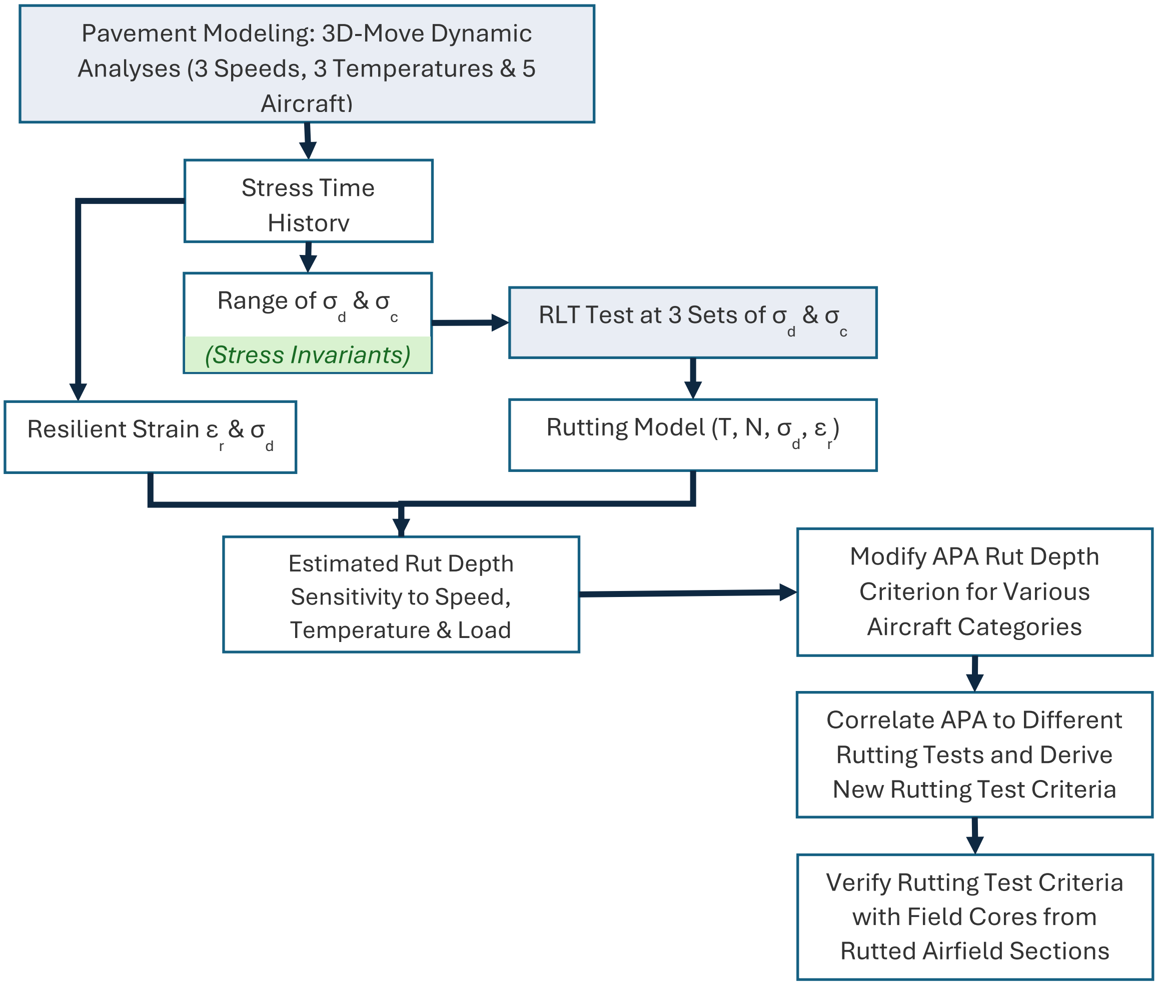
MTOW = maximum takeoff weight.



Source: University of Nevada, Reno.

Figure 6. Air Traffic Mix for San Francisco International (SFO) Airport

The mechanistic-empirical approach and the experimental verification followed in this study are outlined in Figure 7. Following the sequential steps of the methodology, the derived rutting test criteria were based on FAA's current specification for the APA 250 psi/250 lb test (i.e., maximum 10 mm rut depth after 4,000 cycles at design AV) that was selected due to its proven historical performance (Hajj, et al., 2025b). This APA test criterion has shown strong correlation with accelerated airfield testing and actual airfield sections, effectively delineating asphalt mixtures with poor field rutting performance in most cases (Hajj, et al., 2025b).



Source: University of Nevada, Reno

Figure 7. Flowchart of the Mechanistic Empirical Methodology

A series of dynamic modulus (E*) and RLT tests were conducted in the laboratory for the four airfield mixtures associated with the selected airfield pavements (Table 7). The E* tests were performed to obtain the moduli and phase angle values of the AC surface layers in the 3D-Move dynamic analyses. The RLT tests were conducted to develop the rutting performance models for each of the four evaluated airfield asphalt mixtures.

The experimental testing was carried out at an AV level of 5±0.5 percent, which represents the 75th percentile of the in-place density for airfield mat cores. The rational for selecting this AV level was based on detailed analysis of airfield in-place density data that can be found in the referenced appendix (Hajj, et al., 2025a). The four airfield mixtures selected for the mechanistic empirical framework and the characteristics are summarized in Table 8. For each airfield project, RPMLC samples from the surface AC layer were collected and tested in the laboratory for E* and RLT laboratory tests.

The last step of this effort consisted of verifying the established refined test criteria with FMFC samples from additional airfield pavement projects. These projects were selected to include pavements with minimal and significant rutting in the field. The FMFC samples were tested using different laboratory mechanical rutting tests. The rutting test results were then compared against the established threshold values. This step ensured that the selected test criteria could effectively identify airfield asphalt mixtures prone to poor rutting performance in the field.

Table 8. Asphalt Mixture Characteristics for the Evaluated Airfield Projects

Airport	Mixture Type	PG	Gradation	Aggregate Lithology	NMAS, mm
EWR	Modified P-401 Surface	82-22	Mix 2 (PANYNJ Specification Section 321218)	Gneiss from Braen Stone Industries, Sparta, NJ	19
RTS	P-401 Surface (bottom lift)	64-28NV	Grad 2 (401-3.3)	Andesite, Lockwood, NV	12.5
SFO	P-401 Surface	76-22M	Grad 1 (401-3.3)	A.R. Wilson Quarry, Aromas, CA	19
TEB	Modified P-401 Surface	64-22[1]	FAA Mix 3 (PANYNJ Specification Section 321218)	Gneiss from Tilcon New York, Inc., Mt. Hope, NJ	19

[1]Unmodified; M = modified; NV = polymer-modified in accordance with Nevada DOT standard specifications.

Chapter 3. Findings

Final Specimen Test Conditions

As previously mentioned, several specimen types and test parameters were evaluated in this project for implementation in the FAA BMD framework. The recommended specimen parameters and test conditions were tailored to address the following three main aspects: actual airfield pavement conditions, laboratory practice limitations, and practical sampling techniques during production.

Potential challenges associated with the recommended test conditions were assessed across the three main stages of implementation: mix design, initial production (e.g., control strip), and acceptance during final production. The challenges considered included excessive laboratory compaction effort, extended sample preparation and testing times during production, high variability in test results, and inconsistent outcomes between rutting tests.

Based on a thorough analysis and the relationships between varying rutting laboratory mechanical tests, the three AV scenarios outlined in Table 9 were examined for potential inclusion in the FAA AC 150/5370-10H specifications for airfield asphalt mixtures. A detailed comparative analysis was conducted, highlighting the relative pros and cons and potential solutions for each scenario across the three implementation stages.

Based on the experimental results and analyses, testing cut samples at 5 percent AV in the HWTT under wet conditions could lead to stripping failures, compromising appropriate rut depth evaluation. Considering that several agencies currently use the wet HWTT and may lack temperature-controlled chambers for dry testing, a 7±0.5 percent AV level was selected for the final FAA BMD framework. This AV level allows for direct molding of specimens to 75 mm for the APA or to 62 mm for the HWTT, HT-IDT, and IRT.

Table 9. Candidate AV Scenarios and Related Specimen Preparation

Rutting Test (Specimen Height)	Scenario 1: 5±0.5% for APA, HWTT, HT-IDT and IRT	Scenario 2: 5±0.5% for APA, HT-IDT, and IRT (excluding HWTT)	Scenario 3: 7±0.5% for APA, HWTT, HT-IDT, and IRT
APA (75 mm)	Directly molded specimens	Directly molded specimens	Directly molded specimens
HWTT (62 mm)	Cut specimens	Not considered	Directly molded specimens
HT-IDT (62 mm)	Cut specimens	Cut specimens	Directly molded specimens
IRT (62 mm)	Cut specimens	Cut specimens	Directly molded specimens

Rutting Test Results

The experimental plan involved testing both LMLC and RPMLC specimens. This approach was selected to address BMD implementation at the mix design stage and during production. According to the rutting test protocols defined in Table 4, LMLC and RPMLC samples were evaluated at AV levels of 5±0.5 and 7±0.5 percent AV. For the 5±0.5 percent AV level, samples were cut to 62 mm for the HWTT, HT-IDT, and IRT, versus directly molded to 75 mm for the APA. At the 7±0.5 percent AV level, samples were directly molded for all rutting tests.

An exception was made for the fuel-resistant P-404 TPA mixture, which showed a much lower number of gyrations (about four gyrations) to reach the 7 percent AV level. Further investigation of the in-place density of three different P-404 airfield projects (including TPA) indicated a 75th percentile of 3.3 and 6.7 percent AV for the mat and joint cores, respectively, compared to 5.2 and 7.7 percent AV for P-401/P-403 airfield projects. These in-place densities align with the 2.5 percent design AV for P-404 mixtures, rather than 3.5 percent for P-401/P-403 mixtures, making them much easier to compact to 7 percent AV. Consequently, the P-404 TPA project was tested at 3 and 5 percent AV, while the remaining airfield projects were tested at 5 and 7 percent AV (Hajj, et al., 2025a).

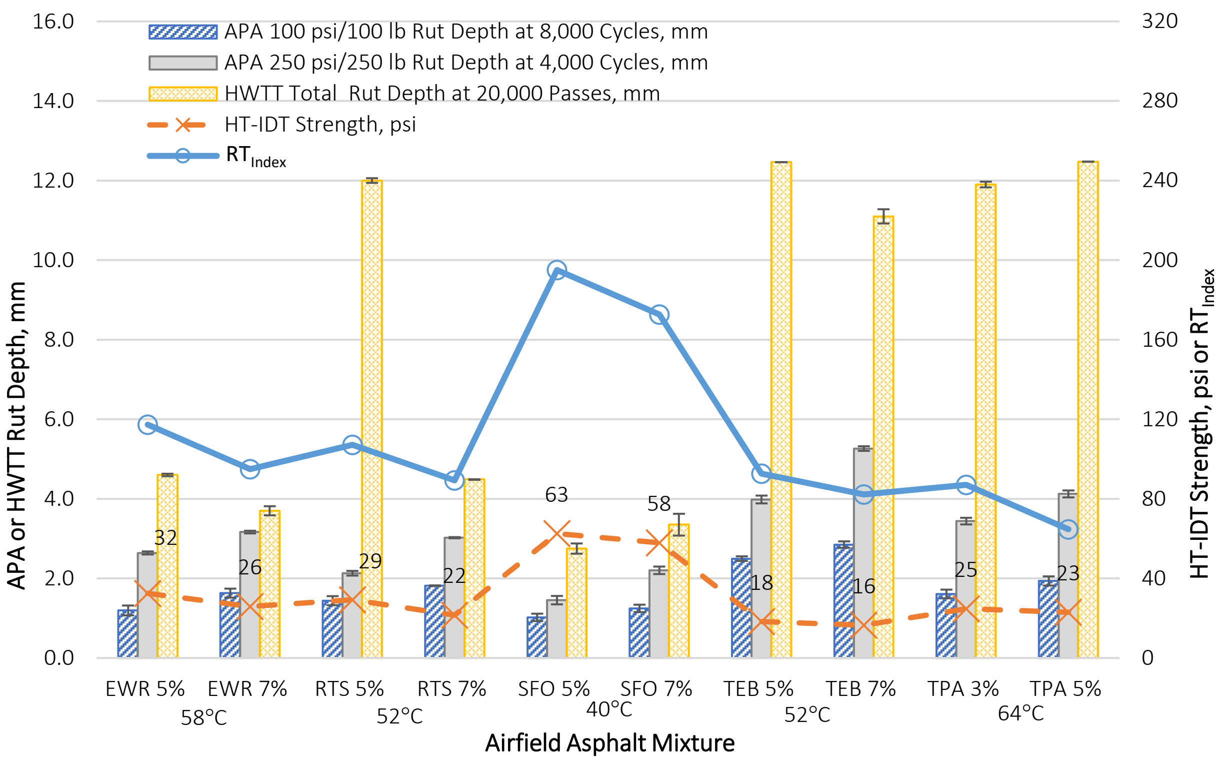
The testing temperatures were determined based on the LTPPBind high PG at surface and 50 percent reliability with no adjustment. The selected temperatures were 40 °C for SFO, 52 °C for RTS and TEB, 58 °C for EWR, and 64 °C for TPA (FHWA, 2024).

LMLC Results

The results of LMLC samples for the different airfield projects are plotted in Figure 8, including the relative AV level, test temperature, and 95 percent confidence interval (CI) bars. The data in Figure 8 show consistent trends between the APA 100 psi/100 lb rut depth at 8,000 cycles and the APA 250 psi/250 lb rut depth at 4,000 cycles, with rut depths ranging from 1.0 to 2.9 mm and 1.5 to 5.3 mm, respectively. Interestingly, the APA 250 psi/250 lb test showed a higher potential to discriminate between the rutting performance of the four different airfield mixtures. The large spread of the APA 250 psi/250 lb test data indicates a higher sensitivity of this test to different asphalt mixture characteristics (e.g., binder PG, NMAS, binder content) and testing temperatures.

Regardless of the binder PG used, the asphalt mixtures were tested at the relative environmental test temperature specific to each project's location. For example, the SFO mixture with a high PG 76 exhibited better rutting resistance at 40 °C compared to the EWR mixture with a high PG 82 tested at 58 °C. The maximum rut depths observed were 2.9 and 5.3 mm for the APA 100 psi/100 lb and APA 250 psi/250 lb tests, respectively, both for the TEB mixture with a neat binder PG 64-22, tested at 7 percent AV and 52 °C.

On the other hand, consistent trends were observed between the HT-IDT strength values and RT_{Index}, as shown in Figure 8, where SFO outperformed all other airfield mixtures at the respective testing temperatures. For each of the evaluated asphalt mixtures, improved rutting performance was noted in all tests conducted at the lower AV level, except for the HWTT. The HWTT results showed inconsistent trends among samples with different AV levels and preparation methods. Specifically, samples at 5 percent AV for three of the projects exhibited unexpectedly higher rut depths after 20,000 passes compared to the 7 percent AV samples, suggesting that sample preparation (i.e., cutting versus directly molding) had a significant impact on the HWTT results.



Source: University of Nevada, Reno

Figure 8. Rutting Test Results for LMLC Airfield Asphalt Mixtures
(Error bars represent the mean plus or minus the 95 percent confidence interval.)

As per the experimental plan in Table 4, the HWTT, HT-IDT, and IRT were conducted on two sample types: samples cut to 62 mm targeting 5±0.5 percent AV and samples directly molded to 62 mm targeting 7±0.5 percent AV. To maintain representative surface texture under the wheel in the HWTT, cut specimens at 5±0.5 percent AV were tested with the cut face placed at the bottom and the uncut surface under the wheel. Nonetheless, relatively higher variability and unorthodox trends were observed when testing cut specimens with the HWTT under wet conditions. The AV gradient in the specimens after cutting and the porewater pressure buildup within the cut specimens may have impacted the HWTT results, highlighting the influence of specimen preparation methods on the HWTT rut depths. It was noticed that stripping failure occurred in many cases due to the potential water infiltration in cut specimens, leading to higher rut depths and irrational trends with AV levels. This issue was encountered with three airfield mixtures (EWR, RTS, and TEB) where samples cut at 5 percent AV showed higher rut depths than those directly molded to 7 percent AV.

To investigate these observations, additional preparation methods with the HWTT using two airfield mixtures (RTS and TEB) and a highway asphalt mixture were studied. The different specimen types examined included the following (Hajj, et al., 2025c; Elias, 2024):

- No cut: HWTT was performed on directly molded samples without any surface cut.
- Top cut: The cut surface of the compacted sample was in immediate contact with the Hamburg wheel (i.e., Hamburg wheel running on the cut surface).
- Top and bottom cut: Both sides of the compacted sample were cut.
- Bottom cut: The uncut surface of the compacted sample was placed under the Hamburg wheel (i.e., Hamburg wheel running on the uncut surface).

The HWTT rut depth at 20,000 passes indicated once again higher rut depths and variability in the test results for bottom-cut samples compared to other specimen types. This led to the unexpected trends of rut depth with AV levels (RPubs by RStudio, 2024). Given the inconsistency of HWTT results with other rutting tests and the unreliable trends of rut depth with AV percent when using cut samples, it was recommended that HWTT should only be used with directly molded samples. This will help avoid the water infiltration and early stripping failure observed in cut samples.

Table 10 summarizes the statistical ranking of the evaluated asphalt mixtures based on the test results and a 95 percent confidence level. This analysis was performed using the Games-Howell test, which provides confidence intervals for group mean differences and indicates whether each pairwise comparison is statistically significant. Games-Howell does not assume an equal variance between the different groups, which was the case in some of the analyzed test results for the airfield mixtures (RPubs by RStudio, 2024). The ranking results at 5 and 7 percent AV confirm again a similar statistical ranking between the HT-IDT strength values and RT_{Index}. In general, the ranking was consistent with the APA test, whereas the HWTT rut depth at 20,000 passes showed the least potential to statistically delineate among the different asphalt mixtures at the 5 percent significance level.

Table 10. Statistical Ranking of LMLC Samples

Ranking at 5% AV[1]					
Mixture, AV (LMLC)	APA 100 psi/100 lb Rut Depth at 8,000 Cycles, mm	APA 250 psi/250 lb Rut Depth at 4,000 Cycles, mm	HWTT Total Rut Depth at 20,000 Passes, mm	HT-IDT Strength, psi	RT_{Index}
EWR 5% (58 °C)	1	3	1	2	2
RTS 5% (52 °C)	1	2	2	2	2
SFO 5% (40 °C)	1	1	1	1	1
TEB 5% (52 °C)	2	3	2	3	3
TPA 3% (64 °C)	1	3	2	3	3

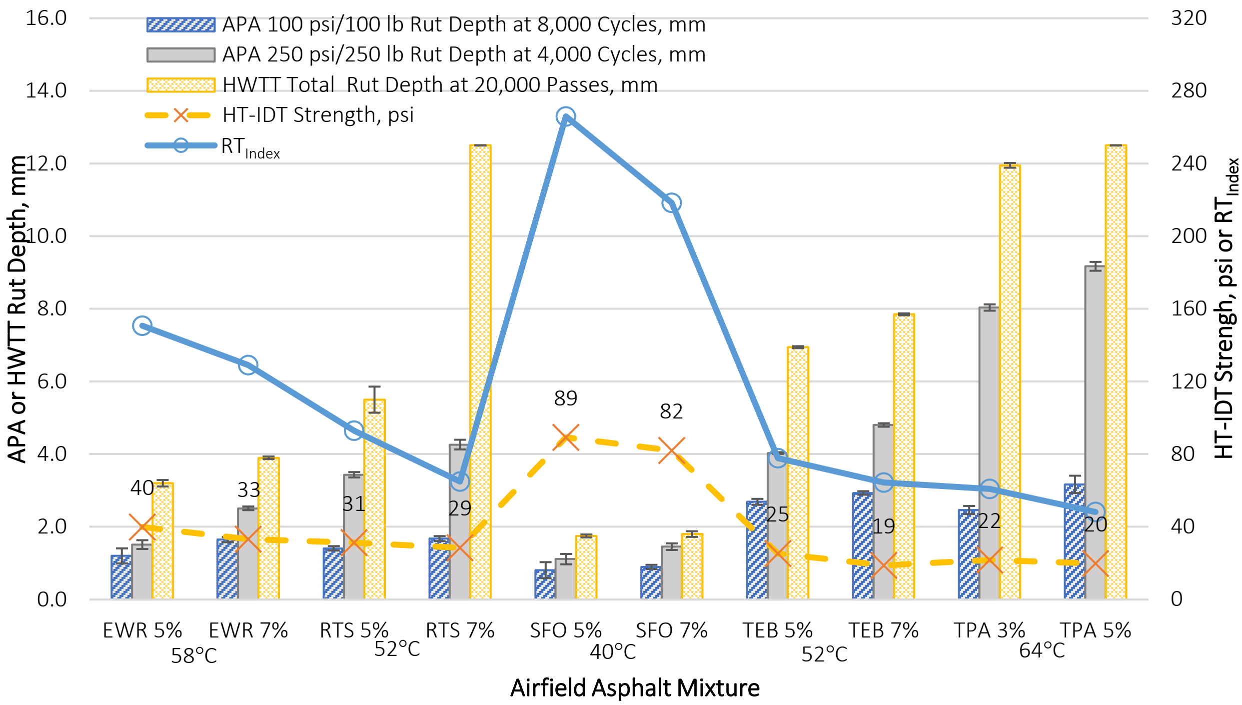
Ranking at 7% AV[1]					
Mixture, AV (LMLC)	**APA 100 psi/100 lb Rut Depth at 8,000 Cycles, mm**	**APA 250 psi/250 lb Rut Depth at 4,000 Cycles, mm**	**HWTT Total Rut Depth at 20,000 Passes, mm**	**HT-IDT Strength, psi**	**RT_{Index}**
EWR 7% (58 °C)	1	2	1	2	2
RTS 7% (52 °C)	2	1	1	2	2
SFO 7% (40 °C)	1	1	1	1	1
TEB 7% (52 °C)	3	3	1	3	3
TPA 5% (64 °C)	2	2	2	2	3

[1]A higher rank (i.e., lower number) indicates statistically better rutting resistance (i.e., lower rut depth, higher HT-IDT strength, higher RT_{Index}).

RPMLC Results

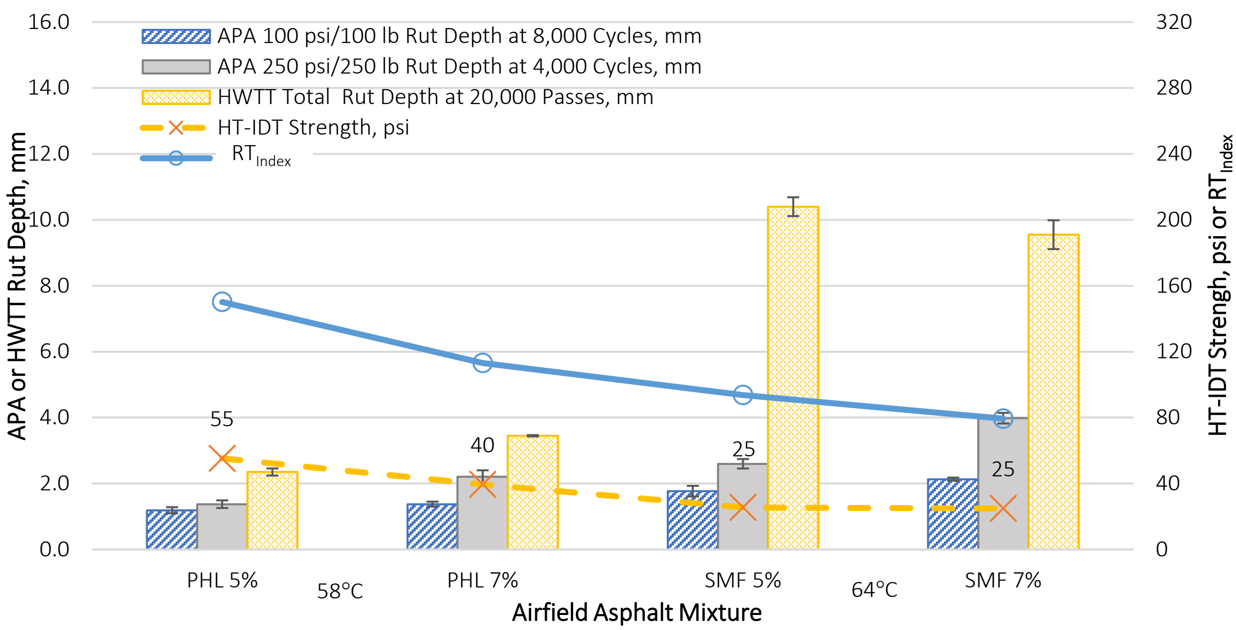
Loose plant mixtures for two additional airfield projects were sampled to expand the testing matrix of RPMLC samples. The experimental plan and protocols for RPMLC sample testing followed those established for LMLC samples, with the exception that cut samples were excluded from the HWTT. Instead, the HWTT was conducted on directly molded RPMLC samples at both 5 and 7 percent AV levels. Consequently, RPMLC samples molded directly to a 62 mm height at 5 percent AV required a greater number of gyrations compared to LMLC samples prepared by cutting from a 165-mm specimen at the same AV level. Specifically, molding RPMLC samples to a 62-mm height and 5 percent AV required 3.2 times the first locking point, whereas LMLC samples only required 1.4 times the first locking point.

Similar to the LMLC results, the RPMLC data presented in Figure 9 and Figure 10 indicate overall consistency across the APA, HT-IDT, and IRT tests, particularly between the two APA tests and the two monotonic tests (note that the error bars represent the mean plus or minus the 95 percent confidence interval). The results confirm that HT-IDT and IRT are suitable for use alongside the APA test or as surrogate rutting tests during production.



Source: University of Nevada, Reno

Figure 9. Rutting Test Results for RPMLC Airfield Asphalt Mixtures (Error bars represent the mean plus or minus the 95 percent confidence interval.)



Source: University of Nevada, Reno

Figure 10. Rutting Test Results for the Additional RPMLC Airfield Asphalt Mixtures (Error bars represent the mean plus or minus the 95 percent confidence interval.)

Excluding the cut samples from the HWTT showed a slight improvement in the correlation between HWTT rut depth after 20,000 passes and the AV level. However, stripping failures observed in some mixtures resulted in higher rut depths, which misled the correlation with other mechanical rutting tests. Consequently, additional test parameters beyond the total rut depth at 20,000 passes were investigated to mitigate the confounding effects of stripping failures. These parameters are discussed in the following sections.

A comparison of the statistical ranking of mixtures in Table 11 suggests that HT-IDT and IRT have the greatest potential to capture statistically significant differences in mixtures' resistance to rutting. In contrast, the HWTT did not effectively distinguish between the evaluated asphalt mixtures, particularly at the lower AV level.

As previously mentioned, the statistical ranking was based on the Games-Howell method, a post-hoc statistical test used to compare group means after an analysis of variance (ANOVA) when variances are unequal or sample sizes differ. The Games-Howell method performs pairwise comparisons with adjusted confidence intervals to control the Type I error rate, providing reliable rankings of group differences. The test is based on Welch's degrees of freedom correction and uses Tukey's studentized range distribution to compute p-values when the evaluated statistical groups do not exhibit equal variances.

Table 11. Statistical Ranking of RPMLC Samples

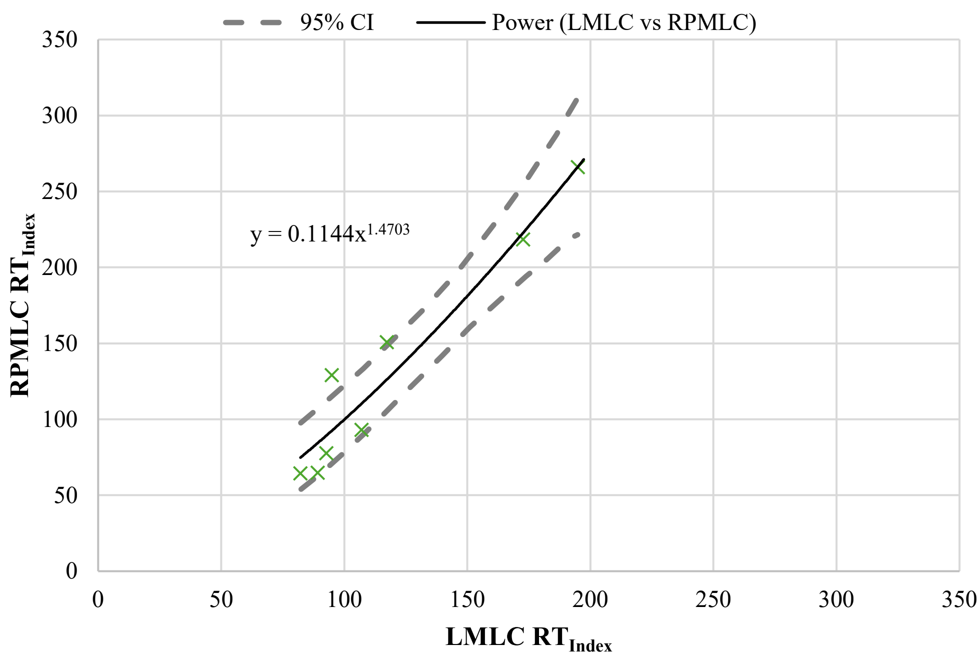
	Ranking at 5% AV | 7% AV[1]				
Mixture, AV (RPMLC)	APA 100 psi/ 100 lb Rut Depth at 8,000 Cycles, mm	APA 250 psi/ 250 lb Rut Depth at 4,000 Cycles, mm	HWTT Total Rut Depth at 20,000 Passes, mm	HT-IDT Strength, psi	RT_{Index}
EWR (58 °C)	1 | 2	1 | 2	1 | 1	3 | 2	2 | 2
PHL (58 °C)	1 | 2	1 | 1	1 | 1	2 | 2	2 | 2
RTS (52 °C)	2 | 2	2 | 2	1 | 2	4 | 3	3 | 3
SFO (40 °C)	1 | 1	1 | 1	1 | 1	1 | 1	1 | 1
SMF (64 °C)	2 | 2	2 | 2	1 | 1	5 | 3	3 | 3
TEB (52 °C)	3 | 2	2 | 3	1 | 1	5 | 4	3 | 4
TPA[2] (64 °C)	3 | 2	3 | 3	1 | 2	5 | 4	4 | 4

[1]A higher rank (i.e., lower number) indicates a statistically better rutting resistance (i.e., lower rut depth, higher HT-IDT strength, higher RT_{Index}).
[2]TPA AV levels consist of 3% and 5%.

Comparative Statistical Analyses

The relationships and statistical analyses among the different rutting mechanical tests, incorporating datasets from both suggested AV levels (i.e., 5 and 7 percent), are presented in the following section. The regression analyses primarily highlight the correlation between the APA 250 psi/250 lb test and different rutting test parameters. All other regressions involving the remaining rutting tests are provided in Appendix E. The comparative analysis was based on the current APA 250 psi/250 lb test criterion, due to the extensive research and satisfactory field performance of this set threshold with accelerated testing facilities and existing airfield sections data (Elias, 2024). Datasets that included LMLC and RPMLC test data at the two AV levels were analyzed using Minitab Statistical Software (Minitab® 17.1.0) to generate power regression models with their corresponding 95 percent confidence bands (Minitab, 2024).

A statistical comparison between LMLC and RPMLC datasets for the various rutting test parameters was conducted first. The results indicate that most of the data points fell within the 95 percent confidence band (see Figure 11 for RT_{Index} as an example; additional graphs are provided in Appendix E). Accordingly, the data points from both LMLC and RPMLC samples were combined into a single dataset for each rutting test parameter.



Source: University of Nevada, Reno

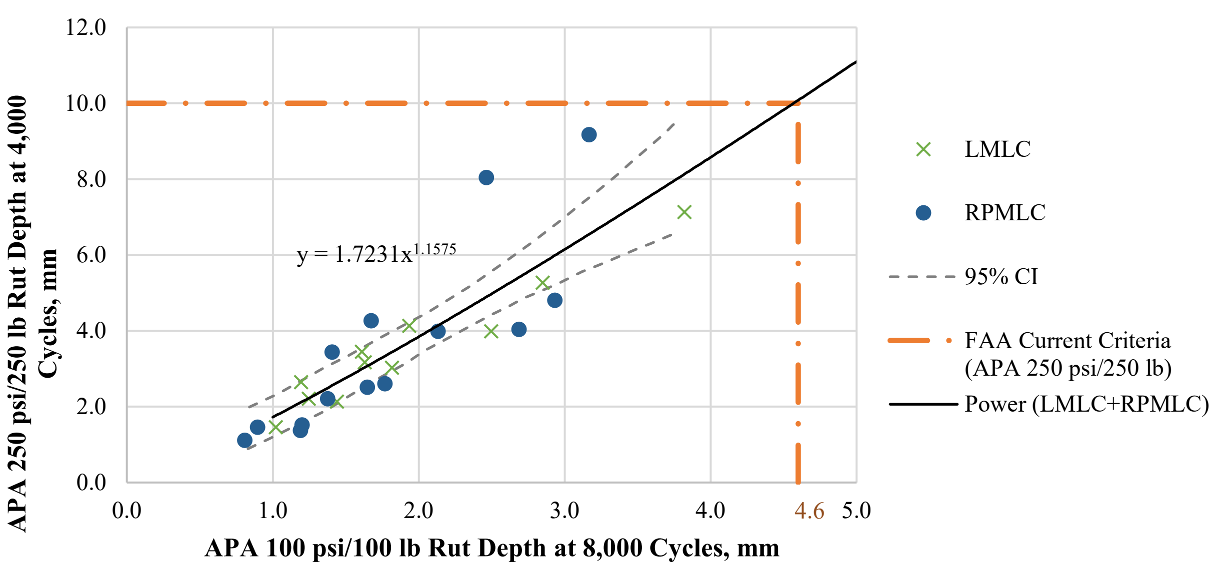
Figure 11. RPMLC RT_{Index} vs. LMLC RT_{Index}

The APA results at both loading conditions (i.e., 100 psi/100 lb and 250 psi/250 lb) showed a strong correlation among their respective rut depths, as well as with the RT_{Index} and HT-IDT, as shown in Figure 12 through Figure 14. Moreover, a robust linear relationship was observed between the two monotonic tests: HT-IDT and IRT (see Figure 15).

The strong correlations among the various rutting mechanical tests confirm the suitability of the surrogate tests for ease of implementation, particularly during production. Robust trends were observed even though the datasets included samples tested at two different AV levels and across a wide range of test temperatures. These findings suggest the potential to derive one test criterion from another, provided the same AV and test temperature are used.

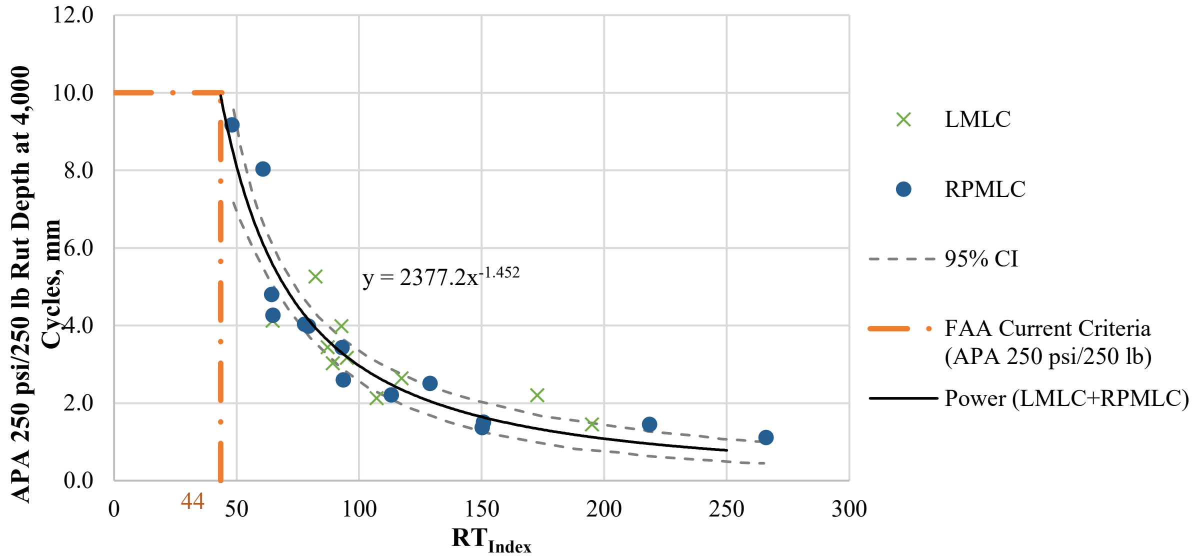
By correlating the current APA 250 psi/250 lb test criterion with other rutting test parameters, new recommended criteria can be established under the same test conditions (i.e., at the design AV and 64 °C). Accordingly, the following thresholds were derived from the existing FAA specification, which limits the APA 250 psi/250 lb rut depth to a maximum of 10 mm after 4,000 cycles:

- **Maximum APA 100 psi/100 lb rut depth:** 4.6 mm after 8,000 cycles. This value is consistent with the current FAA criteria for APA 100 psi/100 lb of 5 mm maximum rut depth. The 4.6 mm derived criterion has a 95 percent CI of [3.9, 6.7] mm. Given that the average coefficient of variation (COV) for the APA 250 psi/250 lb test (9.1 percent) is comparable to that of the APA 100 psi/100 lb COV (10.2 percent), the derived criteria relied on the mean value of 4.6 mm without adjustments for the variability between both tests.
- **Minimum RT_{Index}:** 44, with a 95 percent CI of [34, 47]. Because both tests had an average COV of less than 10 percent (9.1 percent for APA 250 psi/250 lb and 5.5 percent for IRT), the derived criteria relied on the mean RT_{Index} value of 44.
- **Minimum HT-IDT strength:** 8 psi, with a 95 percent CI of [0, 11] psi. Given that both tests had an average COV of less than 10 percent (9.1 percent for APA 250 psi/250 lb and 6.4 percent for HT-IDT), the derived criteria relied on the mean HT-IDT strength value of 8 psi.



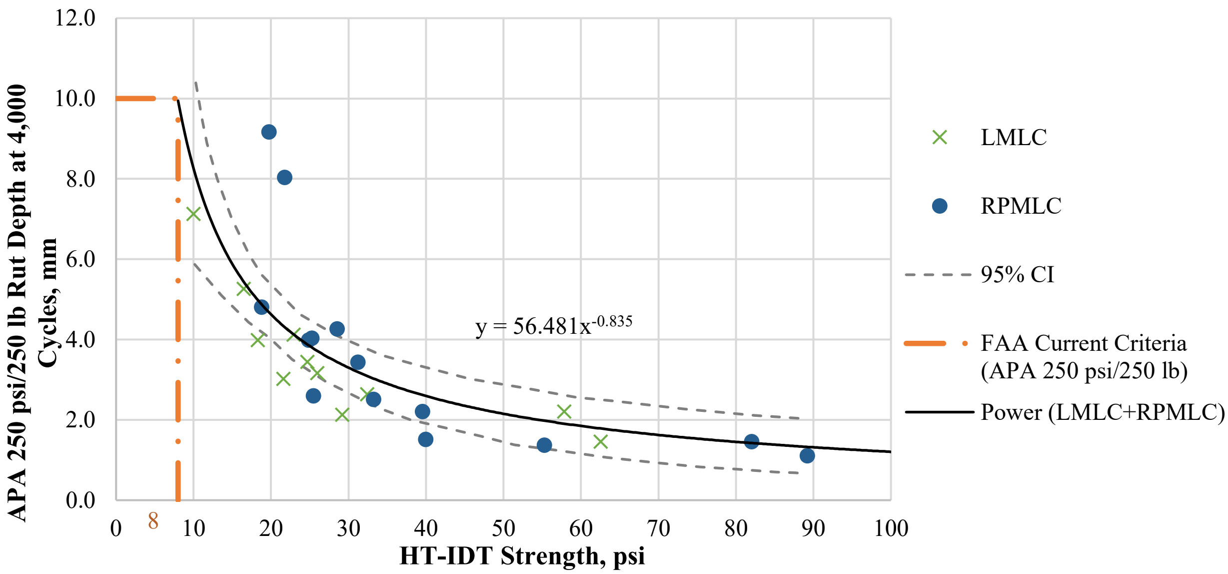
Source: University of Nevada, Reno

Figure 12. APA 250 psi/250 lb vs. APA 100 psi/100 lb



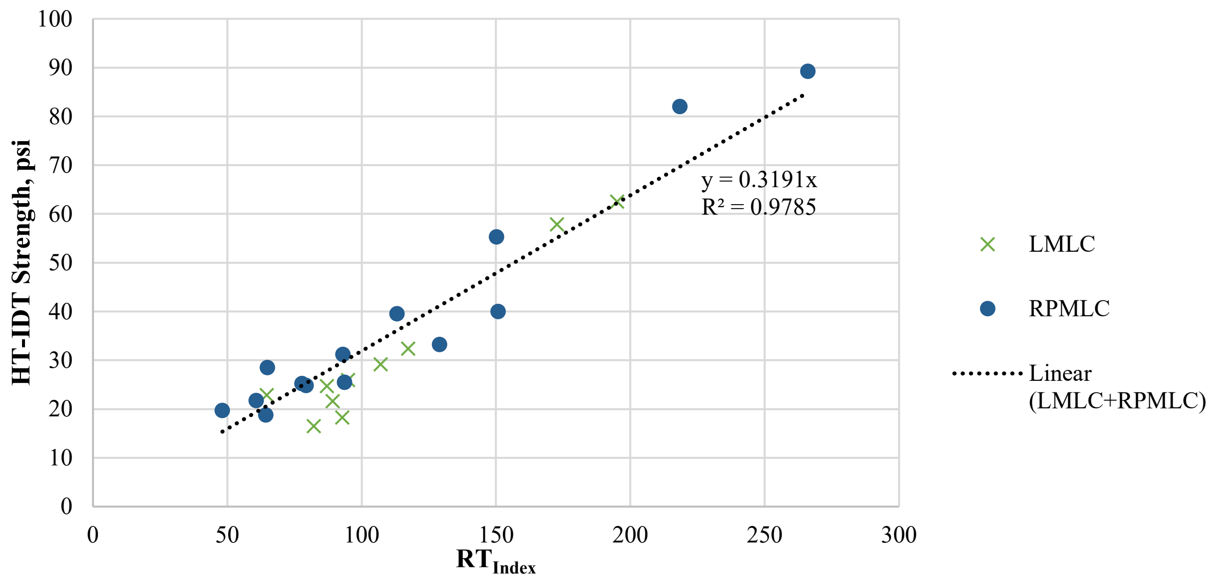
Source: University of Nevada, Reno

Figure 13. APA 250 psi/250 lb vs. RT_{Index}



Source: University of Nevada, Reno

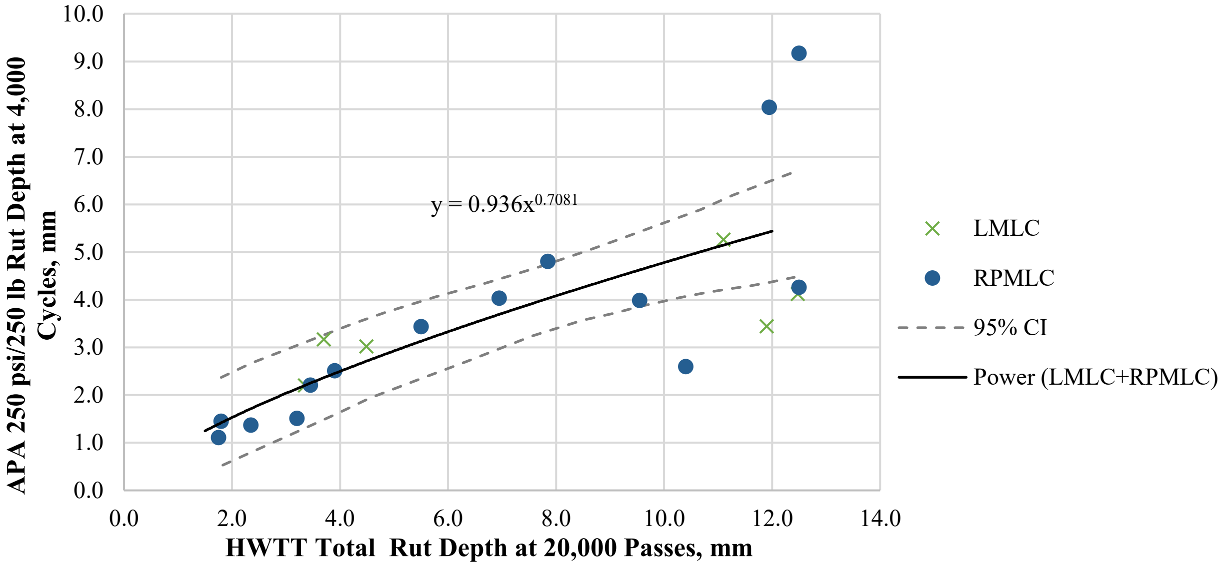
Figure 14. APA 250 psi/250 lb vs. HT-IDT



Source: University of Nevada, Reno

Figure 15. HT-IDT vs. RT_{Index}

In contrast, the HWTT results did not exhibit strong or consistent trends with the other laboratory rutting mechanical tests. Although cut specimens were excluded from the HWTT dataset, directly molded specimens still showed weak correlations with the remaining rutting tests, particularly with the APA test as shown in Figure 16. As an example, an HWTT rut depth of 12 mm after 20,000 passes corresponded to APA 250 psi/250 lb rut depths ranging between 3 and 9 mm. Conversely, an APA rut depth of 4 mm after 4,000 cycles corresponded to HWTT rut depths varying from 8 to 13 mm. In some cases, the data were influenced by stripping failures in the HWTT under wet conditions. Consequently, additional HWTT parameters that could isolate the stripping failure from the mixture's rutting performance were investigated.



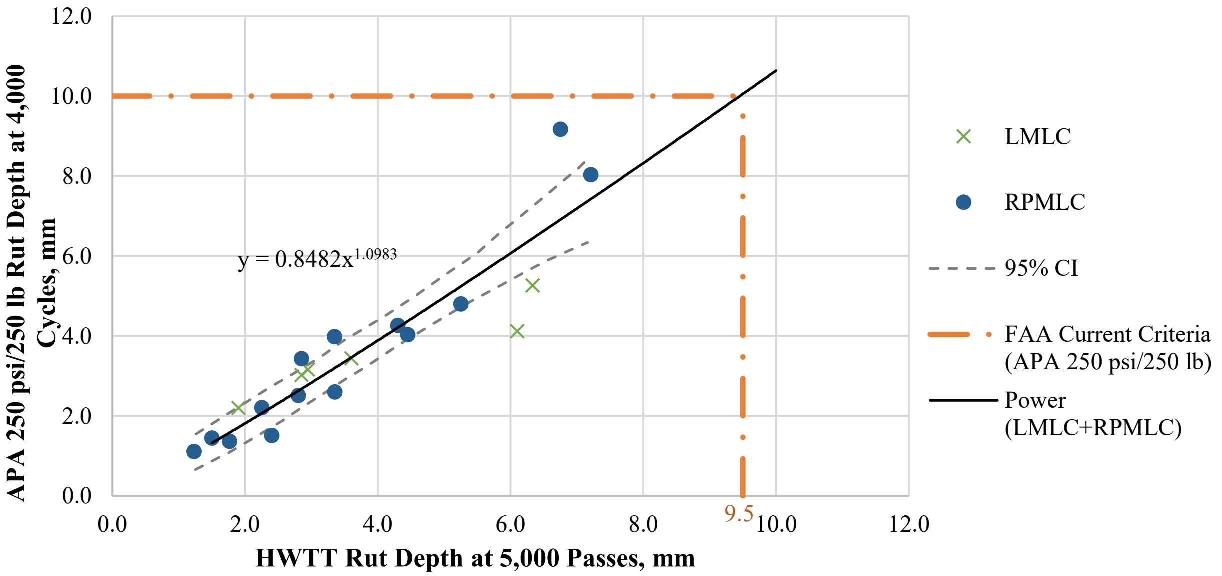
Source: University of Nevada, Reno

Figure 16. APA 250 psi/250 lb vs. HWTT Total Rut Depth at 20,000 Passes

The additional HWTT parameters evaluated included the HWTT-measured rut depth at 3,500, 5,000, and 7,000 passes, as well as the HWTT corrected rut depth (CRD) at 20,000 passes. The 3,500- and 7,000-pass HWTT measurements were selected to establish equivalency in the loading speed with the APA test. Since the APA and HWTT operate at 60 cycles/min and 52 passes/min, respectively, 3,500 and 7,000 HWTT passes correspond approximately to the 4,000 and 8,000 APA cycles referenced in the current FAA APA criteria. The 5,000-pass mark was chosen as an intermediate value between 3,500 and 7,000 passes, ensuring sufficient rutting progression in the HWTT after specimen consolidation and before reaching stripping failure.

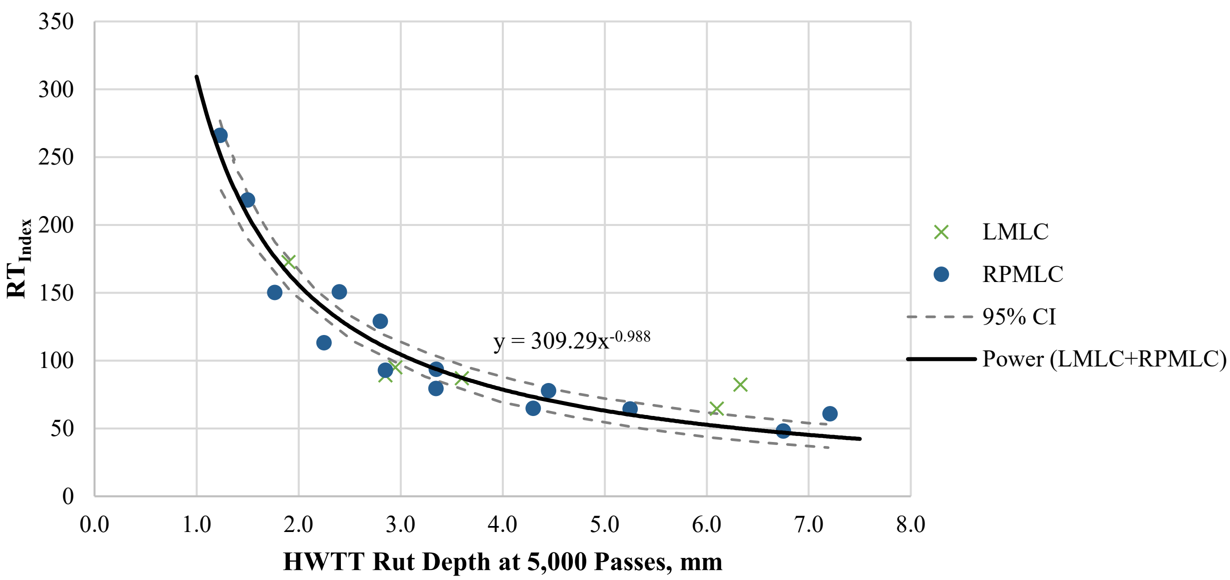
For the HWTT CRD at 20,000 passes, Yin et al. (2020) proposed fitting the entire rut depth curve using a three-parameter deformation model to identify the stripping number (SN). The SN represents the onset of stripping within the mixture and corresponds to the critical number of wheel passes where the curve's curvature changes from negative to positive. Once the SN is determined, the rut depth data are fitted with the Tseng-Lytton model up to the calculated SN. This process allows for extrapolation of the rut depth beyond the SN, providing an estimate of the HWTT CRD at 20,000 passes.

For brevity, this report presents only the regression models involving HWTT rut depth at 5,000 passes. This parameter was selected from the evaluated HWTT candidate parameters due to its lower standard deviation between the observed data values and the fitted values across multiple regression models with different rutting test parameters. The regressions of the HWTT rut depth at 5,000 passes are plotted with the APA 250 psi/250 lb test and the RT_{Index} in Figure 17 and Figure 18, respectively. These new regressions demonstrate a stronger correlation compared to the HWTT rut depth at 20,000 passes, shown in Figure 16.



Source: University of Nevada, Reno

Figure 17. APA 250 psi/250 lb vs. HWTT Rut Depth at 5,000 Passes



Source: University of Nevada, Reno

Figure 18. RT_{Index} vs. HWTT Total Rut Depth at 5,000 Passes

Based on this improved correlation, the 10 mm maximum rut depth criterion for APA 250 psi/250 lb after 4,000 cycles (at design AV and 64 °C) corresponds to a maximum HWTT rut depth of 9.5 mm after 5,000 passes, with a 95 percent CI of [8.1, 11.3] mm. Since the average COV for the APA 250 psi/250 lb test (9.1 percent) is comparable to that of the HWTT rut depth after 5,000 passes (11.3 percent), the derived HWTT criterion relied on the mean value of 9.5 mm without the need for additional adjustments to account for variability between the two tests.

Following the experimental testing plan and analyses at the two AV levels (5 and 7 percent), the 7±0.5 percent AV level was ultimately recommended for all four rutting tests. This AV level allows for directly molded specimens with height dimensions of 75 mm for the APA and 62 mm for the HWTT, HT-IDT, and IRT. The advantages and disadvantages of different percent AV scenarios were evaluated and are summarized in Appendix C, supporting the selection of this single AV level for all four rutting tests. This approach enhances consistency when using alternate rutting tests between the mix design and production stages. Moreover, testing directly molded specimens streamlines sample preparation and testing (a crucial advantage during production) and helps reduce variability in HWTT rut depth results, particularly in cases of stripping failure, which is more prevalent in cut specimens.

Mechanistic Analysis Results

Pavement Modeling: 3D-Move Dynamic Analysis

The 3D-Move Analysis software evaluates the response of a layered medium under dynamic surface loads using a continuum-based finite layer approach (Siddharthan, Krishnamenon, & Sebaaly, 2000). In this study, the pavement system was modeled using a combination of viscoelastic (AC surface layer) and elastic (AC binder/base and unbound layers) horizontal layers, each characterized with a set of uniform properties. The 3D-Move modeling is the first step in the mechanistic framework, as outlined in Figure 7. The critical wheel load is attributed to the main gear, which carries 95 percent of the GAW (FAA, 2024d). Therefore, half of the symmetrical main gear was modeled in 3D-Move for each of the five aircraft, accounting for different loading conditions and gear configurations.

To best simulate actual airfield pavement conditions, 3D-Move dynamic analyses were conducted to estimate the stress invariants used to determine representative confining and deviatoric stress conditions for the RLT test. After conducting the RLT test in the laboratory at these determined stress conditions, the resilient strain responses from 3D-Move were input into laboratory-developed rutting performance models for each of the four evaluated asphalt mixtures.

The parametric analysis matrix summarized in Table 7 involved running 180 simulations in 3D-Move to account for three temperatures, three aircraft speeds, and five aircraft types across four airfield projects. This extensive analysis was performed to examine the influence of key parameters such as pavement temperature, asphalt mixture properties, pavement structure, loading speed, aircraft wheel load and tire pressure, and aircraft gear configuration on the performance of airfield pavements.

The parameters and their corresponding values used in 3D-Move are detailed in Table 12 and Table 13. These values were based either on laboratory experimental data, project specific-data, FAA advisory circular 150/5320-6G—Airport Pavement Design and

Evaluation, actual data from the FAA National Airport Pavement Test Facility (NAPTF), or relevant airfield studies (Gonzalez & Lombaerts, 2023; Daidzic, 2016; FAA, 2021a; FAA, 2024d).

The 3D-Move inputs are grouped into five main categories: analysis type and gear configuration/wheel load in Table 12, and pavement structure, layer properties, and response points in Table 13. The gear configuration parameters were tabulated for all five modeled aircraft under an elliptical uniformly loaded area, which is assumed in the FAA Rigid and Flexible Iterative Elastic Layered Design (FAARFIELD) pass/coverage calculation (FAA, 2021c). It is worth noting that the AC surface layer was modeled as a viscoelastic material based on the experimental testing detailed in subsequent sections, while the AC binder/base layer was modeled as an elastic material with a constant modulus, either set by FAA advisory circular 150/5320-6G or derived from project-specific data (FAA, 2021a).

Response points along the dual tires were selected at key locations: the outer edge, center, and inner edge of the tire, along with the midpoint between the dual tires. These points were analyzed at four depths: at the surface, 50.8 mm (2 inches) below surface, mid-AC surface layer, and bottom of AC surface layer. Thus, each of the 180 runs in 3D-Move evaluated the 16 response points, covering four locations at four different depths to ensure a comprehensive analysis of the pavement responses under varied loading conditions.

Table 12. 3D-Move Inputs for Aircraft Gear Configuration

3D-Move Input	Value				
Analysis Type	Dynamic (Moving Load)				
Aircraft Speed, mph	Refer to Table 7				
Aircraft type	Beechcraft King Air B200	Bombardier CRJ 200	Gulfstream G600	B737–(MAX 8)	B777–(200)
Tire pressure, psi	98	160	188	205	182
Tire load, lb	2,990	11,329	22,468	43,154	43,304
Loaded Area	Elliptical Uniformly Distributed Load	Elliptical Uniformly Distributed Load	Elliptical Uniformly Distributed Load	Elliptical Uniformly Distributed Load	Elliptical Uniformly Distributed Load
L1[1], cm (inch)	0.0	0.0	0.0	0.0	144.8 (57.0)
L2[1], cm (inch)	0.0	0.0	0.0	0.0	144.8 (57.0)
S1[2], cm (inch)	(29.5) 11.6	44.2 (17.4)	47.0 (18.5)	86.4 (34.0)	139.7 (55.0)
Rolling friction coefficient	0.02	0.02	0.02	0.02	0.02
Braking friction coefficient	0.00	0.00	0.00	0.00	0.00

[1]L1, L2: Distances between the main gear axles.
[2]S1: Distance between dual tires of a main gear axle.

Table 13. 3D-Move Inputs for Pavement Structure and Layer Properties

3D-Move Input	Value
Pavement Structure	Pavement plans of modeled airfield projects.
Analysis Temperature	Refer to Table 7.
AC Surface Layer (Viscoelastic)	Laboratory-measured dynamic modulus and phase angle.
AC Binder/Base Layer (Elastic)	
Elastic Modulus, psi	400,000 (based on FAA advisory circular 150/5320-6G) (FAA, 2021a).
Damping Ratio, percent	5
Poisson's Ratio	0.35
Unit Weight, pcf	150
Unbound Base Layer (Elastic)	
CBR	80 (based on FAA advisory circular 150/5320-6G) (FAA, 2021a).
Elastic Modulus, psi	42,205 (estimated using $E = 2555 \times CBR^{0.64}$).
Damping Ratio, percent	5
Poisson's Ratio	0.35
Unit Weight, pcf	145 (NAPTF data) (FAA, 2024d)
Subgrade Layer (Elastic)	
CBR	Natural soil data for each project.
Elastic Modulus, psi	Estimated from CBR value; $E = 2555 \times CBR^{0.64}$.
Damping Ratio, percent	5
Poisson's Ratio	0.40
Unit Weight, pcf	98 (based on NAPTF data) (FAA, 2024d)
Response Points	
Location Relative to the Tire	16 locations in the AC layer.
Depth in the AC Layer	Outer edge of tire, center of tire, inner edge of tire, middle of both tires. Pavement surface, 50.8 mm (2 inches) below pavement surface, bottom of AC surface layer, and middle of AC binder layer.

Mechanistic Responses (Stress Responses and Drucker-Prager Failure Criteria)

For each 3D-Move run, the confining stress (σ_c) and deviatoric stress (σ_d) were calculated using the stress invariants (i.e., octahedral normal and shear stresses) according to Equation 1 through Equation 4 (Hajj, Siddharthan, Sebaaly, & Weitzel, 2007). To identify the critical response point (i.e., location and depth) for each 3D-Move run, the Drucker-Prager failure criterion was used to calculate the factor of safety (FOS) for each response point. The Drucker-Prager failure criterion is used as an alternative to the Mohr-Coulomb failure criterion to model the shear behavior of asphalt mixtures based on cohesion (c) and friction angle (ϕ).

The Drucker-Prager yield or failure envelope is typically plotted on a graph of deviatoric stress (q or σ_d) versus mean normal stress (p) (refer to Equation 5 and Equation 6). The FOS indicates how far the stress state is from the failure envelope with respect to q (Equation 7) (Hajj, Siddharthan, Sebaaly, & Weitzel, 2007). In other terms, the FOS is the ratio between deviatoric stress at failure ($q_{failure}$) and the applied deviatoric stress ($q_{applied}$) at a given mean normal stress ($p_{applied}$). Typical c and ϕ values for asphalt mixtures—541 kPa (78.5 psi) and 17.7 degrees, respectively—were previously determined by Hajj et al. (2007) using the triaxial compression strength test and were adopted in this study.

$$\boldsymbol{\sigma_{oct}} = \frac{1}{3}\left(\boldsymbol{\sigma_{xx}} + \boldsymbol{\sigma_{yy}} + \boldsymbol{\sigma_{zz}}\right) \quad \text{Equation 1}$$

$$\boldsymbol{\tau_{oct}} = \frac{1}{3}\sqrt{(\boldsymbol{\sigma_{xx}} - \boldsymbol{\sigma_{zz}})^2 + \left(\boldsymbol{\sigma_{xx}} - \boldsymbol{\sigma_{yy}}\right)^2 + \left(\boldsymbol{\sigma_{yy}} - \boldsymbol{\sigma_{zz}}\right)^2 + \left(\boldsymbol{\sigma_{xy}}^2 + \boldsymbol{\sigma_{yz}}^2 + \boldsymbol{\sigma_{xz}}^2\right)} \quad \text{Equation 2}$$

$$\boldsymbol{\sigma_d} = \frac{3}{\sqrt{2}}|\boldsymbol{\tau_{oct}}| \quad \text{Equation 3}$$

$$\boldsymbol{\sigma_c} = \boldsymbol{\sigma_3} = \boldsymbol{\sigma_2} = \boldsymbol{\sigma_{oct}} - \frac{\boldsymbol{\sigma_d}}{3} \quad \text{Equation 4}$$

$$\boldsymbol{q} = \sqrt{3\bar{I}_{2D}} \quad \text{Equation 5}$$

$$\boldsymbol{p} = \frac{I_1}{3} \quad \text{Equation 6}$$

$$\boldsymbol{FOS} = \frac{q_{failure}}{q_{applied}} = \left(\frac{6\sin\varphi}{3-\sin\varphi}\right) \times \frac{1}{\left(\frac{q_{applied}}{p_{applied}}\right)} + \left(\frac{6c\cos\varphi}{3-\sin\varphi}\right) \times \frac{1}{q_{applied}} \quad \text{Equation 7}$$

Where:
σ_c = confining stress.
σ_d = q = deviatoric stress.
p = normal stress.

Each of the dynamic analyses conducted using 3D-Move involved the stress-time history for each of the 16 response points. Therefore, for each 3D-Move run, the Drucker-Prager FOS was calculated and used for the two main following applications:

- Critical Time Point Identification (i.e., critical time for each response point): The minimum FOS within the stress time history was identified for each of the 16 response point locations based on 3D-Move runs to determine the respective critical time for each point.
- Critical Response Point Identification (i.e., critical response point for each 3D-Move run): the critical response point was determined by identifying the lowest FOS among all the 16 response points for each 3D-Move run.

Following the calculations, a summary table was generated for each 3D-Move run as illustrated in the example of Table 14 and Figure 19, which correspond to the SFO airfield project under Boeing 777 aircraft loading at 5 mph and 52 °C.

Table 14 summarizes the findings for σ_d, σ_c, bulk stress, $p_{applied}$, $q_{applied}$, and Drucker-Prager FOS for each of the 16 response points. In the specific example presented, the critical condition where the FOS was at its minimum value of 2.0 occurred at the center of the tire, 50.8 mm (2 inches) below the surface. Interestingly, the same critical depth of 50.8 mm (2 inches) below the surface was observed for almost all of the remaining 180 3D-Move runs and has also been identified as a critical depth for the shear failure in AC pavements by previous studies (Epps, et al., 2000; Ulloa Calderon, 2013). Pavement response tables and critical conditions corresponding to all 3D-Move runs are provided in Appendix F.

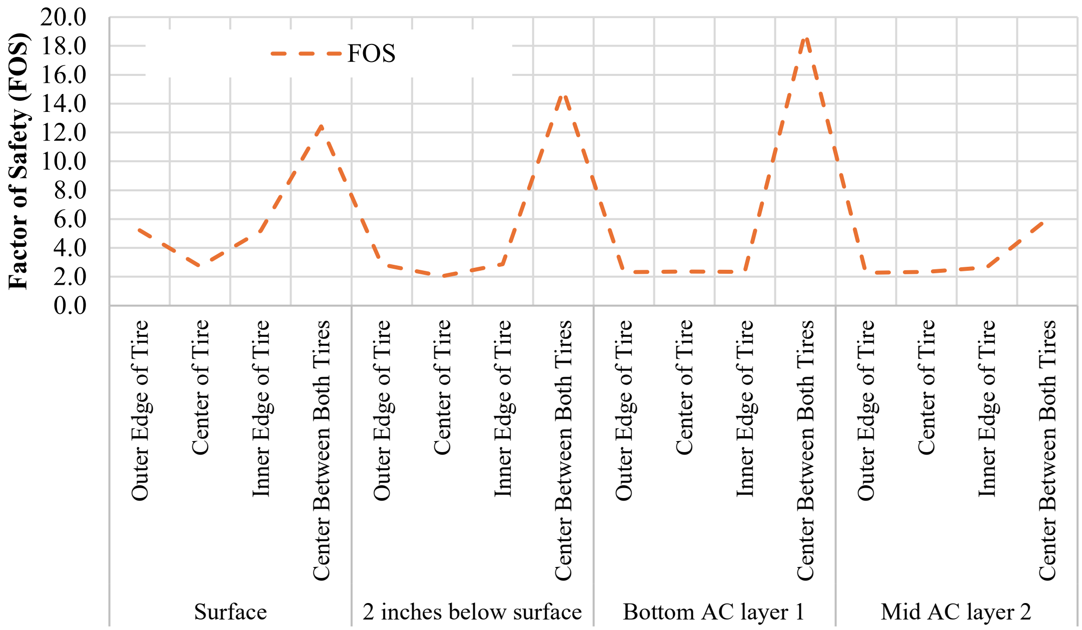
This analysis identified a total of 180 critical conditions (i.e., minimum FOS) corresponding to 45 critical conditions per airfield project/mixture. Figure 20 presents an example box plot illustrating the 45 critical conditions for the TEB airfield project, which included five different aircraft, three speeds, and three temperatures. The box plots for each of the four airfield projects outline the range of FOS values determined under different loading, speed, and temperature conditions. These critical FOS values guided the determination of representative stress conditions (i.e., σ_d and σ_c) for the RLT test.

Table 14. 3D-Move Stress Responses and Critical Condition for SFO Asphalt Mixture Under Boeing 777 at 5 mph and 52 °C Based on Drucker-Prager FOS

Depth	Location	Stress, psi					FOS
		σ_d	σ_c	Bulk	$p_{applied}$	$q_{applied}$	
Pavement Surface	*Outer edge of tire*	40.5	52.2	197.2	65.7	40.5	5.2
	Center of tire	93.1	95.7	380.2	126.7	93.1	2.7
	Inner edge of tire	38.9	38.9	155.6	51.9	38.9	5.2
	Center between both tires	13.7	1.1	17.1	5.7	13.7	12.4
50.8 mm (2 inches) Below Surface	*Outer edge of tire*	72.8	36.2	181.5	60.5	72.8	2.8
	Center of tire	113.4	58.3	288.2	96.1	113.4	2.0
	Inner edge of tire	72.3	35.9	179.9	60.0	72.3	2.9
	Center between both tires	11.4	3.1	20.8	6.9	11.4	15.0
Bottom of AC Layer 1	*Outer edge of tire*	90.1	33.8	191.4	63.8	90.1	2.3
	Center of tire	87.9	30.6	179.7	59.9	87.9	2.4
	Inner edge of tire	89.7	34.2	192.4	64.1	89.7	2.3
	Center between both tires	9.0	2.8	17.4	5.8	9.0	19.0
Middle of AC Layer 2	*Outer edge of tire*	79.1	-6.5	59.6	19.9	79.1	2.3
	Center of tire	77.0	-5.6	60.3	20.1	77.0	2.3
	Inner edge of tire	67.2	-4.5	53.5	17.8	67.2	2.7
	Center between both tires	28.7	-1.0	25.8	8.6	28.7	6.0
Critical Condition: 50.8 mm (2 Inches) Below Surface	**Center of Tire**	**113.4**	**58.3**	**288.2**	**96.1**	**113.4**	**2.0**

Note: Bulk stress = $3\sigma_c + \sigma_d$.

Negative σ_c were observed at some deep locations within the pavement structure caused by tension longitudinal and normal stresses (i.e., σ_x and $\sigma_y < 0$).

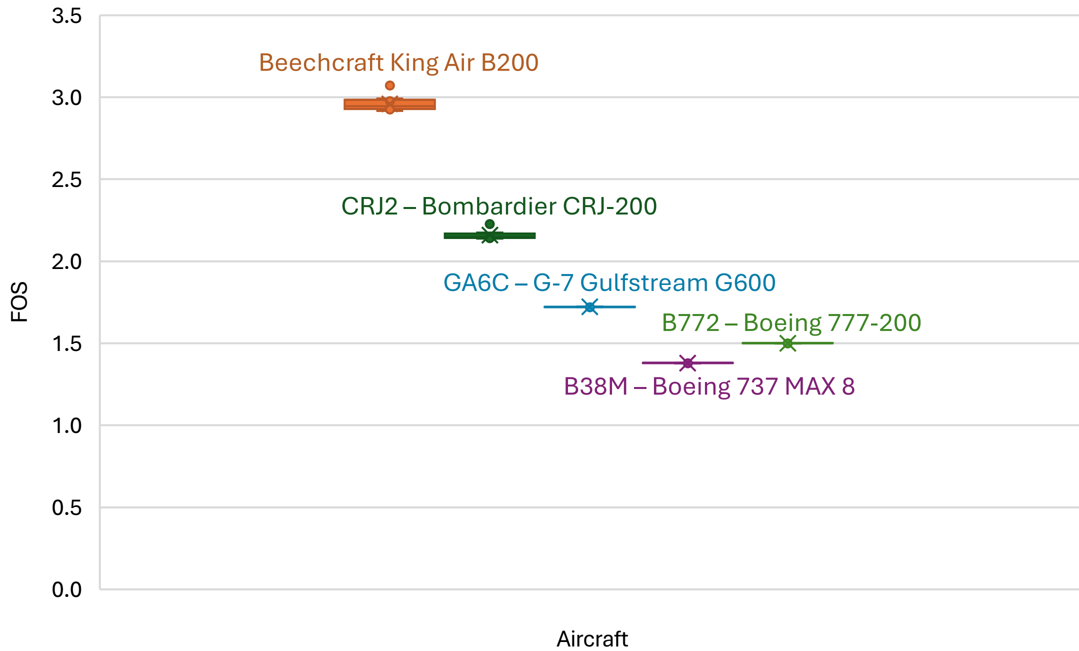


Source: University of Nevada, Reno

Figure 19. Summary of Drucker-Prager FOS for the 16 response points of SFO asphalt mixture under Boeing 777 at 5 mph, and 52 °C.

The RLT test is typically conducted at a single confining and deviatoric stress level across multiple temperatures. However, unlike highway pavements, airfield pavements are subjected to a wider range of loading conditions, varying from a Beechcraft King Air B200 with a 2,990-lb wheel load to a Boeing 737 MAX 8 with a 43,154-lb wheel load. To accommodate these variations, the RLT test in this study was conducted at three different temperatures and three stress states to capture the range of stress conditions observed under all five modeled aircraft. Based on the range of the 180 calculated FOS values, the RLT tests were conducted at the three stress levels shown in Table 15. These levels were selected based on the following two main factors:

- The FOS values ranged from 1.4 to 3.0, which covers the spectrum of critical FOS values observed for the four airfield projects.
- The σ_c and σ_d stress levels were selected while taking into consideration the maximum stress levels that could be applied based on laboratory equipment capabilities.



Source: University of Nevada, Reno

Figure 20. Summary of Drucker-Prager FOS for TEB Asphalt Mixture Under 45 Evaluated Conditions

Table 15. Summary of the Three Selected Stress Load Levels for RLT Testing

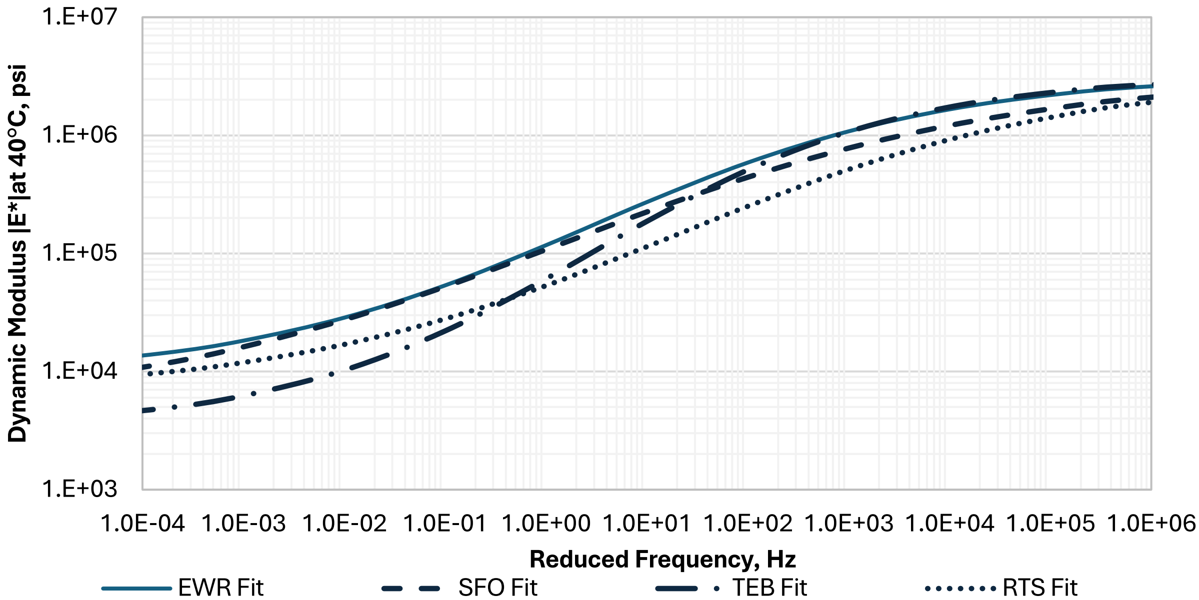
σ_d, psi	σ_c, psi	FOS
65.0	20.0	3.0
105.0	20.0	1.9
150.0	20.0	1.4

Dynamic Modulus (E*)

The dynamic modulus test was conducted in accordance with AASHTO T 378-22, and the corresponding master curve was developed for each asphalt mixture using the sigmoidal function as outlined in AASHTO R 84-17 (AASHTO, 2022c; AASHTO, 2021)(AASHTO, 2022; AASHTO, 2021). The testing temperatures and frequencies were selected based on the binder PG in accordance with AASHTO R 84-17. RPMLC samples were compacted, cored, and cut to a size of 100 mm in diameter by 150 mm in height, with a target AV of 5±0.5 percent on the cut samples.

The developed master curves at 40 °C shown in Figure 21 generally align with the PG of the binders used in each of the four airfield asphalt mixtures. The EWR with the PG 82-22 high polymer-modified binder exhibited the highest stiffness, followed by the SFO mixture with a PG 76-22 polymer-modified binder. The RTS airfield mixture, which used a PG 64-28NV polymer-modified binder, exhibited the softest modulus at medium to high frequencies, simulating AC performance at medium to low pavement temperatures. This result was expected due to the use of a soft base binder as part of the specific polymer-modified binder formulation, which was intended to mitigate brittle behavior and fatigue cracking at low to medium pavement temperatures.

Because the focus of this study was on the rutting aspect of asphalt mixtures, the E* values on the low range of frequencies, simulating critical conditions for rutting mechanism in terms of high pavement temperatures, were closely examined. At low frequencies, the EWR mixture with the PG 82-22 binder demonstrated the highest stiffness, followed by the SFO mixture with PG 76-22, and the RTS mixture with PG 64-28NV. In contrast, the TEB mixture with a neat binder of PG 64-22 exhibited the lowest stiffness (Figure 21).



Source: University of Nevada, Reno

Figure 21. Summary of Dynamic Modulus Master Curves for the Four Airfield Mixtures at a Reference Temperature of 40 °C

Rutting Performance Models

The RPMLC specimens compacted to 5±0.5 percent AV were tested in the RLT setup. A repeated haversine deviatoric stress was applied under a constant confining pressure for 0.1 sec loading time followed by a 0.9 sec rest period while keeping the confining pressure fixed. The resulting resilient (i.e., recoverable) (ε_r) and permanent (ε_p) strains were calculated after each cycle based on the axial deformations of the specimen measured over the middle 101.6 mm (4 inches) of the sample by three linear variable differential transformers (LVDTs) placed 120 degrees apart. The three LVDTs were monitored and recorded separately and then the average reading was used in the calculation of ε_r and ε_p as a function of the number of load cycles. A dummy sample monitored with a temperature probe indicated that samples required 6 hours in the Universal Testing Machine chamber to reach the target temperature. The test was run until 20,000 cycles or when a maximum 6 percent permanent strain was reached, whichever occurred first. The RLT tests were conducted on each of the four airfield mixtures at the three stress levels and three test temperatures summarized in Table 16.

Initially, the study planned to perform the RLT tests at the three temperatures used for the 3D-Move parametric analysis (Table 7). However, the highest temperature (64 °C) was too severe for the asphalt mixtures with softer binders from RTS (PG 64-28NV) and TEB (PG 64-22). Accordingly, the test temperatures were adjusted slightly to better capture the three distinct stages of permanent deformation (i.e., initial, secondary, and tertiary stages). This adjustment also allowed for the calculation of the flow number (FN) under each test condition in accordance with AASHTO T 378-22 (AASHTO, 2022c). Subsequently, the rutting performance model could be developed to fit the permanent deformation secondary stage, up to the calculated FN.

Table 16. RLT Test Conditions

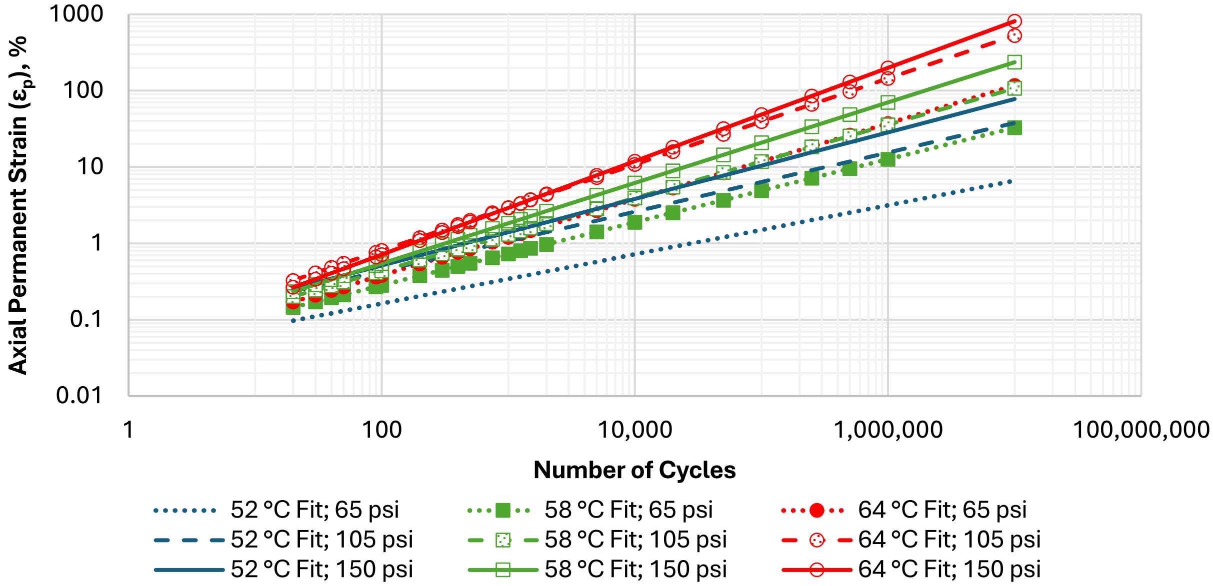
Test Parameter	Airfield Mixture			
	EWR	RTS	SFO	TEB
Temperature, °C	52	40	40[1]	40
	58[1]	52[1]	52	52[1]
	64	58	64	58
Deviatoric | Confining Stress (σ_d | σ_c), psi	65 | 20	65 | 20	65 | 20	65 | 20
	105 | 20	105 | 20	105 | 20	105 | 20
	150 | 20	150 | 20	150 | 20	150 | 20

[1]Environmental baseline testing temperature, LTPPBind Online final PG at surface with 50% reliability, 12.5 mm target rut depth, without grade bumping.

The rutting performance model in AASHTO's *Mechanistic-Empirical Pavement Design Guide* (also known as the MEPDG) is function of AC layer temperature (T) and number of loading repetitions (N) (Ulloa Calderon, 2013). In this study, the model was expanded to include σ_d to account for the varying aircraft loading levels. However, incorporating σ_d as a single independent variable in the regression model did not yield a satisfactory fit with the observed data. To further improve the model, three interaction terms between T, N, and σ_d were identified as statistically significant (based on the p-value at the 5 percent significance level) and were introduced in the final rutting performance model (Equation 8). These interaction terms, typically not considered when testing highway asphalt mixtures, were found necessary when testing airfield asphalt mixtures at the higher state of stresses to represent field conditions.

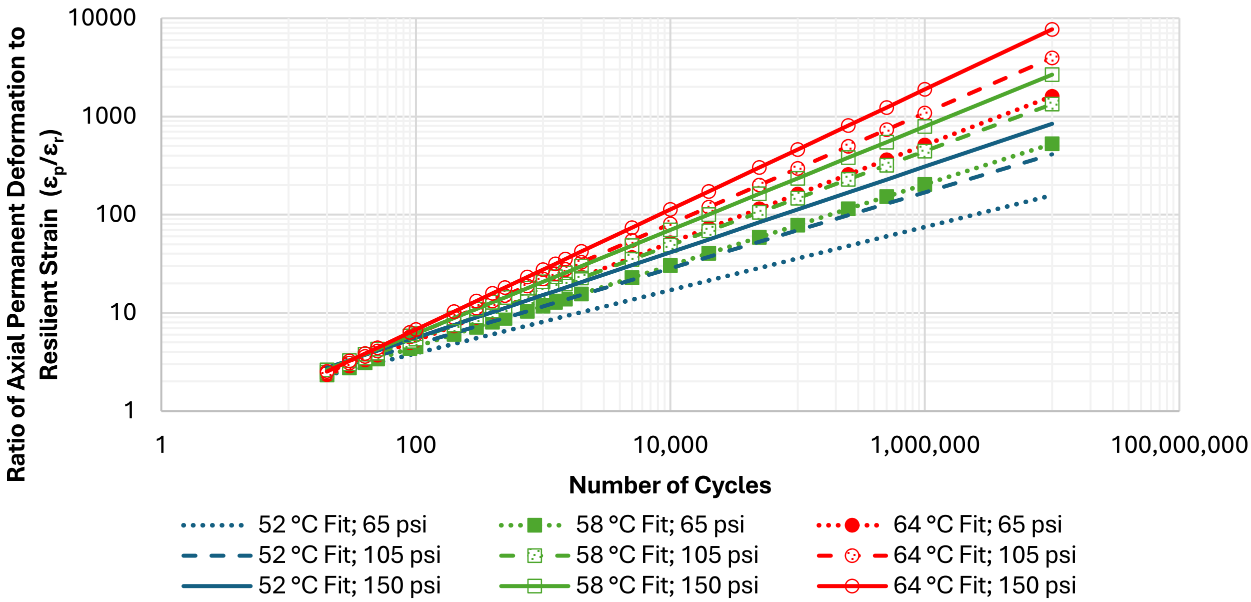
Figure 22 and Figure 23 show, respectively, ε_p and the ratio of ε_p to ε_r as a function of N at the three test temperatures and σ_d levels for the EWR mixture. Higher ε_p values were observed under increased test temperatures and/or higher σ_d values. As shown in the figures, the fitted models do not exhibit similar slopes across the three test temperatures or σ_d levels. This is attributed to the interaction between T and σ_d that influences the rutting behavior. This is particularly evident at the lower loading cycles, where several of the fitted trendlines intersect.

$$\log \frac{\varepsilon_p}{\varepsilon_r} = k_{r1} + k_{r2} \log T + k_{r3} \log N + k_{r4} \log \sigma_d + k_{r5} (\log \sigma_d \times \log T) + k_{r6}(\log N \times \log T) + k_{r7}(\log \sigma_d \times \log N) \quad \text{Equation 8}$$



Source: University of Nevada, Reno

Figure 22. Axial Permanent Deformation Fit for EWR Asphalt Mixture Under Three Temperatures and Three Deviatoric Stresses (65, 105, and 150 psi)

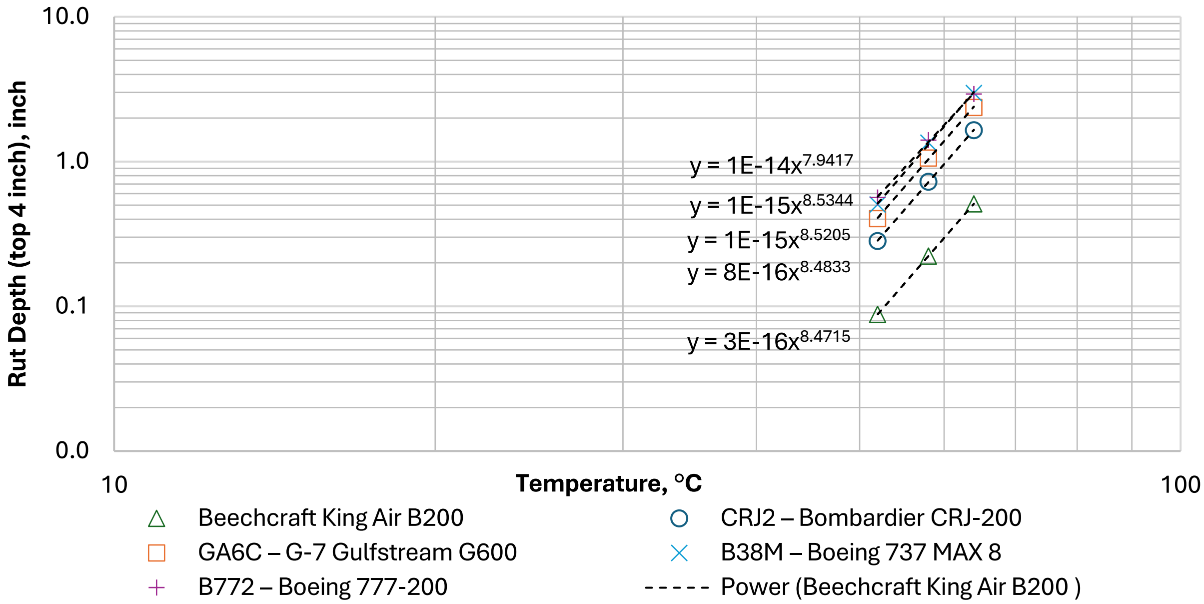


Source: University of Nevada, Reno

Figure 23. Ratio of Axial Permanent Deformation to Resilient Strain Fit for the EWR Asphalt Mixture Under Three Temperatures and Three Deviatoric Stresses

Refined Rutting Test Criteria

Based on the developed rutting performance models, a comprehensive data matrix was created by computing the rut depth (from ε_p) for 45 different conditions per airfield asphalt mixture. These conditions encompassed three temperatures, three speeds, and five aircraft types, aligning with the parametric analysis discussed earlier. The resilient strain required for the rutting performance model to calculate ε_p was obtained for each airfield pavement using 3D-Move responses for a given combination of speed, temperature, and aircraft. This effort aimed to quantify the variation in the calculated rut depth with respect to speed, temperature, and GAW using the laboratory-developed, mixture-specific rutting performance models. As an example, Figure 24 illustrates the effect of temperature on rut depth at 15 mph for five different aircraft in the case of the EWR mixture. Similar relationships were established for the effects of speed and aircraft load on rut depth for each of the four evaluated airfield mixtures. While these relationships have not been field-calibrated, the analysis in this study focused on the relative rut depth sensitivity to different speeds, temperatures, and aircraft loads, which was assessed based on the slope of these power relationships.



Source: University of Nevada, Reno

Figure 24. Effect of Temperature on Rut Depth at 15 mph for the EWR Mixture Under Five Different Aircraft

The developed relationships helped in refining the current APA 250 psi/250 lb test criterion (maximum 10 mm rut after 4,000 cycles) for several key factors. The current FAA test criterion for APA 250 psi/250 lb was adopted as the starting point, as it had been validated previously using an airfield asphalt mixture in two different field experiments (Rushing, Little, & Garg, 2014). The first field trial was under severe conditions (i.e., 32,000-lb wheel load, 325 psi tire pressure, and 43 °C), whereas the second field trial was under moderate conditions (i.e., 45,000 lb-wheel load, 142 psi tire pressure, and 25 °C). The proposed maximum APA rut depth of 10 mm after 4,000 cycles was validated by delineating any mix with field performance worse than the first field trial, where Rushing, Little, and Garg (2014) considered that most airfield pavements do not experience such continuous accelerated loading at high tire pressure and constant elevated temperature. Thus, the APA 250 psi/250 lb test criterion was validated under certain loading and environmental conditions using the FAA's Heavy Vehicle Simulator-Airfields (HVS-A) at a speed of 3 mph.

The study's approach to refining the current APA 250 psi/250 lb criterion for a wide range of climates, load levels, and speeds, along with deriving airfield pavement criteria for new rutting test methods (Table 17) and Table 18), is detailed in the following steps:

1. Start with the current FAA specification of 10 mm maximum rut depth for the APA 250 psi/250 lb test after 4,000 cycles at design AV percent (e.g., 3.5 percent AV) and 64 °C, as a baseline.
2. Refine the 10-mm threshold from design AV to 5 percent AV (the AV selected for mechanistic analysis).

3. Establish equivalency between the APA test temperature and mechanistic analysis pavement temperature. Since mechanistic analysis results are used to refine APA test criteria, a thorough evaluation was conducted to ensure conformity between laboratory and actual pavement temperatures. This evaluation was based on the accelerated pavement testing conducted by Rushing, Little, and Garg (2014) to validate the FAA's current test criteria for the APA 250 psi/250 lb test.
4. Assess the effect of different key parameters (speed, temperature, GAW) on the APA 250 psi/250-lb rut depth (at 5 percent AV) using the developed relationships for each of the four airfield mixtures (as shown in the example in Figure 24). The mechanistic analysis and APA test were not conducted at the same temperature but were based on the temperature equivalency established in step 3.
5. Select the corresponding APA 250 psi/250 lb threshold for each airfield project within each aircraft category.
6. Select one representative APA 250 psi/250 lb threshold per aircraft category, covering the four evaluated asphalt mixtures, based on the 95 percent confidence interval across the four airfield projects.
7. Refine the APA 250 psi/250 lb test criteria from 5 percent AV to the recommended testing AV level (i.e., 7 percent AV) based on the laboratory-developed relationships. Mix-specific correlations were developed in the laboratory between the APA rut depth at 5 and 7 percent AV for the four examined airfield mixtures (Elias, 2024).
8. Account for the variability in the APA 250 psi/250 lb test results by incorporating a 20 percent COV into the set limits.
9. Correlate the APA 250 psi/250 lb test criteria with other rutting mechanical tests (i.e., APA 100 psi/100 lb, HWTT, HT-IDT, and IRT) using laboratory-developed relationships (i.e., Figure 12 through Figure 15 and Figure 17), while considering the variability of the test's interrelationships by using the 95 percent confidence band (Elias, 2024).
10. Account for any significant differences in criteria between LMLC and RPMLC sample types based on laboratory-developed relationships between LMLC and RPMLC sample test data for several AC airfield mixtures. Final thresholds were selected to encompass the range of both sample types (i.e., LMLC and RPMLC).
11. Revise the recommended thresholds of the different rutting parameters based on the rutting test interrelationships.

Table 17. Rutting Test Criteria for Airfield Pavements with Slow or Stationary Aircraft

Pavement Area with Slow or Stationary Aircraft					
Aircraft Gross Weight, lb	Maximum APA 250 psi/ 250 lb Rut Depth after 4,000 Cycles, mm	Maximum APA 100 psi/ 100 lb Rut Depth after 8,000 Cycles, mm	Maximum HWTT Rut Depth after 5,000 Passes, mm	Minimum HT-IDT Strength Value, psi	Minimum RT_{Index}
≤12,500	13.0	9.0	13.0	6	20
≤100,000	6.0	3.0	8.0	12	40
>100,000	4.0	2.0	4.0	25	80

Testing temperature (40 °C–64 °C) = LTPPBind Online final PG at surface, with 50% reliability, 12.5 mm rut depth, without grade bumping.
At 7±0.5% AV; LMLC conditioned for 2 hr at compaction temperature, and RPMLC reheated per set protocol (Hajj, et al., 2025a; Hajj, et al., 2025b).

Table 18. Rutting Test Criteria for All Airfield Pavement Types (No Slow or Stationary Aircraft)

All Pavement Types (no slow or stationary aircraft)					
Aircraft Gross Weight, lb	Maximum APA 250 psi/ 250 lb Rut Depth after 4,000 Cycles, mm	Maximum APA 100 psi/ 100 lb Rut Depth after 8,000 Cycles, mm	Maximum HWTT Rut Depth after 5,000 Passes, mm	Minimum HT-IDT Strength Value, psi	Minimum RT_{Index}
≤12,500	18.0	12.0	17.0	3	10
≤100,000	13.0	9.0	13.0	6	20
>100,000	6.0	3.0	8.0	12	40

Testing temperature (40 °C–64 °C) = LTPPBind Online final PG at surface, with 50% reliability, 12.5 mm rut depth, without grade bumping.
At 7±0.5% AV; LMLC conditioned for 2 hr at compaction temperature, and RPMLC reheated per set protocol (Hajj, et al., 2025a; Hajj, et al., 2025b).

This structured process was designed to ensure the rutting test criteria were robust and accounted for key factors influencing airfield pavement performance across different aircraft and operational conditions. Key sources of variability considered included the following:

- Variability within the APA test results.
- Variability in correlations between the APA and other rutting test methods.
- Differences between LMLC and RPMLC sample types.

Additionally, it is important to note that the laboratory-derived relationships underpinning this methodology were specifically developed for the same airfield projects analyzed in this study. These relationships were described in detail in the above sections (i.e., Figure 12 to Figure 15 and Figure 17).

Regarding aircraft speed, the derived rutting test criteria were refined to account for higher aircraft speed, as the current FAA criterion for APA 250 psi/250 lb test was initially validated using field sections under slow-moving aircraft loads representing slow or stationary aircraft (FAA, 2018). Hence, asphalt mixtures placed at runways and typically trafficked by aircraft at higher speeds can follow the set rutting test thresholds summarized in Table 18, which are categorized by varying GAW for all pavement types (i.e., no slow or stationary aircraft).

The final rutting tests are recommended to be conducted at 7±0.5 percent AV and environmental test temperature, which was previously selected to reflect climatic condition at each airfield project (Hajj, et al., 2025a). The environmental test temperature is equivalent to the LTPP final PG at surface, with 50 percent reliability and 12.5 mm target rut depth without any grade bumping.

The test temperature was bracketed between a minimum of 40 °C and a maximum of 64 °C. The lower limit of 40 °C ensures that excessively low temperatures that may fail to adequately capture the rutting behavior of asphalt mixtures are avoided. Conversely, the upper limit of 64 °C is established to mitigate potential issues with HT-IDT testing or equipment performance at elevated temperatures.

Once the appropriate test temperature for a particular airfield project is established, the rutting test results must meet the criteria outlined in the tables for the applicable maximum GAW (determined from the FAA Airport Data and Information Portal) and aircraft speed conditions. Examples of selected airports, along with the associated test temperatures and GAW categories, are provided in Table 19.

Table 19. Examples of Airports with Testing Temperature and Aircraft Category (FAA, 2024c; FHWA, 2024)

Airport	Testing Temperature[1], °C	Max GAW, lb
Ted Stevens Anchorage International Airport (ANC)	40[2]	>100,000
Ronald Reagan Washington National Airport (DCA)	58	>100,000
Detroit Metro Airport (DTW)	52	>100,000
Newark Liberty International Airport (EWR)	58	>100,000
Philadelphia International Airport (PHL)	58	>100,000
Phoenix Sky Harbor International Airport (PHX)	64[3]	>100,000
Reno Stead Airport (RTS)	52	≤100,000
San Francisco International Airport (SFO)	40	>100,000
Sacramento International Airport (SMF)	64	>100,000
Teterboro Airport (TEB)	52	≤100,000
Tampa International Airport (TPA)	64	>100,000

[1]The test temperature may not be less than 40 °C or exceed 64 °C. The upper limit of 64 °C was set to address potential testing concerns with HT-IDT or equipment limitations at high temperatures. Similarly, the lower limit of 40 °C ensured that excessively low temperatures that may not adequately capture rutting behavior were avoided.

[2]The calculated test temperature was 34 °C and was adjusted up to 40 °C.

[3]The calculated test temperature was 70 °C and was adjusted down to 64 °C.

Chapter 4. Criteria Verification

The recommended rutting test criteria were verified using FMFC cores sampled from various airfield pavement sections. Some of these airfields exhibited more than 25.4 mm (1 inch) of rutting in the field and were designated as "Rutted Airfield Sections," while others showed satisfactory rutting performance and were designated as "Non-Rutted Airfield Sections." FMFC cores were collected and tested in the laboratory to determine whether the recommended criteria could effectively distinguish between asphalt mixtures with poor and good rutting performance in the field. This evaluation also supports assessment of the efficacy of the set rutting criteria when used during the production phase. The FAA William J. Hughes Technical Center in Atlantic City, NJ, assisted in testing several FMFC core samples to verify the recommended test criteria.

Rutted Airfield Pavement Sections

The four rutted airfield projects examined are designated as L, M, N, and O in Table 20, which also includes the construction date, airfield classification, maximum GAW, LTPP climatic zone, and corresponding environmental testing temperature. The rutting test results of the core samples from these projects were evaluated based on the maximum GAW for each airfield and compared to the recommended criteria in Table 17 and Table 18.

All of the airfield pavements had been in place for approximately 6 years or more, representing a higher level of aging than the samples tested to develop the criteria would have experienced under the short-term aging protocol. Consequently, differences in rutting performance were expected when comparing pavements with varying aging levels.

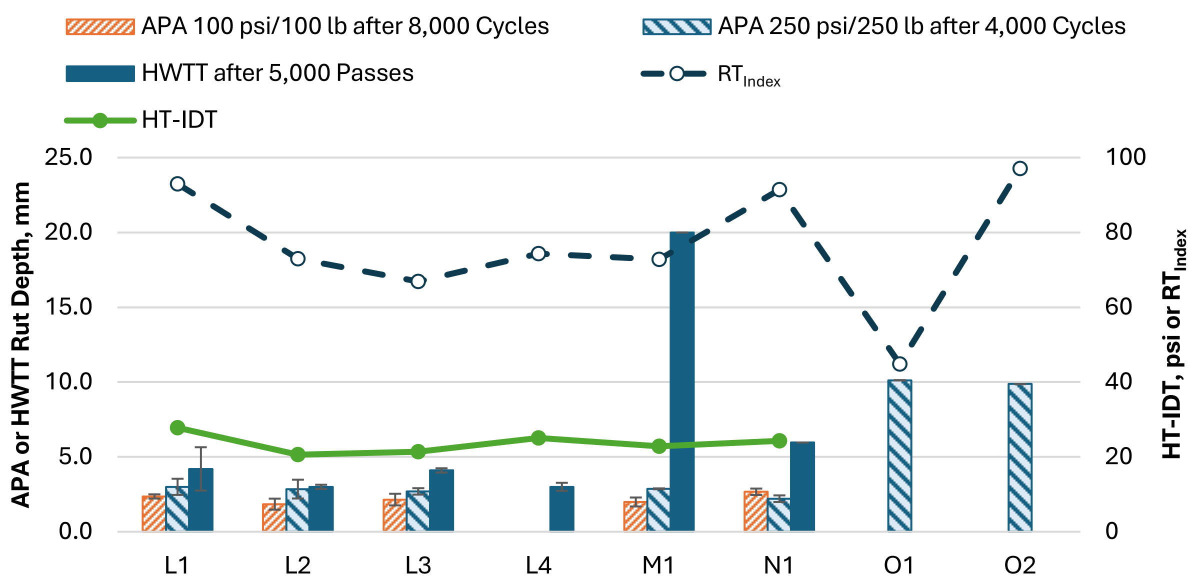
Given that in most airfield construction projects use the same asphalt mixture for both the runway and taxiway pavement sections, the verification effort focused on comparing the rutting test results against the criteria for pavements subjected to slow or stationary aircraft movements. This comparison is appropriate because even high-speed runway sections experience slow turning movements as aircraft approach taxiways at the runway's end. This finding reinforces the practicality of applying the criteria developed for slow or stationary aircraft pavement areas, enabling simpler and more effective implementation.

Table 20. Airfields Selected for Rutted FMFC Core Sampling

Airfield Project	Construction Date	Classification/ Hub	GAW, lb	LTPP Climatic Zone	Environmental Testing Temperature, °C[1]
L	2012/2016	Primary/Medium	> 100,000	Wet-Freeze	40
M	2018	Primary/Large	> 100,000	Wet-Freeze	58
N	2015	Primary/Large	> 100,000	Wet-Freeze	58
O	2015	Primary/Large	> 100,000	Wet-Nonfreeze	64

[1]Test temperature (40 °C–64 °C)=LTPPBind Online final PG at surface, 50% reliability, 12.5 mm rut depth, no grade bumping.

The FMFC test results are shown in Figure 25, where multiple airfield pavement sections corresponding to different mix designs from the same airfield were tested (e.g., L1, L2, L3, L4). The evaluated airfield pavements had been in service for several years and were therefore subjected to field aging. To minimize the effect of field aging in repeated wheel load testing, the cores were flipped and tested with the bottom surface under the wheel of the APA or HWTT. The FMFC test results presented in Figure 25 correspond to the in-place AV of the evaluated cores, which ranged from 1.2 to 4.7 percent. This is consistent with the in-place density data documented in Appendix A of this report, where airfield mat core data showed an average of 4 percent AV (Hajj, et al., 2025a). However, the set of rutting test criteria recommended in Table 17 and Table 18 correspond to 7±0.5 percent AV. Despite that, all FMFC samples had an AV level much less than 7 percent, and many core test data failed to meet the rutting test criteria set at 7 percent AV.



Source: University of Nevada, Reno

Figure 25. Rutted FMFC Core Test Results at In-Place AV

To properly evaluate the FMFC test results against the set rutting criteria, core test data were estimated at 7±0.5 percent AV based on the measured values at in-place AV. Regression models were developed for each of the five rutting test parameters versus AV. The regressions included the full laboratory experimental data collected from LMLC and RPMLC samples across the various airfield asphalt mixtures.

Mix-specific relations were developed between AV and each of the rutting test parameters of several airfield asphalt mixtures based on laboratory experimental testing. The data was then shifted to fit one model encompassing the various tested airfield mixtures. The developed regressions were subsequently employed to derive exponential models to correct each of the five rutting test parameters from in-place AV to the target AV level of 7 percent. This process is similar to the RT_{Index} AV correction described in ASTM D8360 (Equation 9) (ASTM, 2022).

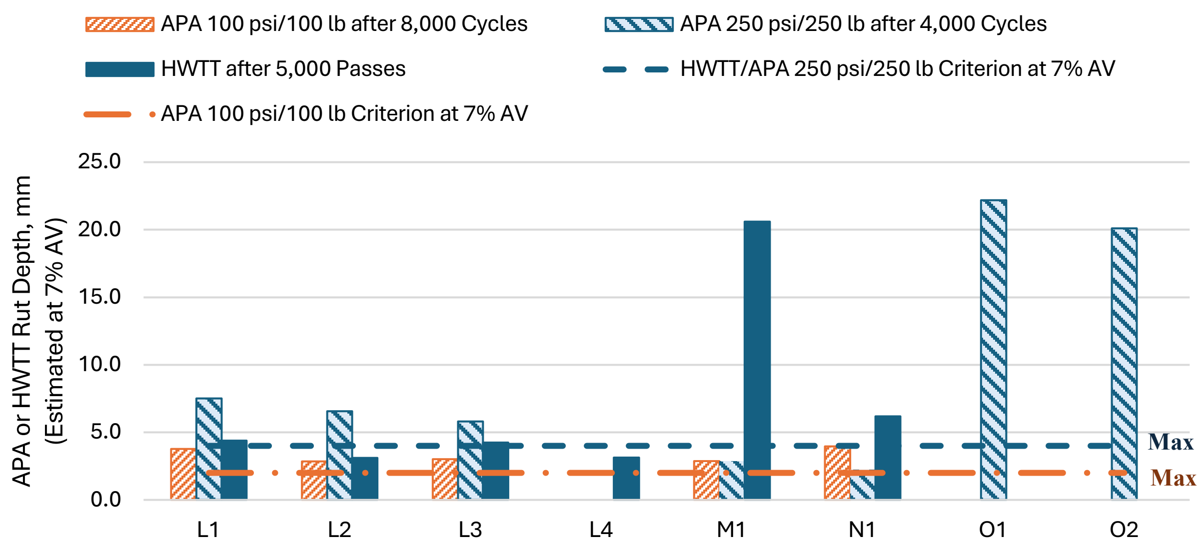
$$\boldsymbol{RT_{Index@7\%} = 0.446e^{0.1186AV} \times RT_{Index}} \quad \text{Equation 9}$$

A similar exponential model was developed for each of the five rutting test parameters, as detailed in Appendix E. The correction equations were then used to adjust the core test results from the in-place AV to the 7 percent AV level during the criteria verification effort.

The different rutting test parameters estimated at 7 percent AV for FMFC core samples are shown in Figure 26 and Figure 27, along with the recommended criteria for slow moving or stationary aircraft. The minimum rut depths estimated at 7 percent AV among evaluated mixtures were 2.8 mm for APA 100 psi/100 lb after 8,000 cycles and 5.8 mm for APA 250 psi/250 lb after 4,000 cycles; both values fail to meet the set criteria. Moreover, the maximum HT-IDT and RT_{Index} estimated at 7 percent AV among all airfield projects corresponded to 17 psi and 75, respectively, as shown in Figure 27. It can be inferred that none of the FMFC test results estimated at 7 percent AV met the rutting criteria suggested in Table 17 for slow or stationary aircraft, with the exception of projects L2 and L4, which passed the HWTT criterion.

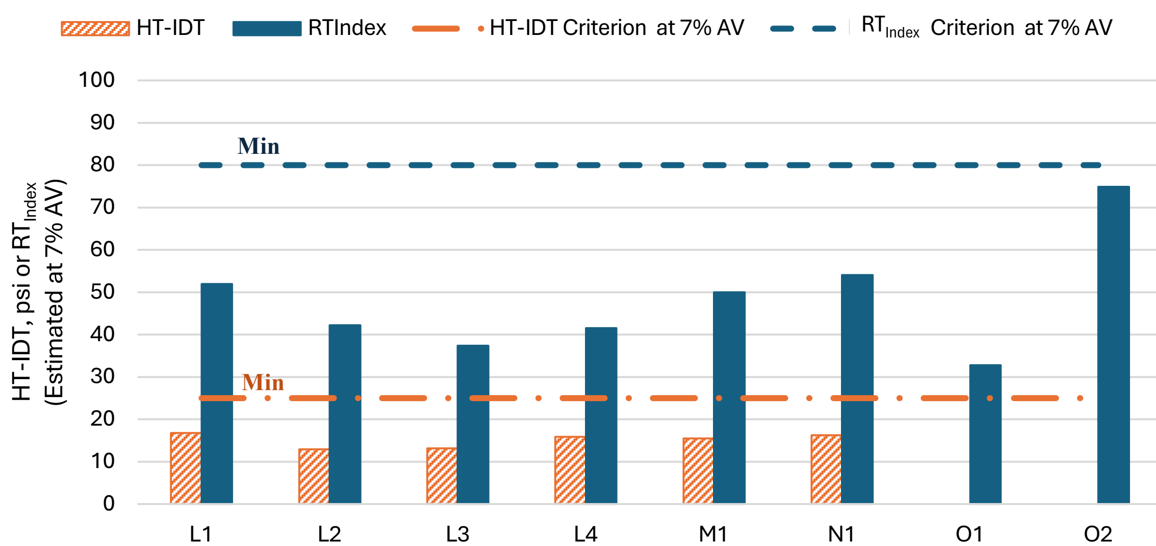
It is important to note again that these FMFC samples were collected from rutted airfield sections, which validates the significance of the set rutting criteria to delineate asphalt mixtures with poor rutting field performance. Conversely, few FMFC test data meet the rutting criteria set for all pavement types in Table 18, suggesting that those airfield sections may have satisfactory rutting in the field under higher runway speeds.

Table 21 summarizes all tested airfield sections, including the binder PG, environmental test temperature, and percent deviation from the set criteria. The tabulated percent deviation from the criteria quantifies the extent to which each core failed to meet the set test thresholds. This measure was computed to assess consistency across different rutting test parameters to avoid acceptance of a poorly performing asphalt mixture that may fail one of the rutting criterion while passing another.



Source: University of Nevada, Reno

Figure 26. Rutted FMFC Core Test Data for APA and HWTT Estimated at 7 percent AV



Source: University of Nevada, Reno

Figure 27. Rutted FMFC Core Test Data for HT-IDT and RT_{Index} Estimated at 7 percent AV

A high consistency between the five rutting test parameters can be inferred from Table 21, where all the poorly performing airfield sections, highlighted in red, failed the different set criteria. The percent deviation from criteria was very similar between the two monotonic tests (i.e., HT-IDT and IRT) and between the two repeated load tests (APA and HWTT). In most cases, the repeated load tests were able to depict the poor rutting performance of the asphalt mixtures, showing higher percent deviations from the criteria compared to monotonic load tests.

Table 21. Percent Deviation of Rutted FMFC Test Results (Corrected to 7 percent AV) from Recommended Criteria

Airfield Section with Poor Field Performance	Binder PG	Test Temperature, °C	% Deviation from Criteria				
			APA 100 psi/ 100 lb	APA 250 psi/ 250 lb	HWTT	HT-IDT	RT_{Index}
L1	PG 58-34	40	-88%	-88%	-10%	-33%	-35%
L2	PG 58-34	40	-42%	-64%	22%	-48%	-47%
L3	PG 58-34	40	-51%	-45%	-6%	-47%	-53%
L4	PG 52-34	40	N/A	N/A	22%	-37%	-48%
M1	PG 76-22	58	-44%	N/A	-415%	-38%	-38%
N1	PG 76-22	58	-98%	N/A	-54%	-35%	-32%
O1	-	64	N/A	-455%	N/A	N/A	-59%
O2	-	64	N/A	-402%	N/A	N/A	-6%

N/A = Not applicable.

Non-Rutted Airfield Pavement Sections

The two non-rutted airfield projects examined are designated as P and Q. Table 22 provides details related to construction date, airfield classification, maximum GAW, LTPP climatic zone, and the environmental testing temperature determined. The field cores selected for laboratory rutting testing were extracted close to the joint area, with an in-place AV within the 7 percent target level. As expected, Figure 28, Figure 29, and Table 23 indicate that the FMFC cores, highlighted in green, met the set test criteria at the relative environmental test temperature for each mixture. Airfield project P barely passed the rutting criteria by 8 and 10 percent for the HT-IDT and IRT, respectively. The two tested airfield projects corresponded to two different GAW categories and thus had different criteria from Table 17.

Table 22. Selected Airfields for Non-Rutted FMFC Core Sampling

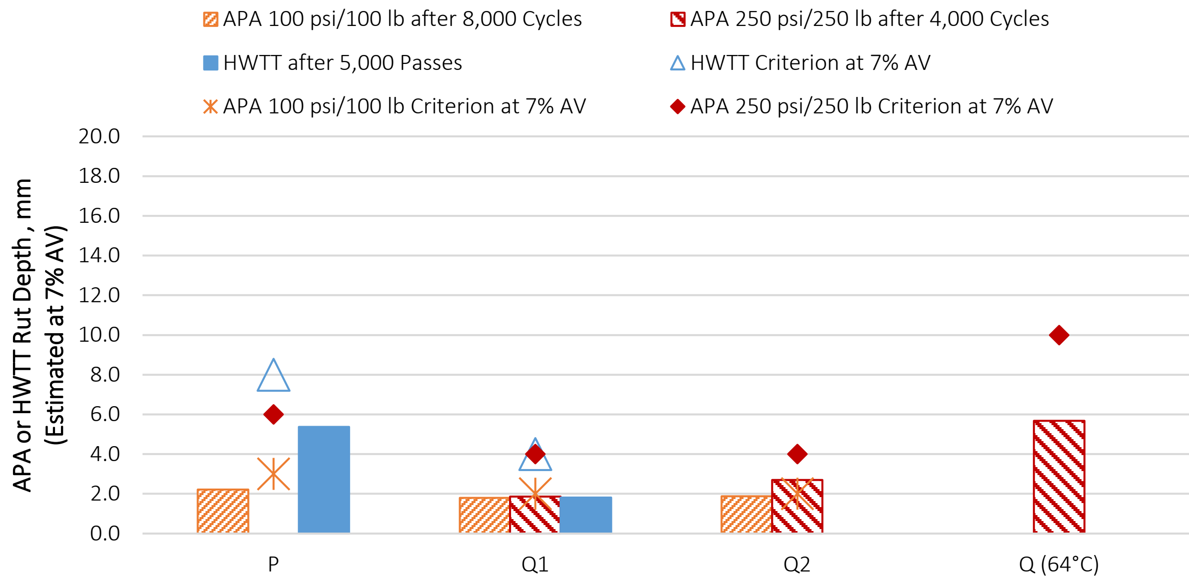
Airfield Project	Construction Date	Classification/ Hub	GAW, lb	LTPP Climatic Zone	Environmental Testing Temperature, °C[1]
P	2022	Reliever/–	≤100,000	Dry-Freeze	52
Q	2023	Primary/Large	>100,000	Dry-Nonfreeze	40

[1]Environmental baseline testing temperature, LTPPBind Online final PG at surface with 50% reliability, 12.5 mm target rut depth, without grade bumping.

Table 23. Percent Deviation of Non-Rutted FMFC Test Results from Recommended Criteria

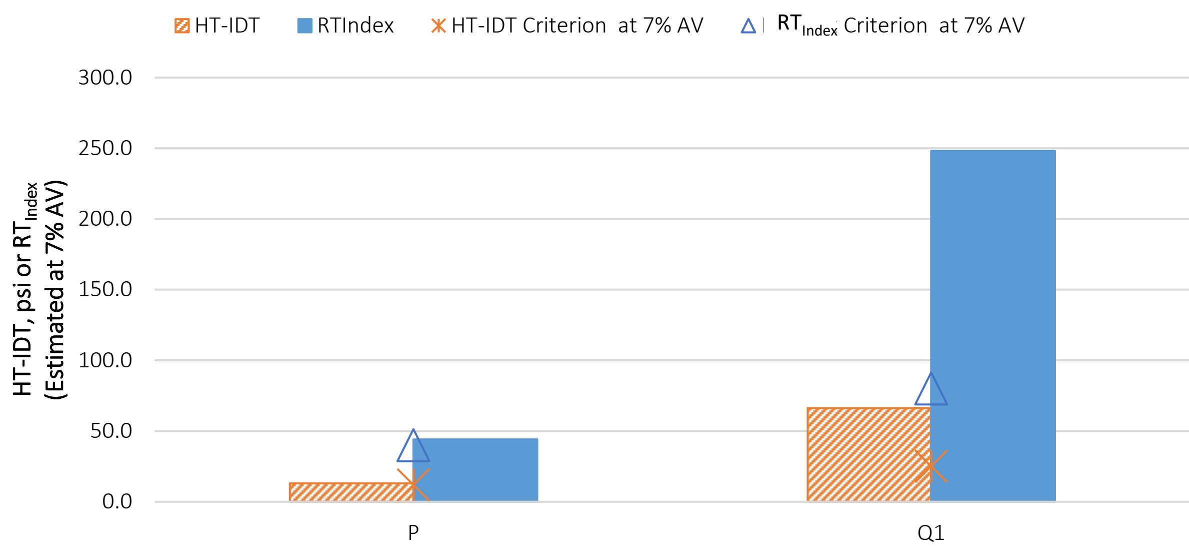
Airfield Project with Good Field Rutting Performance	Binder PG	Test Temperature, °C	% Deviation from Criteria				
			APA 100 psi/ 100 lb	APA 250 psi/ 250 lb	HWTT	HT-IDT	RT_{Index}
P	PG 64-28	52	26%	N/A	33%	8%	10%
Q1	PG 76-22	40	10%	53%	54%	165%	210%
Q2	PG 76-22	40	6%	32%	N/A	N/A	N/A
Q (at design AV, 64 °C)	PG 76-22	64	N/A	43%	N/A	N/A	N/A

N/A = Not applicable or not available.



Source: University of Nevada, Reno

Figure 28. Non-Rutted FMFC Core Test Data for APA and HWTT Estimated at 7 percent AV



Source: University of Nevada, Reno

Figure 29. Non-Rutted FMFC Core Test Data for HT-IDT and RT_{Index} Estimated at 7 percent AV

Chapter 5. Interlaboratory Study

Instructions

The purpose of the abbreviated interlaboratory study (ILS) was to verify that observed repeatability of proposed rutting tests for FAA P-401 asphalt mixtures compared favorably with reported literature for highway asphalt mixtures. For this study, a limited number of laboratories and plant-produced airfield asphalt mixtures (Item P-401) were used. In addition to the three laboratories conducting this research project (WRSC, CAIT, and TTI), the FAA Technical Center and Pavement Technology Inc. (PTI) also participated in the presented ILS. PTI, the equipment manufacturer of the APA, volunteered to participate to help supplement the FAA's APA procedure, which requires higher hose pressures and wheel loads. Table 24 summarizes each laboratory's testing responsibility. Each selected test method had three laboratories involved in testing.

Table 24. Participating Laboratories and Respective Test Procedures

Laboratory	Rutting Test Method				
	APA 100 psi/ 100 lb	APA 250 psi/ 250 lb	HWTT	HT-IDT	IRT
A	Yes	Yes	–	Yes	–
B	Yes	Yes	–	–	–
C	Yes	–	Yes	Yes	Yes
D	–	–	Yes	Yes	Yes
E	–	Yes	Yes		Yes
Number of Participating Laboratories	3	3	3	3	3

– = Did not participate in the testing.

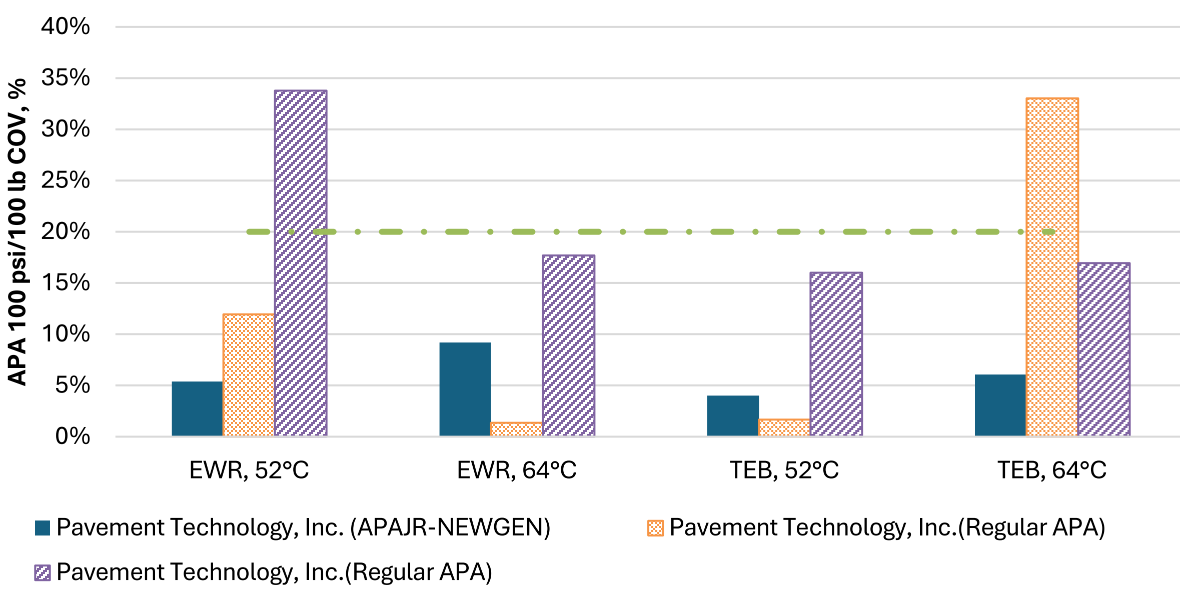
Two different plant-produced P-401 asphalt mixtures were selected for the study: EWR Airport P-401 with PG 82-22 produced from Braen Stone Industries in Sparta, NJ, and TEB Airport P-401 with PG 64-22 produced from Tilcon New York, Inc. in Mt. Hope, NJ. Test specimens were compacted to 7±0.5 percent AV. Four specimens were used for each of the APA and HWTT tests, while three specimens were used for the HT-IDT and IRT tests. All test specimens were produced by the University of Nevada, Reno to minimize the addition of variability to the test results due to specimen preparation at multiple laboratories. AV were determined for each specimen using the measured theoretical maximum specific gravity (G_{mm}) for each of the two asphalt mixtures (ASTM, 2019). Each test method was conducted in accordance with the respective standardized procedure as follows:

- APA: AASHTO T 340-2019.
- HWTT: AASHTO T 324-2022.
- HT-IDT: ASTM D8360.
- IRT: ASTM D8360-22.

Two test temperatures (52 °C and 64 °C) were selected for each of the two asphalt mixtures to cover a wider range of mixture characteristics and stiffness, resulting in a total of 216 samples prepared within the target AV level for the ILS testing. The test temperatures recommended were within the range encountered during the project and were expected to produce four different mixture responses during testing. Detailed instructions for specimen handling, conditioning protocols, testing, and reporting were provided to all participating laboratories prior to shipment of the compacted samples and the start of experimental testing.

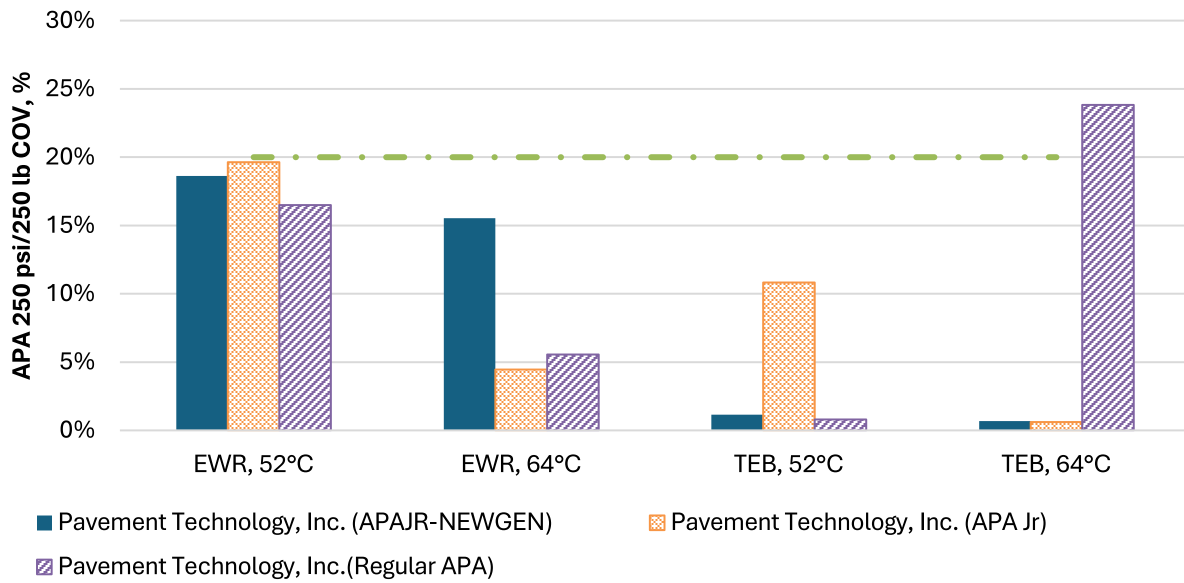
Results

The ILS data were first analyzed by equipment manufacturer and model used during experimental testing. The COV for the two asphalt mixtures at each of the two test temperatures is shown in Figure 30 through Figure 34 for the five different rutting test parameters. This analysis aimed to identify any potential bias associated with a specific equipment manufacturer or model. In some cases, the bar plots indicate high consistency in COV across different test equipment, while in others, a higher COV deviation is observed. However, none of the employed test equipment was associated with any bias toward a consistently higher or lower test result. Furthermore, a dashed line at 20 percent COV level was added to all five bar charts, representing the typical COV of these rutting mechanical tests found in previous highway studies and literature. The 20 percent COV was assumed when developing the test thresholds in Table 17 and Table 18 to account for variability in the rutting mechanical tests prior to recommending the final criterion limits.



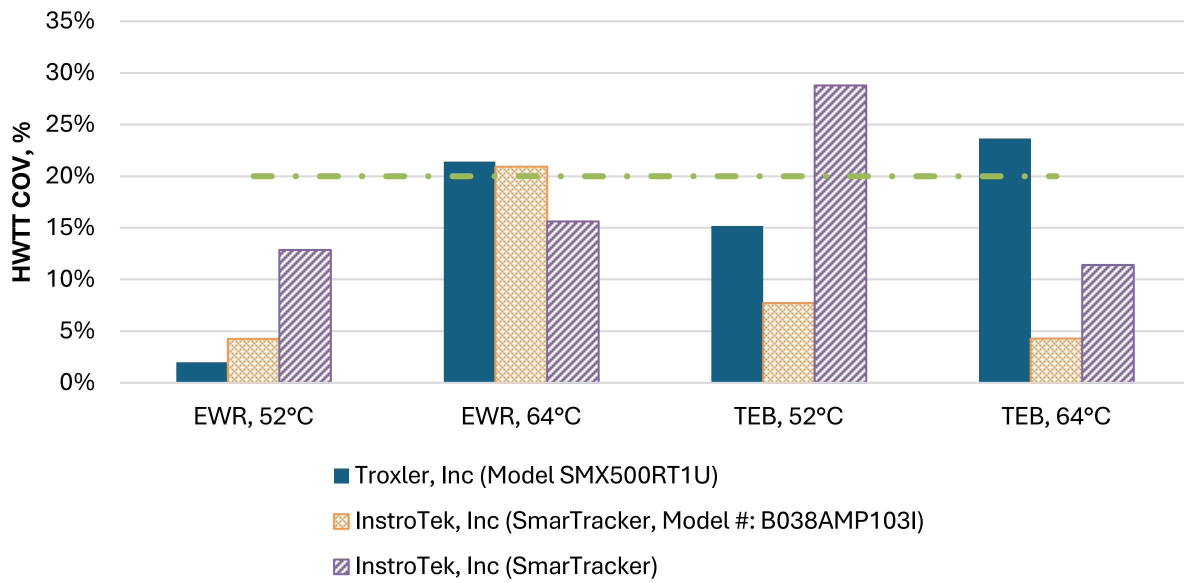
Source: University of Nevada, Reno

Figure 30. APA 100 psi/100 lb COV After 8,000 Cycles per Equipment Model



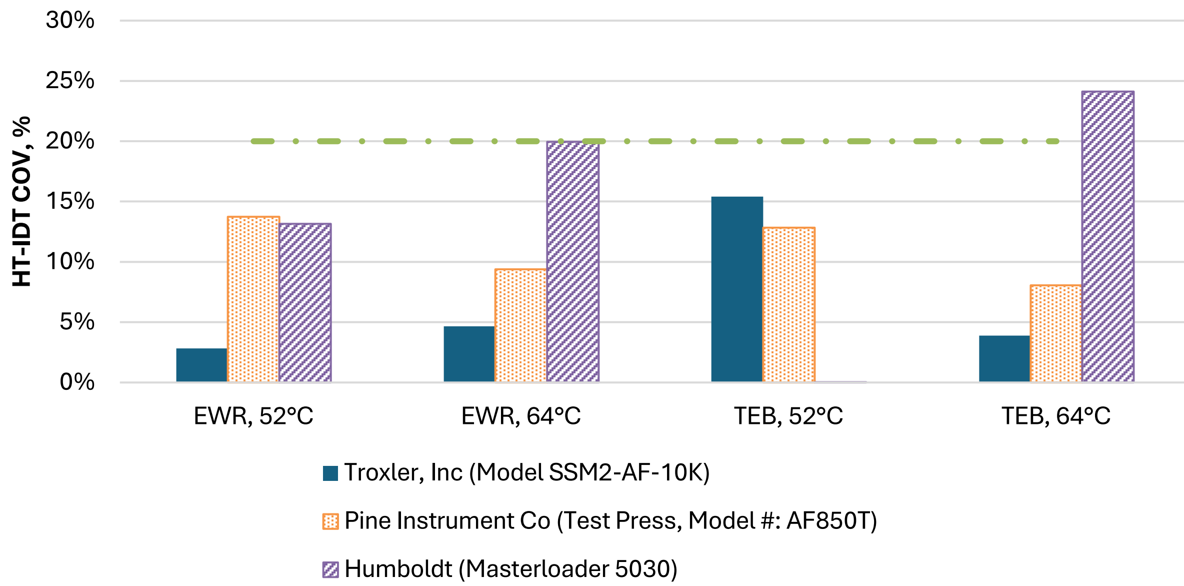
Source: University of Nevada, Reno

Figure 31. APA 250 psi/250 lb COV After 4,000 Cycles per Equipment Model



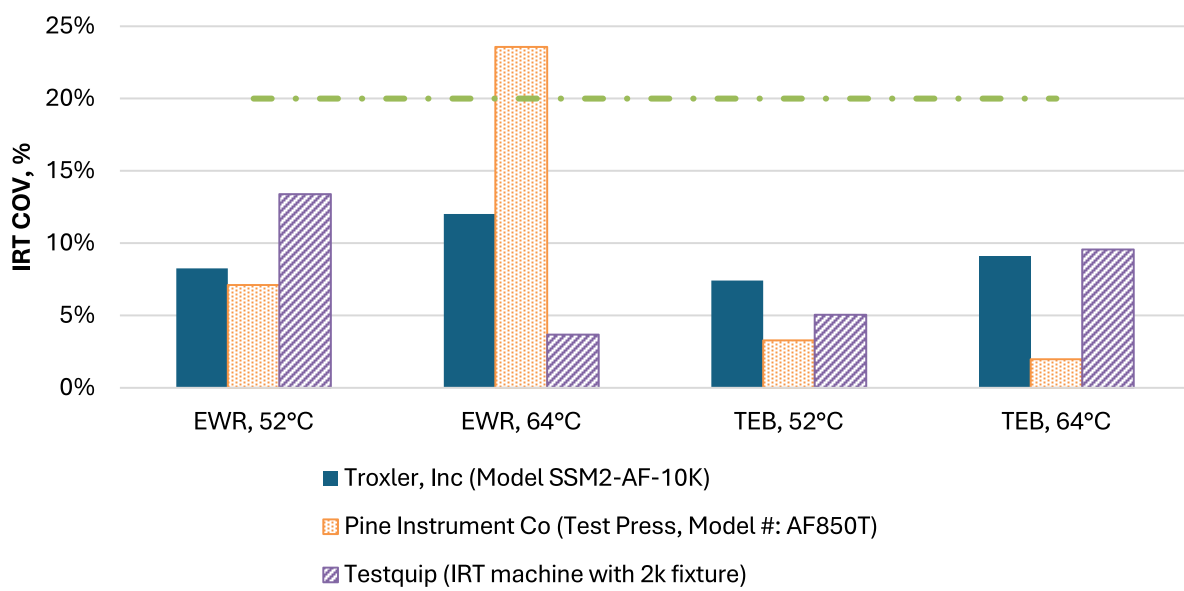
Source: University of Nevada, Reno

Figure 32. HWTT COV After 5,000 Passes per Equipment Model



Source: University of Nevada, Reno

Figure 33. HT-IDT COV per Equipment Model



Source: University of Nevada, Reno

Figure 34. IRT COV per Equipment Model

Subsequently, the ILS data was analyzed based on ASTM E691, *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method* (ASTM, 2021). This ASTM practice presents the statistical techniques with adequate information for developing the precision statement of a certain test method. The main objective was to quantify the repeatability and reproducibility standard deviation for each test parameter, which would further allow for the computation of within-laboratory and between-laboratory COV, respectively.

The within-laboratory and between-laboratory standard deviation, along with the COVs, are tabulated in Table 25 and Table 26, respectively, for each of the four test combinations for the five rutting test parameters. Additionally, a graphical representation using the whisker plots is shown in Figure 35 and Figure 36 for within-laboratory and between-laboratory COV, respectively, with the average and error bars for each rutting parameter.

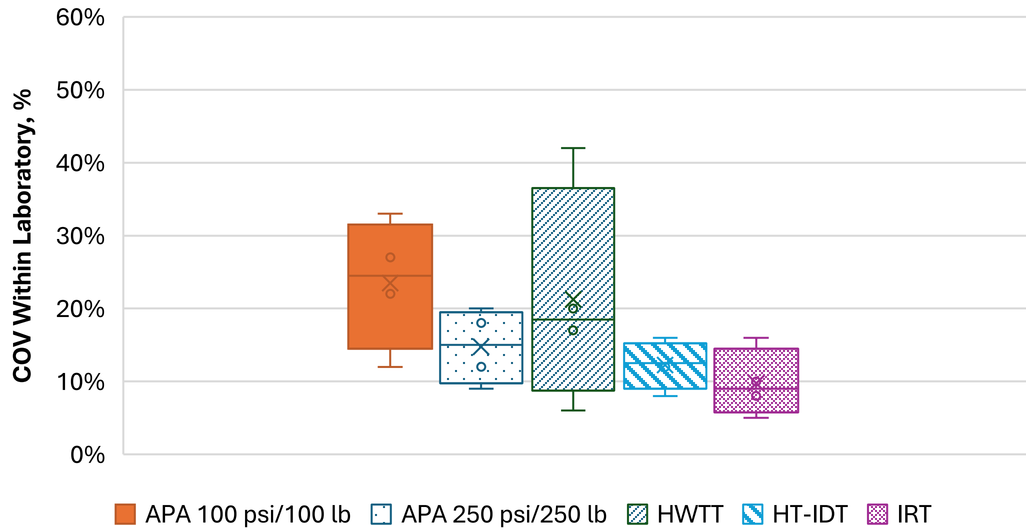
Lower variability, hence lower COV with a maximum average of 14 percent between-laboratory COV, was observed for the monotonic tests compared to the repeated load testing. The minimum within-laboratory and between-laboratory average COV were observed for the IRT with 10 and 13 percent, respectively. Conversely, the APA 100 psi/100 lb rut depth after 8,000 cycles test showed the highest within-laboratory average COV of 24 percent and between-laboratory average COV of 31 percent. The ILS data imply the suitability of using the 20 percent COV limit in the recommended test criteria, based on the average values reported for the four test combinations.

Table 25. Within- and Between-Laboratory Standard Deviations

Test	APA 100 psi/100lb (8,000 Cycles)		APA 250 psi/250 lb (4,000 Cycles)		HWTT (5,000 Passes)		HT-IDT		IRT	
Mix	*Within lab σ*	*Between labs σ*	*Within lab σ*	*Between labs σ*	*Within lab σ*	*Between labs σ*	*Within lab σ*	*Between labs σ*	*Within lab σ*	*Between labs σ*
EWR 52	0.3	0.4	0.3	0.3	0.1	0.5	4.4	5.2	14.4	14.4
EWR 64	0.2	0.9	0.2	0.2	0.3	0.3	2.8	2.8	11.9	11.9
TEB 52	1.0	0.4	1.0	0.7	1.8	1.6	1.4	2.1	2.9	5.7
TEB 64	1.0	1.0	1.0	1.3	1.8	3.1	1.4	1.4	2.9	5.7
Pooled Standard Deviation	0.8	0.7	0.8	0.8	1.3	1.7	2.8	3.2	9.6	10.2

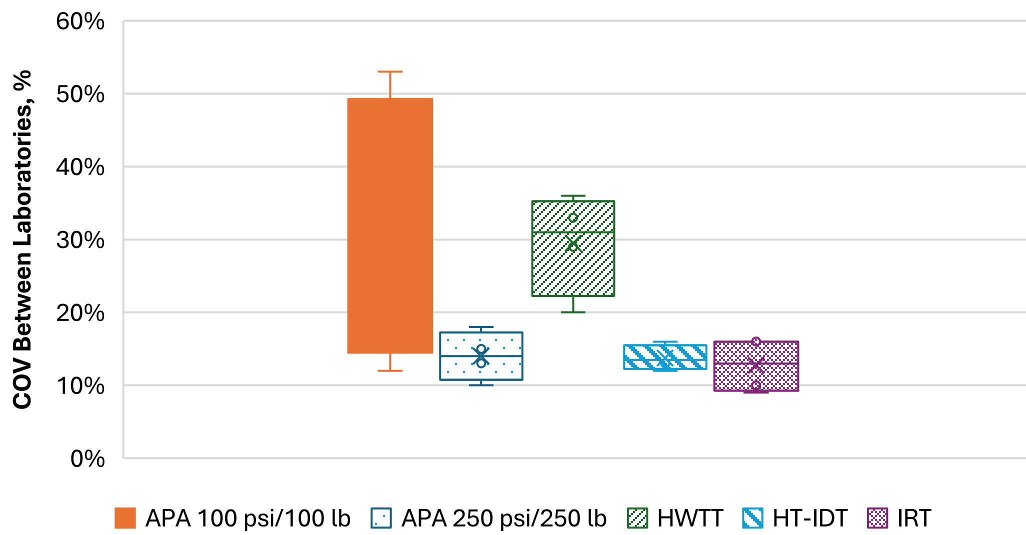
Table 26. Within- and Between-Laboratory COVs

Test	APA 100 psi/100lb (8,000 Cycles)		APA 250 psi/250 lb (4,000 Cycles)		HWTT (5,000 Passes)		HT-IDT		IRT	
Mix	*COV within lab*	*COV between labs*	*COV within lab*	*COV between labs*	*COV within lab*	*COV between labs*	*COV within lab*	*COV between labs*	*COV within lab*	*COV between labs*
EWR 52	27%	38%	18%	18%	6%	33%	12%	14%	10%	10%
EWR 64	12%	53%	9%	10%	20%	20%	13%	13%	16%	16%
TEB 52	33%	12%	20%	13%	42%	36%	8%	12%	5%	9%
TEB 64	22%	22%	12%	15%	17%	29%	16%	16%	8%	16%
Avg.	24%	31%	15%	14%	21%	29%	12%	14%	10%	13%



Source: University of Nevada, Reno

Figure 35. Whisker Plot of Within-Laboratory COV from ILS Data



Source: University of Nevada, Reno

Figure 36. Whisker Plot of Between-Laboratory COV from ILS Data

Chapter 6. Implementation Plan

The value of research is realized primarily through implementation of its findings. The societal benefits of pavement research are often measured by its ability to influence and alter specifications and practices related to pavement design and construction. Economic benefits can be demonstrated through reductions in initial and/or life-cycle pavement costs.

A planned implementation phase is necessary to fully appreciate the economic benefits of research projects. The traditional process for developing a new technology for society includes three phases: research, development, and implementation. Research typically yields several items that are not fully defined or understood, requiring a development effort to move the research into implementation. In many cases, the development and implementation phases merge into one.

Task 7 in the initial project proposal focused on developing an implementation plan that included consideration of some development efforts. Moving research into practice for most airfield pavement-related topics requires a clearly defined and structured development phase. The key efforts to implement the developed BMD method for airfield pavements are outlined below.

Background

BMD for asphalt mixtures is not a new concept. The long-standing Marshall and Hveem procedures, which were widely used from the 1930s into the 2000s, contain elements of BMD. Both methods incorporated the concepts of providing stability and durability, including crack resistance. The concept of BMD as it has emerged in the 2000s focuses on permanent deformation (stability—rutting and shoving) and crack resistance (fatigue cracking, thermal cracking, reflection cracking, etc.). It also includes considerations for durability associated with both asphalt binder aging tests and asphalt mixture aging prior to performance testing.

The goal of BMD is to develop performance-based tests that are sensitive to the rutting and cracking behavior of asphalt mixtures and can be used for mix design, test strip or trial batches, and production QA. This is a difficult if not impossible task when considering the current limitations in workforce availability (both in numbers and technical expertise), as well as time constraints related to sampling, testing, and variability in asphalt mixture production at the plant.

The BMD development and implementation plan described below assumed that existing NAPA-FAA projects would provide the following:

- Test methods for both rutting and cracking tests for use in mix design, test strip/trial batches, and production control.

- BMD specification complete with the mix design method and project production acceptance criteria.

The goals of the development and implementation efforts were as follows:

- Fill gaps in the research program that are needed to provide improved mix design tests and production acceptance criteria.
- Provide a foundation for contractors and material suppliers to understand the financial risks associated with using the BMD specification.
- Allow for more rapid deployment of the BMD specification on airfield pavement projects.

A list of nine prioritized development and implementation efforts was established and depending on the funding available, a detailed implementation plan will be developed and executed for each effort. Efforts 1 and 9 (formation of Working Group and knowledge sharing) are high-priority items that will be needed to implement the findings from the existing research project regardless of any additional development and implementation efforts. The Working Group and knowledge sharing efforts will also consider other ongoing research projects in both the highway and airport paving areas.

The following are the nine high-priority items identified for development and implementation of the BMD specification for airfield asphalt pavements:

1. Establish a working group of FAA engineers, airport authority engineers, consultants, and contractor/material suppliers.
2. Assess test method and construction variability.
3. Evaluate suitability of test methods for mix design, test strip or trial batches, and production control (QC and acceptance).
4. Analyze the impact of mixture component variability on test method output parameters.
5. Investigate correlations among rutting tests and among cracking tests.
6. Examine relationships among test methods, pavement performance, and acceptance criteria.
7. Revise specifications and test methods.
8. Implement accreditation of laboratories and certification of technicians.
9. Promote specification acceptance by engineers and contractors through workshops, conferences, webinars, briefs, videos, etc.

Development and Implementation Efforts

Effort 1: Establish a Working Group of FAA Engineers, Airport Authority Engineers, Consultants, and Contractor/Material Suppliers

A Working Group of FAA engineers, airport authority engineers, consultants, contractors, and materials suppliers that routinely perform work on airports needs to be formed to review the test methods and specifications that are products of this FAA–NAPA research project. This review will identify potential shortfalls in the test methods and specification that may need to be addressed by the development and implementation effort identified below.

The Working Group will also be responsible for reviewing the test plans associated with the development and implementation plans described below, as well as assisting in prioritizing the effort. The results from the development and implementation efforts will be used to revise the test methods and specifications with input from the Working Group, as identified in Effort 7.

The Working Group members will gain knowledge of the test methods and specifications and will assist in using some of the airfield projects they are associated with to help collect field production samples and information, as described in several of the efforts shown below.

Effort 2: Assess Test Method and Construction Variability

Variability of laboratory and field tests must be defined to prepare a specification with acceptance limits that balances buyer (public agency) and seller (contractor) risks. Tests are performed on the following types of samples:

- LMLC—typically for laboratory mix design purposes.
- Plant mixed and laboratory compacted—not reheated—samples typically used for production control testing by contractors and for some of the production acceptance testing by the agency.
- RPMLC—typically used for some of the production acceptance testing by the agency.
- FMFC—typically used for some of the production acceptance testing by the agency.

If all four of these sample types are used for specification purposes, variability must be defined to ensure that a knowledge base is used to balance buyer and seller risk.

When samples are obtained for laboratory mix design, the two major sources of variability are sampling and testing errors. When samples are obtained after plant production, the major sources of variability are sampling, testing, and the additional variability associated with the construction process. Ideally, it is essential to define the variability resulting from the construction process, as this is the only variability that directly impacts performance.

Sampling and testing variability can mask the results used for acceptance on projects; therefore, it is important to minimize sampling and testing variability.

To minimize sampling variability, it is essential to use appropriate test methods and employ certified technicians. However, test method variability can significantly influence the analysis of the production test results. Factors such as sample compaction, temperature control, lag time, dwell time, technician certification, laboratory equipment, and laboratory accreditation are all contributors to overall variability. For some commonly used tests for asphalt mixtures, sampling and testing variability can account for up to 70 percent of the reported variability on production samples.

It is therefore important that test method variability be studied and factors significantly impacting the test results be quantified and controlled. For example, test equipment manufacturers often develop equipment with varying loading frame stiffness or compliance, different loading rates, and different measuring systems, all of which can significantly impact the test results. Temperature control of the samples at the time of testing, as well as dwell and lag times, are all important factors.

Several of the tests used for the BMD effort have defined measures of variability, which can be found in the literature. Additional studies are currently underway. A mini interlaboratory study using RPMLC samples was conducted as part of this research project. These data need to be carefully reviewed, and any gaps should be addressed through further studies and supporting information.

Very little data are available to define construction variability associated with some of the proposed tests. Some limited information from highway research is available. The literature should be thoroughly reviewed, and any identified gaps should be addressed.

Both within- and between-laboratory test method and construction variabilities need to be defined for some tests and sample preparation methodologies. Within-laboratory variability is more commonly reported in the literature than between-laboratory variability. Variability studies are both expensive and time-consuming to perform.

Effort 3: Evaluate Suitability of Test Methods for Mix Design, Test Strip or Trial Batches, and Production Control (QC and Acceptance)

The tests that are recommended from the rutting study include the APA 100 psi/100 lb, APA 250 psi/250 lb, HWTT, IRT, and HT-IDT. Some cracking tests will be recommended from the BMD cracking study.

The suitability of these tests for use in mix design, acceptance of test strip or trial batches, production control, and pay factors needs to be evaluated. Key considerations include equipment availability and cost; laboratory size and space requirements; sample preparation time; temperature control requirements; dwell and lag times and temperatures; sample sizes and the quantity of materials sampled and compacted; testing speed; time required for test result analysis; ease of interpretation of test results; data availability in the laboratory and plant; and the resource needs of contractors and public agencies (technical qualifications and workforce size). Ideally, the tests selected will be performance-based or performance-related and suitable for use at the various stages of asphalt mix design and production.

Effort 4: Analyze the Impact of Mixture Component Variability on Test Method Output Parameters

The sensitivity of the test parameters to changes in mixture components—including source and grade of asphalt binder; use of polymers or other additives in asphalt binders; source and gradation of aggregate (fine graded, coarse graded, stone matrix asphalt, etc.); aggregate shape and surface texture; asphalt binder content; volumetrics of mixtures (including AV, voids in mineral aggregates, and voids filled with asphalt); dwell time; and lag time (sample preparation and aging)—needs to be defined. Limited data are available in the literature on the sensitivity of some of the proposed test methods and their test parameters to mixture variables.

A laboratory study to define the sensitivity of the test methods and their parameters is needed for most of the proposed tests. The results will support asphalt mixture design and guide changes to mixture components during production.

Effort 5: Investigate Correlations Among Rutting Tests and Among Cracking Tests

Acceptance criteria for certain rutting tests are based on several years of use in actual construction projects. Thus, the engineering community generally more confidence in these tests and their acceptance criteria compared to others.

Establishing strong correlations among various rutting (or cracking) tests will allow for the selection of a test method that may be suitable for a variety of uses, as defined in Effort 3. Ideally, a desirable test would be one that can be conducted in a short period of time, requires inexpensive equipment, has a low within- and between-laboratory variability, and is suitable for mix design, test strip and trial batches, and production acceptance and QC. This may be one of the newer tests that, with a strong correlation to an existing established test, would become acceptable to the engineering community.

This study could be performed on samples obtained from laboratory mixing or from plant-produced mixtures. The samples should be compacted in the laboratory to reduce variability. Likewise, to reduce variability it is recommended that laboratory-mixed samples be used.

Samples of materials (asphalt binders and aggregates) should be obtained from the field projects that are used in Effort 6. The samples should be carefully prepared to ensure uniformity in asphalt binder content, aggregate gradation, and AV. Dwell and lag times should be kept the same prior to laboratory testing. Test results could also be utilized to determine within-laboratory variability, as described in Effort 2.

Effort 6: Examine Relationships Among Test Methods, Airfield Pavement Performance, and Acceptance Criteria

This is a large study that will require an extended amount of time. Airport projects need to be identified that include a range of variables, such as facility (runway, taxiway, aprons, etc.), location in different climatic regions, use of different aggregates and asphalt binders, and varying levels of rutting and cracking parameters (both within and outside of the proposed specification limits). The performance of these projects should be evaluated over an extended period to provide the information necessary to establish relationships among the different types of tests, test parameters, and airfield pavement performance. Acceptance criteria for specification use can be developed from these relationships.

This effort can be ongoing as the BMD specification is implemented. A series of field projects need to be identified and materials sampled and tested with the proposed tests. These field implementation projects can be placed into the following four categories:

- Benchmarking Projects—A relatively large number of airfield projects should be identified, with samples of the asphalt mixtures collected and tested. The data generated will provide a good estimate of the quality of the asphalt mixtures being designed, produced, and placed under today's specifications. These projects will provide information that can be used to establish correlations among the rutting (and cracking) tests, as well as insight into the relationships between the test methods, their parameters, and field performance. Extensive field performance measurements are not anticipated in this effort. The effort is aimed at sampling a large number of airfield projects to determine the mixture parameters as measured by the proposed tests.
- Experimental Sections (Demonstration Sections)—This effort will leverage existing scheduled airfield projects and primarily use standard asphalt mixtures on the majority of the project. The implementation team will work with the airport authority, FAA, and contractor to develop a BMD using the project's asphalt binders and aggregates. A relatively short section (equivalent to 1 day's production) will be placed with the designed BMD mixture. Both the demonstration section and the control mixture will be monitored in a reasonable level of detail. Extensive sampling and testing will be performed. Typically, about six airfield demonstration projects per year can be placed and evaluated. A multiple-year effort will be needed.
- Shadow Specification—On this category of projects, a significant portion of the project will be constructed with the BMD material. The BMD specification will be used to control this portion of the project; however, pay factors will not be used. Thus, the contractor has limited risk.

- Pilot BMD Project—The latest BMD specification will be used to design and construct an airfield project. The BMD acceptance process and pay factors will be used on this pilot project. This is an interim step prior to full implementation of the BMD.

Careful planning of Effort 6 can provide information for several other development and implementation efforts.

Effort 7: Revise Specifications and Test Methods

As significant findings are obtained from the development and implementation projects, the BMD specification will be reviewed and revised. The Working Group will work with the development and implementation team to make these changes.

Effort 8: Implement Accreditation of Laboratories and Certification of Technicians

The BMD specification will contain some of the latest test methods, and it is important for laboratory accreditation organizations and technician certification programs to be aware that these test methods are being used for materials acceptance. Whenever possible, these test methods need to be standardized by ASTM International or AASHTO.

National laboratory accreditation programs need to include accreditation processes for the test equipment required for the implemented test methods. Similarly, national certification programs need to incorporate technician certification for these employed test methods. While AASHTO-supported programs may or may not be able to fulfill these tasks, other national accreditation and certification programs can be available to meet these needs.

Effort 9: Promote Specification Acceptance by Engineers and Contractors through Workshops, Conferences, Webinars, Briefs, Videos, etc.

Successful implementation of the new BMD specification for the FAA will require some level of acceptance from the contracting and materials supply industries. If the contracting community does not support or accept parts of the specification, they may increase their bids to account for additional risks, resulting in higher project costs. It is important that the contracting, consulting engineering, and airport authorities in the United States become familiar with the specification, understand its background, recognize its benefits, and realize its risks.

The Working Group identified in Effort 1 will assist with the introduction of the new BMD specification. Workshops, a training program, conferences, videos, webinars, and other knowledge-sharing methods will be needed to provide confidence to those responsible for implementation of the specification. This will be a continuous effort over the next several years.

Chapter 7. Conclusions and Recommendations

The primary goal of this project was to establish representative rutting test protocols that are tailored to actual airfield pavements and derive proper test criteria for airfield asphalt mixtures. Four different rutting test methods, which can serve as part of the FAA BMD during mix design as well as during production, were explored in this study.

One of the key elements required to establish test acceptance criteria is to have laboratory testing protocols that best simulate actual field conditions for airfield pavements. These test parameters involved selecting proper compaction method, AV level, specimen size and preparation method (cutting, coring, etc.), loose-mixture aging temperature and duration, compacted specimen conditioning, test temperature, and load level and rate reflecting actual flexible airfield pavement conditions. Proper testing protocols were classified into the four main categories shown in Table 4 and selected based on in-place data from actual airfield sections, common test methods, and previous research findings.

Robust correlations were observed between the APA test at both conditions—100 psi/100 lb and 250 psi/250 lb—as well as between the APA test, HT-IDT, and IRT. The HWTT rut depth after 20,000 passes did not hold clear trends with the rest of the laboratory rutting mechanical tests. Additional parameters from the HWTT that could isolate the stripping failure from the mixture rutting characterization were investigated. The new regressions of the APA and IRT tests with the HWTT rut depth at 5,000 passes demonstrated stronger correlation compared to the HWTT total rut depth at 20,000 passes. Based on experimental test results at the two AV levels (i.e., 5 and 7 percent), the 7±0.5 percent AV level was recommended for the four rutting tests. This level allows for preparing specimens by directly molding to either 75 mm for APA or 62 mm for HWTT, HT-IDT, and IRT.

Furthermore, a mechanistic-empirical approach was developed to refine the current FAA rutting test criterion for the APA 250 psi/250 lb test to consider aircraft speed and load. New test criteria were established for different rutting test methods: APA 100 psi/100 lb, HWTT, HT-IDT, and IRT at environmental test temperature and 7-percent AV level. The mechanistic framework included the modeling of four airfield sections using 3D-Move software to compute the pavement responses under varying key parameters critical to pavement performance.

Mechanistic analyses involving three pavement temperatures, three traveling speeds, and five aircraft load levels were completed. The state of stresses generated by 3D-Move were used to determine the deviatoric and confining stresses that simulate airfield pavement loading conditions. Subsequently, the RLT test was performed using three sets of σ_d and σ_c, encompassing the broad spectrum of computed stress states for the modeled airfield pavements. Rutting performance models were then developed for the four airfield mixtures, enabling quantification of the sensitivity of rutting to different speeds, temperatures, and load levels. Using the results from the mechanistic-empirical analysis, two sets of rutting test criteria were established, with one for airfield pavements with slow or stationary aircraft (Table 17) and the other for all airfield pavement types (Table 18).

The set of recommended rutting test criteria were verified using laboratory experimental testing of field cores sampled from actual airfield pavement sections showing either satisfactory or poor field rutting performance. The field core test results were adjusted from in-place to 7-percent AV level, based on laboratory-developed regression models using an array of LMLC and RPMLC airfield asphalt mixtures.

The recommendations of this project include the revised specifications for P-401/P-403 airfield asphalt mixtures (refer to Appendix G), as summarized in Table 27. Two AV correction models for each of the monotonic tests that are recommended to be implemented for QC and acceptance are presented in Equation 10 and Equation 11. The summary of the main testing protocols and thresholds to be considered in the final P-401/P-403 specifications are summarized in Table 28.

$$\boldsymbol{RT_{Index@7\%}} = \mathbf{0.449}\boldsymbol{e}^{\mathbf{0.1141}\boldsymbol{AV}} \times (\boldsymbol{RT_{Index}}) \quad \text{Equation 10}$$

$$\boldsymbol{HT - IDT_{Index@7\%}} = \mathbf{0.526}\boldsymbol{e}^{\mathbf{0.0917}\boldsymbol{AV}} \times (\mathbf{HT - IDT}) \quad \text{Equation 11}$$

Table 27. Implementation Stages of Rutting Mechanical Tests

Stage	Designer	Owner Acceptance	Contractor QC	Independent Assurance
Mix Design (LMLC)	APA 250 psi/250 lb, APA 100 psi/100 lb, HWTT, HT-IDT, or IRT	Review only	N/A	N/A
Control Strip (RPMLC)	N/A	APA 250 psi /250 lb, APA 100 psi /100 lb, HWTT, HT-IDT, or IRT	APA 250 psi /250 lb, APA 100 psi /100 lb, HWTT, HT-IDT, or IRT	APA 250 psi /250 lb, APA 100 psi /100 lb, HWTT, HT-IDT, or IRT
Production (RPMLC)	N/A	HT-IDT or IRT	N/A[1]	HT-IDT or IRT

N/A=not applicable.
[1]Contractors can elect to run HT-IDT or IRT as part of their process control.

Table 28. Final Recommendations for Consideration in P-401/P-403 Specifications

Rutting Tests	APA 250 psi/250 lb, APA 100 psi/100 lb, HWTT, HT-IDT, and IRT
Implementation Stage	Refer to Table 27
Specimen Type and AV%	Directly molded samples to 7±0.5% AV (Table 4)
Specimen Size	APA: 150 mm by 75±2 mm; HWTT, HT-IDT, IRT: 150 mm by 62±1 mm (Table 4)
Loose Mixture Conditioning	Protocols developed by the research team for LMLC and RPMLC (Refer to Appendix B and Table 4)
Compacted Mixture Conditioning	Refer to Table 4
Test Temperature	LTPPBind Online environmental PG (no grade bumping), 12.5 mm rut depth, 50% reliability, at surface (40 °C–64 °C) (Table 4)
Test Load Level/Rate	Refer to Table 4
Test Criteria	Refer to Table 17

Future Work

To expand the implementation of the BMD framework into the production phase (i.e., QC and acceptance), a new project phase is recommended. This phase should include pilot projects for airfield mixtures designed and produced in accordance with the revised proposed final P-401/P-403 specifications. These pilot projects align with the suggested implementation plan in this study and are intended to evaluate the revised specification and its practical application in real-world conditions.

Asphalt mixtures from these pilot projects will be tested, and the respective pavements will be monitored and evaluated against the proposed test criteria. These pilot projects will help establish key testing parameters such as test frequencies during production and other essential procedures and practices. Additionally, they will help identify common challenges associated with the recommended test protocols and thresholds during production.

The follow-up phase with trial projects will complete the primary development goals outlined in the implementation plan. This includes assessing test method and production variability and expanding the understanding of relationships between different test methods, airfield pavement performance, and acceptance criteria. Furthermore, this phase will provide an opportunity to evaluate the revised specification and its practical application in real-world conditions. As part of the future recommended phase, the performance of airfield pavement sections sampled during this project's experimental testing should be monitored over additional service years. This long-term monitoring will enhance the correlation between the established laboratory mechanical test criteria and actual field performance.

References

AASHTO. (2021). *AASHTO R 84, Standard Practice for Developing Dynamic Modulus Master Curves for Asphalt Mixtures Using the Asphalt Mixture Performance Tester (AMPT)*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2022a). *AASHTO R 30, Standard Practice for Laboratory Conditioning of Asphalt Mixtures*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2022b). *AASHTO R 47, Standard Practice for Reducing Samples of Asphalt Mixtures to Testing Size*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2022c). *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)*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2022d). *Draft Proposed Changes by NCAT to AASHTO MP 46-22, Standard Specification for Balanced Mix Design*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2023a). *AASHTO T 324, Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO. (2023b). *AASHTO T 340, Standard Method of Test for Determining Rutting Susceptibility of Hot Mix Asphalt (HMA) Using the Asphalt Pavement Analyzer (APA)*. Washington, DC: American Association of State Highway and Transportation Officials.

Advanced Asphalt Technologies, LLC. (2011). *NCHRP Report 673: A Manual for Design of Hot-Mix Asphalt with Commentary*. Washington, DC: Transportation Research Board. doi:https://doi.org/10.17226/14524

Alabama DOT. (2022). *ALDOT-458: High Temperature Indirect Tensile Test for HMA*. Alabama Department of Transportation.

ASTM. (2017). *ASTM Standard D6931, Indirect Tensile (IDT) Strength of Asphalt mixtures*. West Conshohocken, PA: ASTM International. doi:10.1520/D6931-17

ASTM. (2019). *ASTM Standard D2041, Standard Test Method for Theoretical Maximum Specific Gravity and Density of Asphalt Mixtures*. West Conshohocken, PA: ASTM International. doi:10.1520/D2041-03A

ASTM. (2020). *ASTM Standard D6926, Standard Practice for Preparation of Asphalt Mixture Specimens Using Marshall Apparatus*. West Conshohocken, PA: ASTM International. doi:10.1520/D6926-20

ASTM. (2021). *ASTM Standard E691, Standard Test Method for Conducting an Interlaboratory Study to Determine the Precision of a Test Method*. West Conshohocken, PA: ASTM International. doi:10.1520/E0691-21

ASTM. (2022). *ASTM Standard D8360, Standard Test Method for Determination of Rutting Tolerance Index of Asphalt Mixture Using the Ideal Rutting Test*. West Conshohocken, PA: ASTM International. doi:10.1520/D8360-22

Christensen, D. W. (2008). *Airport Asphalt Pavement Technology Program Project 04-02: PG Binder Grade Selection for Airfield Pavements*. Airport Asphalt Pavement Technology Program.

Christensen, D. W., & Bonaquist, R. F. (2007). Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing. *Transportation Research Circular E-C124*.

Daidzic, N. E. (2016). Optimization of Takeoffs on Unbalanced Fields Using Takeoff Performance Tool. *International Journal of Aviation Aeronautics and Aerospace, 3*(3).

Elias, N. G. (2024). *Advancing FAA Asphalt Mix Design: Evaluation of Rutting Mechanical Tests for Balanced Mix Design*. Ph.D. Dissertation, University of Nevada, Reno.

Elias, N. G., Hand, A. J., Sebaaly, P., Hajj, E. Y., Piratheepan, M., & Gibson, S. (2022). Local Agency Transition to Balanced Mix Design. *International Journal of Pavement Engineering, 23*(13), 4792–4802.

Epps, J. A., Leahy, R. B., Mitchell, T., Ashmore, C., Seeds, S., Alavi, S., & Monismith, C. L. (2000). WesTrack: The Road to Performance-Related Specifications. In World of Asphalt Pavements. *International Conference, 1ST, 2000*. Sydney, New South Wales: Australian Asphalt Pavement Association.

FAA. (2018). *AC 150/5370-10H – Standard Specifications for Construction of Airports*. Washington, D.C.: Federal Aviation Administration. Retrieved from https://www.faa.gov/airports/resources/advisory_circulars/index.cfm/go/document.current/documentNumber/150_5370-10

FAA. (2021a). *AC 150/5320-6G – Airport Pavement Design and Evaluation*. Washington DC: Federal Aviation Administration, AAS-100, Office of Airport Safety & Standards–Airport Engineering Division.

FAA. (2021b). *CY 2021 Enplanements at All Airports (Primary, Non-primary Commercial Service, and General Aviation)*. Retrieved from https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/passenger/cy21_all_enplanements

FAA. (2021c). *FAA Rigid and Flexible Iterative Elastic Layer Design (FAARFIELD) Version 2.0*. Retrieved from Airport Design Software: https://www.faa.gov/airports/engineering/design_software

FAA. (2022). *Airport Pavements*. Retrieved from YouTube: https://www.youtube.com/watch?v=yWihlqUvBos

FAA. (2024a). *AC 150/5300-13B – Airport Design*. Washington, DC: Federal Aviation Administration.

FAA. (2024b). *Aeronautical Information Services*. (Federal Aviation Administration) Retrieved from https://nfdc.faa.gov/nfdcApps/services/ajv5/airportDisplay.jsp?airportId=EWR

FAA. (2024c). *Airport Data and Information Portal*. (Federal Aviation Administration) Retrieved from Search Facility Repository: https://adip.faa.gov/agis/public/#/airportSearch

FAA. (2024d). *NAPTF*. Retrieved from https://www.airporttech.tc.faa.gov/NAPTF.

FHWA. (2024). *LTPPBind Online, LTPP InfoPave Tools*. (Federal Highway Administration) Retrieved from https://infopave.fhwa.dot.gov/Tools/LTPPBindOnline

Gonzalez, J., & Lombaerts, T. (2023). *Calculation and Analysis of the Rejected Takeoff Speed of a Commercial Airliner Under Various Environmental Conditions*. 2023 Regional Student Conferences.

Hajj, E. Y., Aschenbrener, T. B., & Nener-Plante, D. (2022a). *Positive Practices, Lessons Learned, and Challenges When Implementing Balanced Design of Asphalt Mixtures: Site Visits*. Pavement Engineering & Science Program, University of Nevada, Reno.

Hajj, E. Y., Aschenbrener, T., & Nener-Plante, D. (2022b). Examples of Successful Practices with State Implementation of Balanced Design of Asphalt Mixtures. *Transportation Research Record, 2676*(5), 44–66.

Hajj, E. Y., Elias, N. G., Khanal, B., Alrajhi, A., Hand, A., Bennert, T., . . . Duval, J. (2025a). *Technical Memo 1: Analysis of In-Place Density Data from Airfield Projects*. Airport Asphalt Pavement Technology Program.

Hajj, E. Y., Elias, N. G., Khanal, B., Alrajhi, A., Hand, A., Bennert, T., . . . Duval, J. (2025b). *Balanced Mix Design: Rutting Performance Tests, Phase I Interim Report*. Airport Asphalt Pavement Technology Program.

Hajj, E. Y., Elias, N. G., Khanal, B., Alrajhi, A., Hand, A., Bennert, T., . . . Duval, J. (2025c). *Technical Memo 2: Recommendations of Air Void Level for Flexible Airfield Pavements Mechanical Tests*. Airport Asphalt Pavement Technology Program.

Hajj, E. Y., Hand, A. J., Chkaiban, R., & Aschenbrener, T. B. (2019). *Index-Based Tests for Performance Engineered Mixture Designs for Asphalt Pavements*. Federal Highway Administration.

Hajj, E. Y., Siddharthan, R. V., Sebaaly, P. E., & Weitzel, D. (2007). Laboratory-Based Unified Permanent Deformation Model for Hot-Mix Asphalt Mixtures. *Journal of Testing and Evaluation, 35*(3), 272–280.

Minitab. (2024). *LLC. Minitab® Statistical Software*. Retrieved from https://www.minitab.com/en-us/products/minitab/

NAPA. (2023). *Performance Test Resources*. (National Asphalt Pavement Association) Retrieved from https://www.asphaltpavement.org/expertise/engineering/resources/bmd-resource-guide/performance-test-resources

NCAT. (2023). *Using the Hamburg Wheel Track Test for Balanced Mix Design*. National Center for Asphalt Technology, Auburn University, Auburn, AL. Retrieved from https://www.eng.auburn.edu/research/centers/ncat/newsroom/2020-fall/hamburg.html.

RPubs by RStudio. (2024). *Games-Howell Post-Hoc Test*. Retrieved from https://rpubs.com/aaronsc32/games-howell-test

Rushing, J. F., & Garg, N. (2017). Using the Asphalt Pavement Analyzer as a Mixture Performance Test to Select Appropriate Binder Grades for Airport Pavements. *Journal of Transportation Engineering, Part B: Pavements, 143*(3), 04017010.

Rushing, J. F., Little, D. N., & Garg, N. (2012). Asphalt Pavement Analyzer Used to Assess Rutting Susceptibility of Hot-Mix Asphalt Designed for High Tire Pressure Aircraft. *Transportation Research Record, 2296*(1), 97–105.

Rushing, J. F., Little, D. N., & Garg, N. (2014). Selecting a Rutting Performance Test for Airport Asphalt Mixture Design. *Road Materials and Pavement Design, 15*(1), 172–194.

Schwartz, C. W., Elkins, G. E., Li, R., Visintine, B. A., Forman, B., Rada, G. R., & Groeger, J. (2015). *Evaluation of LTTP Climatic Data for Use in Mechanistic-Empirical Pavement Design Guide Calibration and Other Pavement Analysis.* Federal Highway Administration. Washington, DC: Federal Highway Administration.

Siddharthan, R. V., Krishnamenon, N., & Sebaaly, P. E. (2000). Finite-Layer Approach to Pavement Response Evaluation. *Transportation Research Record, 1709*(1), 43–49.

Song, I., & Garg, N. (2010). High Tire Pressure and Temperature Effects on Hot Mix Asphalt Concrete Permanent Deformation Using Customized Asphalt Pavement Analyzer (APA). *In Proc. 2010 FAA Worldwide Airport Technology Transfer Conference.* Atlantic City, NJ: Federal Aviation Administration and American Association of Airport Executives.

TRB. (2022). *Transportation Research Circular E-C280: Glossary of Terms for Balanced Design of Asphalt Mixtures.* Washington, DC: Transportation Research Board.

Ulloa Calderon, A. E. (2013). *Development of a Mechanistic-Based Approach to Evaluate Critical Conditions of Hot Mix Asphalt Mixtures.* University of Nevada, Reno.

Wang, H., Li, M., Garg, N., & Zhao, J. (2020). Multi-Wheel Gear Loading Effect on Load-Induced Failure Potential of Airfield Flexible Pavement. *International Journal of Pavement Engineering, 21*(6), 805–816. Retrieved from https://doi.org/10.1080/10298436.2018

West, R., Rodezno, C., Leiva, F., & Yin, F. (2018). *Development of a Framework for Balanced Mix Design.* Washington, DC: Final Report to the National Cooperative Highway Research Program (NCHRP), Project NCHRP 20-07/Task 406.

White, G. (2016). Shear Stresses in an Asphalt Surface Under Various Aircraft Braking Conditions. *International J. of Pavement Res. and Tech, 9*, 89 – 101.

Yin, F., Chen, C., West, R., Martin, A. E., & 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)*, 281–291.

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