Refinement of P-401 Gradation Bands and Development of Predictive Models for Airfield Asphalt Mixture Performance and Volumetrics

by

Trung Tran

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 02, 2026

Keywords: Airfield Asphalt Mixtures, P-401 Gradation Limits, Aggregate Gradation, Volumetric Properties, Laboratory Performance Tests, Prediction Model

Approved by

Dr. Nam Tran, Chair, Research Professor, National Center for Asphalt Technology
Dr. Chen Chen, Co-Chair, Assistant Research Professor, National Center for Asphalt Technology
Dr. Fan Yin, Research Professor, National Center for Asphalt Technology
Dr. Benjamin Bowers, Associate Professor, Auburn University
Dr. Bertram Zinner, Associate Professor, Auburn University

ABSTRACT

Many airport runways, taxiways, and aprons are designed using asphalt mixtures following the P-401 specification from Advisory Circular 150/5370-10H. This specification mandates a preliminary mix design based on volumetric criteria, specifically 3.5% air voids and a minimum Voids in Mineral Aggregate (VMA). Final mix approval requires passing the moisture susceptibility using the Tensile Strength Ratio (TSR) test, and rutting resistance using either the Asphalt Pavement Analyzer (APA) or the Hamburg Wheel Tracking Test (HWTT).

Compared to Superpave (AASHTO M323), P-401 imposes more restrictive gradation controls, requiring a full gradation limit across all sieve sizes and a higher VMA threshold. While this enhances mix quality, it introduces substantial challenges for designers in meeting both the VMA and full gradation limits simultaneously.

The study focuses on two primary objectives: 1) evaluating the effects of adjusting gradation outside the P-401 gradation limits without changing volumetric properties on the laboratory performance and proposing a more flexible gradation limit that preserves overall mixture performance. 2) evaluating the effects of gradation change on the volumetric properties and laboratory performance of airfield asphalt mixtures and developing prediction models to quantify these relationships.

The experimental program began with the selection of 13 P-401 mix designs representing diverse aggregate sources across four Long Term Pavement Performance (LTPP) climate zones (dry/no-freeze, dry/freeze, wet/no-freeze, and wet/freeze). Preference was given to designs compacted at 75 blows or gyrations, as this compaction level requires more stringent design requirements. Each selected mix was redesigned using measured aggregate properties to meet P-401 specifications with a gradation closer to the gradation limits, referred to as In-Spec 1. Three

additional gradation designs were then developed for each mix: Out-Spec 1 outside the gradation limits with a similar VMA with In-Spec 1; In-Spec 2 centered within the limits, and Out-Spec 2 further outside the limits, both with VMA 1–2% difference from In-Spec 1. All four gradation designs shared the same binder type and content to isolate the effect of gradation. Laboratory performance was evaluated using the APA or HWTT for rutting resistance, TSR for moisture susceptibility, Cantabro test for durability, Florida permeability test, Disk-Shaped Compact Tension (DCT) test for low-temperature cracking resistance, and the Indirect Tensile Asphalt Cracking Test (IDEAL-CT) and Illinois Flexibility Index Test (I-FIT) for intermediate-temperature cracking resistance.

Comparing In-Spec 1 and Out-Spec 1 showed that adjusting the gradation outside the P-401 gradation limits while maintaining volumetric properties and binder content did not significantly affect the laboratory performance of P-401 mixtures. The lower limits of the 12.5 mm NMAS mix design were revised by a 2% reduction in the No. 16, No. 30, and No. 50 sieves.

Comparing all four gradation designs showed that moving the gradation further from the maximum density line (MDL) toward either the upper or lower limits consistently increased VMA and air voids. Moisture susceptibility, permeability, durability, and low-temperature cracking resistance remained unaffected. Coarser gradations improved intermediate-temperature cracking resistance. Moreover, coarser gradations maintained or significantly improved the rutting resistance. Prediction models were developed to predict changes in VMA, rut depth, CT_{Index}, and Flexibility Index (FI) as functions of changes in gradation and aggregate properties. The developed models were deployed as a web-based decision-support tool to assist mix designers in evaluating the volumetric and laboratory performance of gradation adjustments before laboratory testing.

ARTIFICIAL INTELLIGENCE (AI) USE DISCLOSURE STATEMENT

In the preparation of this dissertation, the following Artificial Intelligence (AI) tools were used: Google Gemini Pro, Claude, and Grammarly-AI Assistant. This tool was used primarily to assist in developing prediction models and grammar checks. The author acknowledges full responsibility for the intellectual content of this work and has ensured that all AI-assisted sections have been reviewed and revised for accuracy and appropriate academic style. All AI-generated content was reviewed and validated for relevance, appropriateness, and accuracy before incorporation into the final document to maintain scholarly integrity of this research.

ACKNOWLEDGMENTS

I extend my heartfelt gratitude to the individuals who have played a significant role in the completion of this dissertation. Dr. Tran, thank you for your unwavering support and guidance throughout my Ph.D. journey. You have always set a good example as a researcher and have provided invaluable direction for my future career. Dr. Chen, you have not only served as my co-advisor but also supported me as a caring friend and brother. You have always looked out for me and provided support during the stressful time.

I would like to thank Dr. Fan, Dr. Bowers, and Dr. Zinner for serving as committee members and for their significant contributions to the completion of my dissertation.

I want to express my appreciation to the research team from the University of Nevada and Heritage Research Group for their contribution to the project. Three P-401 mix designs (IND, SBN, and CMH) were developed and evaluated for laboratory performance at Heritage Group, except for the APA test, which was performed by NCAT. Four P-401 mix designs (RNO, TUS, KBVS, and PSC) were developed and tested for laboratory performance at the University of Nevada, except for the DCT test, which was performed by NCAT. I would like to extend my gratitude to Bill Pine from Heritage Group for his support with all the mix designs. I have learned a great deal from you, not only about mix design, but also the importance of attention to detail.

I also want to thank Josue, a graduate student at NCAT, for his assistance with the web-based prediction tool and Tianhao Yan and Bertram Zinner for reviewing the model development procedure and results.

I would like to express my sincere thanks to the entire NCAT family for their assistance, support, and encouragement. Without your collective help, this project would not have been completed.

Lastly, I would like to express my sincere gratitude to my family, who have consistently supported and believed in me. I am profoundly grateful for their love and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AASHTO	American Association of State Highway and Transportation Officials
AC	Asphalt Content
ALDOT	Alabama Department of Transportation
ANN	Artificial Neural Network
APA	Asphalt Pavement Analyzer
ARZ	Above the Restriction Zone
ASTM	American Society for Testing and Materials
AFQM	Augmented Full Quadratic Model
BDL	Bradley International Airport
BMD	Balance Mix Design
BRZ	Below the Restriction Zone
CAA	Coarse Aggregate Angularity
CA	Coarse Aggregate
CMH	John Glenn Columbus International Airport
CMHB	Coarse Matrix High Binder
CMOD	Crack Mouth Opening Displacement
CUW	Chosen Unit Weight
DCT	Disk-Shaped Compact Tension
DEM	Discrete-Element Modeling
DTR	Decision Tree Regression
COE	Corps of Engineers

ESALs	Equivalent Single Axle Loads
FA_c	Fine Aggregate Coarseness
FAA	Fine Aggregate Angularity
FA_f	Fine Aggregate Fineness
FE	Flat and Elongated
FI	Flexibility Index
GBR	Gradient Boosting Regression
GEP	Gene Expression Programming
Gsb	Bulk Specific Gravity
HMA	Hot Mix Asphalt
HWTT	Hamburg Wheel Tracking Test
IDEAL-CT	Indirect Tensile Asphalt Cracking Test
IDOT	Illinois Department of Transportation
IND	Mix Indiana International Airport
I-FIT	Illinois Flexibility Index Test
ITS	Indirect Tensile Strength
JMF	Job Mix Formula
LA	Los Angeles (Abrasion)
LTPP	Long-Term Pavement Performance
LR	Linear Regression
LUW	Loose Unit Weight
LVDT	Linear Variable Differential Transducer
MAE	Mean Absolute Error

MAPE	Mean Absolute Percentage Error
MBV	Methylene Blue Value
MDL	Maximum Density Line
MDQ	Huntsville Executive Airport
MEPDG	Mechanistic-Empirical Pavement Design Guide
MLR	Multiple Linear Regression
MSE	Mean Squared Error
ML	Machine Learning
NAPA	National Asphalt Pavement Association
NCAT	National Center for Asphalt Technology
NMAS	Nominal Maximum Aggregate Size
ODOT	Oklahoma Department of Transportation
OLS	Ordinary Least Squares
OGFC	Open-Graded Friction Courses
KBVS	Skagit Regional Airport
QC	Quality Control
PCS	Primary Control Sieve
PSC	Tri-Cities Airport
PGD	Punta Gorda Airport
PVC	Polyvinyl Chloride
RAP	Reclaimed Asphalt Pavement
RMSE	Root Mean Squared Error
RNO	Reno-Tahoe International Airport

RUW	Rodded Unit Weight
RVS	Tulsa Riverside Airport
RF	Random Forest
SAT	San Antonio International Airport
SAV	Mix Savannah/Hilton Head International Airport
SBN	South Bend International Airport
SCB	Semi-Circular Bending
SCS	Second Control Sieve
SBS	Styrene-Butadiene-Styrene
SE	Sand Equivalent
SER	Standard Error of Regression
SHAP	Shapley Additive Explanations
SHRP	Strategic Highway Research Program
SMA	Stone Matrix Asphalt
SGC	Superpave Gyratory Compactor
SVR	Support Vector Regression

CHAPTER 1 INTRODUCTION

1.1 Problem Statement

Many airport runways, taxiways, and aprons are designed with asphalt mixtures following the P-401 specification in Advisory Circular 150/5370-10H, “Standard Specifications for Construction of Airports.” [1]. According to the P-401 specification, the mix design is developed based on volumetric criteria, requiring 3.5% air voids and a minimum Voids in Mineral Aggregate (VMA) that varies with Nominal Maximum Aggregate Size (NMAS). Laboratory testing specimens can be compacted using either the Marshall or the Superpave gyratory compaction method. The mix design is then evaluated for moisture susceptibility using the Tensile Strength Ratio (TSR). In addition, airfield asphalt mixtures serving aircraft with a gross weight over 60,000 pounds must be tested with the Asphalt Pavement Analyzer (APA) or Hamburg Wheel Tracking Test (HWTT) to ensure adequate rutting resistance. In terms of aggregate, the P-401 specification uses crushed aggregate (stone, gravel, and/or slag), fine aggregate (screenings and/or natural sand), and mineral filler, as needed. These materials must meet the requirements related to particle angularity, shape, toughness, soundness, and cleanliness. Note that while natural (uncrushed) sand can be used to achieve the required gradation or improve the workability of a mixture, it is limited to a maximum of 15% of the total aggregate to prevent rutting under heavy aircraft loading.

The P-401 specification and Superpave AASHTO M323 [2] differ in several ways, particularly in their gradation requirements. P-401 mix designs use a fine gradation for low permeability and require a VMA of 1% above the Superpave specification to accommodate more binder for improved durability. The P-401 gradation also requires a full gradation limit with both upper and lower limits at all sieve sizes, while the Superpave gradation only controls four sieve sizes. These differences in the gradation limits are shown in **Table 1**.

Table 1. Gradation and VMA requirements in the P-401 specification and the AASHTO M323 Superpave specification [1, 2]

Sieve Size	Passing Percentage (%)					
	19.0 mm NMAS		12.5 mm NMAS		9.5 mm NMAS	
	P-401	Superpave	P-401	Superpave	P-401	Superpave
1 inch (25.0 mm)	100	100	--		--	
3/4 inch (19.0 mm)	90–100	90–100	100	100	--	
1/2 inch (12.5 mm)	68–88	Max 90	90–100	90–100	100	100
3/8 inch (9.5 mm)	60–82		72–88	Max 90	90–100	90–100
No.4 (4.75 mm)	45–67		53–73		58–78	Max 90
No.8 (2.36 mm)	32–54	23-49	38–60	28-58	40–60	32-67
No.16 (1.18 mm)	22–44		26–48		28–48	
No.30 (600 µm)	15–35		18–38		18–38	
No.50 (300 µm)	9–25		11–27		11–27	
No.100 (150 µm)	6-18		6-18		6-18	
No.200 (75 µm)	3–6	2-8	3–6	2-10	3–6	2-10
VMA Requirement						
Min VMA (%)	14.0	13.0	15.0	14.0	16.0	15.0

Although the full gradation limits have contributed to producing qualified airfield asphalt mixtures, it also creates significant challenges for mix designers. The following are challenges encountered in designing P-401 asphalt mixtures:

- Adjusting gradation to meet VMA requirements often results in sieve violations, and fixing one sieve issue can cause others or lead to VMA or/and gradation failures. To comply with full gradation limits, mix designers have tried solutions like reprocessing fine aggregates or sourcing materials from distant locations, which can raise costs and delay projects.
- The P-401 gradation has stricter No.200 sieve limits than Superpave, raising challenges for mix design and plant production. To address it, contractors may use washed screenings, natural sand, or waste dust, but the resulting blend gradation often still approaches the gradation limits

in the fine aggregate fraction. These challenges highlight the need to relax current gradation controls to accommodate local aggregates without reducing performance.

- Full gradation limits are applied during laboratory design, whereas for Quality Control (QC), the requirements are less strict and focus on specific sieves. Furthermore, the acceptance of the P-401 mix design is based on lab air voids, field density, pavement grade, and roughness, rather than the strict requirement of a full gradation limit. These distinctions raise the question of whether a full gradation limit is necessary during the mix design phase.

Overall, these challenges have led mix designers and practitioners to question whether gradation requirements could be modified to be less restrictive without compromising mixture performance. Clear guidance is needed on assessing how gradation changes influence volumetric properties and laboratory performance test results during the mix design process.

1.2 Research Objectives

Based on the problems described above, two objectives were established for the study:

- Evaluate the effect of gradation outside the P-401 gradation limits while maintaining the volumetric properties and binder content. Based on the laboratory test results, the study proposes a more flexible P-401 gradation limit without compromising the performance.
- Evaluate the effect of the gradation change on the volumetric properties and laboratory performance of P-401 asphalt mixtures. Based on the test results, prediction models were developed to estimate the influence of gradation adjustments on the volumetric properties and laboratory performance of airfield asphalt mixtures.

1.3 Organization of Dissertation

The dissertation includes seven chapters. Chapter 1 describes the problems with the current gradation requirements in the P-401 specification and defines two objectives. Chapter 2 presents

a literature review on the evolution of aggregate gradation requirements in P-401. It summarizes the previous findings on the impact of aggregate properties and gradation on the volumetric properties and performance of asphalt mixtures. The chapter also reviews existing prediction models developed to estimate these properties and outlines the Bailey method for gradation adjustment. Chapter 3 explains the research methodology and summarizes the laboratory performance tests used in the study. Chapter 4 focuses on revising the P-401 gradation limits requirements. Chapter 5 focuses on the effect of gradation change on volumetric properties and laboratory performance of airfield mixtures. Chapter 6 presents prediction models to quantify the relationship between gradation changes and volumetric properties and laboratory performance. Chapter 7 presents the conclusions of the study and recommendations for future study.

CHAPTER 2 LITERATURE REVIEW

This chapter conducts a literature review focusing on the evolution of the aggregate gradation limits in the P-401 specification. It summarizes previous findings on the effects of aggregate properties and gradation on the volumetric and performance properties of asphalt mixtures and presents developed prediction models to quantify these effects. Finally, it describes the Bailey method, which was used to adjust gradation in the study.

2.1 Evolution of Aggregate Gradation Limits for Airfield Asphalt Mix Design

This section reviews the historical and technical evolution of aggregate gradation, especially focusing on gradation limits in airfield asphalt mixture design, tracing their development from empirical practice and packing theory to Marshall stability-based approaches that established the foundation of modern airfield specifications. It examines the applicability of Superpave mix design to airfield pavements, with particular attention to compaction equivalency. The review further analyzes progressive refinement of gradation limits through successive editions of Item P-401 in FAA AC 150/5370-10. Key developments discussed include the evolution of gradation limits, enhanced control of aggregate properties, and the transition from Marshall stability-based acceptance criteria to volumetric requirements and performance-based validation methods.

2.1.1 Early Foundations: Gradation Laws and Aggregate Packing Theory

The development of the gradation ideal in asphalt mixtures emerged from early empirical attempts to link aggregate size distribution to pavement stability and durability. One of the earliest documented concepts came in the 1903 patent of Frederick J. Warren [3]. In the patent, Warren claimed that aggregates of varying sizes, when “homogeneously mingled,” would form a dense, interlocking mass with less than 20% voids, a concept later formalized as VMA. Warren specified an early gradation limit with 50–80% retained between 0.25 and 3.0 inches, 10–49% passing 0.25

inches, and 1–3% described as “impalpable powder,” corresponding approximately to material passing the modern No. 200 sieve. Translated into modern terms, this corresponds to 100% passing 3 inches, 10–49% passing 0.25 inches, and 1–3% passing No.200 sieve. The key conceptual advance was the recognition that stability was derived from aggregate interlock rather than binder stiffness. By reducing void space through controlled size distribution, Warren argued that softer asphalt could be used because the aggregate skeleton, not the binder, carried the load.

Another earlier study, conducted by Clifford Richardson, helped to inform a methodology for selecting gradation and asphalt content (AC) [4]. Richardson distinguished between two principal mixture types: surfacing mixtures and asphaltic concrete. Surfacing mixtures were essentially sand mixes, typically characterized by 100% passing the No. 10 sieve and approximately 15% passing the No. 200 sieve, with relatively high asphalt contents ranging from 9 to 14%. In contrast, asphaltic concrete mixtures incorporated coarser aggregate fractions and introduced VMA with emphasizing that VMA must be properly adjusted to accommodate the appropriate asphalt content to ensure durability. **Table 2** describes the asphaltic concrete mixture properties specified by Clifford Richardson.

Table 2.Asphaltic Concrete Mixture Properties Specified by Clifford Richardson [4]

Parameter	Value
1.5 inch, % passing	100.0
1.0 inch, % passing	83.6
0.50 inch, % passing	50.1
0.25 inch, % passing	40.3
No.8, % passing	36.8
No.200, % passing	5.2
VMA, %	13.2
Asphalt, % by weight aggregate	7.4
Asphalt, % by weight of mixture	6.9

Another advancement in asphalt mixture design occurred in 1924 when Charles Hubbard and Fredrick Field [5] developed the Hubbard Field Method, which included volumetric concepts such as air voids and voids in the aggregate skeleton, and used the Hubbard Field Stability test to evaluate the effect of aggregate gradation. Their work reflected the belief that mixture performance was governed by an aggregate packing law, and substantial effort was devoted to identifying an "ideal" gradation that would provide a skeleton with adequate stability for traffic and have sufficient space for a minimum amount of asphalt binder and air voids.

In the 1940s, Nijboer [6], was the first to discover that the densest packing of aggregate occurred for a gradation that plotted as a straight line with a slope of 0.45 on a log percent passing versus log sieve size plot, which later became a fundamental tool for evaluating aggregate gradation in asphalt mixture design.

Overall, these early studies collectively established the foundation for modern gradation concepts, highlighting the critical role of aggregate skeletons in controlling mixture stability and durability.

2.1.2 Marshall Method and the Origin of FAA P-401 Gradation Limits

The origin of airfield asphalt mixture design emerged during World War II, when the United States Army Corps of Engineers (COE) was tasked with designing pavements capable of supporting unprecedented aircraft wheel loads. Existing highway design methods, which had never contemplated loads of this magnitude, were insufficient to ensure the stability of airfield pavements [7]. The Corps evaluated both Hubbard–Field and Marshall mix design methods and ultimately adopted the Marshall method due to its practicality and reproducibility under field conditions. Controlled laboratory and field experiments were conducted at the COE Waterways Experiment Station in Vicksburg, Mississippi. Test sections with strong aggregate bases were built to isolate

asphalt surface behavior. Asphalt mixtures (¾-inch maximum aggregate size, generally following Fuller's curve) were produced with varying gradations and asphalt contents and compacted using Marshall procedures. Pavements were trafficked with 15,000, 37,000, and 60,000 lb wheel loads representative of military aircraft. Results showed that aggregate gradation, especially the coarse fraction (retained on the No. 10 sieve), directly influenced Marshall Stability. Stability changed little below about 35–40% coarse aggregate, increased between roughly 40% and 65%, and declined above 65%, indicating an overly coarse structure. Based on these findings, the COE established gradation limits with an upper limit of at least 40% coarse aggregate and a lower limit of no more than 65% coarse aggregate, as shown in **Table 3** and **Figure 1**. Rather than prescribing a single target gradation, the research established gradation envelopes within which acceptable stability could be achieved.

Table 3. 1948 Gradation Limits for Airfield Asphalt Mixtures [7]

Sieve Size	Gradation Limits Per Cent Passing: Asphaltic Concrete	Stone Filled Sand Asphalt	Sand Asphalt
3/4-in.	100	100	
1/2-in.	76–100	92–100	
No. 4	50– 80	72–100	100
No. 10	35– 60	55– 82	75–100
No. 20	22– 49	40– 66	50– 82
No. 40	12– 38	28– 52	35– 65
No. 80	7– 26	16– 35	18– 44
No. 200	3– 12	5– 16	8– 20

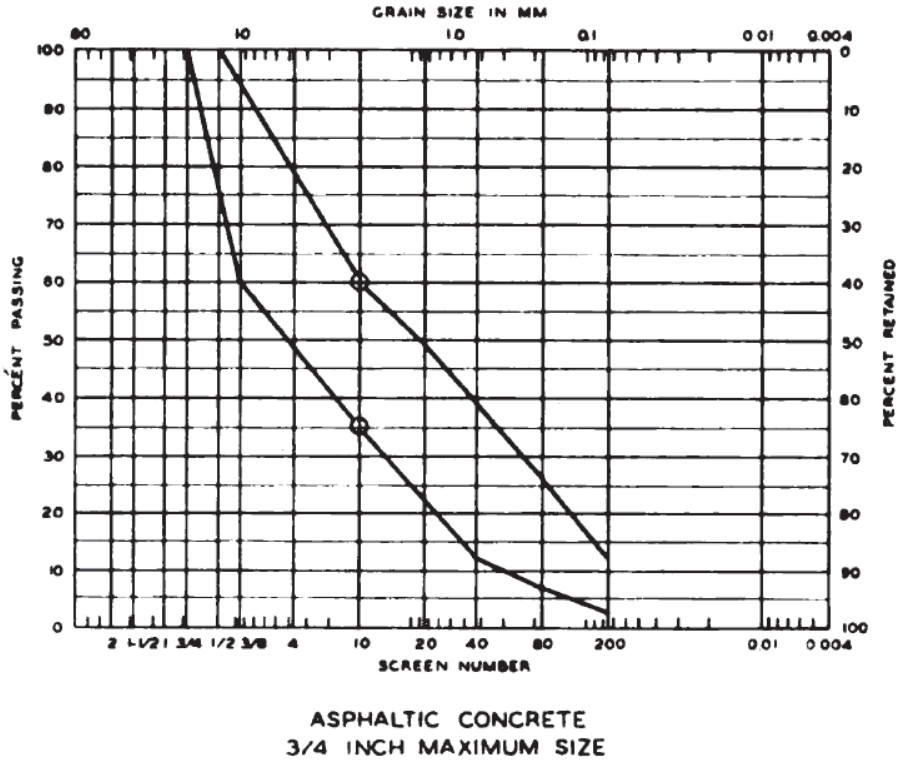

Figure 1. Gradation Limits for ¾-Inch Maximum Size Asphaltic Concrete Mixture [7]

In terms of compaction efforts, initially, the standard compaction procedure was 50 blows per face with a 10 lb Marshall hammer [7]. As aircraft loadings and tire pressures increased, the COE developed a modification. For pavements expected to receive tire pressures from 100 psi to 250 psi, the compactive effort was increased from 50 to 75 blows with a Marshall hammer [8]. These early gradation developments by COE served as the technical foundation for what would later be codified as Item P-401.

2.1.3 Applicability of Superpave Gradation Specifications for Airfields

Following the implementation of the Superpave design method for highway pavements, the Federal Aviation Administration conducted a formal evaluation to determine whether Superpave mix design, including gradation control and compaction procedures, was applicable to airfield asphalt pavements. **Figure 2** shows the FAA gradation limits, the Superpave control points, the restricted zone, and two aggregate blends. Although many FAA gradations could be plotted within

Superpave control points and restricted zones, certain Superpave-compliant gradations fell outside established P-401 gradation limits. The authors believed that FAA mixtures had an acceptable performance record and wanted to ensure that switching to a gyratory design would provide the "same" mixture.

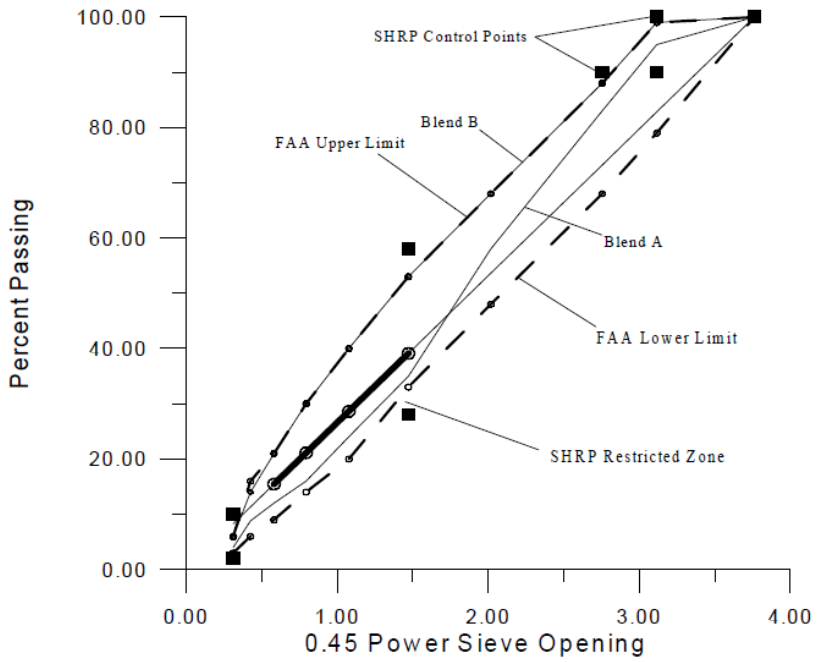

Figure 2. An Aggregate Blend Plotted within P-401 Gradation Limits and Superpave Control Points [9].

Cooley et al. [10] investigated the implementation of Superpave mixtures for airfields. The results indicate that no single gyration level in the Superpave gyratory compactor (SGC) exactly matches the compaction effort of the Marshall hammer; instead, equivalent ranges vary with aggregate type and gradation. Statistical analysis of materials from ten visited airfields determined that 32 to 40 gyrations in the SGC provided a compaction effort equivalent to 50 blows per face of the Marshall hammer. For higher traffic applications, 43 to 55 gyrations were found to be equivalent to 75 blows per face of the Marshall hammer. These findings highlight that historical

Superpave airfield designs often utilized N_{design} values that were too high, leading to lower ultimate in-place densities and a lack of sufficient asphalt binder for long-term durability.

Cooley et al. [11] recommended utilizing a single gradation limit for each maximum aggregate size that effectively consolidates the historical Item P-401 requirements. Because airfield pavements are highly susceptible to durability issues like age hardening, cracking, and raveling caused by air and water infiltration, the study advised against the coarser gradations typically allowed in highway Superpave designs, which tend to be more permeable. Specifically, the researchers recommend increasing the lower control point on the No. 8 sieve by 5% to minimize permeability potential and ensure a finer, more durable mix. These recommended gradation limits, ranging from 1-1/2 inch (37.5 mm) to dust (0.075 mm), were designed to produce a stable, rut-resistant, and long-lasting surface suitable for various aircraft tire pressures, as shown in **Table 4**.

Table 4. Adaptation of Superpave Gradation Requirements for Airfield Asphalt Mixtures [11]

Sieve Size	Percentage by Weight Passing Sieves			
U.S. (mm)	1½"max	1" max	3/4" max	½" max
1-1/2 (37.5)	100	---	---	---
1 (25.0)	86-98	100	---	---
¾ (19.0)	68-93	76-97	100	---
½ (12.5)	57-81	67-87	77-98	100
3/8 (9.5)	49-69	58-80	68-89	77-98
No. 4 (4.75)	34-54	42-62	50-70	58-78
No. 8 (2.36)	22-42	29-48	35-55	40-60
No. 16 (1.18)	13-33	19-40	23-34	27-47
No. 30 (0.600)	8-24	12-30	16-34	18-36
No. 50 (0.300)	6-18	8-22	12-28	11-25
No. 100 (0.150)	4-12	6-17	7-20	6-18
No. 200 (0.075)	3-6	3-6	3-6	3-6

2.1.4 The refinement of Item P-401 in FAA AC 150/5370-10 Series

The inaugural version of AC 150/5370-10 [12], released in 1974, and Change 1 [13], released in 1977, relied primarily on the Marshall Mix Design method, which focused on empirical measures of stability and flow. **Table 5** and **Table *6*** present the gradation limits for pavements subjected to

aircraft gross weights greater than 60,000 pounds and those with gross weights less than 60,000 pounds, respectively. For each loading level, three gradation limits corresponding to NMAS sizes of 1 inch, 3/4 inch, and 1/2 inch were developed for airfield pavements. The gradation limits specified for aircraft with gross weights of 60,000 pounds or greater were generally tighter than those specified for aircraft with gross weights below 60,000 pounds. The stricter gradation limits reflected the need to improve rutting resistance under heavier aircraft loading conditions.

Table 5. Percentage by Weight Passing Sieves in AC 150/5370-10 (Gross Weight of 60,000 pounds or Greater) [12]

Sieve	Percentage by Weight Passing Sieves		
Size	1" maximum	3/4" maximum	1/2" maximum
1 in. (25.0 mm)	100	----	----
3/4 in. (19.0 mm)	84-100	100	----
1/2 in. (12.5 mm)	74-88	82-96	100
3/8 in. (9.5 mm)	68-82	75-89	79-93
No. 4 (4.75 mm)	53-67	59-73	59-73
No. 8 (2.36 mm)	40-54	46-60	46-60
No. 16 (1.18 mm)	30-44	34-48	34-48
No. 30 (600 μm)	20-34	24-38	24-38
No. 50 (300 μm)	13-25	15-27	15-27
No. 100 (150 μm)	9-17	8-13	8-18
No. 200 (75 μm)	3-6	3-6	3-6
Bitumen percent: Stone or gravel	4.5-7.0	5.0-7.5	5.5-8.0
Slag	6.0-9.0	6.5-9.5	7.0-10.0

Table 6. Percentage by Weight Passing Sieves in AC 150/5370-10 (Gross Weight Less than 60,000 pounds) [12]

Sieve	Percentage by Weight Passing Sieves		
Size	1" maximum	3/4" maximum	1/2" maximum
1 in. (25.0 mm)	100	------	-------
3/4 in. (19.0 mm)	82-100	100	-------
1/2 in. (12.5 mm)	70-90	82-100	100
3/8 in. (9.5 mm)	60-82	68-90	82-100
No. 4 (4.75 mm)	42-70	50-79	56-88
No. 8 (2.36 mm)	32-62	39-69	43-77
No. 16 (1.18 mm)	24-54	29-59	32-65
No. 30 (600 μm)	18-45	20-49	23-55
No. 50 (300 μm)	13-35	14-38	15-42
No. 100 (150 μm)	7-22	8-25	8-27
No. 200 (75 μm)	3-8	3-8	4-9
Bitumen percent: Stone or gravel	4.5-7.0	5.0-7.5	5.5-8.0
Slag	6.0-9.0	6.5-9.5	7.0-10.0

The release of AC 150/5370-10A in 1989 [14], particularly following Change 12 in 1999 [15], represented a significant technical development of the P-401 specification. Unlike the earlier edition, which used three gradation limits, this version formalized a system with four distinct gradation limits, classified by NMAS: 1-1/4 inch, 1 inch, 3/4 inch, and 1/2 inch, as shown in **Table 7**. Moreover, the 1989 version removed the separation of gradation limits based on aircraft gross weight and used a unified gradation framework classified by NMAS. The specification also systematically shifted the entire gradation limits downward at mid-range sieves, producing mixes with stronger coarse aggregate skeletons that responded to rutting failures observed under heavy wide-body aircraft loads. It also introduced minimum requirements for the percentage of coarse aggregate particles with fractured faces, recognizing the importance of angular particles. For

pavements designed to support aircraft with gross weights greater than 60,000 pounds, the coarse aggregate was typically required to have at least 70% of particles with one fractured face and 85% with two or more fractured faces. For pavements subjected to aircraft with gross weights less than 60,000 pounds, the requirements were slightly less stringent, typically 50% with one fractured face and 65% with two or more fractured faces.

Table 7. Percentage by weight passing sieves in AC 150/5370-10A includes changes thru CHG 12 [15]

AGGREGATE - BITUMINOUS PAVEMENTS				
Sieve Size	Percentage by Weight Passing Sieves			
	1-1/4"max	1"max	3/4"max	1/2"max
1-1/4 in. (30.0 mm)	100	--	--	--
1 in. (24.0 mm)	86-98	100	--	--
3/4 in. (19.0 mm)	68-93	76-98	100	--
1/2 in. (12.5 mm)	57-81	66-86	79-99	100
3/8 in. (9.5 mm)	49-69	57-77	68-88	79-99
No. 4 (4.75 mm)	34-54	40-60	48-68	58-78
No. 8 (2.36 mm)	22-42	26-46	33-53	39-59
No. 16 (1.18 mm)	13-33	17-37	20-40	26-46
No. 30 (0.600 mm)	8-24	11-27	14-30	19-35
No. 50 (0.300 mm)	6-18	7-19	9-21	12-24
No. 100 (0.150 mm)	4-12	6-16	6-16	7-17
No. 200 (0.075 mm)	3-6	3-6	3-6	3-6
Asphalt percent: Stone or gravel Slag	 4.5-7.0 5.0-7.5	 4.5-7.0 5.0-7.5	 5.0-7.5 6.5-9.5	 5.5-8.0 7.0-10.5

In the subsequent versions, AC 150/5370 10B (2005) [16], AC 150/5370-10C (2007) [17], AC 150/5370-10D (2008) [16], AC 150/5370-10E (2009) [17], and AC 150/5370-10F (2011) [18], the gradation limits remained largely unchanged from the previous version. However, these specifications introduced an additional Item P-403, "Plant Mix Bituminous Pavements (Base, Leveling, or Surface Course). The introduction recognized the alignment of design requirements with aircraft loading severity. P-401 mixtures, which required strict controls, were increasingly reserved for high-performance surface courses. While Item P-403 was intended primarily for the base and leveling course, particularly for pavement designed to support aircraft weighing greater than 12,500 pounds. Although P-403 mixtures could also be used as a surface course, this

application was generally limited to pavements designed for aircraft with gross weights less than or equal to 12,500 pounds. In these specifications, attention shifted to controlling natural sand. An excessive amount of natural sand within the fine aggregate fraction was a primary cause of rutting in airfield asphalt, leading to limiting natural sand to no more than 15% by weight of the total aggregate blend.

The publication of AC 150/5370-10G in 2014 [19] was a transformative moment for Item P-401, as it moved back to three gradation limits with NMAS sizes of 1 inch, ¾ inch, and ½ inch. This revision was consistent with the FAA recommendation that the aggregate size should not exceed 1/4 of the total lift thickness, thereby maintaining the relationship between aggregate size and lift thickness. In addition, the mix design requirements were refined by specifying a target air void content of 3.5%, replacing the broader range of 2.8% to 4.2% used in the previous version. Moreover, the minimum requirement for the percentage of coarse aggregate with one fractured face for aircraft with a gross weight greater than 60,000 pounds increased from 70% to 75%.

The current standard, AC 150/5370-10H, released in 2018 [1], represents the most comprehensive technical update in the history of the specification. AC 150/5370-10H removed the requirement for Marshall stability and flow. Instead, the specification introduced the APA and the HWTT for rutting testing and the TSR for moisture susceptibility. Under AC 150/5370-10H, the update primarily modified Gradations 1 and 2, corresponding to NMAS of ¾ inch and ½ inch, respectively. For these gradations, the lower and upper limits of the No. 50 sieve and up to the NMAS increased by approximately 5 percent on average. This adjustment moved the lower boundary of the gradation limits closer to the maximum density line (MDL), effectively encouraging the use of finer-graded asphalt mixtures. Although this modification may result in a slight increase in material costs, it can facilitate construction by improving mixture workability.

Finer mixtures typically contain higher binder content and compact more easily, which may help contractors meet the strict density requirements specified for P-401 mixtures [20]

2.2 Impact of Aggregate Properties on Volumetric Properties

Many studies have evaluated the effects of aggregate properties, especially the Superpave consensus properties, on the volumetric properties of asphalt mixtures. Superpave specification AASHTO M323 specifies the minimum VMA based on the NMAS. Beyond this baseline requirement, at each NMAS level, parameters such as the percentage of crushed coarse aggregate and Fine Aggregate Angularity (FAA) can also affect VMA [21, 22].

Among consensus properties, aggregate angularity has demonstrated a strong correlation with VMA. Angular aggregates generally produce higher VMA than their rounded counterparts, as their rough texture increases inter-particle friction and prevents the aggregates from packing tightly together [23]. Park et al. [24] investigated the relationship between FAA and VMA using four types of aggregates: crushed river gravel, crushed granite, crushed limestone, and uncrushed rounded natural sand. The findings indicated that mixtures with crushed aggregates exhibited significantly higher VMA compared to those with uncrushed, rounded natural sand. Notably, incorporating 30% rounded sand, expressed as a proportion of the fine aggregate portion, led to a sharp decrease in VMA. These findings demonstrated that higher FAA values were typically associated with increased VMA, a relationship further supported by a strong statistical correlation between the two variables ($R^2 = 0.70$). Purcell and Cross [25] developed coarse and fine gradations using 100% crushed limestone. To evaluate the effect of FAA and isolate the effect of gradation on VMA, the gradation was held constant by progressively replacing portions of the crushed limestone with increasing percentages of natural sand. Test results showed that the natural sand mixture with a lower FAA provided lower VMA than the crushed limestone mixtures.

Beyond angularity, aggregate shape plays a significant role in VMA. Cubical aggregates tend to resist compaction more than flat or elongated particles; therefore, mixtures containing a greater proportion of cubical particles typically lead to higher VMA [26].

2.3 Impact of Aggregate Properties on Mixture Performance

2.3.1 Impact of Consensus Properties of Aggregates on Mixture Performance

2.3.1.1 Impact of CAA and FAA on Mixture Performance

Aggregate angularity, including Coarse Aggregate Angularity (CAA) and FAA, is widely recognized as a major factor affecting the rutting performance of asphalt mixtures [27, 28]. By ensuring a high degree of internal friction within the aggregate skeleton, angularity increased mixture stability, and minimum CAA and FAA values are therefore specified in Superpave mix design requirements[29, 30].

The influence of CAA on permanent deformation has been consistently demonstrated across multiple studies. Leon and Charles [31] indicated that as the CAA value increased, the rutting resistance of HMA also improved. Similarly, Souza et al. [32], through a combination of laboratory experiments and finite-element modeling, found that higher CAA enhanced the interlock between aggregate particles, thereby improving the deformation resistance under load. This finding was further supported by Pan et al. [33], who found that CAA provided better interlocking and mechanical properties, which minimized lateral flow and shear deformation under traffic loading. In addition, Park et al. [24] proved that the effect of CAA on the rutting performance of HMA mixtures was more significant when a soft binder was utilized. In the study, the mixtures were prepared with five levels of CAA (100%, 85%, 70%, 50%, and 35%) and three asphalt binders (PG 64-22, PG 76-22, PG 82-22), and their rutting performance was investigated using creep and APA tests. A consistent trend was observed for all binder grades, whereby higher CAA levels

improved rutting resistance. However, this effect was most pronounced with softer binders, while the differences in permanent strain and rut depth between CAA levels became negligible with the stiffest binder (PG 82-22). Therefore, the research concluded that an extremely stiff binder overshadowed the effect of aggregate angularity on rutting resistance.

The beneficial effect of CAA on rutting resistance was also confirmed for Warm Mix Asphalt (WMA). Park et al. [24] investigated the influence of limestone aggregates with five CAA levels on the rutting resistance of WMA. The results demonstrated a strong linear relationship between increased CAA and improved rutting resistance.

In contrast to its relatively straightforward effect on rutting, the influence of CAA on cracking resistance is more complex. Regarding fatigue cracking, although angular particles can promote crack formation by creating stress concentration, they also require a higher binder content, which helps reduce cracking by enhancing viscoelastic energy dissipation [32]. Mahmoud et al. [34] assessed the effect of aggregate angularity on crack initiation, crack propagation, and crack path of 600 asphalt Superpave and Coarse Matrix High Binder (CMHB) using the Semi-Circular Bending (SCB). Findings revealed the contrasting trend between CMHB and Superpave mixes. CMHB mixes containing higher angular aggregates exhibited higher peak loads prior to failure than those with lower angularity. In contrast, Superpave mixes containing lower angularity resulted in higher average peak loads. Using the same mix types but assessed through the IDEAL-CT test, Saadeh and Asmar [35] reported that for both Superpave and CMHB mixtures, lower angularity aggregate improved the cracking resistance.

Similar to the CAA, the FAA also significantly influences the rutting and cracking resistance of asphalt mixtures. Park and Lee [36] and Topal and Sengoz [37] both found that mixtures incorporating finer aggregates with higher FAA values exhibited superior rutting

resistance. Ramli et al. [38] compared the rutting resistance and stability of mixtures containing crushed granite (FAA= 46%) and natural sand (FAA=37%) using the Wheel Tracking Test and the Marshall Test. The mixture with crushed granite demonstrated better stability and lower rut depths than those with natural sand, confirming that higher FAA significantly improved mixture stability and rutting resistance. Iuele and Vaiana [39] further supported this trend by using a wheel-tracking machine to assess slab rutting. Johnson et al. [40] evaluated the effectiveness of the Superpave FAA criterion in predicting the rutting performance of asphalt concrete using APA and dynamic modulus tests. Results showed a positive correlation between the FAA of the aggregate blend and rutting resistance.

Despite this general trend, some research proved that the FAA may not be a sufficient indicator of the shear strength and rutting resistance of fine aggregates in asphalt mixtures [41, 42]. A study by Bahia and Stakston [43] found that higher FAA values did not necessarily guarantee improved shearing resistance. In addition, rutting performance is influenced by multiple factors beyond FAA, including aggregate gradation, asphalt binder content, and the compatibility between aggregates and binder [44]. For example, Purcell and Cross [25] reported that mixture gradation could outweigh the influence of FAA, with fine-graded mixtures outperforming coarse-graded ones regardless of FAA level. Roque et al. [45] suggested a combined assessment using FAA and measures of toughness or direct shear strength to provide a more reliable indication of the rutting resistance of asphalt mixtures.

Regarding the effect of the FAA on cracking, Park and Lee [36] evaluated the effect of FAA on the fatigue performance of the asphalt mixtures by using two types of sand. The results showed that quarry sand, with higher FAA, rougher, and stronger particles, provided better fatigue resistance than mining sand with lower FAA.

2.3.1.2 Impact of FE on Mixture Performance

Flat and elongated (FE) particles are defined as coarse aggregates whose maximum-to-minimum dimension ratio exceeds a specified threshold, most commonly 5:1, measured either by weight or particle count. FE particles are typically unfavorable because they decrease aggregate packing efficiency, load transfer, and make the material more prone to particle breakage during compaction and traffic loading. The Superpave specifications developed under the Strategic Highway Research Program (SHRP) restrict FE particles with a ratio greater than 5:1 to no more than 10% by weight in aggregate blends used for traffic levels over 3 million equivalent single axle loads (ESALs). Supporting this guideline, Pan et al. [33] found that when the proportion of FE particles with a 5:1 ratio was well below the 10% threshold, their presence in the coarse aggregates had no significant effect on the permanent deformation behavior of HMA mixtures. Kandhal and Parker [46] suggested a more conservative 2:1 FE particle ratio to enhance resistance to permanent deformation and fatigue cracking in HMA pavements. This recommended ratio is stricter than the current specifications of 5:1 in the United States and 3:1 in China.

Beyond specification thresholds, studies have shown that increasing FE content leads to greater aggregate breakdown during construction and traffic, which affects the rutting susceptibility of compacted HMA mixes [47, 48]. Wang [49] showed that FE particles adversely impacted the development of compaction forces and load transfer efficiency within asphalt pavements. Similarly, Chen et al. [50] found that FE particles provided lower rutting resistance compared to cubical particles in asphalt mixtures.

2.3.1.3 Impact of SE Value on Mixture Performance

Clay content, often expressed as the sand equivalent (SE) value, represents the percentage of clay materials present in the portion of aggregate that passes through a No.4 sieve. The presence of

harmful clays is a significant concern in asphalt mix design, as these materials tend to swell upon exposure to moisture, which can disrupt the bond between aggregates and asphalt binder [51-53]. This expansion weakens the internal structure of the mixture, potentially leading to premature failures such as rutting or stripping in asphalt pavements [51]. Afsha et al.[54] conducted a study to examine the effects of clay contamination on the rutting performance of asphalt mixtures using the HWTT. A total of 21 asphalt mixture samples were prepared with varying levels of clay contamination, incorporating both inactive clays (such as calcium carbonate and dolomite) and active clays (such as bentonite and natural clays), as identified through the methylene blue value (MBV) test. The findings indicated that clay contamination with an MBV exceeding 6 mg/g significantly deteriorated rutting and moisture damage resistance in asphalt mixtures.

2.3.2 Impact of Source Properties of Aggregate on Mixture Performance

2.3.2.1 Impact of Toughness on Mixture Performance

Toughness is the ability of coarse aggregates to withstand abrasion and mechanical breakdown during handling, construction, and service life. It is typically quantified using the Los Angeles (LA) abrasion test, in which a standard coarse aggregate sample, generally using aggregate retained on the No. 8 sieve, is subjected to tumbling, repeated impact, and grinding with steel balls. The result is reported as the percentage of mass lost due to mechanical degradation, with acceptable values usually ranging between 30% and 45%. Higher LA values indicate more friable aggregates that are prone to breakdown during production and compaction, while lower LA values suggest more durable aggregates that enhance pavement skid resistance and resist wear from tire chains [29]. Dahir [55] showed that aggregate hardness, toughness, and mineralogy were key factors in determining the wear resistance and skid resistance of bituminous surfaces. These properties are

particularly important for OGFC and SMA mixtures, where the absence of a substantial fine aggregate matrix means that coarse particles are subjected to elevated contact stresses [56].

Amirkhanian et al. [57] investigated the effect of aggregate source with different LA values (55,48,30, and 28) on the strength of Marshall specimens. The laboratory results showed that specimens prepared with LA abrasion values of 28 exhibited significantly higher dry and wet Indirect Tensile Strength (ITS) values than those with LA values of 55, 48, and 30. However, a study by Masad et al. [58] revealed that aggregate texture, rather than toughness or hardness, has the strongest correlation with asphalt mixture rutting resistance. Supporting the statement, Hassan et al. [59] found a weak linear correlation between LA values and asphalt mixture performance, while shape parameters showed stronger linear relationships. These findings suggest that while toughness remains an important quality indicator, it should be incorporated with other aggregate properties to ensure the performance of asphalt mixtures.

2.3.2.2 Impact of Soundness on Mixture Performance

Beyond toughness and abrasion resistance, aggregates must withstand degradation from environmental factors, particularly the cyclical effects of moisture exposure and temperature fluctuations associated with wetting-drying and freeze-thaw cycles. When the asphalt binder coating remains intact, these environmental stresses typically have minimal impact on the asphalt mixture. However, weak or soft particles that degrade under compaction make it easier for moisture to enter, leading to raveling, stripping, and, in severe cases, rutting. To preserve the long-term performance of asphalt pavements, it is crucial to select durable and sound aggregate [56].

Based on the literature review, **Table 8** summarize the effect of aggregate properties on the performance of asphalt mixtures.

Table 8. Effect of Aggregate Properties on Asphalt Mixture Performance

Aggregate Property	Performance Category	Impact on Performance
CAA	Rutting Resistance	Higher CAA →Higher rutting resistance (better interlock and mechanical stability) [24, 27, 28, 31-33]
	Cracking Resistance	Mixed effects (depending on mix type) [32, 34, 35]
FAA	Rutting Resistance	Higher FAA → Higher rutting resistance (more internal friction) [30, 36-40] FAA alone is an insufficient indicator of rutting resistance [25, 41, 42, 44, 45]
	Fatigue Cracking Resistance	Higher FAA → Higher fatigue resistance (with strong, rough particles)[36]
FE	Rutting Resistance	High FE → Lower rutting resistance and load transfer [47-50]
SE	Rutting Resistance	Higher clay content (low SE) →Lower rutting resistance and stripping [51-54]
	Moisture Damage	Higher clay content (low SE) → Lower moisture resistance (reduce binder–aggregate bond) [51-53].
Toughness	Rutting Resistance	Low LA loss → Higher strength and load resistance [29]
	Durability	Low LA loss → Higher skid and wear resistance (more durable aggregates) [29, 55]
Soundness	Durability and Rutting	Durable and sound aggregate → Higher durability and rutting resistance [56]

2.4 Impact of Aggregate Gradation on Volumetric Properties of Asphalt Mixtures

Gradation is widely recognized as the dominant factor impacting the VMA of asphalt mixtures. Early studies, conducted in the 1950s by Dr. Norman W. McLeod [60], highlighted key gradation factors influencing VMA. The findings indicated that shifting the aggregate gradation curve away from the MDL led to higher VMA. Increasing the fine aggregate content enlarged the voids between coarse particles, thereby increasing the VMA. Substantially reducing fine aggregate content produced an “open” gradation, which also increased VMA. In contrast, the addition of mineral fillers significantly reduced VMA. In 1959, McLeod expanded his work on the volumetric properties of asphalt mixtures by establishing a relationship between critical minimum VMA and

NMAS in dense-graded mixtures. Hudson and Davis [61] introduced a numerical method to estimate VMA from aggregate gradation. Their approach was based on calculating ratios between the percentage of material passing successive sieves and included adjustments for differences between rounded and angular aggregates. They identified several key factors influencing VMA: (1) the degree of aggregate compaction, (2) the ratio of percent passing between adjacent sieve sizes, (3) the overall size range between fine and coarse aggregates, and (4) the particle shape.

Different aggregate gradation types, including dense-graded, OGFC, and SMA, yield different VMA ranges. OGFC mixtures, which contain minimal fine aggregate and rely on a stone-on-stone contact skeleton with high interconnected air voids, are generally designed with VMA exceeding 25 [62]. SMA is a gap-graded mixture in which coarse aggregate forms the primary load-bearing skeleton with fine aggregate and filler occupying the interstitial voids, typically targeting a minimum VMA of 17% to 18% [63]. Dense-graded mixtures, which incorporate a continuous distribution of aggregate sizes to achieve relatively tight packing, are designed with minimum VMA values ranging from 11% to 16% [2, 64]. These differences were illustrated by Koch [65], who compared the VMA of 9.5 mm NMAS dense-graded mix and 9.5 mm NMAS SMA mix using the same aggregate source. The results showed that SMA had significantly higher VMA (18 %) than the dense-graded mix (15.2 %) despite similar air voids. The difference is attributed to the gap-graded structure and high binder/filler content in SMA.

The distance of the aggregate gradation from the MDL on the 0.45 power chart also shows a correlation with VMA. MDL represents the gradation with the tightest packing and lowest VMA, and deviations from it, whether toward a finer or coarser gradation, generally result in increased VMA[66, 67]. Similarly, Bessette [68] found that increasing the distance from the MDL increased the VMA of asphalt mixtures for both coarse and fine-graded mixtures. Aschenbrener and

MacKean [23] evaluated 101 gradation designs in Colorado and found that VMA was most closely related to the spacing between the No. 8 and No. 200 sieves, and to an MDL drawn from the origin through the percent passing the NMAS. While gradation was confirmed as a significant influence on VMA, a practical statistical correlation between the two could not be established. Huber and Shuler [69] investigated the effect of aggregate type and gradation on VMA. The research found that by moving gradations farther away from the MDL, VMA values initially decreased, then increased for both crushed limestone and uncrushed gravel.

The percentage of aggregate passing the individual sieves, particularly the finer sieves, has been shown to significantly influence VMA. Kandhal et al. [70] identified that the percent passing No. 4 and smaller sieves were the most practical predictive variables for VMA. In a study of OGFC mixtures, Kandhal and Mallick [62] evaluated the effect of the percentage passing the No.4 sieve on the VMA using varying percentages passing the No.4 sieve (15%, 25%, 30%, and 40%). The results show that increasing the percentage passing the No.4 sieve from 15% to 40% reduced VMA progressively from 26.3% to 23.9%. However, when the fine content exceeded approximately 30% passing, the VMA did not decrease further because stone-on-stone contact was lost and the aggregate skeleton collapsed.

The amount of aggregate passing the No. 200 sieve also substantially influences the VMA of HMA mixtures. Lower contents of No. 200 material result in increased VMA, whereas higher contents lead to reduced VMA. Moreover, VMA in gradations positioned on the fine side of the MDL is more sensitive to variations in No. 200 content than those on the coarse side [23]. Xie et al [71] evaluated the effect of dust content on the VMA of eight SMA mixtures with a 4.75 mm NMAS. The results showed that increasing the dust content from 9 % to 12 % slightly decreased the VMA, with values reducing from 18.1 % to 17.1 % for granite mixtures and from 14.4 % to

14.1 % for limestone mixtures. For SMA mixtures, National Asphalt Pavement Association (NAPA) guidelines found that the VMA increased by decreasing the percentage of aggregate passing the No.4 sieve or by decreasing the percentage passing the No. 200 sieve [63].

VMA is sensitive to the gradation around the primary control sieve (PCS) [72]. According to the Bailey method [73], decreasing the percentage of aggregate passing PCS in coarse-graded mix and SMA mixtures increased VMA. For fine-graded mixes, increasing the percentage of aggregate passing PCS increases VMA.

NMAS is another gradation parameter that directly affects VMA. Khalifa [56] concluded that as aggregate size increased, both VMA and air voids decreased at any given asphalt content. Liu et al. [74] tested SMA mixtures with NMAS from 5 to 30 mm and reported a consistent inverse relationship between NMAS and VMA. For example, the 5 mm NMAS SMA had a VMA of 18-20%, whereas the 30 mm NMAS SMA mixtures had a lower VMA of about 14%. The Superpave mix design system reflects this relationship by specifying higher minimum VMA requirements for mixtures with smaller NMAS values[75]. However, a study conducted at Iowa State University suggested that specifying minimum VMA requirements solely based on NMAS may not be a realistic approach. The researchers found that factors such as the fineness modulus, which reflect the overall shape of the gradation curve, and the proportions of crushed coarse and fine aggregates, which influence surface texture, are stronger indicators of critical VMA levels. They also cautioned that the current minimum VMA criteria might exclude asphalt mixtures that could still provide satisfactory field performance [22].

Since the relationship between gradation and VMA has been well established, modeling and simulation approaches have been developed to predict VMA based on gradation characteristics. Shen and Yu [76] used discrete-element modeling (DEM) to establish the

relationship between aggregate gradation and VMA of HMA. The authors introduced a new classification system for dense-graded aggregates with coarse, medium, and fine-graded mixes and developed a gradation weighting factor, which quantified the change in void volume caused by adding aggregate. Based on the gradation weighting factor, a theoretical method for estimating VMA was proposed and validated using 11 mix designs from various states. The predicted VMA values closely matched the design values, with most differing by less than 1%. Using a 12.5 mm NMAS as an example, the study demonstrated that the proposed approach enabled fast, accurate VMA estimation, reducing design time and improving early-stage performance evaluation of HMA mixtures. Pouranian [77] proposed a linear-mixture packing model to estimate VMA changes caused by gradation variations. The model introduced the concept of central particle size to characterize the aggregate skeleton and achieved a high correlation with measured VMA. Zhang et al. [78] presented a theoretical–empirical model to predict VMA ranking without testing. They derived equations for skeleton porosity and VMA based on measured VCA and percentages passing No.4 and No.8 sieves. The predicted VMA order for AC-13 mixes matched measured values, showing potential for rapid mix design screening.

Based on the literature review, **Table 9** summarizes the effect of gradation on VMA of asphalt mixtures.

Table 9. Effect of Gradation on VMA

Gradation Factor	Impact on VMA
Gradation Type	OGFC > SMA > Dense Graded [2, 62-65].
Distance from MDL	Increase the distance → Increase VMA [60, 66-69]
Percent Passing Specific Sieves	Passing No.4, No.8, No.16, No.30, No.50, No.100, and No.200 are key predictors. [70] Lower No. 200 sieve → Increase VMA [23, 60, 63, 71] SMA and OGFC: Lower No.4 sieve → Increase VMA [62, 63] Coarse-graded and SMA: Lower PCS passing → Increase VMA Fine-graded: Increase PCS passing → Increase VMA [73]
NMAS	Smaller NMAS → Increase VMA [74, 79]

2.5 Impact of Aggregate Gradation on Performance of Asphalt Mixtures

Gradation is one of the most important aggregate characteristics influencing the performance of asphalt mixtures, including rutting resistance, cracking, durability, and permeability [80, 81]. The following sections review the current state of knowledge regarding the influence of gradation on rutting, cracking resistance, durability and permeability with particular attention to gradation type, specific aggregate sieve size, Bailey method parameters, and NMAS.

2.5.1 Rutting Resistance

Many studies have been conducted to evaluate the influence of aggregate gradation on the rutting resistance of asphalt mixtures. Golalipour et al. [82] concluded that aggregate gradation was the main factor affecting the rutting resistance. Similarly, Stakston and Bahia [43] indicated that without proper aggregate gradation, asphalt mixtures fail to achieve adequate rutting resistance, even when designed with high-quality materials. Among asphalt mix types, SMA provides the best rutting resistance as a result of stone-on-stone contact provided by coarse aggregate skeleton, followed by dense-graded mixtures [83-86].

Within dense-graded mixtures, the position of the gradation curve relative to its upper and lower bounds also affects rutting performance. Sangsefidi et al. [87] evaluated the impact of lower-

bound, mid-range, and upper-bound gradations on the permanent deformation of HMA. The results indicated that lower-bound gradation exhibited the best rutting resistance, followed by mid-range, while upper-bound gradation showed the poorest performance. This further confirmed that stone-on-stone contact among coarse aggregates in the lower bound gradation enhanced load-bearing capacity and reduced rutting potential.

Some research showed that the position related to the restricted zone of aggregate gradation could impact the rutting resistance of the asphalt mixture. Earlier Superpave specifications classified gradations according to their position relative to a restricted zone, distinguishing between gradations passing above the restricted zone (ARZ, equivalent to fine-graded), through the restricted zone (TRZ), and below the restricted zone (BRZ, equivalent to coarse-graded), as shown in **Figure 3**. BRZ gradations generally develop a better coarse aggregate skeleton and internal friction than ARZ gradations and are therefore expected to provide better rutting resistance for the asphalt mixture [88, 89]. However, other studies have reported contradictory findings, showing more rutting susceptibility for coarser-graded BRZ mixtures than finer-graded TRZ and ARZ mixtures [90-93].

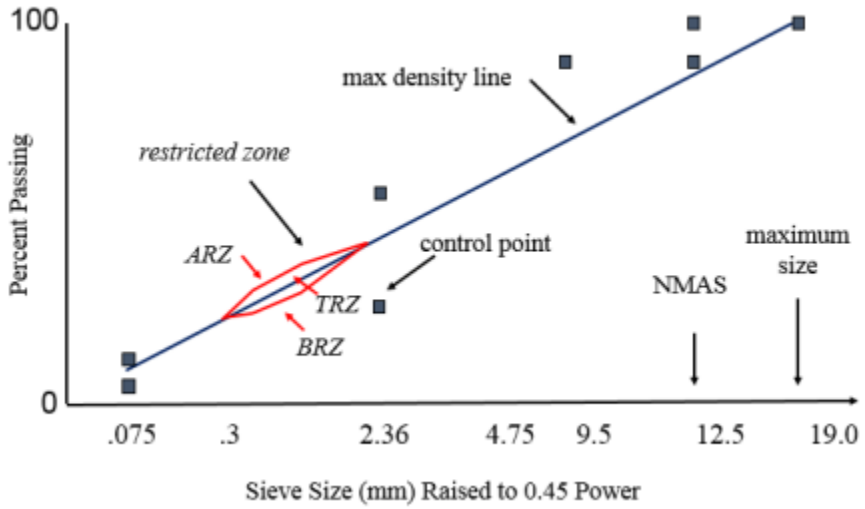

Figure 3. Superpave Gradation Types Based on Restriction Zone [94]

Moreover, the impact of restricted zones on rutting performance was also found to vary between aggregate types. Kandhal and Mallick [95] evaluated the effect of the restricted zone on mixtures containing different aggregate types using the APA test. The results showed that for granite and limestone, BRZ generally showed the highest rut depth, followed by ARZ and then TRZ. While for gravel, ARZ generally showed the highest rut depth, followed by TRZ and BRZ. Chowdhury et al. [96] reported similar variability, with BRZ producing the highest rut depths for river gravel and natural sand mixtures, while no significant differences between gradation types were observed for granite and limestone.

Due to the inconsistent trend, many studies have suggested that the restricted zone is an unreliable criterion for characterizing gradation to ensure adequate rutting resistance. Hand et al. [97] conducted a comprehensive evaluation of 21 Superpave mixtures and found that acceptable rutting performance could be achieved across all three gradation types, with laboratory results suggesting that ARZ and TRZ gradations may even outperform BRZ in some cases. This was further supported by Test track sections from 2009 to 2011 at NCAT, which demonstrated that mixtures violating restricted zones could still provide good rutting resistance [98]. Moreover, Kim et al. [90] showed that overall coarseness or fineness of the gradation likely impacted the rutting performance rather than the restricted zone. These findings highlight the limited reliability of the restricted zone in predicting rutting resistance. The restricted zone was later eliminated from the AASHTO specification.

Following the removal of the restricted zone, aggregate gradation was classified as coarse-graded or fine-graded based on NMAS and percentage passing at PCS. Hasan et al. [99] evaluated the effect of coarse and fine gradation, classified by a PCS of 2.36mm, on rutting resistance. The HWTT results showed that coarse-graded mixtures experienced significantly better rutting and

stripping resistance than fine-graded mixtures. Pan et al. [100] used PCS of 2.36 mm to design two gradations with different percentages of coarse aggregate between 2.36mm and 4.75mm. HWTT results demonstrated that the mixture designed with more coarse aggregate improved rutting resistance due to enhanced stone-on-stone contact between coarse aggregate structures.

Another key factor impacting the rutting performance of asphalt mixtures is the percentage of aggregate passing through certain sieves, particularly the No.4 sieve [101, 102]. Chen and Liao [103] demonstrated that maintaining an optimal range of aggregate passing the No.4 sieve could provide sufficient rutting resistance. For example, with a 9.5 NMAS mm mixture, the percentage passing No.4 of approximately 30% promoted effective interlock among coarse particles, whereas increasing this value to around 45% could disrupt the aggregate structure and reduce rutting resistance. In a 19 mm NMAS mixture, a percentage passing No.4 between 40% and 42% could provide an optimum internal resistance between coarse aggregates. Lv et al. [104] corroborated these findings using HWTT data from 26 mixtures, reporting that rutting depth increased progressively as the percentage passing the No.4 sieve rose from 20% to 35%, with values above 41% associated with significantly higher rutting susceptibility and stripping.

In addition to the No.4 sieve, the percentages of aggregate passing other sieves also play a significant role in the rutting performance of asphalt mixtures. Zhang et al. [105] determined that the aggregates retained on No.4 and No.8 sieves contributed over 50% of the rutting resistance, while the aggregates retained on No.16, No.30, and No.50 sieves provided more than 50% of the structural strength. Teklu [106] indicated that a small change in passing No.8 resulted in a considerable change in rutting values. Gary Thompson [107] investigated the effect of changes in the percentage of aggregate passing No.16, No.30, and No.50 sieves on the rutting susceptibility of four 12.5 mm NMAS coarse aggregate mixtures. The results showed that decreasing the

percentage passing the No. 16 sieve increased predicted rut depth. Increasing the percentage passing the No. 30 sieve slightly increased the rut depth, while a decrease in the percentage passing the No. 50 sieve led to a slight reduction in the rut depth.

The Bailey method provides a systematic framework for characterizing aggregate gradation using parameters linked to rutting performance. In the method, the Coarse Aggregate (CA) ratio is calculated by dividing the percentage of aggregate passing the half sieve and retained on the PCS by the percentage of aggregate retained on the half sieve, and the half sieve is defined as half of the NMAS [108, 109]. The CA ratio helps quantify the coarse aggregate interlock, which directly influences rutting resistance and stability [110]. According to Chen Ai-wen [111], the CA ratios between 0.4 and 0.6 ensured the optimal stability of the asphalt mixtures. Quan et al. [104] further evaluated the effect of different CA ratios on the rutting resistance by designing five gap-graded gradations with CA ratios of 1/7, 1/3, 1/1, 3/1, and 11/1 corresponding to the 9.5- 4.75 mm/ 13.2-9.5 mm ratio. The mixtures were designed with the same asphalt content of 6.2% and evaluated using the HWT test at 50°C. The results showed that a CA ratio of 1/1 produced the highest rut depth, likely because of a uniform distribution between the smaller aggregate (9.5-4.75 mm) and the larger aggregate (13.2-9.5 mm), which hindered the formation of a stable aggregate skeleton. The mixture with a CA ratio of 1/7 had a lower rut depth than those with CA ratios of 11/1 and 3/1, indicating that the skeleton created by the 13.2-9.5 mm aggregate outperformed that of the smaller aggregate (9.5-4.75 mm). Validation from the NCAT test track study conducted by the Oklahoma Department of Transportation (ODOT) further confirmed a strong correlation between the CA ratio and HWTT rut depth data, with an R^2 value of 0.769 [112]. Because of a good correlation with rut depth, the CA ratio could be used to predict the rutting performance of the asphalt mixtures [106].

Besides the CA ratio, the Fine Aggregate Coarseness FA_c ratio is another Bailey method parameter impacting the rutting resistance of the asphalt mixture. The FA_c is defined as the ratio of the fine portion of the fine aggregate fraction to the total fine aggregate fraction. It ranges from 0 to 1, where a higher FA_c value indicates a finer gradation [108, 109]. Lv et al. [104] evaluated the effect of the FA_c ratio on the rutting resistance of 13 mm NMAS dense-graded mixtures, with a similar effective asphalt film thickness, to isolate the gradation effect. In the study, the FA_c ratio was calculated as the ratio between the percentage of aggregate passing 0.6 mm and 2.36 mm. It was observed that higher FA_c ratios resulted in lower rut depths, supporting the conclusion that using finer gradation in the fine aggregate fraction improves rutting resistance. Similarly, Thompson [107] confirmed an inverse relationship between FA_c with rut depth. However, the research suggested that the CA ratio was not a reliable indicator of rutting performance for asphalt mixtures.

Another factor influencing the rutting performance of asphalt mixtures is NMAS. In general, increasing NMAS improves the rutting resistance of asphalt binder [113]. Christensen et al. [114] applied the AASHTO Mechanistic-Empirical Pavement Design Guide (MEPDG) to assess the effect of NMAS and found that increasing NMAS from 9.5 mm to 12.5 mm reduced predicted rutting depth from 50% to 30%. Similarly, Hafeez et al. [115] investigated the permanent deformation behavior of SMA mixtures with NMAS values of 9.5 mm, 12 mm, 19 mm, and 25.4 mm across a range of test temperatures of the HWTT test. The findings indicated that rutting susceptibility was influenced by both NMAS and temperature, with larger NMAS improving rutting resistance. Another research by Liu et al.[74] found the same conclusion that SMA designed with larger NMAS showed better rutting resistance. In contrast, Hand et al. [97] indicated that NMAS did not significantly impact the rutting resistance of the asphalt mixtures.

2.5.2 Cracking Resistance

Aggregate gradation parameters, including coarseness or fineness of aggregate gradation, directly impact the cracking resistance of asphalt mixtures [116]. In terms of low-temperature cracking, an increase in fine aggregate content negatively impacts the cracking resistance. This is because finer particles require more asphalt binder for adequate coating, causing less free asphalt in the mixture [117]. Therefore, mixtures designed with coarse gradation with thicker asphalt film thickness would provide better resistance to contraction and deformation at low temperatures [99, 118]. Nian et al. [119] compared the low-temperature cracking resistance of two different dense graded gradations, AC-16 and AC-13, using splitting, three-point beam bending, and a mesoscopic model. The results confirmed that the coarser aggregate and larger NMAS of the AC-16 mixture improved low-temperature cracking resistance compared to finer AC-13 mixtures.

Regarding intermediate-temperature cracking resistance, Saadeh and Asmar [35] evaluated the sensitivity of IDEAL-CT test results of two gradation types the Superpave gradation and CMHB gradation, the latter featuring a coarser aggregate structure achieved by increasing coarse aggregate percentage and reducing fines. Testing with both granite and limestone mixtures demonstrated that coarser gradation exhibited higher CT_{Index}, demonstrating better intermediate-temperature cracking.

However, using a coarser gradation adversely impacts fatigue cracking resistance. A comparison of HMA and SMA mixtures using the indirect tensile fatigue test found that SMA mixtures with coarser gradation had lower fatigue resistance than finer-graded HMA mixtures [120]. Similarly, Liu et al. [121] evaluated five HMA mixtures with varying gradations using a uniaxial penetration test and reported that mixtures containing less coarse aggregate and higher fine content exhibited better fatigue resistance.

Lv et al. [122] investigated the influence of different aggregate gradation curves on the fatigue performance of asphalt mixtures. Four gradation types were evaluated: Type A, with gradation near the upper limit and a higher proportion of fine aggregates; Type B, representing a median gradation; Type C, with gradation near the lower limit and a higher proportion of coarse aggregates; and Type D, an "S-shaped" gradation. Type D gradation was modified from Type B by reducing fine particles (<2.36 mm), increasing coarse particles (>2.36 mm) and more medium-sized aggregates (0.6–9.5 mm), and reducing the aggregate at both ends (9.5–19 mm, 0.075–0.6 mm). The results indicated that Type D exhibited the longest fatigue life, benefiting from its stable aggregate structure, where an adequate amount of fine aggregate effectively fills the voids between coarse particles. Type A, with a finer gradation, also demonstrated better fatigue performance than Types B, while Type C which contained a higher percentage of coarse aggregates, showed the shortest fatigue life.

Similarly, Saleh [123] investigated the fatigue performance of two types of dense-graded HMA: AC10-80/100 and AC14-60/70. The AC10-80/100 mix featured a 10 mm NMAS and a softer 80/100 penetration grade binder, while the AC14-60/70 mix used a 14 mm NMAS and a stiffer 60/70 penetration grade binder. Fatigue behavior was evaluated using four-point bending tests under constant strain at various strain levels. The results indicated that the finer gradation and softer binder in the AC10-80/100 mix provided significantly better fatigue resistance than the coarser, stiffer AC14-60/70 mix. This finding was further supported by Sousa et al. [124] and Chowdhury et al. [96], who also indicated that fine-graded gradation provided better fatigue performance than coarse-graded mixtures.

In contrast, a study by Hasan et al [125] observed a different trend. This study evaluated the fatigue performance of coarse- and fine-graded asphalt mixes, constructed side by side in a

two-lane, two-mile section in each direction and shared the same performance grade (PG) binder, Reclaimed Asphalt Pavement (RAP) content, and aggregate source. Laboratory-compacted samples were tested for dynamic modulus and four-point bending beam fatigue, and the results were analyzed using the viscoelastic continuum damage (VECD) model. Results indicated that the coarse mix exhibited a higher dynamic modulus and better fatigue performance compared to the fine mix.

Another gradation factor indicating the influence on the cracking resistance of asphalt mixtures is NMAS. In terms of low-temperature cracking resistance, the impact of NMAS has yielded mixed results. Jiangcai et al. [126] reported that decreasing NMAS increased the temperature-induced shrinkage strain of asphalt mixtures, suggesting that larger NMAS mixtures could help control temperature-related cracking. Similarly, a study by Moon et al. [127] found that 12.5 mm NMAS mixtures had better low-temperature cracking resistance than 9.5 mm NMAS mixtures due to a stiffer response and lower relaxation capacity at low temperatures. However, studies from Liu et al. [74] and Christensen and Bonaquist [128] reported that an increase in NMAS resulted in a decrease in the low-temperature resistance. Another study investigating the impact of HMA overlay gradation on crack resistance at low temperatures indicated that smaller NMAS improved crack resistance, but emphasized that NMAS influence was not as strong as that of fine aggregate content [129].

Regarding fatigue performance, the evidence more consistently favours smaller NMAS. Smaller NMAS generally results in higher fatigue life and better anti-cracking performance [130, 131]. Shen and Carpenter [132] quantified this relationship, finding that each increase in NMAS size corresponded to a 15% reduction in fatigue life. A comparative analysis conducted by Abo-Qudais and Shatnawi [133] compared mixtures with NMAS values of 12.5 mm, 19 mm, and 25

mm, confirming that the smallest NMAS yielded the highest fatigue life. Another research comparing the fatigue performance of 12.5 mm and 9.5 mm NMAS mixtures using flexible index (FI) obtained from the Illinois Flexibility Index Test (I-FIT) indicated that 9.5 mm NMAS provided a higher FI, indicating better fatigue performance [134]. Similarly, a study by Del Pilar Vivar, and Haddock [116] found that mixtures designed with a 9.5 mm NMAS tend to provide longer fatigue life than mixtures prepared with a 19 mm NMAS, further confirming the better fatigue cracking resistance of smaller NMAS mixtures.

In addition to improving fatigue life, using smaller NMAS improves the reflect cracking resistance of HMA mixtures [135]. Moreover, Christensen and Bonaquist [136] indicated that mixtures with a smaller NMAS were generally easier to compact, leading to better field performance compared to those with a larger NMAS.

The percentage of aggregate passing on certain sieves also impacts the cracking resistance. A study by Sreedhar and Coleri [137] found that a higher percentage aggregate passing No.200 reduced the asphalt film thickness around aggregate particles, leading to lower FI values and increased susceptibility to fatigue cracking. For reflective cracking, Abbas et al. [138] conducted a full-scale laboratory investigation and developed a predictive model through multi-linear regression analysis. The model identified the percentages passing the 3/4-inch, No. 8, No. 50, and No. 100 sieves as significant predictors of the CT_{Index}, with the percentage passing the 3/4-inch sieve demonstrating the most substantial influence, as evidenced by p-values below 0.05.

Bailey method gradation parameters have also been linked to cracking performance. Garcia [139] applied the Bailey method to evaluate how variations in aggregate gradation influence the volumetric properties and mechanical behavior of asphalt mixtures. The test results showed that ITS is positively influenced by FA_c and Fine Aggregate Fineness FA_f ratios, whereas the CA

ratio exhibited a negative correlation. Haryanto and Takahashi [140] investigated the impact of the CA ratio and the FA_c ratio on the cracking resistance of HMA mixtures using the SCB test. The results showed that as the CA ratio and FA_c ratio increased, accumulated energy required to create a new crack decreased, indicating a lower resistance to cracking.

2.5.3 Durability

Aggregate gradation influences the durability of asphalt mixtures by affecting resistance to raveling and susceptibility to moisture damage. Previous studies indicate that NMAS, gradation type, and the percentage passing specific sieve sizes are the primary factors controlling durability performance.

NMAS impacts the raveling resistance of the asphalt mixture. Reducing the NMAS from 9.5 mm NMAS or 12.5 mm to a smaller size of 4.75 mm can significantly reduce the raveling of the mixtures [141]. Likewise, the SMA mixture designed with a smaller NMAS exhibited better raveling resistance [74].

In addition to its effect on raveling resistance, NMAS also significantly affects the moisture resistance of asphalt mixtures [142]. In general, smaller NMAS reduces the moisture sensitivity of asphalt mixtures. Ji et al. [143] compared the stripping ratio from HWTT between dense-graded mixtures with NMAS of 13 mm and 20 mm. The results showed that 13 mm NMAS provided a lower stripping ratio than 20 mm NMAS, indicating better moisture damage resistance and water stability. This trend was consistent across different mixture types. For SMA mixtures, Liu et al. [74] found that reducing NMAS lowered moisture sensitivity. For OGFC mixtures, Watson et al. [144] and Xie et al. [145] concluded that 9.5 mm NMAS OGFC mixtures had higher tensile strengths after moisture and freeze/thaw conditioning compared with 12.5 mm NMAS OGFC.

Using finer gradation also improves the raveling resistance of asphalt mixtures. Tran [146] evaluated the durability of 12.5 mm NMAS, 9.5 mm NMAS OGFC, and 12.5 mm SMA mixtures using the Cantabro test. The result showed that SMA designed with finer gradation provided the highest raveling resistance, followed by 9.5 mm NMAS OGFC and then 12.5 mm NMAS OGFC, which had the lowest raveling resistance. Similarly, Bennert and Cooley [147] indicated the better durability of the OGFC mixture with a fine gradation than a coarse gradation. Watson et al [144] also confirmed the improvement in the durability of the OGFC mixture by using finer gradation and smaller NMAS size.

Besides improving raveling resistance, finer gradation provides better moisture resistance than coarser gradation. Coarser gradation has more interconnected voids, leading to a greater propensity for water to enter the voids but not drain through freely. Consequently, the mixture experiences an increase in pore pressure during loading, leading to the disruption of the asphalt film from the aggregate surface [148]. Khan et al [149] evaluated the moisture resistance of upper (fine), middle, and lower (coarse) gradation types using the Modified Lottman Test (AASHTO T-283 [150]), measuring the Tensile Strength Ratio (TSR) of both conditioned and unconditioned specimens. The results showed that the upper gradation exhibited better moisture as a result of its dense gradation, followed by the middle gradation and then the lower gradation. Similarly, Aodah et al. [151] reported that the mixtures with upper gradation exhibited better TSR than those with middle or lower gradation. An additional mechanism contributing to the poorer moisture performance of coarser mixtures is aggregate breakdown during compaction, which leaves uncoated particle surfaces that are more susceptible to water absorption and stripping.

The percentage of aggregate passing certain sieves influences the durability of the asphalt mixtures. Chadbourn et al. [67] reported that the No.8 sieve and smaller sieve sizes had a

pronounced effect on the VMA, which was directly related to the mixture durability and resistance to moisture damage. For the same NMAS, the gradation with the highest percentage passing the No.4 sieve provided the lowest Cantabro loss, indicating the best raveling resistance. This behavior can be attributed to the role of finer particles in filling voids and increasing the number of contact points between neighbouring particles, thereby enhancing cohesive bonding and aggregate-binder adhesion within the mixture [152].

2.5.4 Permeability

Aggregate gradation parameters, including coarseness or fineness of aggregate gradation, NMAS, and percentage aggregate passing of certain sieves, directly impact the permeability of asphalt mixtures.

The proportion of coarse and fine aggregate influences the permeability in asphalt mixtures. Coarse-graded mixtures generally exhibit higher permeability than fine-graded mixtures, as the greater coarse aggregate content produces a higher volume of interconnected air voids that facilitate water flow through the pavement structure [153, 154]. Research conducted in Kansas confirmed that for any given NMAS, coarse-graded mixtures consistently provided greater permeability than fine-graded mixtures [155]. Similarly, a survey by Brown et al. [156] indicated that coarse-graded Superpave mixtures have higher permeability than conventional dense-graded mixtures with a similar air void content. Additionally, studies by Westerman [157] and Choubane et al. [158], demonstrated that coarse-graded Superpave mixtures became permeable once the air void contents exceeded 6 percent.

NMAS is another key factor influencing permeability, with larger NMAS mixtures generally being more permeable than smaller NMAS mixtures at the same air void content [154, 159]. For example, at the same air void content of 6%, dense graded mixtures designed with 25

mm NMAS provided the highest permeability of $1200 * 10^{-5}$cm/sec, followed by 19 mm and then 12.5 mm NMAS mixtures, with the permeability of $140 * 10^{-5}$cm/sec and $40 * 10^{-5}$cm/sec, respectively. 9.5 mm NMAS mixtures had the lowest permeability of $6 * 10^{-5}$cm/sec [160]. A similar trend has been observed for OGFC mixtures, where increasing NMAS leads to higher permeability [161, 162]. Momm et al. [163] reported that OGFC mixtures with a 19 mm NMAS showed the greatest permeability, followed by those with 12.5 mm and 9.5 mm NMAS. This finding aligns with the results of Hasan et al. [164], who also observed a higher permeability of 12.5 mm NMAS mixtures than 9.5 mm NMAS mixtures. Moreover, SMA mixture also shows a similar relationship, where decreasing the NMAS results in reduced permeability [74].

Permeability in asphalt mixtures is influenced by the percentage of material passing specific sieves**.** Kandhal and Mallick [62] evaluated mixtures with varying percentages passing the No.4 sieve using a falling head permeameter and found a clear inverse relationship: as the percentage passing the No.4 sieve decreased, permeability increased due to the larger void spaces associated with coarser gradations. Notably, a significant increase in permeability was observed when the percentage passing the No. 4 sieve decreased from 30% to 15%. Khosla and Sadasivam [165] recommended using a low proportion of No.4 size aggregates combined with higher proportions of No.8 and No.16 size aggregates to achieve low-permeability mixtures. Additionally, increasing the proportions of ½-inch and 3/8-inch aggregates can help ensure the discontinuity of smaller voids and further reduce permeability.

Hainin et al. [166] developed a model to describe the influence of in-place density, gradation, lift thickness, and design compaction effort (N_{des}) on the permeability of Superpave pavements, as shown in **Equation 1**. The model shows that permeability increases with a higher CA ratio, and a greater percentage passing the 12.5 mm and No.30 sieves.

$$Ln(k) = -19.2 + 5.96Ln(CL) + 1.47(CA\ Ratio) + 0.078(P_{12.5}) + 0.0485(P_{1.18}) + 0.00928(N_{des}) - 0.0124(Ave.\ Thickness) \quad (1)$$

Where:

Ln (k) = Natural log of permeability

Ln (CL) = Natural log of air voids from Corelok

CA Ratio = coarse aggregate ratio

$P_{12.5}$ = Percent passing 12.5 mm sieve

$P_{1.18}$ = Percent passing 1.18 mm sieve

N_{des} = Design compaction effort of mix design

Ave. Thickness = Average thickness of a given core (lift thickness at location of core)

Based on the literature review in this Chapter, **Table 10** summarizes the effect of gradation on the performance of asphalt mixtures.

Table 10. Impact of Gradation on Asphalt Mixture Laboratory Performance

Gradation Parameter	Performance Type	Impact on Laboratory Performance
Gradation Type	Rutting resistance	SMA > Dense-graded [83-86] Coarse-graded > Fine-graded [99, 100] Lower-bound > Mid- bound > Upper-bound [87]
	Intermediate-temp cracking	Coarser gradation → Increase cracking resistance [35]
	Fatigue cracking	Finer gradation → Increase fatigue life [120-123] Coarse gradation→ Increase fatigue cracking resistance [125]
	Low-temp cracking	Coarse gradation → Increase low-temp cracking resistance (thicker asphalt film) [99, 117-119]
	Durability	Finer gradations → Increase raveling and moisture resistance [146-149, 151]
	Permeability	Coarse gradations → Increase permeability [153-158]
Restricted Zone (ARZ, TRZ, BRZ)	Rutting resistance	Mixed findings- vary by aggregate types [88-93, 97, 98]
Percent Passing Certain Sieve	Rutting resistance	Passing No.4, No.8, No.16, No.30, and No.50 sieves greatly affect rutting [101-104, 106, 107]
	Cracking resistance	Significant influence on fatigue and reflective cracking performance [138] High dust content (No.200) → Decreasing fatigue resistance [137]
	Permeability	Lower passing No.4 sieve, higher passing of No.8 and No.16 sieves → Higher permeability [62, 165] Higher passing 12.5 mm and No. 30 sieves → Higher permeability [166]
	Durability	Highest passing the No.4 → the best raveling resistance [67, 152]
Bailey Method	Rutting resistance	Strong correlation with rut depth [104, 110, 111] Higher FA_c → Increase rutting resistance [104, 107]
	Cracking resistance	Higher CA and FA_c ratios → Reduce crack resistance [140] ITS positively influenced by FA_c and FA_f; and CA ratio negatively correlated with ITS[139]
NMAS	Rutting resistance	Larger NMAS → Increase rut resistance [74, 113-115] Insignificantly affects rutting resistance [97]

	Low-temp cracking	Mixed findings [74, 126-128]
	Fatigue cracking	Smaller NMAS → Increase fatigue life [116, 130-134]
	Reflect cracking	Smaller NMAS → Increase the reflect cracking resistance [135]
	Durability (raveling)	Smaller NMAS → Increase raveling resistance [74, 141]
	Moisture damage	Smaller NMAS → Increase moisture resistance [74, 143-145]
	Permeability	Larger NMAS → Increase permeability [154, 159], [161, 162]
CA Ratio	Permeability	Higher CA ratio → Increase permeability [166]

2.6 Prediction Models for Volumetric and Laboratory Performance of Asphalt Mixtures

2.6.1 Prediction Models for Volumetric Properties

Since the relationship between gradation and VMA has been well established, modeling and simulation approaches have been developed to predict VMA from gradation characteristics, aiming to reduce reliance on extensive laboratory testing during the mix design process. **Table 11** below summarizes the prediction models developed to predict VMA using parametric models.

Table 11. Summary of Regression Model for VMA Prediction

Authors	Prediction Formula	Parameter Definitions
Zang [167]	$VMA=10.7-0.0193X_1+0.04138X_2-0.06001X_3+0.03735X_4-0.04294X_5-0.01393X_6-0.05843X_7+0.0385X_8+0.2596X_9-0.1064X_{10}-0.1317X_{11}+0.0856X_{12}-165.5X_{13}+164.4X_{14}$	X_1–X_{12} = Percent aggregate passing at sieves: 26.5 mm, 19 mm, 16 mm, 13.2 mm, 9.5 mm, 4.75 mm, 2.36 mm, 1.18 mm, 0.6 mm, 0.3 mm, 0.15 mm, and 0.075 mm X_{13} = Bulk specific gravity of aggregate. X_{14} = Apparent specific gravity of aggregate.
Hudson and Davis [61]	$VMA=VMA_0 \cdot F_1 \cdot F_2 \cdots F_n$	VMA_0 = Void percentage of the base aggregate (typically 26- 32%) F_i = Interference factor between size groups n = Number of gradation size groups added incrementally.
Shen and Yu [76]	$VMA = \frac{\Sigma_{i=1}^{n} f_{vi} V_{ai}}{\Sigma_{i=1}^{n} (1+f_{vi}) V_{ai}}$	f_{vi} = Interference factor between base aggregate and size group *i*, refers to the ratio of the changed void volume after adding a new aggregate size to the bulk volume of the

		new aggregate V_{ai} = Volume fraction of aggregate retained on sieve *i*. n = Number of gradation size groups.
Vavrik, Pine and Huber [109]	VMA = −24.6 + 20.1CA^2 − 3.8CA − 191.1FAC^{2+} 181.0FAC + 87.3FAf^2 − 36.6FAf	CA = Coarse Aggregate Ratio = ratio of % passing half sieve to total coarse fraction. FA_C = Fine–Coarse Ratio = ratio of % passing secondary control sieve to % passing primary control sieve. FA_f = Fine–Fine Ratio = ratio of % passing tertiary control sieve to % passing secondary control sieve.
Liu [168]	VMA = 21.435 − 13.06FAC + 3.317CA	
Li [169]	VMA=49.48 − 147.395FAC −12.866CA^2 + 84.094FAC^2 −35.603FAf^2 + 63.997CAFA_C −27.799CAFA_f + 7.257$FA_f FA_C$	
Inoue, Gunji and Akagi [170]	Representative diameter: $Sd = d_r m d_0 n$ Two-aggregate void ratio: $e = me_s + ne_p - \frac{1}{4k}(\text{mlog}\, r - \frac{1}{2}\text{nlog}\, r - \frac{1}{2}\text{nlog nlog}\, r\, \frac{n}{m} - kn\text{log}\ n\text{log}\ Sd)$ Final VMA: VMA $= \frac{e}{1+e}$	d_r, and d_0 = Particle diameters of coarse and fine fractions, respectively m, n = Volume fractions of coarse and fine particles (m + n = 1). e_s, and e_p = Void ratios of coarse and fine aggregates, respectively r = Sieve size ratio (coarse/fine). k = particle shape factor
Yu and Standish [171] applied by Pourania and Haddock [172]	$V_i^T = \sum_{j=1}^{M-1} \bar{V}_{Lj} X_j + \bar{V}_{Mi} \sum_{j=M}^{N} X_j + \sum_{j=N+1}^{n} \bar{V}_{Sj} X_j$	V_i^T = Specific volume of the entire particle system. $\bar{V}_{Mi}$ = Specific volume of particles within the control size range $\bar{V}_{Lj}$ = Specific volume of particles larger than the control size range $\bar{V}_{Sj}$ = Specific volume of particles smaller than the control size range X_j = Volume ratio of each particle group M and N = Define the control size range
Li and He [173]	VMA= $V_{CA}^{DRC} - (1 - V_{CA}^{DRC}) \frac{\rho_c (P_{2.36} - P_{0.075})}{(1 - P_{2.36})\rho_f} - (1 - V_{CA}^{DRC}) \frac{\rho_c P_{0.075}}{(1 - P_{2.36})\rho_m}$	V_{CA}^{DRC} = Voids in coarse aggregate (from rodded test) ρ_c,ρ_f, and ρ_m e = Density of coarse aggregate, fine aggregate, and filler $P_{2.36}$ and $P_{0.075}$ = Percent passing 2.36 mm and 0.075 mm sieve

Besides using a parametric model, machine learning (ML) has been increasingly developed to predict volumetric properties of asphalt mixtures. Othman, K [174] employed Deep Artificial

Neural Networks (ANN) to predict VMA during the Marshall mix design process. The models were trained and tested on 102 hot mix asphalt samples, with the aggregate gradation serving as the sole input, specifically using the percentage passing sieves from 3/8 inch to No. 200. Several ANN architectures were evaluated by testing different activation functions (linear, logistic, tanh) and hidden-layer configurations, with model performance assessed using the correlation coefficient on the testing set. Overall, linear activation functions produced higher correlation values than the other activation functions; however, the best performance (correlation = 0.66) was achieved with an architecture consisting of four hidden layers, each with seven neurons, using the tanh activation function.

Shaikh, Sadiya et al. [175] introduced data-driven predictive models for optimizing pavement design by estimating VMA. The models were developed using a dataset of 960 samples, with 80% allocated for training and 20% for testing. Key input variables included aggregate source, NMAS, aggregate surface area, binder type, and compaction levels. The study employed four ML algorithms: Linear Regression (LR), Decision Tree Regression (DTR), Random Forest Regression (RFR), and Gradient Boosting Regression (GBR). Among these, the GBR model was identified as the most effective, demonstrating high accuracy with the following performance metrics: coefficient of determination R^2 of 0.9739, Root Mean Square Error (RMSE) of 0.2347, Mean Absolute Error (MAE) of 0.1817, and Mean Absolute Percentage Error (MAPE) of 1.1703. The ensemble models, GBR and RFR, significantly outperformed LR and DTR, demonstrating their reliability in making accurate predictions.

2.6.2 Prediction Models of Rut Depth

Several prediction methods have been explored to predict the rut depth of asphalt mixtures. Ezzat and Abed [176] developed a multiple linear regression (MLR) model to predict rutting depth in

asphalt mixtures modified with Styrene-Butadiene-Styrene (SBS) and Polyvinyl Chloride (PVC) polymers. The model utilized experimental data from HWTT tests, with four primary input variables: the number of loading cycles (N), aggregate gradation (Ga), expressed as the percentage passing the No. 8 sieve, asphalt binder viscosity (g), and testing temperature. All predictors were identified as statistically significant through a stepwise regression analysis. The final prediction equation was presented in **Equation 2**. Model performance, assessed using R^2 and the Standard Error of Regression (SER), yielded an R^2 of 0.69 and an SER of 0.35, indicating good agreement between measured and predicted rut depths. Additional validation using Q–Q plots, residual histograms, scatter plots, and the Kolmogorov–Smirnov test confirmed that residuals followed a normal distribution and that the model reliably captures permanent deformation trends.

$$\log(RD) = 0.787\log(N) - 7.094\log(Ga) - 1.429\log(g) + 2.831\log(T) - 9.856 \quad (2)$$

Majidifard et al [177] developed a rut depth prediction model for asphalt mixtures using Gene Expression Programming (GEP). The model was formulated based on 96 HWTT results with 13 input variables representing mixture composition and test conditions. These variables included mix type, useful temperature interval, binder high-temperature grade, asphalt content, asphalt binder replacement, NMAS, RAP content, RAS content, gradation type, aggregate type, crumb rubber content, test temperature, and number of wheel passes. The dataset was split into learning (70%), testing (15%), and validation (15%) subsets. Model performance was evaluated using the correlation coefficient, RMSE, and MAE. These results show that the GEP model accurately captured nonlinear relationships between mixture variables and rutting behavior, with a correlation coefficient of 0.75, an RMSE of 1.81, and an MAE of 1.27.

Ma [178] predicted the rutting resistance index of HMA, which was the function of the number of wheel passes and rut depth, using the Random Forest (RF) algorithm. The model

utilized 4321 observations from 25 different mix types, with ten input variables that describe the material composition and properties, including mix type, aggregate properties, asphalt binder PG, asphalt content, anti-stripping additives, RAP content, NMAS, and modified binder identification. After data pre-processing and hyperparameter optimization, the model performance was evaluated using 3-fold cross-validation, yielding an acceptable prediction accuracy with an R^2 of 0.64 and an RMSE of 2739.3. The RF model was also compared to a convolutional neural network model built on the same dataset; the RF model was found to have better accuracy and lower error, as the model achieved an R^2 less than 0.3 and an average RMSE of 4070.6.

Lang et al. [179] developed an ML model to predict rut depth using the Illinois DOT mixture performance database, which contains 8117 HWTT data samples from 3772 asphalt mix designs collected between 2010 and 2020. The rutting model utilized 32 input variables, including mix type (AC/SMA), NMAS, pavement layer, N-design gyrations, binder PG, polymer modification, recycled material content, aggregate properties, air voids, and blend G_{sb}. After evaluating multiple algorithms, the RF model without local trend adjustment provided the best rut-depth prediction with R^2 of 0.789, MSE of 3.337, and MAE of 1.069. These results confirm that the RF model provided robust and accurate predictions of rut depth across diverse mix designs and testing conditions.

Erten and Remzi [180] developed ML models to predict Marshall stability and flow using two datasets, each containing 17 physical and mechanical input variables such as aggregate gradation, binder properties, volumetric parameters, and specific gravity measures. Among prediction models, Extra Trees consistently provided the highest predictive accuracy, achieving near-perfect performance for flow (MAE $\approx 4.06\times10^{-15}$, RMSE $\approx 4.97\times10^{-15}$, Accuracy $\approx$ 99.999%) and strong performance for stability (MAE = 109.52, RMSE = 150.67, Accuracy = 90.45%).

Model interpretability using Shapley Additive Explanations (SHAP) values revealed that variables such as void content, fine aggregate fractions, binder softening point, and VMA were the most influential predictors.

Tong et al. [181] analyzed 648 asphalt mixture specimens from the 2020 Virginia Accelerated Pavement Testing Program to develop models for predicting Balanced Mixture Design (BMD) performance indices, with rutting susceptibility evaluated by the Asphalt Pavement Analyzer (APA) rut test. Four algorithms were evaluated, including LR, RF, Extreme Gradient Boosting (XGB), and Support Vector Regression (SVR). Model inputs included two categorical variables ("Design" and "Type"), volumetric properties, RAP content, three gradation parameters (percentage passing at No. 200, No. 30, and No. 8 sieves), and binder viscosity expressed as log (|G*|). The model was validated using 5-fold cross-validation. The resulting R^2 values for LR, RF, XGB, and SVR were -0.03, 0.42, 0.56, and 0.77, respectively.

2.6.3 Prediction Models of Cracking

Several prediction methods have been developed to predict the cracking performance of asphalt mixtures. Cooper Jr et al. [182] developed a predictive model for intermediate-temperature cracking resistance, using the SCB critical strain energy release rate (J_c) as the response variable. The model used 24 input variables, which described gradation, volumetric, and binder properties. After eliminating non-significant and multicollinear variables using Pearson correlation and variance inflation factor (VIF) screening, the final cracking model was developed using a nonlinear ANN. The resulting equation expresses J_c as a function of percentage passing No. 200 sieve, dust-to-asphalt ratio, film thickness, and polymer content. The performance of the model showed excellent accuracy, with an R^2 of 0.93 and an RMSE of 0.003 kJ/m². When validated using an independent test set, the model showed acceptable accuracy with an R^2 of 0.81 and an RMSE of

0.015 kJ/m². These results demonstrate that the ANN model reliably predicted SCB cracking performance from mixture design parameters and could be used as an efficient screening tool during mix design.

Lang et al [179] developed a two-step machine learning framework for predicting FI from the I-FIT test, using an extensive Illinois DOT database. The cracking model utilized 32 input variables that described mixture characteristics, including mix type (AC or SMA), NMAS, pavement layer, number of gyrations, aggregate properties, recycled material content, binder PG, polymer modification, and volumetric characteristics, such as air voids and blend aggregate G_{sb}. Among the evaluated algorithms, the RF regressor achieved the best FI prediction performance, yielding an R^2 of 0.797, an MSE of 15.396, and an MAE of 2.618. Moreover, SHAP value analysis for the FI model identified asphalt content and aggregate content retained on the 9.5 mm and 0.6 mm sieves as the most significant features influencing FI predictions.

Nguyen et al [183] developed ML models to predict the CT_{Index} of asphalt mixtures using a database of 107 experimental samples, with eight input variables, including percentage passing No.4 and No.200 sieves, binder content, penetration at 25°C, flash point, softening point, RAP, and rejuvenator content. Three model types were evaluated, including Multivariable Adaptive Regression Spines, RF, and K-Nearest Neighbors. The models were trained on 70% of the data and tested on the remaining 30% with R^2, RMSE, MAE, and Mean Absolute Percentage Error (MAPE) as performance metrics. The RF model was identified as the most effective and stable predictor, achieving the highest performance metrics on the test set: R^2 of 0.9953, RMSE of 8.7535, MAE of 7.2763, and MAPE of 0.0795. A SHAP sensitivity analysis revealed that No.4 had the most significant effect on CT_{Index}, followed by RAP and binder content.

Rad, Sina Mousavi, et al. [184] predicted the CT_{Index} under Long-Term Oven-Aged (LTOA) conditions using the cracking resistance of 34 plant-produced asphalt mixes in Oklahoma. The input features of models included continuous variables, including gradation, VMA, Voids Filled with Asphalt (VFA), dust-to-binder ratio, asphalt content, theoretical maximum specific gravity G_{mm}, and RAP content, and categorical variables including NMAS, binder PG, and mix type. A range of parametric and non-parametric ML models were assessed. The non-parametric algorithms consistently outperformed the parametric ones, with the XGBoost regressor identified as the best-performing model, with an R^2 of 80.8, RMSE of 11.000, and Mean MAE of 7.706. Moreover, the residual of XGBoost model showed no discernible pattern, further confirming the robustness and reliability.

Tong et al. [181] used 648 asphalt mixture specimens from the 2020 Virginia Accelerated Pavement Testing Program to develop models for predicting CT_{Index}. The ML algorithms employed were LR, RF, XGB, and SVR. The input features for the models, selected using Pearson correlation coefficients ($r < 0.8$), included two categorical variables ("Design" and "Type"), volumetric properties, RAP content, three gradation parameters (percentage passing at No.200, No.30, and No.8 sieves), and binder viscosity expressed as log|G*|. The model performance was validated using 5-fold cross-validation. The test results showed that LR achieved an R^2 of 0.63 and MAPE of 0.53; RF achieved an R^2 of 0.88 and MAPE of 0.17; XGB achieved an R^2 of 0.92 and MAPE of 0.19; and SVR achieved an R^2 of 0.88 and MAPE of 0.23.

Mirzaiyanrajeh et al. [185] developed prediction models for low-temperature fracture energy (G_f) of asphalt mixtures using an experimental database of 852 Disk-Shaped Compact Tension (DCT) test results from 71 mixtures. The input variables, selected based on their availability during the mix design process, were categorized into three groups: Group A (traffic

level, binder PG, NMAS, binder content, RAP), Group B (percentage passing 3/8-inch, No.4, and No.200), and Group C (VMA, asphalt film thickness, G_{mm}, G_{mb}, and G_{sb}). Three model types were evaluated: Augmented Full Quadratic Model (AFQM), ANN, and Self-Validated Ensemble Modeling (SVEM). For the combination of all input variables (Group A+B+C), the SVEM technique was selected as the preferred prediction method because it provided comparable precision (R^2 of 0.93 for training and R^2 of 0.87 for testing) to other model approaches, while also having a lower computational cost and time consumption. A sensitivity analysis concluded that the most effective factors were the design traffic level, low-temperature binder grade, the percentage passing 3/8-inch sieve, and VMA.

2.7 Bailey Method

2.7.1 An overview of The Bailey Method

Over the years, many mix design methods have been developed to ensure asphalt mixtures that meet both design specifications and performance requirements, with the Superpave method now representing the most widely adopted approach in the United States [186]. However, despite aggregates accounting for approximately 95% of the total mixture weight [187], the Superpave specification provides no formal guidance on selecting aggregate gradation or understanding its relationship with volumetric properties and performance. Consequently, gradation selection within the Superpave framework remains largely a trial-and-error process driven by practitioner experience rather than systematic design principles. This highlights the need for a more systematic approach to designing aggregate structure and gradation.

This section introduces the Bailey method, which is used to develop or adjust aggregate gradation in HMA based on principles of aggregate packing and interlock. Originally developed by Robert D. Bailey in the early 1980s, the method was later refined by Dr. Bill Vavrik of the

ERES Division of Applied Research Associates, Inc., and Mr. Bill Pine of Heritage Research Group [109]. The Bailey method applies to all dense-graded mixtures (coarse-graded and fine-graded mixtures) and SMA, regardless of NMAS.

2.7.2 Coarse and Fine Aggregate

In general, coarse aggregate is defined as particles larger than No.4 sieve, while fine aggregate refers to particles smaller than No.4 but larger than No. 200 sieve. However, the Bailey method redefines these categories to better reflect the functional roles of coarse and fine aggregates in achieving interlock. In this context, coarse aggregate consists of larger particles that create voids, while fine aggregate comprises smaller particles that fill those voids. Therefore, a single sieve size is insufficient to distinguish between coarse and fine aggregate.

To address this, the Bailey method introduces the PCS, which serves as the dividing point between the coarse and fine aggregate fractions in the combined blend gradation. The PCS is determined based on NMAS, but it is important to note that the Bailey method defines NMAS as one sieve larger than the first sieve that retains more than 15%, which is different from the Superpave, which uses a 10% threshold [188] . The PCS is calculated using **Equation 3.**

$$PCS=0.22*NMAS \quad (3)$$

Table 12 below presents the NMAS, calculated PCS, and the closest standard sieve normally used in the US.

Table 12 Primary Control Sieves Used in the Bailey Method [109, 189]

NMAS	Calculated PCS (0.22*NMAS)	PCS
37.5 mm (1-1/2 inch)	8.250 mm	9.5 mm (3/8 inch)
25 mm (1inch)	5.500 mm	4.75 mm (No.4)
19.00 mm (3/4 inch)	4.180 mm	4.75 mm (No.4)
12.5 mm (1/2 inch)	2.750 mm	2.36 mm (No.8)
9.5 mm (3/8 inch)	2.090 mm	2.36 mm (No.8)
4.75 mm (No.4)	1.045 mm	1.18 mm (No.16)

The PCS also categorizes each aggregate stockpile as either coarse or fine aggregate. If the individual aggregate has a percentage of aggregate passing at PCS of the combined blend equal to or less than 49.9%, it is considered coarse aggregate. Otherwise, if the passing percentage is equal to or higher than 50%, it will be considered as fine aggregate [109].

2.7.3 Evaluating the Aggregate by Volume

Current mix design practices typically combine aggregates by weight, which does not account for differences in aggregate packing. For example, aggregates with the same mass but different bulk specific gravity (G_{sb}) can produce different volumes, leading to variations in void content and particle arrangement. Therefore, combining aggregates by volume offers a more accurate representation of packing characteristics and better control over the proportions of coarse and fine aggregates.

To assess aggregate packing, unit weight tests, in accordance with AASHTO T19 [190], including the Loose Unit Weight (LUW) and Rodded Unit Weight (RUW) tests, are conducted on full gradation for each aggregate stockpile. The LUW, measured for coarse aggregate (CA LUW), represents the amount of aggregate that fills a unit volume with no compaction effort. The LUW is calculated by dividing the weight of the aggregate by the volume of the bucket (e.g., lb/ft³ or

kg/m³). From the LUW and the dry G_{sb}, the volume of voids in the coarse aggregate can be determined. This condition reflects the volume of aggregate interlock without external compaction. In the Bailey method, the designer selects a target volume of coarse aggregate, or degree of interlock, based on the desired mix type: coarse-graded or fine-graded. The selected LUW value is referred to as the Coarse Aggregate Chosen Unit Weight (CA CUW). A coarse-graded mix is a dense-graded mixture in which the volume of coarse aggregate is equal to or exceeds the CA LUW, indicating the beginning of coarse aggregate interlock. This mix type relies on both the coarse and fine aggregate fractions for load bearing, and the CA CUW value is typically set at approximately 100% of the CA LUW.

A fine-graded mix is a dense-graded mixture in which the coarse aggregate volume is less than the CA LUW. This condition indicates that the coarse aggregate fraction (above PCS) is dispersed and "floating" within the fine aggregate fraction (below PCS). In this case, the fine aggregate primarily carries the load. To achieve this structure, the Bailey method recommends a CA CUW value of 80% or less of the CA LUW, ensuring that the coarse aggregate fraction remains separated and well-distributed within the fine aggregate fraction.

To further evaluate aggregate packing under compaction, the RUW is measured for each stockpile, both fine (FA RUW) and coarse (CA RUW), using AASHTO T19 [190], Unit Weight and Voids in Aggregate. This test reflects the increase in aggregate packing that results from applied compaction effort. The RUW (lb/ft³ or kg/m³) is determined by dividing the weight of the aggregate by the volume of the bucket.

Because fine aggregate plays a key role in supporting loads within a dense-graded mixture, 100% of the FA RUW is typically used to ensure it effectively fills the voids created by the coarse fraction. When combining multiple fine and coarse aggregate stockpiles, both the coarse and fine

aggregate blends are commonly selected at 100% by volume. This will impact the gradation of coarse (above PCS) and fine (below PCS) fractions.

In the Bailey method, SMA mixtures are defined as gap-graded mixes where the volume of coarse aggregate exceeds the CA RUW, indicating a high level of coarse aggregate interlock. As a result, the CA CUW is set to be greater than 100% of the CA RUW. In SMA mixtures, the coarse aggregate is primarily responsible for load bearing, while the fine aggregate plays a minimal structural role. Therefore, the FA CUW is selected as 100% of the FA LUW [108, 109]. **Table 13** below summarizes the recommended CA CUW values for different mix types in the Bailey method.

Table 13. Recommended CA CUW for Each Mix Type [108]

Mix Type	Min% CA LUW	Max% CA LUW	Min% CA RUW	Max%CA RUW
Fine-Graded	60	80	--	--
Coarse Graded	95	105	--	--
SMA	--	--	110	125

2.7.4 Evaluation of Aggregate Packing

After determining the percentage by weight of each individual aggregate in the combined blend, the aggregate packing of the blend is analyzed. It is important to note that, in addition to gradation, the degree of packing is influenced by several other factors, including compaction effort, aggregate shape, surface texture, and aggregate strength [109]:

- Gradation: Gradation is the dominant factor that affects aggregate packing. A well-graded mixture with a proper balance of coarse and fine aggregates packs more efficiently than a uniform or single-sized aggregate.

- Type and amount of compaction effort: Different compaction types and efforts, for example, 50-blow Marshall compaction and 75 Superpave gyrations, can orient aggregate structures differently, therefore influencing packing density and interlock.
- Shape of the aggregate: Aggregate shape plays a critical role. Rounded aggregate creates less interlock than cubical and angular aggregate. Aggregates with higher FE are prone to breaking down during compaction, negatively impacting the packing degree.
- Surface texture of the aggregate: Microtexture increases friction, limiting the ability of aggregates to slide and rearrange during compaction.
- Aggregate strength: Stronger aggregates resist degradation during mixing and compaction. On the other hand, weaker aggregates are more susceptible to fracture, breakdown, change the gradation, and affect the intended packing structure.

Within the Bailey Method, aggregate packing is addressed through four interrelated principles: (a) CA Volume, (b) CA ratio, (c) FA_c ratio, and (d) FA_f ratio. These principles provide a structured framework for understanding how changes in aggregate gradation and packing affect mix type, VMA, and compactability [108, 109].

2.7.4.1 Principle No.1- CA Volume

The first Bailey principle focuses on determining the desired CA volume in the combined blend, which is controlled through CA CUW. This selection not only influences the volume of the coarse fraction but also indirectly affects the volume of fine aggregate in the mixture. For fine- and coarse-graded mixes, the CA CUW is expressed as a percentage of the CA LUW, whereas for SMA mixtures, it is expressed as a percentage of the CA RUW.

For coarse-graded mix and SMA mixtures, increasing CA volume by raising CA CUW makes the combined blend coarser, corresponding to a decreasing percentage of aggregate passing PCS. This results in an increase in CA interlock and VMA. Although the volume of the fine

aggregate fraction decreases, the increase in CA interlocking redirects the compaction energy to the coarse fraction, resulting in less compaction in the fine aggregate fraction. Consequently, the actual volume of fine fraction achieved is higher than the design, but at a lower density. Generally, a 4% change in PCS leads to approximately a 1% change in VMA [108, 109].

For fine-graded mixes, decreasing CA volume increases the volume of fine aggregate (increases the percentage of aggregate passing PCS). The fine aggregate, which has voids between them, will replace the equivalent amount of the coarse aggregate, increasing VMA. Additionally, the VMA increases by having more particles to orient with the same type and amount of compaction effort. As a general guideline, a 4% change in PCS typically leads to an approximate 1% change in VMA [108, 109].

2.7.4.2 Principle No.2- CA Ratio

The CA ratio is used to characterize the aggregate size distribution of coarse aggregate fraction (above PCS). This distribution influences the aggregate packing of both coarse and fine fractions. To better evaluate this packing character, the concept of the half sieve, which is defined as half of the NMAS, is introduced. The particles passing through the half sieve are called “interceptors”. The interceptor aggregates are too large to fit into the voids created by the larger aggregates. The larger aggregates retained on the half sieve are called “pluggers”. The CA ratio is calculated as the ratio of “interceptors” to “pluggers” using the **Equation 4**.

$$\text{CA ratio} = \frac{(\%\ \text{Passing Half Sieve} - \%\ \text{Passing PCS})}{(100\% - \%\ \text{Passing Half Sieve})} = \frac{\text{"Interceptors"}}{\text{"Plugger"}} \quad (4)$$

Table 14 summarizes the recommended CA ratios based on the mix type and NMAS. In general, lower NMAS reduces the recommended CA ratio. This is because a lower NMAS corresponds to a narrower range of coarse aggregate sizes, making the blend easier to compact.

The mixture with higher NMAS has a wider range of aggregate sizes, which can be considered as continuous gradation.

Changing the CA ratio will affect the VMA of the asphalt mixtures. In coarse-graded and SMA mixtures, increasing the CA ratio introduces a greater proportion of interceptor aggregates into the coarse fraction. These particles reduce compaction efficiency and limit the packing of the fine aggregate fraction, leading to an increase in VMA. Typically, a 0.2 increase in CA ratio results in a 0.5% to 1.0% increase in VMA. However, when the CA ratio approaches the upper limit of the recommended range, the mix becomes "balanced", meaning that neither interceptors nor pluggers dominate the coarse fraction. In this condition, compaction becomes more difficult. On the other hand, if the CA ratio is too low, an excess of pluggers may lead to segregation, reducing mixture uniformity and performance [108, 109].

Moreover, SMA also has a lower range of the CA ratio, as shown in **Table 14**. SMA provides a higher coarse aggregate interlock than a coarse-graded gradation mixture and therefore requires being more compact. In a fine-graded mix, this ratio is referred to as the old CA ratio and has minimal effect on VMA. However, it still affects the segregation susceptibility of the asphalt mixture.

Table 14. Recommended CA Ratio [108, 109]

	Coarse Graded		SMA		Fine Graded	
NMAS	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**
37.5mm (1-1/2 inch)	0.80	0.95	NA	NA	0.80	1.425
25.0mm (1inch)	0.70	0.85	0.45	0.60	0.70	1.275
19.0mm (3/4 inch)	0.60	0.75	0.35	0.50	0.60	1.125
12.5mm (1/2 inch)	0.50	0.65	0.25	0.40	0.50	0.975
9.5mm (3/8 inch)	0.40	0.55	0.15	0.30	0.40	0.825
4.75mm (No.4)	0.30	0.45	0.05	0.20	0.30	0.675

2.7.4.3 Principle No.3- The Overall Fine Fraction

The aggregate fraction (below PCS) contains a sufficient range of aggregate sizes to be considered as a blend itself, including coarse and fine portions. The coarse portion of the fine aggregate fraction creates voids that will be filled by the fine portion of the fine aggregate fraction. If the aggregate retained on PCS is removed, the amount of aggregate passing PCS is 100%, and the percentage retained on the next smaller sieve is greater than 15%. In that situation, the PCS can be treated as NMAS. To break down coarse and fine portions of the fine fraction, the Second Control Sieve is introduced, which is determined using **Equation 5**. Principle No.3 is used to characterize the aggregate packing of the fine fraction and is presented using the following **Equation 6**.

$$SCS = 0.22 * PCS \quad (5)$$

$$FA_c = \frac{\%\ \text{Passing SCS}}{\%\ \text{Passing PCS}} \quad (6)$$

When FA_c ratio increases, the packing between fine fraction improves due to the increase in the volume of fine portion (below SCS). As a result, the VMA decreases. The recommended

value of FA_c The ratio is between 0.35 and 0.50. Higher values generally indicate an excessive amount of the fine portion of the fine aggregate is included in the mixture, which is often caused by an excessive amount of natural sand. The FA_c above 0.5 should be avoided, as it makes a gradation fall in the restricted zone. Conversely, too low FA_c The ratio makes the gradation not uniform, leading to compaction problems [108, 109].

2.7.4.4 Principle No.4- Fine Portion of Fine Fraction

The fine portion of the fine fraction will fill the voids created by the coarse portion of the fine fraction (passing SCS). The fine portion of the fine fraction has enough particle sizes and can be considered as a blend itself. If the aggregate retained on SCS is removed, the amount of aggregate passing SCS is 100%, and the percentage retained on the next smaller sieve is normally greater than 15%. In that situation, the SCS can be viewed as NMAS and the maximum size. To breakdown coarse and fine portions of the fine part of the fine fraction, the Tertiary Control Sieve (TCS) is introduced, which is determined using **Equation 7**. Principle No.4 is used to characterize the aggregate packing of the fine portion of the fine fraction and is presented using the following **Equation 8.**

$$TCS = 0.22 * SCS \quad (7)$$

$$FA_f = \frac{\%\ \text{Passing TCS}}{\%\ \text{Passing SCS}} \quad (8)$$

When FA_f ratio increase, the fine part of a fine portion (minus SCS) is finer, which decreases VMA. The recommended value of FA_f ratio is between 0.35 and 0.50 [108, 109].

2.7.5 Fine-Graded Mixes

For the fine-graded mixes, the volumetric properties are primarily influenced by the fine fraction (passing PCS). The particles above the PCS are generally spaced too far apart to form a structural

skeleton and create voids. Therefore, the Bailey method focuses on fine-fraction analysis to better understand aggregate packing. The fine fraction is viewed as a primary blend with coarse aggregate, which creates voids, and fine aggregate, which fills them. The new NMAS is normally determined as the original PCS. From that, the new half sieve, new PCS, new SCS, and TCS (if available) are determined. The new CA ratio, new FA_c ratio, and new FA_f ratio are determined based on the new sieves. **Table 15** and **Figure 4** below illustrate the new sieves and ratios.

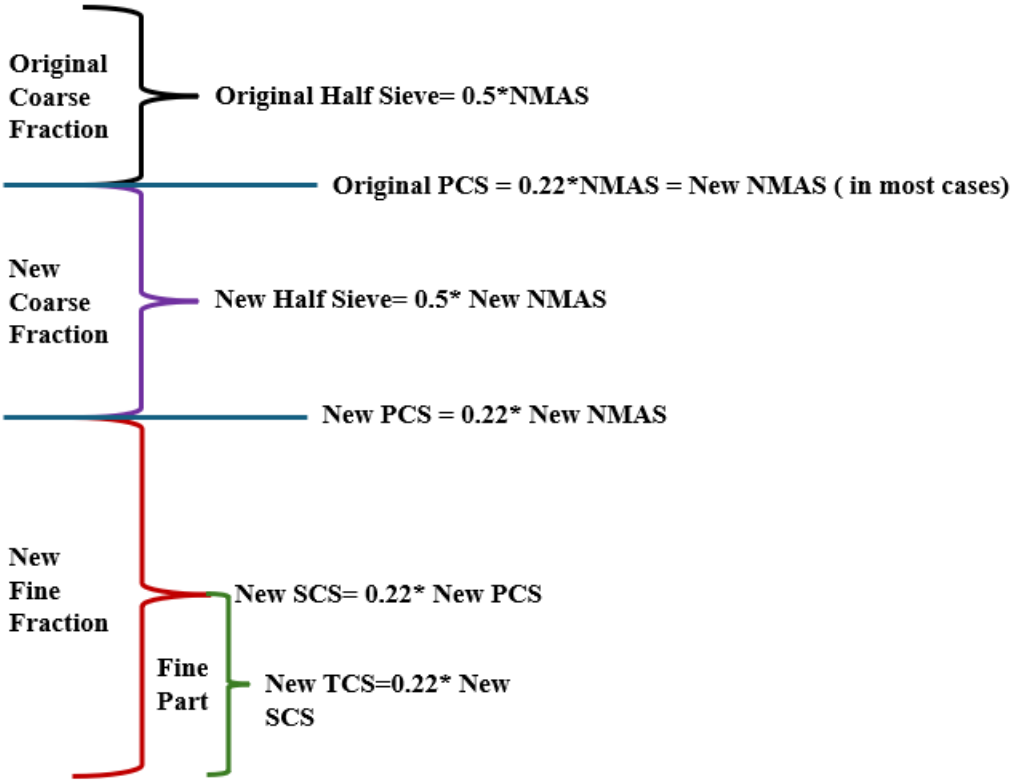

Figure 4. New Sieves of Fine-Graded Mix

Table 15. New Sieves and Bailey Ratios of Fine-Graded Mix [108]

Property	37.5 mm	25.0 mm	19.0 mm	12.5 mm	9.5 mm	4.75 mm
Original PCS	9.5 mm	4.75 mm	4.75 mm	2.36 mm	2.36 mm	1.18 mm
New Half Sieve	4.75 mm	2.36 mm	2.36 mm	1.18 mm	1.18 mm	0.600 mm
New PCS	2.36 mm	1.18 mm	1.18 mm	0.600 mm	0.600 mm	0.300 mm
New CA Ratio	$\frac{4.75-2.36}{9.5-4.75}$	$\frac{2.36-1.18}{4.75-2.36}$	$\frac{2.36-1.18}{4.75-2.36}$	$\frac{1.18-0.600}{2.36-1.18}$	$\frac{1.18-0.600}{2.36-1.18}$	$\frac{0.600-0.300}{1.18-0.600}$
New SCS	0.600 mm	0.300 mm	0.300 mm	0.150 mm	0.150 mm	0.075 mm
New FA_c Ratio	$\frac{0.600}{2.36}$	$\frac{0.300}{1.18}$	$\frac{0.300}{1.18}$	$\frac{0.150}{0.600}$	$\frac{0.150}{0.600}$	$\frac{0.075}{0.300}$
New TCS	0.150 mm	0.075 mm	0.075 mm	—	—	—
New FA_f Ratio	$\frac{0.150}{0.600}$	$\frac{0.075}{0.300}$	$\frac{0.075}{0.300}$	—	—	—
Note	Sieves shown refer to the % passing for the blend below the original PCS.					

CHAPTER 3 RESEARCH METHODOLOGY

This chapter presents the research methodology of the study, which includes three sections: 1) experimental design, 2) laboratory tests, and 3) analysis of test results. The experimental design was structured into two complementary parts, each addressing a distinct objective of the study. The first part evaluated the feasibility of modifications to the current P-401 gradation limits. The second part examined the effects of gradation adjustments on both volumetric properties and laboratory performance. The data generated in this phase were used to develop prediction models that quantify the influence of gradation changes on VMA and laboratory performance indicators, including rutting and intermediate- temperature cracking.

3.1 Experimental Design

The two-part experimental plan consisted of four sequential steps, as shown in **Figure 5**. In Step 1, existing P-401 mix designs representative of a range of aggregate sources and climatic conditions were selected, with particular focus on mixtures that were difficult to simultaneously satisfy both the current gradation limits and VMA requirements. For each selected mix design, a mixture complying with the current P-401 gradation specifications was developed and designated as In-Spec 1.

Step 2 involved adjusting aggregate gradation and differing between the two parts. In Part 1, an additional mixture, Out-Spec 1, was developed from In-Spec 1 by shifting the gradation outside the gradation limits while maintaining similar VMA (±0.5%) and binder content (±0.2%) at the same design air void of 3.5%. In Part 2, two additional gradations were developed: In-Spec 2, adjusted further within the gradation limits, and Out-Spec 2, adjusted further outside the gradation limits. Both were prepared with the same binder content as In-Spec 1 and Out-Spec 1, but with VMA values differing by approximately 1–2% from those mixtures.

In Step 3, all mixtures were subjected to a comprehensive laboratory test covering moisture susceptibility, rutting resistance, durability, permeability, cracking resistance, and FAA of the combined aggregate blend. Finally, in Step 4, the test results were compared between two gradations in Part 1 to revise the current P-401 gradation limits of 12.5 mm NMAS mixes and across all four gradations in Part 2 to evaluate the effects of gradation changes on volumetric properties and laboratory performance. Subsequently, prediction models were developed to quantify these relationships. Details of each step are described in the following sections.

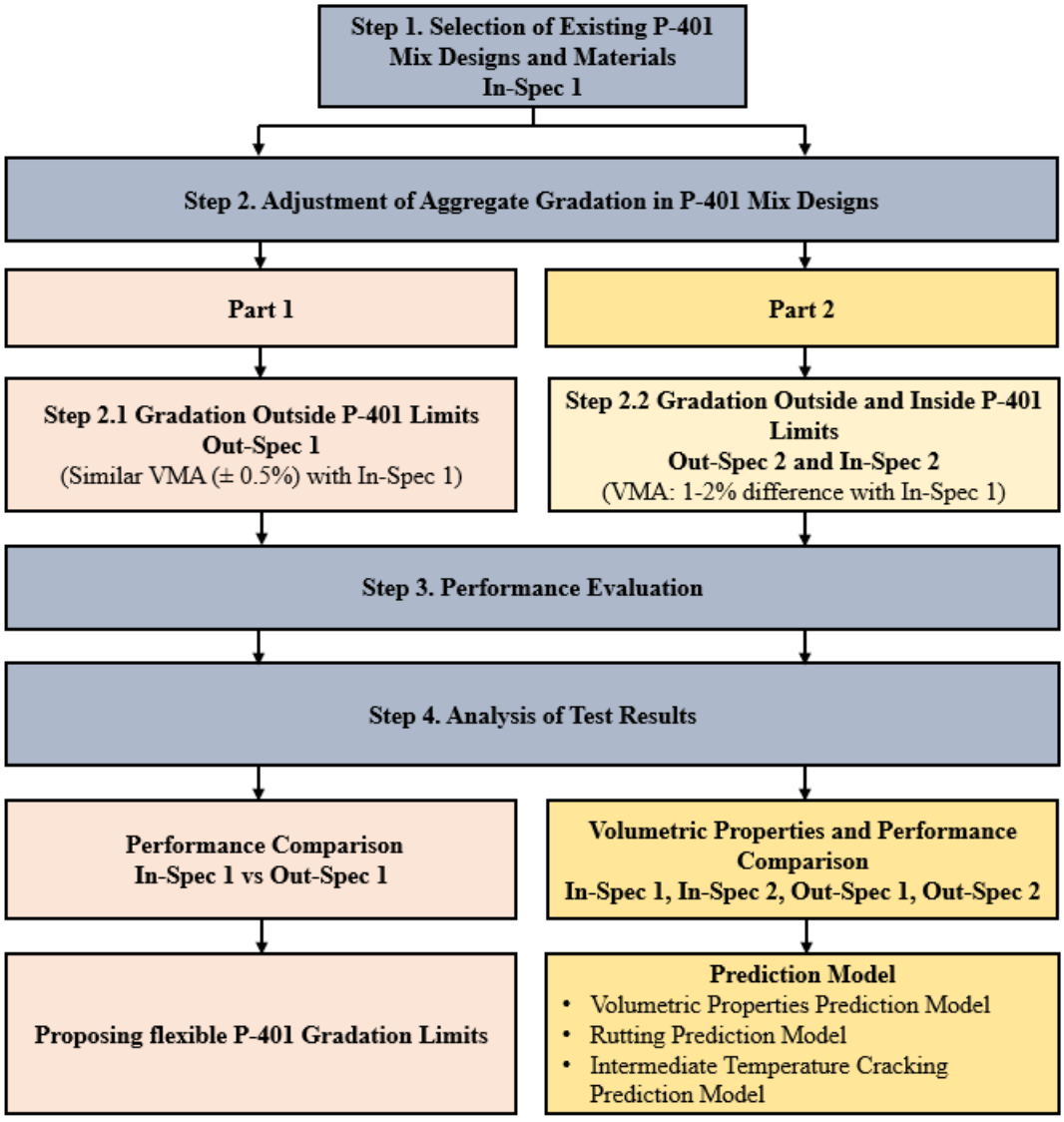

Figure 5. Experimental Plan

3.1.1 Selection of Existing P-401 Mix Designs and Materials

Step 1 involves selecting existing airfield mix designs for laboratory evaluation based on the following requirements:

- The mix designs had difficulty meeting the P-401 gradation limits and VMA requirements. Since the research objective required the selected gradation to be adjusted outside the limits while maintaining VMA, only mix designs whose gradations were positioned near the P-401 limits at certain sieve sizes or could feasibly be adjusted toward those limits were considered.
- The materials represented different aggregate sources used in four Long-term Pavement Performance (LTPP) climatic zones (dry/no-freeze, dry/freeze, wet/no-freeze, and wet/freeze).
- The P-401 specifications currently require 75 blows or gyrations for airports serving aircraft greater than 60,000 lbs. and 50 blows or gyrations for airports serving lighter aircraft. In the study, the selection of mix designs prioritized those with 75 blows or gyrations, as they presented stricter design challenges.

Table 16 summarizes the information of the 13 selected designs, including airport name, location, aggregate type, binder type, and gradation type. **Figure 6** presents the geographic locations of the 13 selected mix designs. Among 13 selected mix designs, three mix designs (KBVS, PSC, and RNO) are from the dry freeze climate zone, two mix designs (TUS and SAT) are from the dry no freeze climate zone, four mix designs (RVS, MDQ, PGD, and SAV) are from the wet no freeze climate zone, and four mix designs (SBN, IND, BDL, and CMH) are from the wet freeze climate zone. Note that all the mixes were designed with 12.5 NMAS, except for TUS, which was designed with 19 mm NMAS.

Table 16. P-401 Mix Design Summary

Airport	State	Aggregate Type	Gradation Type	Binder Type
San Antonio International Airport (SAT)	Texas	Basalt, Limestone, and Natural Sand	12.5 NMAS-Fine Graded	PG 76-22
Huntsville Executive Airport (MDQ)	Alabama	Limestone and Couch Sand	12.5 NMAS-Fine Graded	PG 67-22
Punta Gorda Airport (PGD)	Florida	Limestone and Sand	12.5 NMAS-Fine Graded	PG 76-22
Tulsa Riverside Airport (RVS)	Oklahoma	Limestone	12.5 NMAS-Fine Graded	PG 70-28
Mix Savannah/Hilton Head International Airport (SAV)	Georgia	Granite	12.5 NMAS-Fine Graded	PG 76-22
Mix Indiana International Airport (IND)	Indiana	Regular Limestone	12.5 NMAS-Fine Graded	PG 64-22
South Bend International Airport (SBN)	Indiana	Dolomitic Limestone	12.5 NMAS-Fine Graded	PG 76-28
John Glenn Columbus International Airport (CMH)	Ohio	Limestone and Natural Sand	12.5 NMAS-Fine Graded	PG 76-22
Tucson International Airport (TUS)	Arizona	Limestone and Quartzite	19.5 NMAS-Fine Graded	PG 76-22
Reno-Tahoe International Airport (RNO)	Nevada	Basalt	12.5 NMAS-Fine Graded	PG 64-28
Skagit Regional Airport (KBVS)	Washington	Limestone	12.5 NMAS-Fine Graded	PG 70-22
Tri-Cities Airport (PSC)	Washington	Limestone	12.5 NMAS-Fine Graded	PG 64-28NV
Bradley International Airport (BDL)	Connecticut	Traprock	12.5 NMAS-Fine Graded	PG 76-22

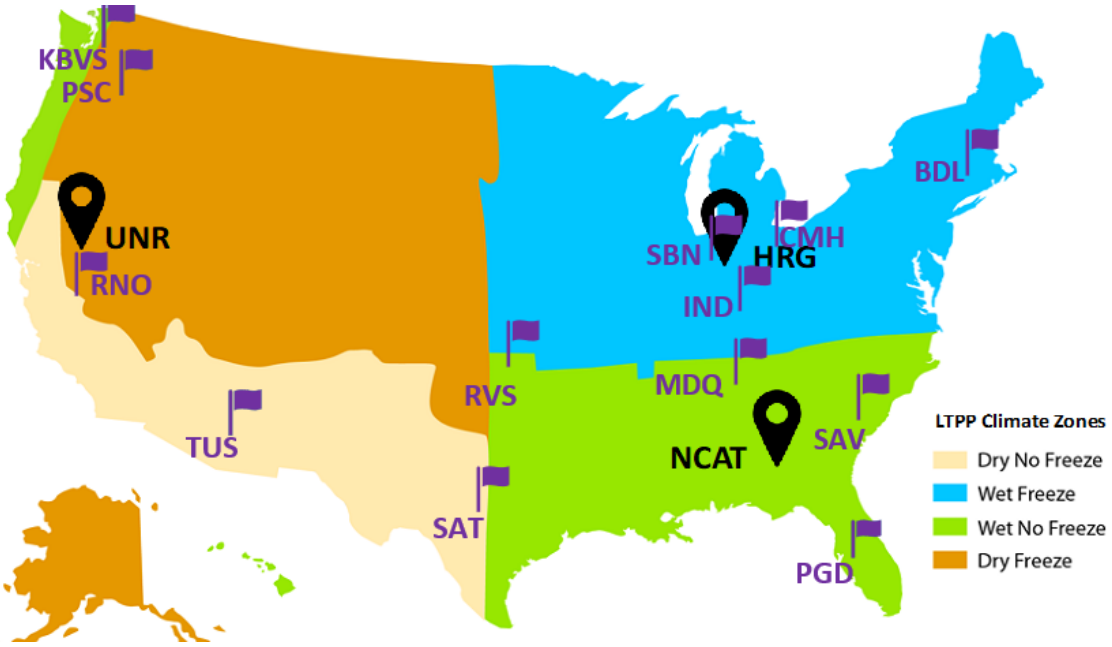

Figure 6. Geographic Locations of the Selected P-401 Mix Designs

3.1.2 Adjustment of Aggregate Gradation in P-401 Mix Designs

The second step involved adjusting aggregate gradations to produce the gradation designs required for each selected P-401 mix design.

In part 1, In-Spec 1 was adjusted using the Bailey method to produce a mixture with similar volumetric properties (i.e., VMA within ± 0.5% of In-Spec 1) but located outside the current P-401 gradation limits. This modified mixture was labeled as Out-Spec 1. Note that Out-Spec 1 was designed with a similar binder content (± 0.2%) and the same binder type as In-Spec 1 to eliminate the effect of binder on performance.

In part 2, two additional gradations were also developed. In-Spec 2 was designed with a gradation centered within the P-401 gradation limits, placing it farther within those limits than In-Spec 1. Out-Spec 2 was designed with a gradation positioned even further outside the gradation limits than Out-Spec 1. Both In-Spec 2 and Out-Spec 2 were designed to have a VMA difference of approximately 1–2% compared to In-Spec 1. Consequently, each mix design developed four

gradation designs: two inside the gradation limits (In-Spec 1 and In-Spec 2) and two outside the limits (Out-Spec 1 and Out-Spec 2). Again, all mix designs used a similar binder type and content to eliminate the effect of binder on performance.

Figure 7 and **8** illustrate the gradation adjustments of Out-Spec 1, Out-Spec 2, and In-Spec 2 when In-Spec 1 is close to the upper or lower P-401 limits, respectively. In both cases, whether In-Spec 1 was positioned near the upper or lower P-401 limits, moving the gradation outside the gradation limits resulted in these gradation curves being further away from the MDL. When the In-Spec 1 gradation was near the upper P-401 limit, the gradation became finer from In-Spec 2 to Out-Spec 2. Conversely, when In-Spec 1 gradation was near the lower limits, the gradation became coarser moving from In-Spec 2 to Out-Spec 2.

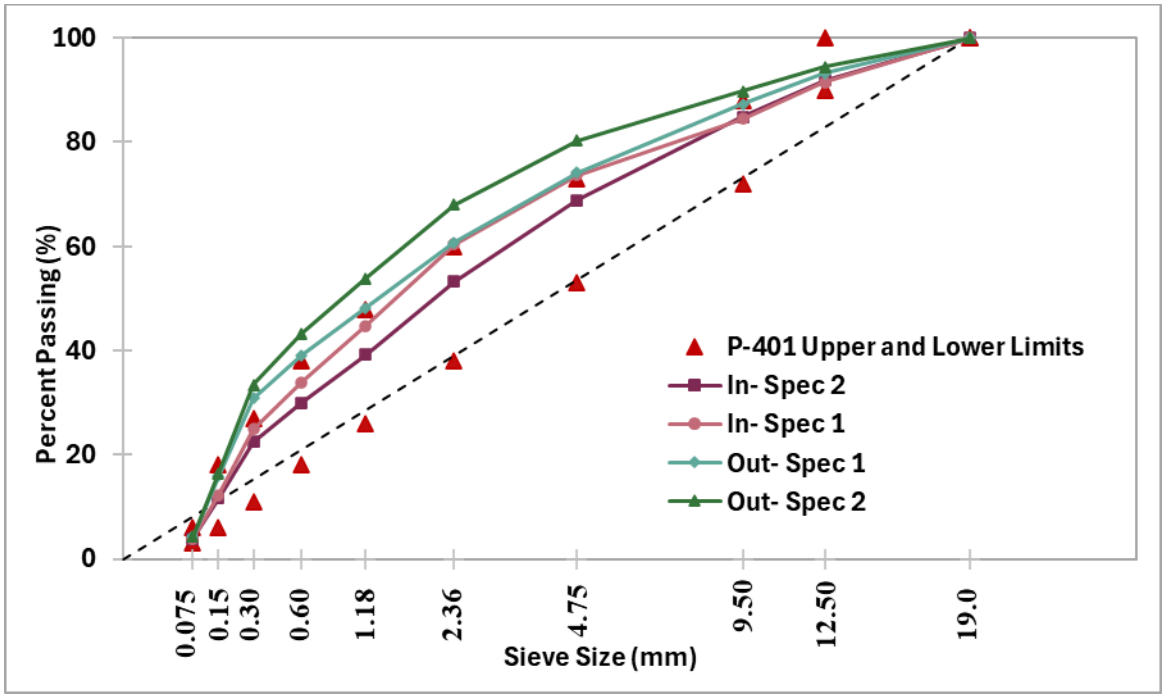

Figure 7. Illustration of Adjustments of Gradation Close to Upper Limits

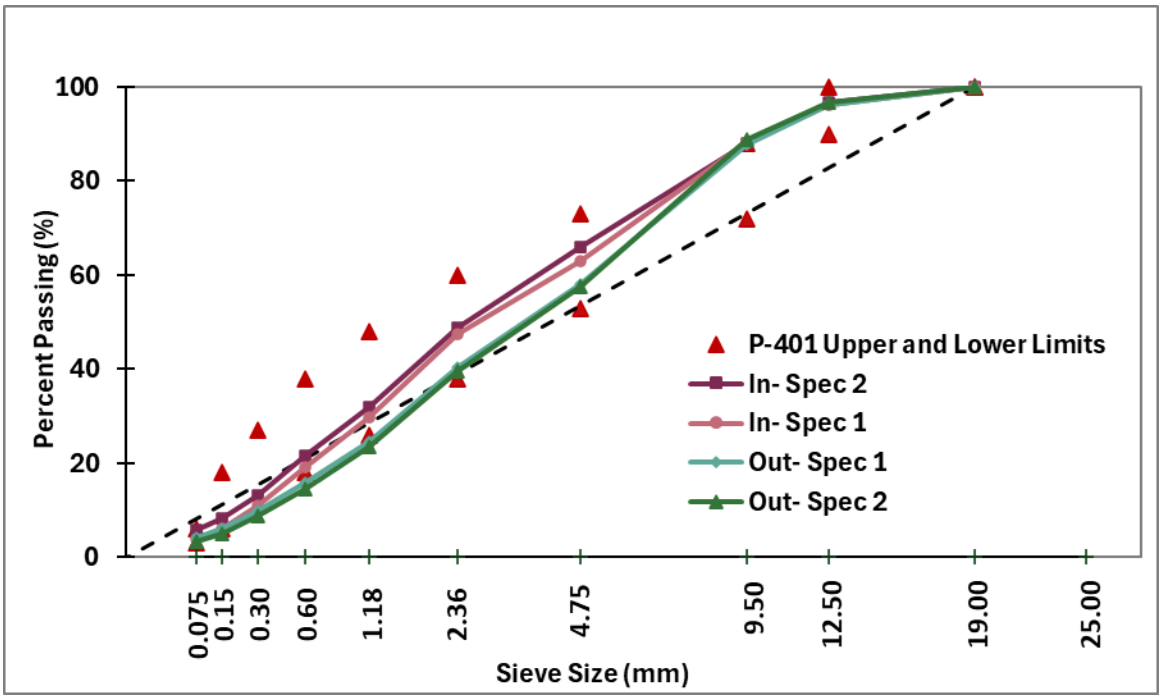

Figure 8. Illustration of Adjustments of Gradation Close to Lower Limits

3.1.3 Laboratory Performance and Aggregate Blend Evaluation

A comprehensive laboratory test plan was conducted to characterize the properties of both In-Spec and Out-Spec mixtures for each mix design. APA, HWTT, and TSR, as required in the current P-401 specification, were used to evaluate the rutting resistance and moisture susceptibility of the asphalt mixtures. Additional tests were performed to evaluate low-temperature cracking resistance (DCT), intermediate-temperature cracking resistance (Indirect Tensile Asphalt Cracking Test, IDEAL-CT, and Illinois Flexibility Index, I-FIT), durability (Cantabro test), and permeability (Florida permeability test). **Table 17** summarizes the laboratory mixture tests used in the study, including the test standard, test parameters, and specimen air voids. Note that the mixtures were short-term aged for 2 hours at the compaction temperature. Specimen air voids varied by performance test. Cracking test specimens were compacted to 5.0 ± 0.5% air voids to align with projects from the Airfield Asphalt Pavement Technology Program (AAPTP) [191]. Specimens for

rutting, durability, permeability, and moisture susceptibility tests were compacted to 7.0 ± 0.5% air voids, in accordance with specification recommendations to represent a worst-case scenario.

Table 17. Laboratory Performance Test

Performance	Mixture Test	Test Standard	Test Parameter	Air Void (%)
Rutting	APA	AASHTO T340 at 250 psi hose pressure	Rut Depth (mm)	7.0 ± 0.5
	HWTT	AASHTO T324		7.0 ± 0.5
Moisture	TSR	ASTM D4867	TSR Ratio	7.0 ± 0.5
Intermediate-Temperature Cracking	I-FIT	AASHTO T393	Flexibility Index (FI)	5.0 ± 0.5
	IDEAL-CT	ASTM D862	CT_{Index}	5.0 ± 0.5
Low-Temperature Cracking	DCT	ASTM D7313	Fracture Energy (J/m^2)	5.0 ± 0.5
Durability	Cantabro	AASHTO T401	Mass Loss (%)	7.0 ± 0.5
Permeability	Florida Permeability Test	FM 5-565	Permeability Coefficient ($10^{\wedge-5}$cm/s)	7.0 ± 0.5

3.2 Laboratory Tests

3.2.1 Moisture Susceptibility

The moisture resistance of the asphalt mixture was evaluated using the TSR test per ASTM D4867 [192], as shown in **Figure 9** . The test required six cylindrical specimens, typically 150 mm in

diameter and 95 ± 5 mm in height, compacted to an air void content of 7.0 ± 0.5%. These specimens were divided into two groups: the conditioned group (wet) and the unconditioned group (dry/control).

The unconditioned group was stored at room temperature without any additional treatment. Prior to testing, these samples were soaked in a water bath at 25 ± 1.0°C for 20 minutes. For the conditioned group, specimens were vacuum-saturated to achieve 70–80% saturation. In regions subjected to freeze–thaw conditions, freeze–thaw conditioning was applied. This involved a freeze cycle, where the saturated samples were frozen at -18°C (0°F) for 16 hours, followed by a thaw cycle in which the samples were submerged in a 60°C water bath for 24 hours. After thawing, the specimens were conditioned at 25 ± 1.0°C by soaking in a water bath for 1 hour before testing. Both conditioned and unconditioned specimens were tested for indirect tensile strength by applying a load diametrically across the specimen at a constant deformation rate of 50 mm/min until failure. The maximum load was recorded, and the tensile strength of each specimen was calculated using **Equation 9**. The TSR was calculated using **Equation 10** with a minimum threshold is 80%.

$$S_t = \frac{2P}{\pi t D} \quad (9)$$

where:

S_t= tensile strength (psi), P = maximum load, t = specimen thickness, D = specimen diameter.

$$TSR = \frac{\text{Average Strength of Conditioned Specimens}}{\text{Average Strength of Unconditioned Specimens}} * 100 \quad (10)$$

Figure 9. Tensile Strength Ratio (TSR)

3.2.2 Rutting Resistance

The rutting resistance of the asphalt mixtures was evaluated using either the APA, according to AASHTO T 340 [193] or the HWTT following AASHTO T324 [194].

For the APA test, six cylindrical specimens (150 mm in diameter and 75 ± 2 mm in height) were prepared for testing with a three-wheel APA device, as shown in **Figure 10**. These specimens were compacted to 7.0 ± 0.5% air voids. Prior to testing, the samples were conditioned at 64°C for a minimum of 6 hours and a maximum of 24 hours. During the test, a 250 lb wheel load with a hose pressure of 250 psi was applied. The test continued until 8,000 passes were completed or the rutting depth exceeded 14 mm. Rut depth progression was recorded using a Linear Variable Differential Transducer (LVDT). Lower rut depths indicate better rutting resistance. According to P-401 specifications, the allowable maximum rut depth is 10 mm after 4,000 passes.

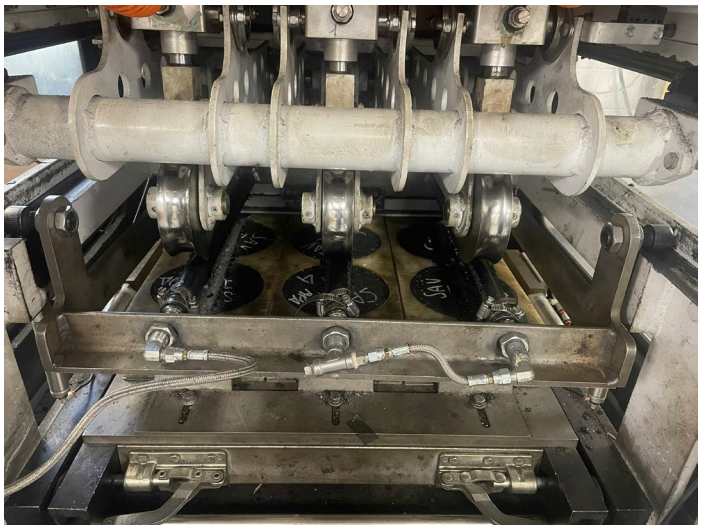

Figure 10. Asphalt Pavement Analyzer (APA)

Figure 11 presents the HWTT equipment used in this study. For the HWTT, cylindrical specimens were cut horizontally in half, and both halves were oriented face-to-face in the test mold, with the trimmed flat surfaces in contact to form a rectangular cross-section test specimen. Like the APA samples, HWTT specimens were also compacted to 7.0 ± 0.5% air voids, presenting the worst-case scenario. For each gradation design, two sets of samples were prepared and conditioned in water at 50°C for 45 minutes. After conditioning, a 158 ± 1.0 lb steel wheel reciprocated over the samples at a rate of 52 ± 2 passes per minute, continuing until 20,000 passes were completed. As per P-401 specifications, the maximum allowable rut depth is 10 mm after 20,000 passes.

Figure 11. Hamburg Wheel Tracking Test (HWTT)

3.2.3 Durability

The Cantabro test, following AASHTO T401 [195] was used to evaluate the raveling resistance and durability of the asphalt mixture, as shown in **Figure 12** . Cylindrical specimens of 150 mm in diameter and 115 ± 1 mm in height at a target air void of 7.0 ± 0.5% were used for evaluation. Each mix design required at least 3 replicates. Before testing, the specimens were conditioned in an environmental chamber at 25 °C for a minimum of 4 hours. The specimens were then placed in the drum without steel balls, rotating for 300 revolutions at a speed of 30- 33 rpm. After testing, the specimen was carefully removed. The Cantabro loss was calculated as the difference between the initial and final weight divided by the initial weight, as presented in **Equation 11**. A lower Cantabro loss value indicates better durability and raveling resistance.

Figure 12. Cantabro Test

$$\text{Cantabro Loss} = \frac{M_{initial} - M_{final}}{M_{initial}} \cdot 100 \quad (11)$$

Where:

$M_{initial}$ = Initial mass of the specimen, g.

M_{final} = Final mass of the specimen, g.

3.2.4 Permeability

The Florida permeability test was conducted per FM 5-565 [196] to assess the permeability of the asphalt mixtures using the falling-head permeability apparatus on 6-inch cylindrical specimens, as shown in **Figure 13**. Each test specimen was prepared by trimming approximately 1 inch from both the top and bottom surfaces of the compacted sample. For each gradation design, three replicate specimens were tested at an air void content of 7.0 ± 0.5% to represent a worst-case scenario for permeability. Prior to testing, the specimens were fully submerged in a water tank for at least 1 hour at ambient temperature to achieve saturation. Subsequently, the specimen was placed on top of the pedestal plate and assembled with the remaining part, including a graduated

cylindrical upper cap and a sealing tube with a membrane. The membrane was then inflated to seal the sides of the specimen throughout the whole test process. Next, water was added to the graduated cylinder to a level above the upper timing mark, and then allowed to flow through the saturated specimen, and the interval of time taken to reach a known change in the head was recorded. During the testing, the inflated latex membrane sealed the sides of the specimen, ensuring that water flow was restricted to the vertical direction only. The recorded time interval was used to calculate the coefficient of permeability (k) based on Darcy's law, as shown in **Equation 12**. The mixtures with higher k values had higher permeability than those with lower values, which represents a greater risk of water infiltration and is generally undesirable for pavement performance.

$$k = \frac{aL}{At} \cdot \ln\frac{h_1}{h_2} \cdot t_c \qquad (12)$$

Where,

k = coefficient of permeability, cm/s; a = inside cross-sectional area of the buret, cm^2

L = average thickness of the test specimen, cm; A = average cross-sectional area of the test specimen, cm^2; t = elapsed time between h_1 and h_2, s; h_1= initial head across the test specimen, cm; h_2= final head across the test specimen, cm; t_c= temperature correction for viscosity of water. A temperature of 20°C (68°F) is used as the standard.

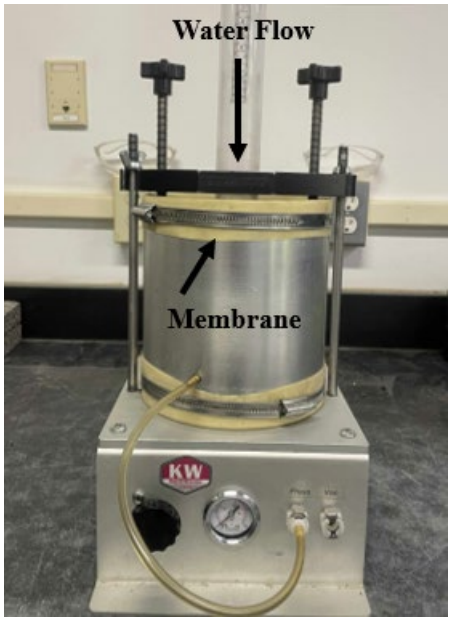

Figure 13. Florida Permeability Test

3.2.5 Intermediate-Temperature Cracking Resistance

The cracking resistance at an intermediate temperature was evaluated using two cracking tests: the IDEAL-CT test, following ASTM D8225 [197], as shown in **Figure 14** and the I-FIT following AASHTO T 393 [198] , as shown in **Figure 15.**

The IDEAL-CT is designed to characterize the cracking resistance of an asphalt mixture by calculating the Cracking Tolerance Index CT_{Index} which is based on load-displacement behavior under monotonic loading. The test was conducted using cylindrical specimens with a diameter of 150 mm and a height of 62 ± 2 mm, and compacted at a target air void of 5.0 ± 0.5%. A minimum of 4 replicate specimens were prepared for each gradation design. Before testing, the specimens were conditioned in an environmental chamber at 25 ± 1°C. A compressive load was then applied vertically along the diameter of the specimen at a constant rate of 50mm/min until complete failure occurred. The load versus displacement data were recorded during the test. The CT_{Index} was calculated using the **Equation 13** with a higher value indicating greater resistance to intermediate-temperature cracking.

$$CT_{Index} = \frac{G_f * |m| * D}{t * P_{max}} \quad (13)$$

Where:

G_f = Fracture Energy (Joules)

|m| = Absolute value of post-peak slope (N/mm)

D = Specimen diameter (mm)

t = Specimen thickness (mm)

P_{max} = Maximum load (N)

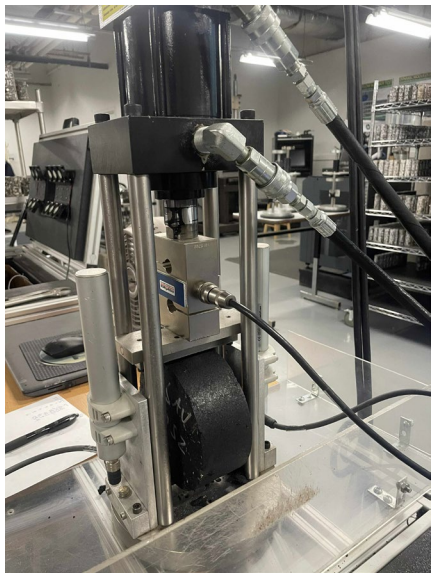

Figure 14. Indirect Tensile Asphalt Cracking Test (IDEAL-CT)

For the I-FIT test, a cylindrical asphalt specimen with a diameter of 150 mm and a height of 160± 1 mm was first cut in half horizontally to obtain two disc-shaped specimens with a thickness of 50 ± 1 mm. Each disc was then cut in half vertically along the diameter to produce two semi-circular bending (SCB) samples. A notch with a depth of 15 ± 0.5 mm was introduced at the center of the flat edge of each SCB sample to facilitate controlled crack initiation during testing. Similar to the IDEAL-CT test, the SCB sample was targeted at an air-void content of 5.0 ± 0.5%. A minimum of four replicates was used for each gradation design. Prior to testing, all SCB

specimens were conditioned at 25 ± 1°C for at least 2 hours. Each notched SCB sample was placed on a three-point bending fixture with a support span of 120 mm. A vertical load was applied at the midpoint of the curved face (directly above the notch) at a constant displacement rate of 50 mm/min until the specimen fractured. Load-displacement data were recorded continuously throughout the test. The Flexibility Index (FI) was calculated based on the fracture energy (G_f) and the post-peak slope (m) of the load-displacement curve using **Equation 14,** with a higher FI value, indicating greater flexibility and improved resistance to crack propagation

$$FI = \frac{G_f * A}{|m|} \qquad (14)$$

Where:

G_f = Fracture energy (J/mm^2)

|m| = Absolute value of post-peak slope (N/mm)

A = Calibration constant (empirically derived)

Figure 15. Illinois Flexibility Index Test (I-FIT)

3.2.6 Low-Temperature Cracking Resistance

The DCT test was conducted following ASTM D7313 [199] to evaluate the low-temperature cracking resistance of the asphalt mixture. A cylindrical specimen, 150mm in diameter and 160 ± 1 mm in height, was cut in half horizontally to obtain two disk-shaped specimens with a thickness of 50 ± 1 mm. A notch was cut at the center of one edge to control cracking initiation, and two loading holes were drilled near the top of the specimen to allow load application through steel pins. A minimum of 4 replicates was required for each gradation design. Similar to IDEAL-CT and I-FIT, the replicates were tested at an air void content of 5.0 ± 0.5%. **Figure 16** and **17** present the DCT test setup and the testing sample, respectively. Before testing, the specimen was conditioned for 2-8 hours at a temperature of 10°C above the low-temperature performance grade of the asphalt binder. The specimen was then mounted in a testing fixture, and a monotonic tensile load was applied via the steel pin at a constant displacement rate of 0.017 mm/s. Load and Crack Mouth Opening Displacement (CMOD) was recorded during the test. The fracture energy is calculated by dividing the total work of fracture (area under the load-CMOD curve) by the ligament area (area of uncracked section), as shown in **Equation 15**. Higher fracture energy values indicate improved resistance to low-temperature cracking.

$$G_f = \frac{AREA}{B(W-a)} \quad (15)$$

where:

G_f= fracture energy (J/m²),

AREA = area under load-$CMOD_{fit}$ curve (N·m)

B = specimen thickness (m), and

W – a = initial ligament length (m).

Figure 16. Disk-Shaped Compact Tension (DCT) Test

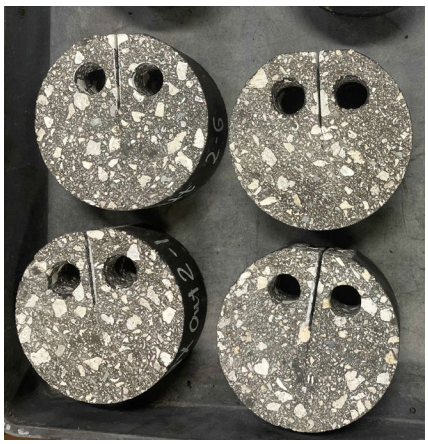

Figure 17. DCT Sample

3.2.7 Fine Aggregate Angularity Test

The FAA test was conducted in accordance with AASHTO T 304-22 [200] to measure the uncompacted void content of the fine aggregate fraction. For this study, testing was performed on the blended gradation of each mix design rather than on individual aggregate stockpiles. Method A was selected because it measures the void content under standardized conditions that reflect the particle shape and texture of a fine aggregate in the blend. First, the aggregate blend was batched based on individual aggregate stockpile percentage and then washed over a No.200 sieve to remove dust. The washed blend was dried and sieved into different size fractions following AASHTO T27 [201]. Four size fractions were collected, and the mass of each fraction was as follows:

- 2.36 mm (No. 8) to 1.18 mm (No. 16): 44 g
- 1.18 mm (No. 16) to 600 µm (No. 30): 57 g
- 600 µm (No. 30) to 300 µm (No. 50): 72 g
- 300 µm (No. 50) to 150 µm (No. 100): 17 g

The four size fractions were combined and thoroughly mixed to create a sample of 190g ± 0.2g. The mixture was then allowed to flow freely through a funnel from a fixed height into a calibrated 100 mL cylindrical measure without any vibration or compaction, as shown in **Figure 18**. After the measure was filled, excess material was struck off in a single smooth motion using a spatula with a straight edge held horizontally. The mass of the filled measure was recorded, and the procedure was repeated for a second measurement. The average of the two test results was reported. The uncompacted void content, expressed as a percentage, was calculated using **Equation 16**. Blends with higher uncompacted void content are generally more angular and rough-textured, while those with lower values indicate more rounded and smoother particles.

$$\mathrm{U} = \frac{(\mathrm{V} - \mathrm{F/G})}{\mathrm{V}} \times 100 \qquad (16)$$

where:

- U= Uncompacted voids, % (reported as FAA value)
- V= Volume of the cylindrical measure, mL
- F= Net mass of fine aggregate in the measure, g
- G= Bulk dry specific gravity of the blended fine aggregate fraction

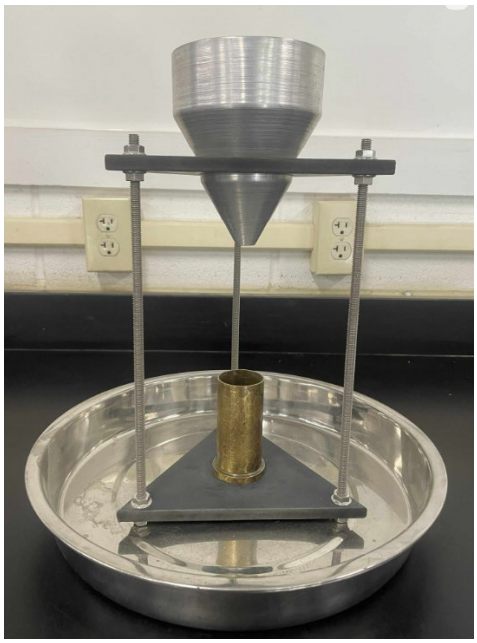

Figure 18. Fine Aggregate Angularity (FAA) Test

3.3 Analysis of Test Results

For Part 1 of the experimental design, the laboratory test results of In-Spec 1 and Out-Spec 1 were compared to assess the effect of gradation variation while maintaining volumetric properties. If the Out-Spec 1 provides performance similar to or better than the In-Spec 1, the P-401 gradation limits of 12.5 mm NMAS mixes would be revised to be more flexible, while maintaining volumetric properties and performance consistent with the current P-401 specification requirements. More detail was presented in Chapter 4.

For Part 2 of the experimental design, the volumetric properties and performance were compared among four gradations (In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2) of each mix design and across the full dataset to identify consistent trends. The analysis explored the impacts of gradation changes on volumetric properties and laboratory performance. Based on these results, prediction models were developed to quantitatively characterize the effects of gradation change on mixture behavior. It should be noted that these models were not intended to directly predict the

absolute values of VMA or performance for a given mixture. Instead, the models were formulated to predict changes in volumetric properties and performance of a new mixture relative to existing mixtures, using differences in gradation and aggregate properties between them as inputs. These included volumetric properties and performance-prediction models focused on rutting and intermediate-temperature cracking. More detail was presented in Chapter 5.

CHAPTER 4 PROPOSED REVISIONS TO P-401 GRADATION LIMITS

This chapter presents the detailed experimental design for Part 1 and the analysis of all laboratory test results for In-Spec 1 and Out-Spec 1. Based on the analysis results, a more flexible P-401 gradation limit was proposed. The chapter is organized into four sections: (1) experimental design, (2) In-Spec 1 and Out-Spec 1 design summary, (3) test results and data analysis, and (4) revising P-401 gradation limits.

4.1 Experimental Design

Figure 19 presents the experimental design for Part 1, which includes four steps. The first step focused on selecting and redesigning existing P-401 mix designs. After selecting the 13 airport mix designs, the raw materials for each design were sampled and characterized. Typically, the measured aggregate properties (gradation and G_{sb}) did not align with the Job Mix Formula (JMF) due to stockpile variability over time. Therefore, a baseline mixture, referred to as In-Spec 1, was developed for each selected mix design using measured aggregate properties and complying with the current P-401 specifications. If the JMF gradation was already positioned near the upper or lower limits, the In-Spec 1 gradation was maintained near those same limits. However, when the JMF gradation was centered within the limits, In-Spec 1 was intentionally adjusted toward either the upper or lower limit. This adjustment was guided by engineering judgment, with the primary objective of avoiding significant deviations in VMA from the JMF design. Positioning the gradation closer to the gradation limits allows for greater flexibility in future designs to move it outside the limits without affecting volumetric properties. Note that, to maintain confidentiality and present the results objectively without directly associating performance outcomes with specific airports, the airport mixes were anonymized and labeled as Mix 1 through Mix 13. Two of the

selected mix designs (Mix 1 and Mix 2) were designed close to the upper limits, while the remaining 11 mix designs (Mix 3 to Mix 13) were designed close to the lower limits.

Step 2 focused on designing Out-Spec 1. First, the preliminary Out-Spec 1 was designed by adjusting the In-Spec 1 to be outside the P-401 gradation limits while maintaining volumetric properties using the Bailey method. For the upper limit mix designs, the gradation adjustments targeted the upper limits on all sieves (except the No. 200 sieve). For the lower limit mix designs, the adjustment focused on the lower limits at No. 16 and smaller sizes. This approach avoided coarsening the gradation on the No. 8 sieve and larger sieve sizes, which prevented an increase in interconnected air voids and, therefore, helped to reduce mixture permeability, improving durability [11]. Then, the volumetric properties and rutting resistance of the preliminary Out-Spec 1 were verified through laboratory measurements. The design process was repeated until the measured VMA was within ±0.5 % of the corresponding In-Spec 1 and met the P-401 rutting requirement. Note that the same binder content (± 0.2%) was used to isolate the effect of gradation change on the laboratory performance. Once these criteria were satisfied, the preliminary design was accepted as the final Out-Spec 1.

Step 3 was conducted to characterize the properties of In-Spec 1 and Out-Spec 1. Initially, the full testing matrix (shown in **Table 17**) was intended to be applied to all mix designs. However, after analyzing the results of six mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11), no significant differences were observed in TSR, permeability, or Cantabro tests. As a result, these tests were removed from the testing plan. The reduced testing matrix included only cracking tests (IDEAL-CT, I-FIT, and DCT) and rutting tests (APA/HWTT) and applied to the remaining mix designs **(**Mix 2, Mix 4, Mix 5, Mix 7, Mix 8, Mix 12, and Mix 13).

Step 4 focused on revising the current P-401 gradation limits. The laboratory results of In-Spec 1 and Out-Spec 1 were compared to assess how shifting the gradation outside the current P-401 gradation limits while maintaining volumetric properties and binder content affects mixture performance. As both mixtures were designed with comparable VMA and binder content, any observed performance differences should be attributed solely to differences in gradation. A Student's T-test was conducted at a 0.05 significance level to determine whether these differences were statistically significant. Additionally, the difference between In-Spec 1 and Out-Spec 1 was compared with the allowable difference between the two test results, known as the single-operator d2s limit [202], specified by the relevant testing standard or research. However, the d2s limit was established from identical mix designs and therefore represents only the variability attributable to the test method. When applied to mixtures of different gradation designs, the observed difference reflects both test method variability and design differences. Consequently, this comparison indicates only whether the difference between In-Spec 1 and Out-Spec 1 is within the range of testing variability.

In the study, 12 mix designs at 12.5 NMAS, including two upper limit mixes and 10 lower limit mixes, were used for gradation revision. If the findings support increased flexibility, the gradation of Out-Spec 2, with its sieves and the percentage passing outside the gradation limits, would be used to modify the P-401 upper and lower limits.

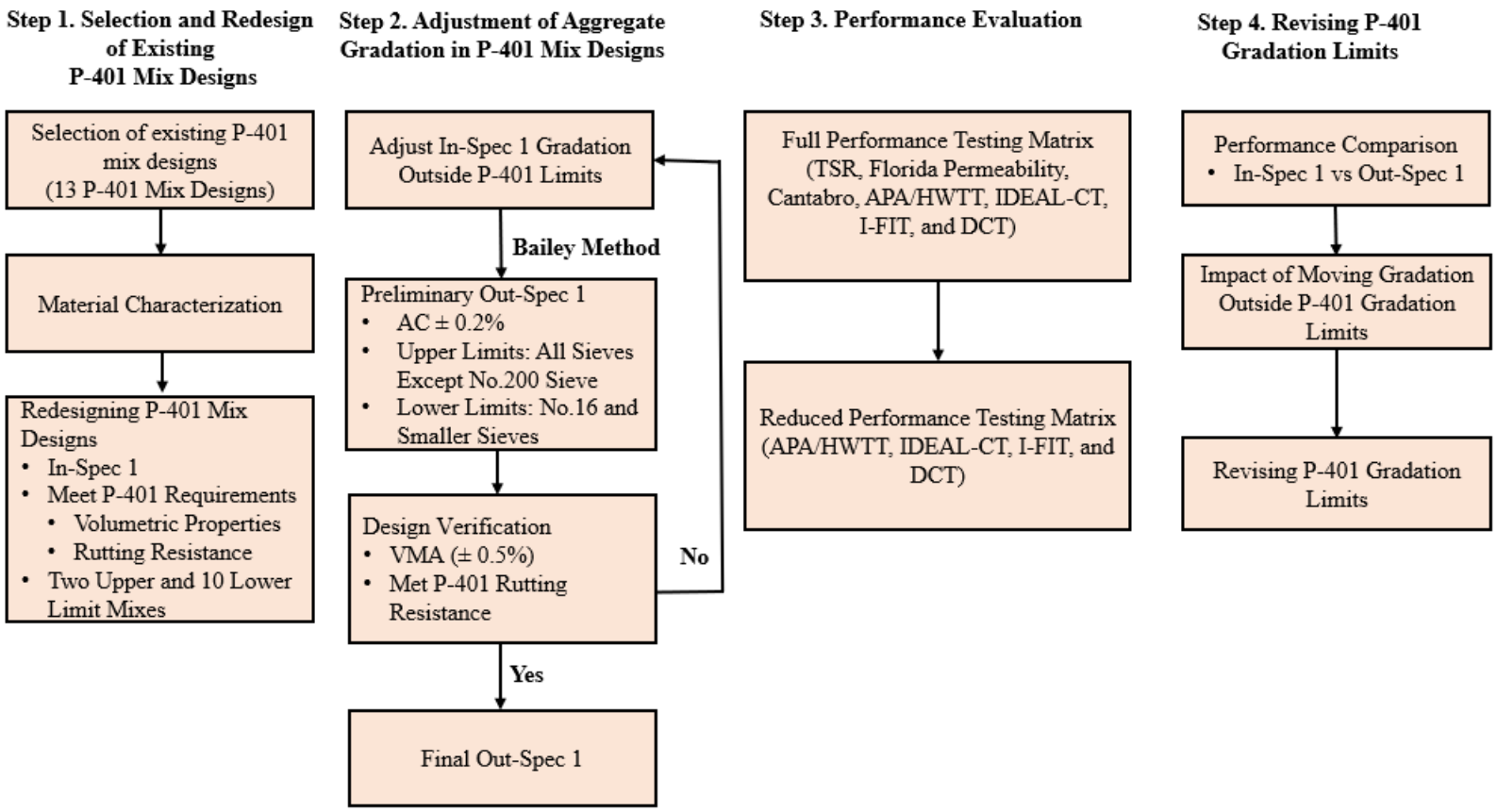

Figure 19. Experimental Design (Part 1) for Revising P-401 Gradation Limits

4.2 In-Spec 1 and Out-Spec 1 Design Summary

Table 18 summarizes the gradation design, FAA, and volumetric properties of In-Spec 1 and Out-Spec 1 of 13 mix designs. Sieve sizes of Out-Spec 1 that fall outside the P-401 gradation limits were highlighted. Across all mixture pairs, binder content was maintained within ±0.2% and VMA was maintained within ±0.5%, which aligns with the volumetric control strategy described in Section 4.1. All mixtures met the minimum VMA requirement of 15% for 12.5 mm NMAS and 14% for 19 mm NMAS. Additionally, the effective binder contents (P_{be}) were comparable between the paired mixtures. FAA values did not show systematic differences between In-Spec 1 and Out-Spec 1, indicating that the gradation adjustments did not substantially alter FAA characteristics.

Table 18. Design Summary of In-Spec 1 and Out-Spec 1

Mix ID	1		2		3		4		5		Gradation Limits and Volumetric Requirements	6		Gradation Limits and Volumetric Requirements
Sieve	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1		IS 1	OS 1	
3/4‘’	100	100	100	100	100	100	100	100	100	100	100	98	98	90-100
1/2"	91	93	98	98	97	93	96	96	92	92	90-100	88	89	68-88
3/8"	85	87	88	88	88	77	88	88	84	84	72-88	81	80	60-82
#4	73	74	67	66	60	56	63	58	69	66	53-73	59	53	45-67
#8	60	61	55	55	45	46	47	40	47	43	38-60	39	34	32-54
#16	45	48	45	46	30	28	30	25	30	27	26-48	24	21	22-44
#30	34	39	37	39	21	18	19	16	19	17	18-38	16	14	15-35
#50	25	31	27	30	13	10	11	10	11	9	11-27	10	8	9-25
#100	12	16	14	16	7	6	6	6	7	6	6-18	7	6	6-18
#200	3.9	4.3	6.0	5.9	4.6	3.7	3.9	4.3	5.4	4.6	3.0-6.0	6.9	4.7	3.0-6.0
FAA (%)	46.2	47.2	47.4	47.3	45.4	46.2	44.0	44.4	44.1	44.4		45.3	44.2	
AC (%)	7.1	7.2	6.2	6.0	5.9	6.1	5.7	5.7	6.8	6.9	5.0-7.5	5.3	5.3	4.5-7
Va (%)	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5	3.5
VMA (%)	15.5	15.6	16.7	16.5	16.2	16.7	15.2	15.1	16.5	16.8	≥ 15	14.6	14.6	≥ 14
Pbe (%)	5.4	5.5	5.7	5.6	5.0	5.1	5.0	5.0	5.8	5.9		4.9	4.8	

*Notes

IS 1: In-Spec 1; OS 1: Out-Spec 1

Table 18. Design Summary of In-Spec 1 and Out-Spec 1 (Continued)

Mix ID	7		8		9		10		11		12		13		Gradation Limits and Volumetric Requirements
Sieve	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	IS 1	OS 1	
3/4‘’	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
1/2"	98	95	94	94	97	90	93	90	98	98	94	93	93	93	90-100
3/8"	88	84	77	74	88	78	87	82	88	88	74	71	87	86	72-88
#4	58	58	53	53	65	60	69	62	61	57	55	54	66	61	53-73
#8	38	38	40	39	43	41	49	42	42	38	39	38	43	38	38-60
#16	26	25	29	23	27	25	32	26	28	25	26	25	30	25	26-48
#30	18	17	21	16	18	16	21	16	18	16	18	17	20	17	18-38
#50	11	10	14	11	12	10	13	10	11	10	13	12	13	12	11-27
#100	7	7	9	9	8	7	6	5	6	6	8	8	9	9	6-18
#200	5.9	5.5	6.3	6.3	5.8	4.9	4.0	3.3	4.6	4.3	4.0	4.5	6.3	6.1	3.0-6.0
FAA	44.0	44.0	45.2	47.4	47.2	45.8	45.0	46.0	45.0	45.0	48.3	49.5	45.0	45.0	
AC (%)	6.3	6.5	6.0	6.0	6.2	6.1	5.8	5.7	5.7	5.8	5.9	5.8	6.1	6.1	5.0-7.5
Va (%)	3.5	3.6	3.5	3.6	3.5	3.5	3.5	3.5	3.6	3.6	3.5	3.5	3.5	3.5	3.5
VMA (%)	15.3	15.2	15.3	15.3	15.4	15.4	15.5	15.5	15.1	15.2	17.0	17.0	15.7	15.8	≥ 15
Pbe (%)	5.1	5.1	5.0	4.9	4.9	4.9	5.2	5.2	4.9	4.9	5.4	5.5	5.2	5.3	

*Notes

IS 1: In-Spec 1; OS 1: Out-Spec 1

Figure 20, 21 and **22** summarize the binder content, VMA, and P_{be} of In-Spec 1 and Out-Spec 1. The data points aligned closely along the equality line, indicating that the binder content and volumetric properties are similar between the two gradation designs.

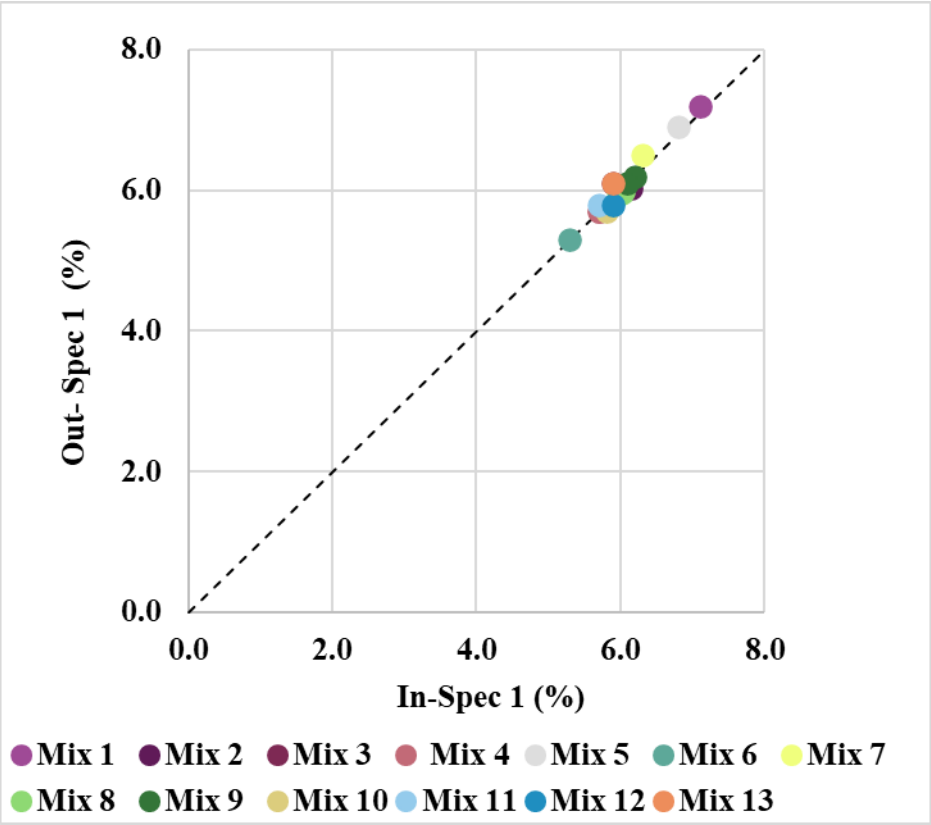

Figure 20. Asphalt Contents of In-Spec 1 and Out-Spec 1

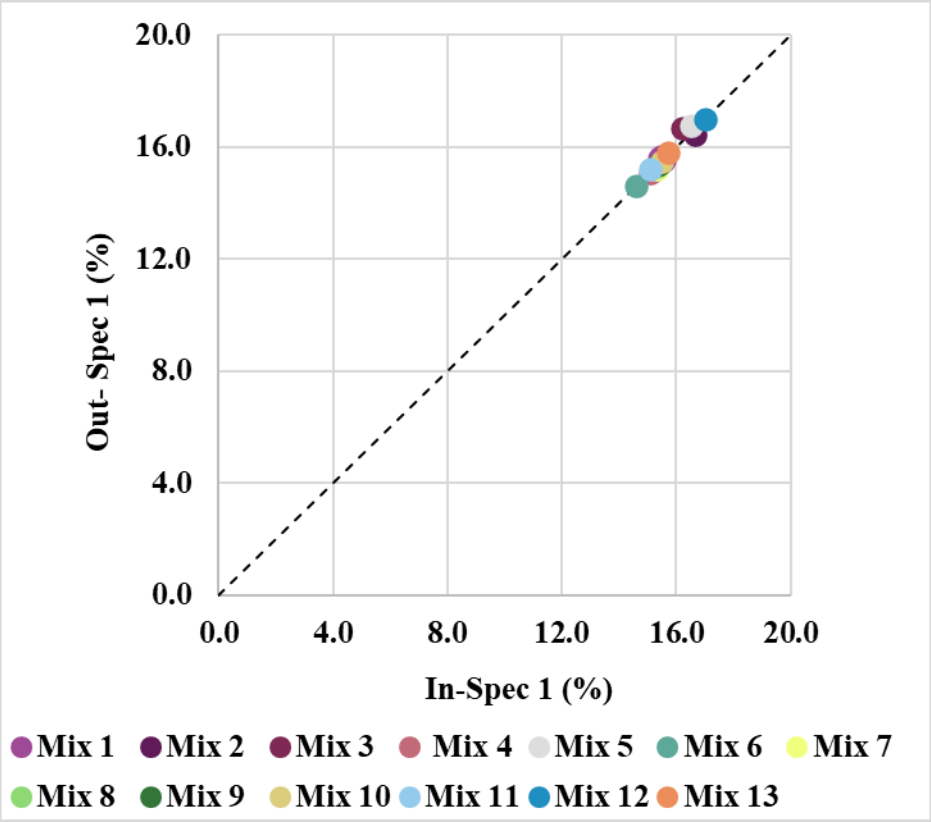

Figure 21. VMA of In-Spec 1 and Out-Spec 1

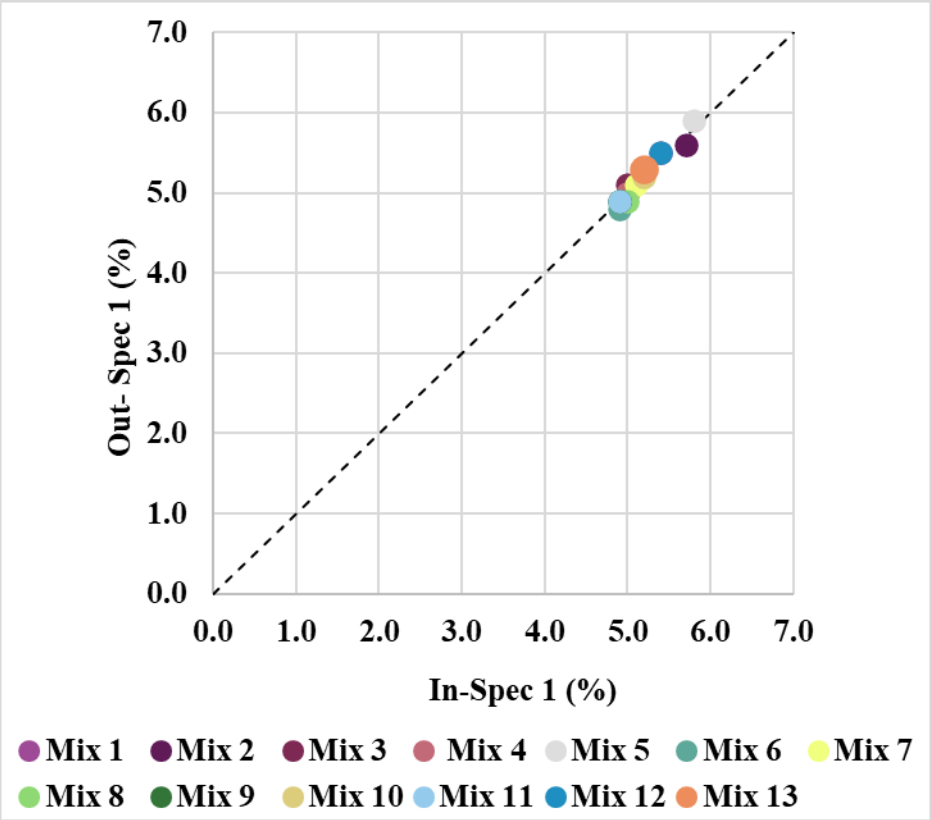

Figure 22. P_{be} of In-Spec 1 and Out-Spec 1

4.3 Test Results and Data Analysis

This section summarizes TSR, Florida permeability, and Cantabro loss results for six mix designs, as well as APA, HWTT, IDEAL-CT, I-FIT, and DCT results for 13 mix designs. Test results were presented using plot charts with error bars representing ±1 standard deviation. Student's t-test at a 0.05 significance level and d2s limit were conducted to assess if the difference between In-Spec 1 and Out-Spec 1 is statistically and practically significant. The comparison serves as the basis for evaluating potential revisions to P-401 gradation limits.

4.3.1 Moisture Susceptibility Evaluation

Figure 23 presents the TSR test results of In-Spec 1 and Out-Spec 1 for six mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11), and **Table 19** summarizes the TSR difference between In-Spec 1 and Out-Spec 1. All In-Spec 1 and Out-Spec 1 exceeded the minimum TSR threshold of

0.8, indicating sufficient moisture resistance required by P-401 specification. Since a single TSR value was determined for each mixture, the differences between In-Spec 1 and Out-Spec 1 were evaluated against the single operator d2s limit (d2s = 0.093 [203]). For five of the six mix designs, the TSR differences were smaller than the d2s limit. Only Mix 9 showed a difference exceeding the d2s limit. Since all mixes were fine-graded, and moisture susceptibility is influenced more by binder and aggregate properties than by minor gradation adjustments [204, 205], the outcome was expected. Consequently, the TSR test was discontinued for the remaining mix designs.

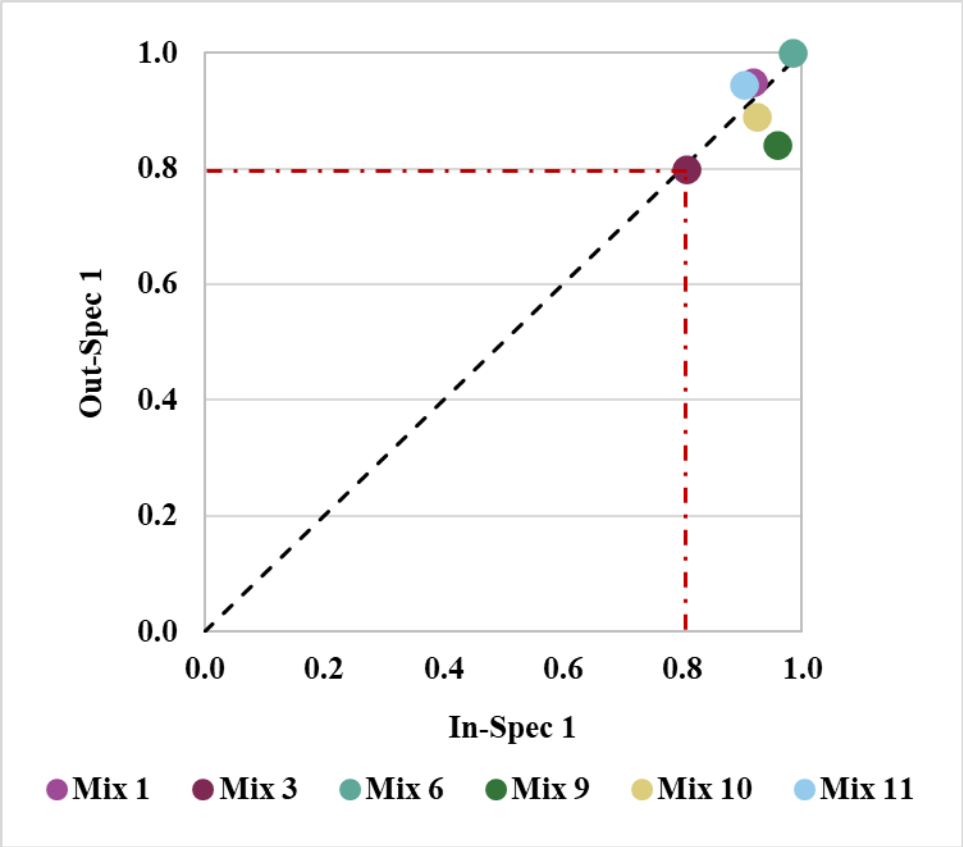

Figure 23. TSR Test Results of In-Spec 1 and Out-Spec 1

Table 19. TSR Difference Results

Mix ID	d2s = 0.093	Exceeding the Single Operator d2s Limit?
In-Spec 1 → Out- Spec 1 (Upper Gradation Limits)		
Mix 1	0.03	No
In-Spec 1→ Out- Spec 1 (Lower Gradation Limits)		
Mix 3	-0.01	No
Mix 6	0.05	No
Mix 9	-0.12	Yes
Mix 10	-0.03	No
Mix 11	0.04	No

4.3.2 Permeability Evaluation

Figure 24 presents the permeability test results for In-Spec 1 and Out-Spec 1 of six mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11), and **Table 20** presents the corresponding p-values. Out-Spec 1 generally exhibited higher permeability than In-Spec 1. Four of the six mix designs indicated no significant difference in permeability between Out-Spec 1 and In-Spec 1 ($p > 0.05$). Two other mixes (Mixes 9 and 10) showed a statistically significant difference ($p < 0.05$). However, the permeability coefficients of all mixtures were less than 125×10^{-5} cm/s, confirming that the mixtures were impermeable [156, 206]. These results indicate that gradation shifts outside the current P-401 gradation limits did not cause permeability issues. Therefore, the permeability was removed for the remaining mix designs.

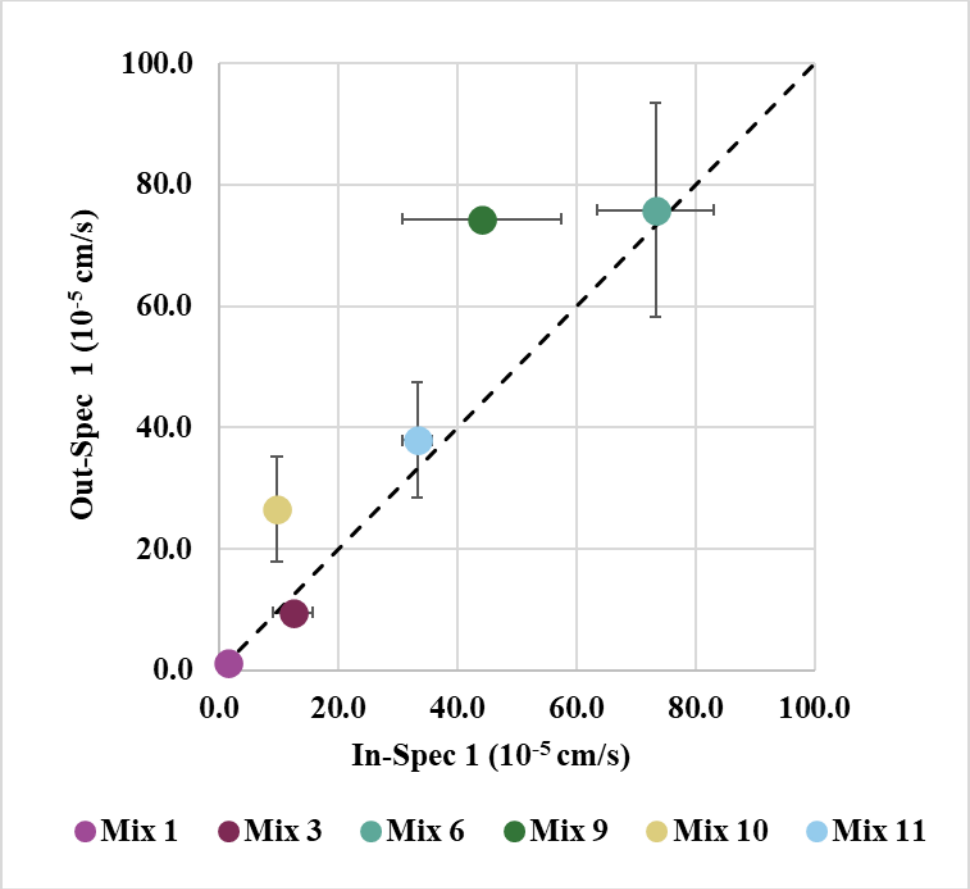

Figure 24. Permeability Test Results of In-Spec 1 and Out-Spec 1

Table 20. P-Values of Permeability Test Results

Mix ID	Observed Difference	P- Value	Statistical Significance?
	In-Spec 1→ Out- Spec 1 (Upper Gradation Limits)		
Mix 1	-0.3	0.30	No
	In-Spec 1→ Out- Spec 1 (Lower Gradation Limits)		
Mix 3	-2.8	0.27	No
Mix 6	2.5	0.88	No
Mix 9	30.3	0.01	Yes
Mix 10	16.9	0.00	Yes
Mix 11	4.6	0.19	No

4.3.3 Durability (Cantabro Mass Loss)

Figure 25 presents the Cantabro mass loss for In-Spec 1 and Out-Spec 1 of six mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11) and **Table 21** presents the corresponding p-values.

The mass loss of all mix designs was less than the maximum threshold of 6.3% used by the Virginia Department of Transportation (VDOT) [207]. Among six mix designs, five indicated no statistically significant difference between In-Spec 1 and Out-Spec 1 ($p > 0.05$). There was an exception for Mix 9; however, the difference was only 0.1%, which was considered not a practical difference. These results indicate that adjusting gradation outside the gradation limits when controlled volumetric properties did not adversely affect mixture durability. Based on these results, the Cantabro test was not performed for the remaining mix designs.

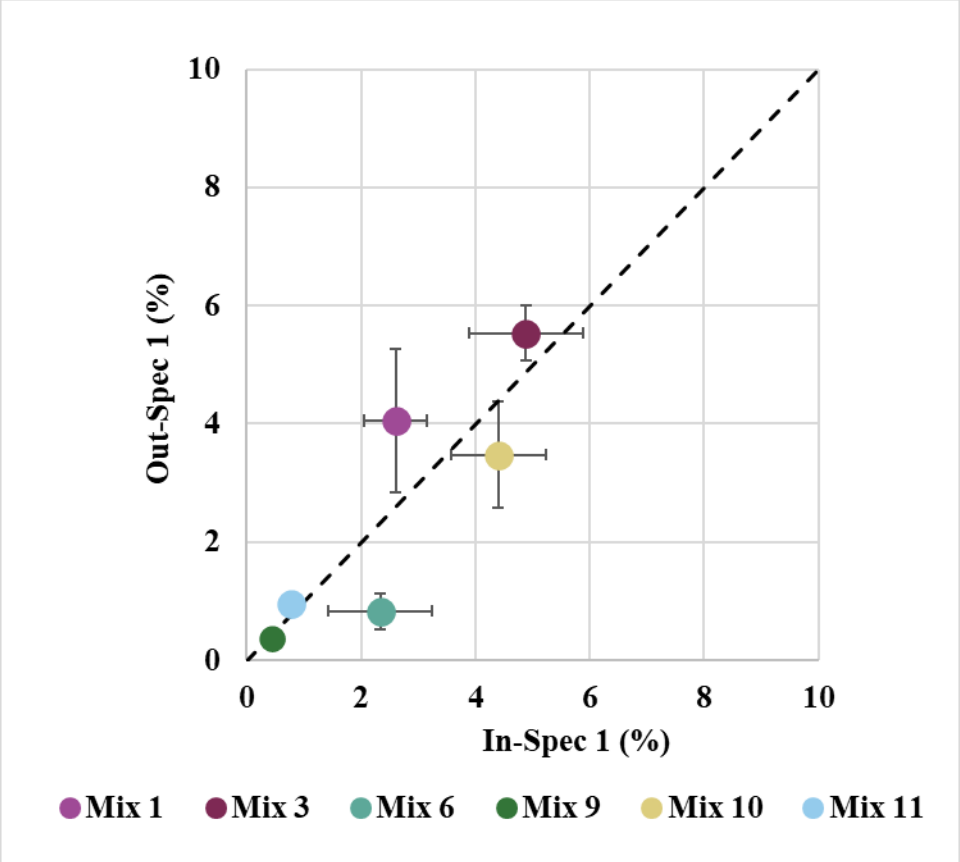

Figure 25. Cantabro Mass Loss of In-Spec 1 and Out-Spec 1

Table 21 P-Values of Cantabro Mass Loss

Mix ID	Observed Difference	P-Value	Statistical Significance?
In-Spec 1 → Out- Spec 1 (Upper Gradation Limits)			
Mix 1	1.4	0.16	No
In-Spec 1→ Out- Spec 1 (Lower Gradation Limits)			
Mix 3	0.7	0.22	No
Mix 6	-1.3	0.11	No
Mix 9	-0.1	0.02	Yes
Mix 10	-0.9	0.25	No
Mix 11	0.2	0.30	No

4.3.4 Rutting Resistance

Figure 26 presents the APA/HWTT test results for In-Spec 1 and Out-Spec 1 of 13 mix designs, and **Table 22** presents the corresponding p-values. Nine mix designs were tested using the APA test, while four were tested using the HWTT. All Out-Spec 1 had the rut depth less than 10 mm after 4000 passes in APA or 20000 passes in the HWTT, meeting the rutting resistance requirement in P-401. Eight of 13 mix designs had lower rut depths in Out-Spec 1 than in In-Spec 1, confirming improved rutting resistance. However, 12 of the 13 mix designs indicated no statistically significant differences in rut depth ($p > 0.05$). There was an exception for Mix 9; however, the difference was 0.7 mm, which was not considered a practical difference. In summary, the rutting resistance of Out-Spec 1 did not statistically or practically differ from that of In-Spec 1, indicating that the gradation adjustments when controlled volumetric properties did not adversely affect rutting resistance.

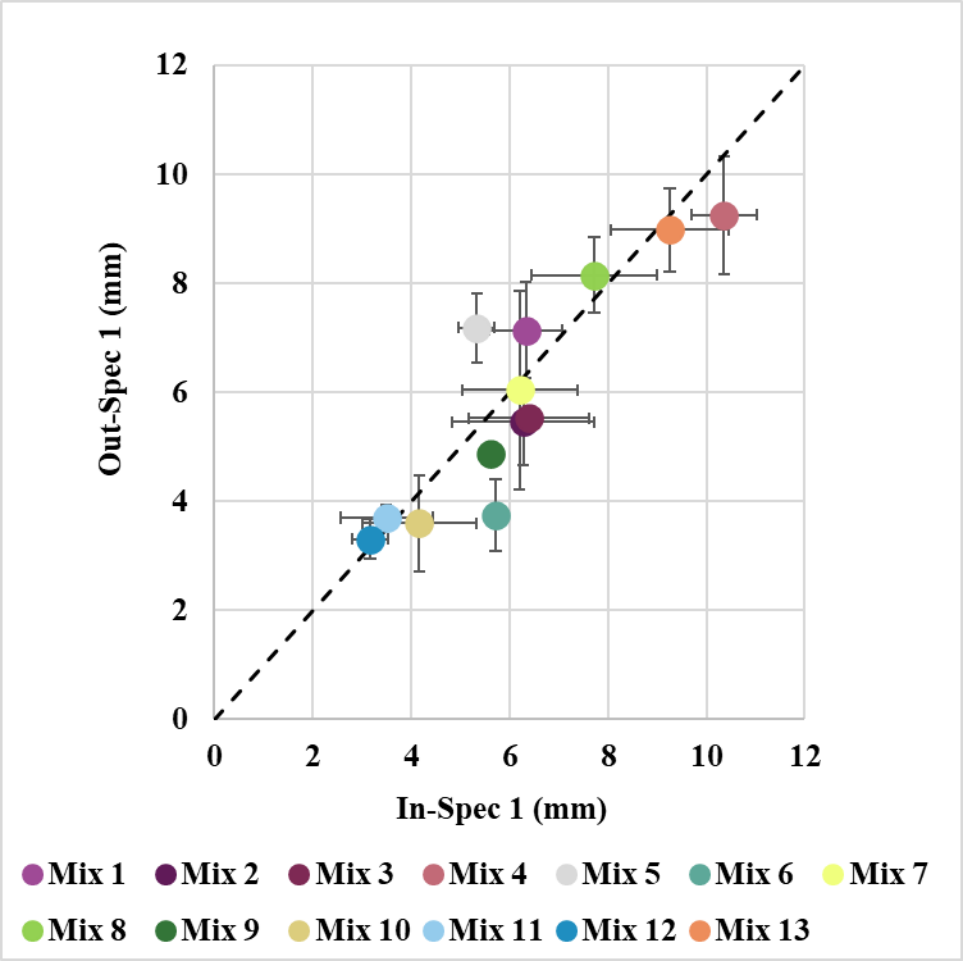

Figure 26. APA/HWTT Test Results of In-Spec 1 and Out-Spec 1

Table 22. P-Values of Rut Depths

Mix ID	Observed Difference	P-Value	Statistical Significance?
In-Spec 1 → Out- Spec 1 (Upper Gradation Limits)			
Mix 1 (APA 250 psi)	0.8	0.29	No
Mix 2 (APA 250 psi)	-0.8	0.45	No
In-Spec 1→ Out- Spec 1 (Lower Gradation Limits)			
Mix 3 (APA 250 psi)	-0.9	0.32	No
Mix 4 (APA 250 psi)	-1.1	0.25	No
Mix 5 (APA 250 psi)	1.9	0.10	No
Mix 6 (HWTT)	-2.0	0.15	No
Mix 7 (APA 250 psi)	-0.2	0.90	No
Mix 8 KBVS (HWTT)	0.4	0.72	No
Mix 9 (HWTT)	-0.7	0.00	Yes
Mix 10 (APA 250 psi)	-0.6	0.54	No
Mix 11 (APA 250 psi)	0.2	0.74	No
Mix 12 (APA 250 psi)	0.2	0.66	No
Mix 13 KBVS (HWTT)	-0.3	0.76	No

4.3.5 Intermediate-Temperature Cracking Resistance

Figure 27 presents the CT_{Index} values of In-Spec 1 and Out-Spec 1 of 13 mix designs, and **Table 23Table *24*** presents the corresponding d2s and p-values. The single-operator d2s limit between the two test results was 51.1%, as reported in an interlaboratory study by the Virginia Transportation Research Council [208]. Of 11 lower limit mix designs, eight showed a higher CT_{Index} value for Out-Spec 1 than for In-Spec 1, indicating improved intermediate-temperature cracking resistance. Notably, four of these improvements were statistically significant ($p < 0.05$). Three other lower limit mix designs (Mix 7, Mix 8, and Mix 13) showed slightly lower CT_{Index} values for Out-Spec 1; however, these reductions were not statistically significant ($p > 0.05$).

The two upper limit mix designs (Mix 1 and Mix 2) showed decreases in CT_{Index} for the Out-Spec 1, indicating reduced intermediate-temperature cracking resistance. This reduction was statistically significant for Mix 1, but not for Mix 2.

Importantly, none of the 13 mix designs showed differences between In-Spec 1 and Out-Spec 1 that exceeded the d2s limits, indicating no practically significant differences. In summary, the IDEAL-CT results revealed two opposite trends depending on the direction of the gradation shift relative to the P-401 gradation limits. Adjusting the gradation outside the lower limits provided equal or improved intermediate-temperature cracking resistance. Conversely, shifting gradation outside the upper limits tended to reduce cracking resistance.

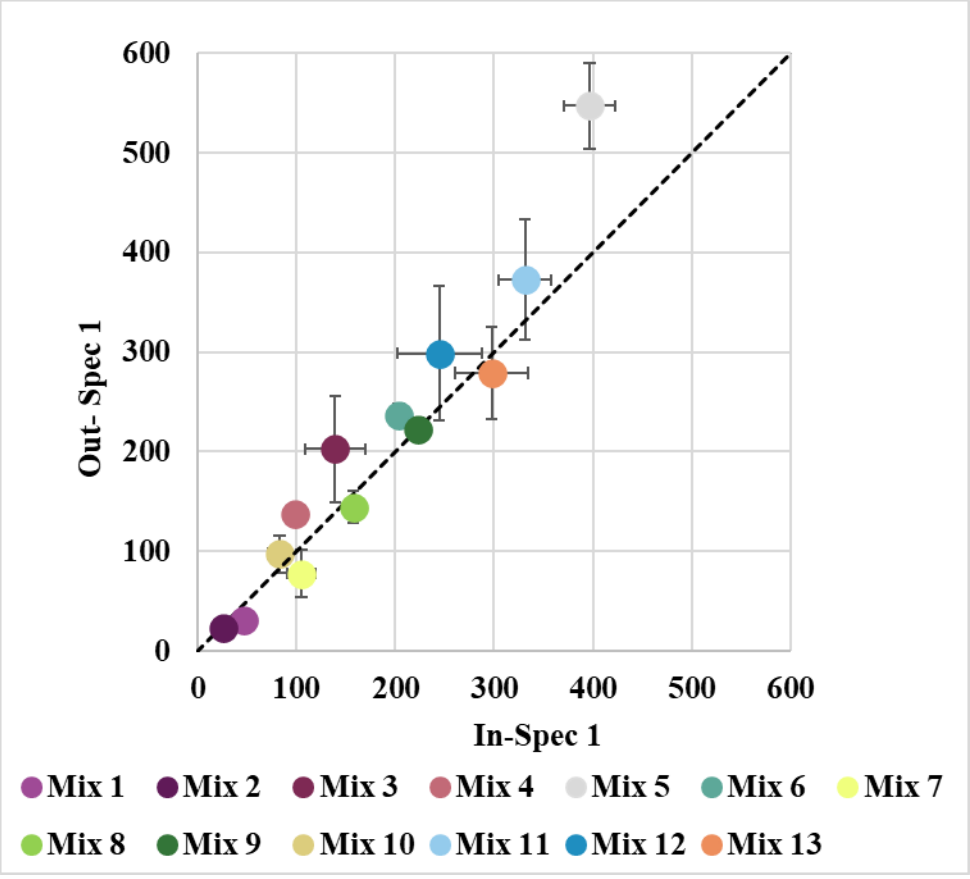

Figure 27. CT_{Index} Results of In-Spec 1 and Out-Spec 1

Table 23. D2S and P-Values of CT_{Index} of In-Spec 1 and Out-Spec 1

Mix ID	Observed Difference	Single Operator d2s (51.1% [208])	Exceeding Single Operator d2s Limit?	P-Value	Statistical Significance?
	In-Spec 1 --> Out- Spec 1 (Upper Gradation Limits)				
Mix 1	-16.1	19.8	No	0.00	Yes
Mix 2	-3.0	12.6	No	0.43	No
	In-Spec 1 --> Out- Spec 1 (Lower Gradation Limits)				
Mix 3	63.2	87.4	No	0.02	Yes
Mix 4	39.6	60.2	No	0.02	Yes
Mix 5	150.5	241.1	No	0.00	Yes
Mix 6	32.6	112.4	No	0.01	Yes
Mix 7	-27.6	46.8	No	0.11	No
Mix 8	-13.2	77.2	No	0.19	No
Mix 9	0.2	113.9	No	0.99	No
Mix 10	14.6	46.1	No	0.24	No
Mix 11	41.1	179.9	No	0.28	No
Mix 12	53.9	138.8	No	0.31	No
Mix 13	-18.5	147.3	No	0.56	No

Figure 28 presents the FI values of In-Spec 1 and Out-Spec 1 of 13 mix designs, and **Table 24** presents the corresponding d2s and p-values. AASHTO T 393 [198] indicated that the d2s between the two test results should not exceed 75.9% of their average. Among 11 lower limit mix designs, ten showed higher FI values for Out-Spec 1 than for In-Spec 1. Notably, three of these increases were statistically significant ($p < 0.05$). Mix 8 showed a lower FI value for Out-Spec 1, but the difference was not statistically significant ($p > 0.05$).

Both upper limit mix designs (Mix 1 and Mix 2) showed slightly higher FI for Out-Spec 1 than for In-Spec 1; however, the increases were not statistically significant ($p > 0.05$). Moreover, none of the 13 mix designs showed differences between In-Spec 1 and Out-Spec 1 that exceeded the d2s limits, indicating no practically significant differences. In summary, FI values showed that

adjusting gradation outside the limits while controlling volumetric properties maintains or improves the intermediate-temperature cracking resistance.

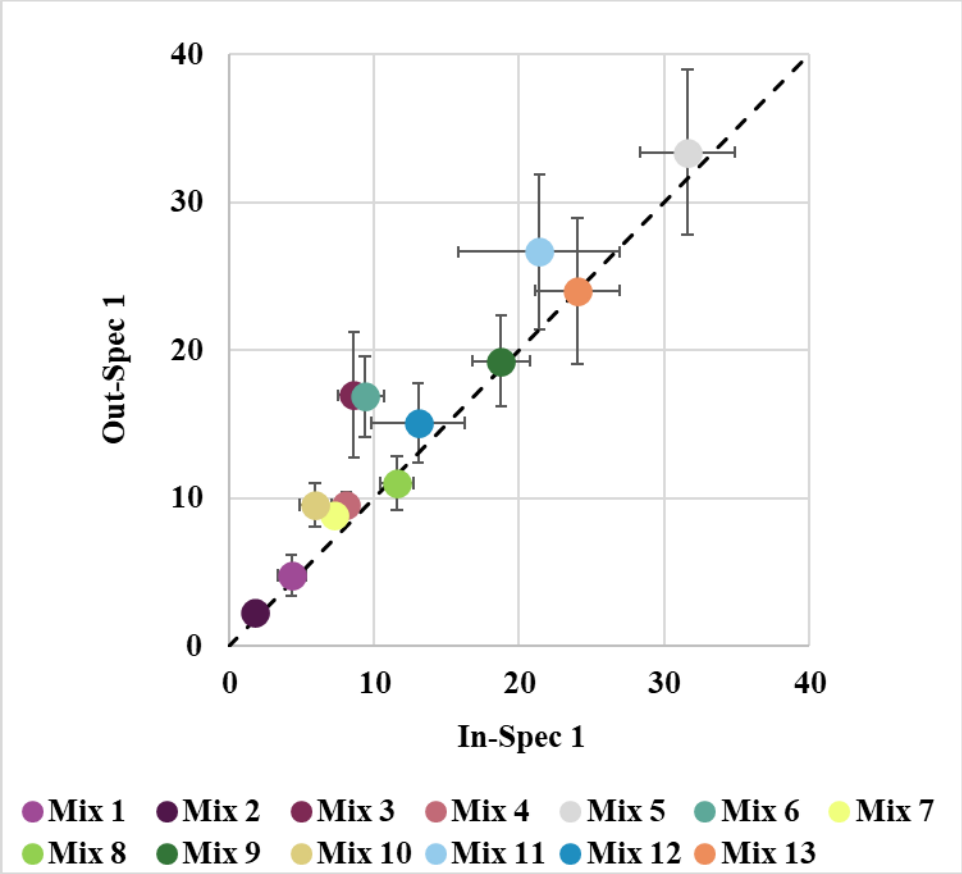

Figure 28. FI Values of In-Spec 1 and Out-Spec 1

Table 24. D2S and P-Values of FI of In-Spec 1 and Out-Spec 1

Mix ID	Observed Difference	Single Operator d2s (75.9% [198])	Exceeding Single Operator d2s Limit?	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Gradation Limits)					
Mix 1	0.5	3.5	No	0.45	No
Mix 2	0.4	1.5	No	0.07	No
In-Spec 1 vs. Out- Spec 1 (Lower Gradation Limits)					
Mix 3	8.4	9.7	No	0.00	Yes
Mix 4	1.4	6.7	No	0.14	No
Mix 5	1.8	24.7	No	0.51	No
Mix 6	7.5	10.0	No	0.01	Yes
Mix 7	1.6	6.1	No	0.35	No
Mix 8	-0.6	8.6	No	0.61	No
Mix 9	0.5	14.4	No	0.79	No
Mix 10	3.6	5.9	No	0.01	Yes
Mix 11	5.3	18.2	No	0.22	No
Mix 12	2.0	10.7	No	0.14	No
Mix 13	2.7	17.2	No	0.40	No

4.3.6 Low-Temperature Cracking Resistance

Figure 29 presents the DCT test results for In-Spec 1 and Out-Spec 1 of 13 mix designs, and **Table 25** presents the corresponding d2s and p-values. ASTM D7313 [199] indicates that the d2s limit of fracture energy between the two test results is 318.151. Nice mix designs exhibited higher fracture energy for Out-Spec 1 than those for In-Spec 1, indicating improved low-temperature cracking resistance. However, the differences in low-temperature cracking resistance were not statistically significant ($p > 0.05$) and were less than the d2s limit. Overall, DCT results indicate that shifting the gradation outside the gradation limits while maintaining volumetric properties did not adversely affect the low-temperature cracking resistance.

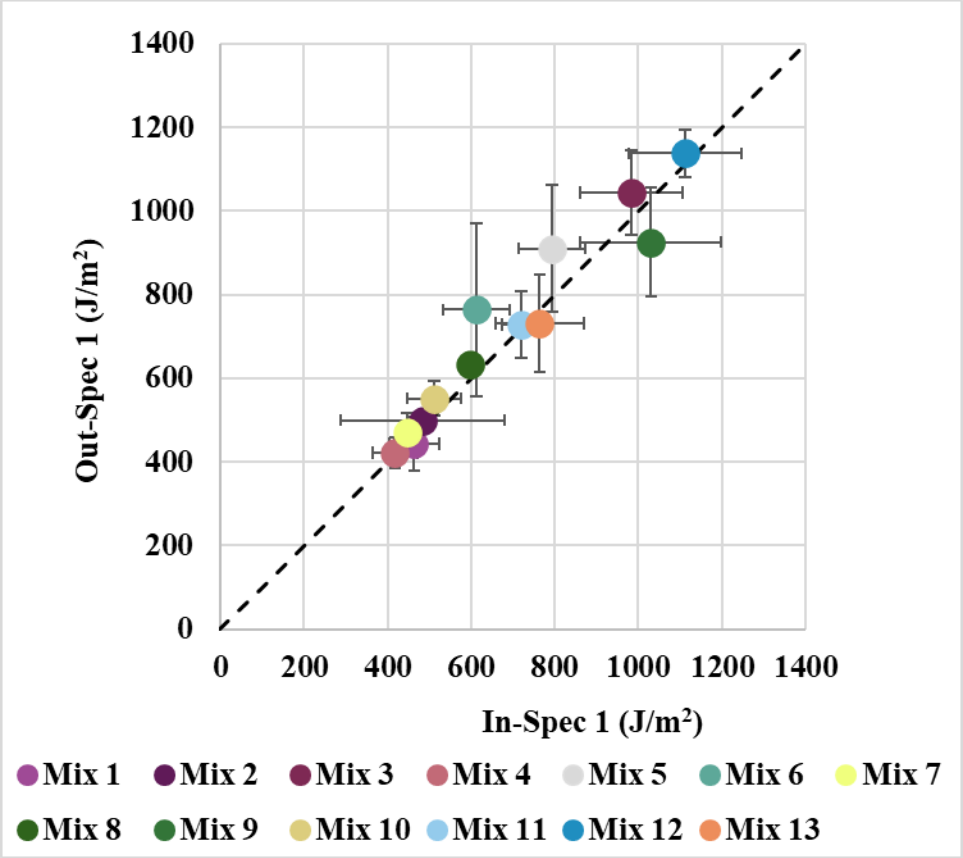

Figure 29. Fracture Energy from DCT of In-Spec 1 and Out-Spec 1

Table 25. D2S and P-Values of DCT Test Results of In-Spec 1 and Out-Spec 1

Mix ID	Observed Difference	Single Operator d2s (318.151 [199])	Exceeding Single Operator d2s Limit?	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Gradation Limits)					
Mix 1	-20.6	318.151	No	0.62	No
Mix 2	15.7	318.151	No	0.49	No
In-Spec 1 vs. Out- Spec 1 (Lower Gradation Limits)					
Mix 3	60.5	318.151	No	0.37	No
Mix 4	4.4	318.151	No	0.88	No
Mix 5	116.4	318.151	No	0.18	No
Mix 6	151.3	318.151	No	0.25	No
Mix 7	24.4	318.151	No	0.30	No
Mix 8	-84.4	318.151	No	0.45	No
Mix 9	-105.0	318.151	No	0.45	No
Mix 10	39.8	318.151	No	0.34	No
Mix 11	8.0	318.151	No	0.87	No
Mix 12	25.7	318.151	No	0.75	No
Mix 13	-32.5	318.151	No	0.69	No

4.3.7 Summary

In this chapter, the laboratory performance of In-Spec 1 and Out-Spec 1 was measured and compared to evaluate the effects of moving gradation outside the gradation limits while maintaining the volumetric properties and binder content. Based on the testing results, the following conclusions could be obtained:

- With respect to moisture susceptibility, permeability, durability, rutting resistance, and low-temperature cracking resistance, gradation shifts outside the P-401 gradation limits did not produce significant changes in laboratory performance. All mixtures satisfied the applicable performance requirements, and the differences observed between paired In-Spec 1 and Out-Spec 1 were neither statistically nor practically significant across these performance indicators.

- Intermediate-temperature cracking resistance showed a directional trend. For mixtures near the lower gradation limits, moving the gradation outside the band generally preserved or enhanced cracking resistance, with several statistically significant increases observed in CT_{Index} and FI. Conversely, for mixtures near the upper limits, CT_{Index} values typically declined when gradation shifted beyond those limits, while FI results showed no statistically significant change.

In summary, it was feasible to adjust the gradation outside the P-401 gradation limits for 12.5 mm NMAS mix designs without significantly affecting the mixture properties.

4.4 Revising P-401 Gradation Limits

Of the 13 mix designs evaluated in the study, 12 were designed with a 12.5 NMAS, and one was a 19 mm NMAS. Since the gradation limit adjustment was focused on 12.5 mm NMAS, the 12 corresponding mix designs, including two upper limit mixes and 10 lower limit mixes, were used. The gradation design of Out-Spec 1 was used to modify the upper and lower P-401 limits.

4.4.1 Upper Limit Evaluation

The P-401 upper limits were adjusted based on two upper limit mix designs (Mix 1 and Mix 2). **Table 26** summarizes the number of mixes and percentage deviations from upper limits by sieve size. As shown, the gradation was adjusted primarily on No.4, No.8, No.30, and No.50 sieves. For No.4 and No.8 sieves, one mix exceeded the upper limit by 1%. For the No. 30 sieve, both mixes exceeded the upper limit by 1%. At No.50 sieve, one mix exceeded by 3%, and the other exceeded by 4%. Based on the results, the most notable deviations were concentrated at the No. 30 sieve (1% exceedance) and the No. 50 sieve (3% to 4% exceedance), which could theoretically serve as a basis for upper adjustments to these limits.

Table 26. Number of Mixes and Percentage Deviations from Upper Limits by Sieve Size

Sieve Sizes	Percent Changes from Upper Limits/ Number of Mixes		
	1%	3%	4%
No.4	1 Mix	0 Mix	0 Mix
No.8	1 Mix	0 Mix	0 Mix
No.30	2 Mixes	0 Mix	0 Mix
No.50	0 Mix	1 Mix	1 Mix

Although upper limit mix designs exhibited acceptable rutting resistance and durability, their intermediate-temperature cracking resistance was inadequate. For example, the CT_{Index} values of Out-Spec 1 and In-Spec 1 of Mix 1 and Mix 2 were less than the minimum threshold of 50 used by the Alabama Department of Transportation [209], and their FI values were less than the minimum threshold of 8 used by the Illinois Department of Transportation [210]. In addition, from a practical standpoint, contractors have not reported difficulties in meeting the P-401 upper limit. Therefore, it was recommended to retain the current upper limits specified in the P-401 gradation limits.

4.4.2 Lower Limit Evaluation

The lower limits were adjusted based on the gradation of 10 lower limit mixes. **Table 27** summarizes the number of mixes and percentage deviations from lower limits by sieve size.

Table 27. Number of Mixes and Percentage Deviations from Lower Limits by Sieve Size

Sieve sizes	Percent Changes from Lower Limits/ Number of Mixes		
	1%	2%	3%
No.16	6 Mixes	0 Mix	1 Mix
No.30	4 Mixes	5 Mixes	0 Mix
No.50	6 Mixes	1 Mix	0 Mix
No.100	1 Mix	0 Mix	0 Mix

As shown, the gradation was adjusted at different sieve sizes, with a primary focus on the No. 16, No. 30, and No. 50 sieves. For the No. 16 sieve, seven mix designs fell below the lower limit, with six exceeding it by 1% and one by 3%. For the No. 30 sieve, nine mixes fell outside the lower limits, with four mixes exceeding by 1% and five mixes by 2%. At the No.50 sieve, seven mixes were moved outside the lower limit, with six mixes exceeding it by 1% and one mix by 2%. Finally, for the No.100 sieve, one mix was adjusted outside the limit by 1%. Performance evaluation confirmed that shifting aggregate gradations outside the P-401 lower limits did not adversely affect moisture resistance, permeability, durability, rutting resistance, or low-temperature cracking resistance. Intermediate-temperature cracking resistance was maintained or improved. Based on the results, it was recommended to reduce the lower limit at sieves No. 16, No. 30, and No. 50 by 2%.

Table 28 presents the current and revised P-401 gradation limits for the 12.5 NMAS mix design.

Table 28. Current and Revised P-401 Gradation Limits for 12.5 mm NMAS Mix Design

Sieve (%)	Current P-401 Gradation Limits	Revised P-401 Gradation Limits
3/4‘’	100	100
1/2"	90-100	90-100
3/8"	72-88	72-88
#4	53-73	53-73
#8	38-60	38-60
#16	26-48	**24**-48
#30	18-38	**16**-38
#50	11-27	**9**-27
#100	6-18	6-18
#200	3.0-6.0	3.0-6.0

4.4.3 Engineering Rationale for Proposed Adjustments to Lower Limits

Reducing the lower limits by 2% for sieves No. 16, No. 30, and No. 50 provides both practical and performance benefits. Under the current specification, natural sand is added to meet the lower limits at these sieves. Natural sand gradation typically contains about 20% more material passing the No. 16, No. 30, and No. 50 sieves than manufactured sand. As a result, incorporating roughly 10% natural sand into the blend can increase the percentage passing through these sieves by about 2%, helping meet the lower limit requirement. However, replacing manufactured sand with natural sand generally reduces FAA and lowers VMA. This makes it more challenging to meet both the minimum VMA requirement and achieve adequate rutting resistance. Reducing the lower limits allows greater flexibility in selecting fine aggregates, thereby facilitating increased use of manufactured sand. This approach helps maintain or enhance VMA while preserving rutting resistance.

Additionally, the cracking resistance of asphalt mixtures is primarily controlled by binder properties and the effective binder content. Future cracking performance criteria will likely require higher effective binder volumes than those typically used in current airport mix designs. Increasing VMA is an effective approach to accommodate more binder. Thus, reducing the lower limits at No. 16, No. 30, and No. 50 sieves supports the development of airport mix designs with higher VMA and improved potential for cracking resistance.

CHAPTER 5 EFFECT OF GRADATION CHANGE ON MIXTURE VOLUMETRIC AND PERFORMANCE PROPERTIES

This chapter presents Part 2 of the experimental design, which evaluated the effect of gradation change on volumetric properties and laboratory performance of airfield asphalt mixtures. The findings provide quantitative insight into relationships between gradation change and volumetrics, and laboratory performance, and were used to support the predictive modeling developed in Chapter 6. It includes four sections: (1) experimental design, (2) design summary, (3) test results and data analysis, and (4) summary.

5.1 Experimental Design

Figure 30 describes the experimental design, including three steps: 1) adjust In-Spec 1 gradation, 2) performance evaluation, and 3) VMA and performance comparison. Step 1 presented the gradation adjustment of In-Spec 1. 10 In-Spec 1, including two upper limit mixes (Mix 1 and Mix 2) and eight lower limit mixes (Mix 3, Mix 4, Mix 5, Mix 6, Mix 7, Mix 8, Mix 12, and Mix 13), were used for gradation adjustments. The selection was based on the available material and the need to maintain representation of aggregate sources and climate areas. The outcome of this step is two additional gradation designs: Out-Spec 2 and In-Spec 2. While Out-Spec 2 had the gradation further outside the gradation limits, In-Spec 2 was centered within the gradation limits. Gradation adjustments were guided by the Bailey method. Both mixtures were designed to have a similar binder content (± 0.2%) and a VMA within 1-2% of the In-Spec 1. This approach isolated the influence of gradation changes on volumetric and performance responses.

In step 2, comprehensive performance tests were conducted to evaluate the performance of In-Spec 2 and Out-Spec 2. Initially, a full performance testing matrix was planned for all 10 mix designs. However, test results from Mix 1 and Mix 3 showed no significant difference in durability

and generally ensured impermeability and moisture resistance. Therefore, these tests were removed from the testing plan for the remaining mix designs. Mix 1 and Mix 3 received the full testing matrix, while the remaining eight mix designs were evaluated using a reduced matrix consisting of APA/HWTT, IDEAL-CT, I-FIT, and DCT.

In Step 3, the effect of gradation changes on volumetric properties and laboratory performance was evaluated. Volumetric properties of four gradations (In-Spec 1, In-Spec 2, Out-Spec 1, Out-Spec 2) were compared for each mix design, using the mean value. Regarding the effect on performance, a Games-Howell test at the 0.05 significance level was conducted to evaluate whether the performance difference between In-Spec and Out-Spec mixtures was statistically significant. The statistical groupings are indicated by capital letters above the error bars. The gradation design, with the same statistical grouping, indicates no significant difference.

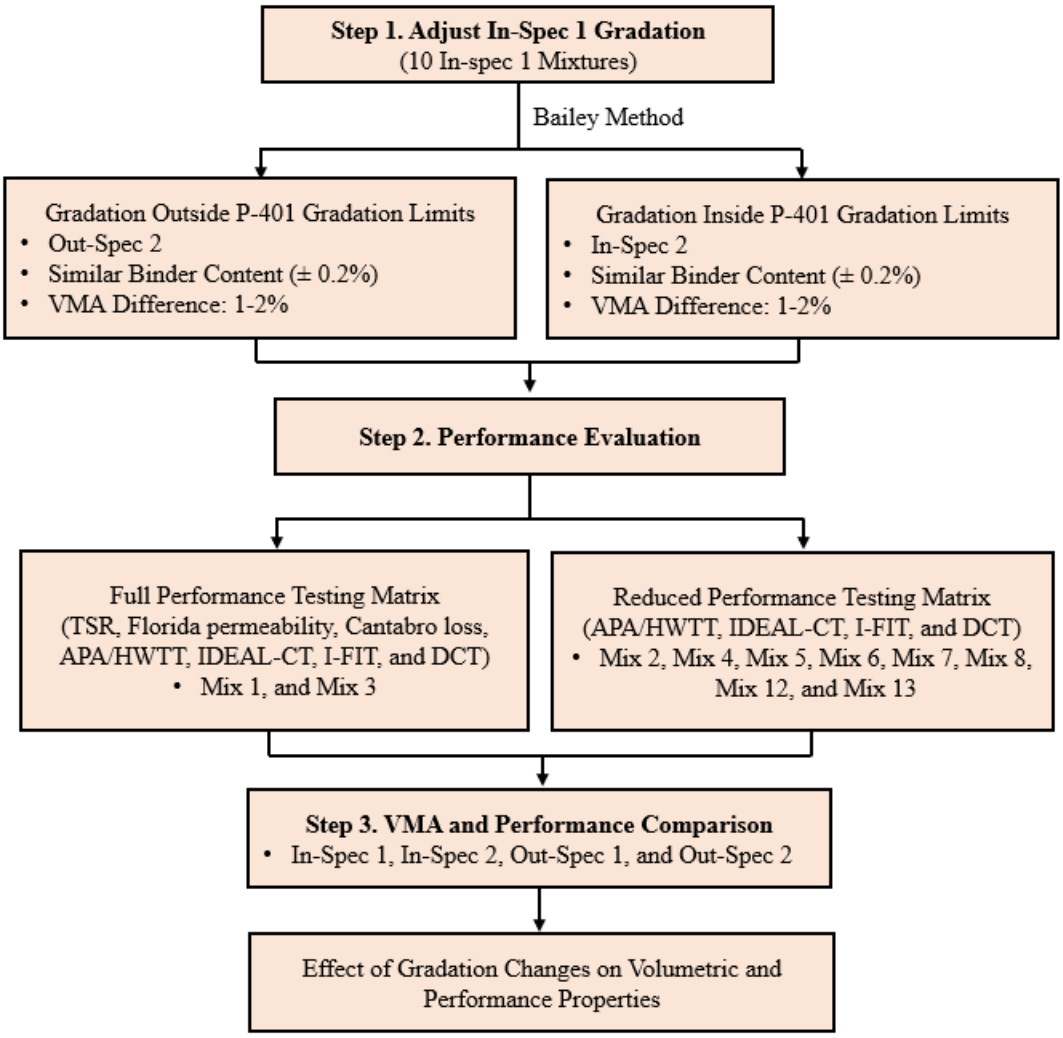

Figure 30. Experimental Design (Part 2) for Evaluating the Effect of Gradation Changes on Volumetric and Performance Properties

5.2 Design Summary

Table 29 summarizes the gradation design, FAA, AC, and volumetric properties of In-Spec 2 and Out-Spec 2 of ten mix designs. As shown, In-Spec 2 gradations were adjusted toward the center of the P-401 gradation limits, whereas Out-Spec 2 gradations were positioned further outside the gradation limits relative to Out-Spec 1 gradations. Similar asphalt binder (±0.2%) and effective binder content were maintained between In-Spec 2 and Out- Spec 2. Note that, unlike the In-Spec

1 and Out-Spec 1, the volumetric properties of the In-Spec 2 and Out-Spec 2 changed as a result of the gradation adjustments.

Table 29. Design Summary of In-Spec 2 and Out-Spec 2

Mix ID	1		2		3		4		5	
Sieve	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**
3/4‘’	100	100	100	100	100	100	100	100	100	100
1/2"	92	94	97	98	95	93	97	97	94	91
3/8"	85	90	88	90	83	77	88	89	87	83
#4	69	80	59	71	57	48	66	58	70	62
#8	53	68	46	60	46	35	49	39	46	41
#16	39	54	38	50	34	22	32	23	29	26
#30	30	43	31	43	25	14	22	15	19	15
#50	23	33	23	32	15	8	13	9	11	7
#100	12	16	11	17	6	5	8	5	8	4
#200	3.9	4.3	4.4	5.9	3.4	3.2	5.8	3.5	6.0	3.1
FAA (%)	46.2	47.3	46.8	47.2	43.9	46.4	44.6	44.7	42.8	43.3
AC (%)	7.1	7.2	6.2	6.0	5.9	6.1	5.7	5.7	6.8	6.9
Va (%)	1.5	5.1	1.4	4.7	1.7	4.6	2.5	4.6	2.8	4.6
VMA (%)	13.9	17.7	14.6	17.6	14.9	17.8	13.9	15.9	15.8	18.1
Pbe (%)	5.6	5.6	5.6	5.7	5.2	5.2	4.8	4.9	5.7	6.1

*Notes

IS 2: In-Spec 2; OS 2: Out-Spec 2

Table 29. Design Summary of In-Spec 2 and Out-Spec 2 (Continued)

Mix ID	6		7		8		12		13	
Sieve	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**	**IS2**	**OS2**
3/4‘’	99	98	100	100	100	100	100	100	100	100
1/2"	88	90	96	99	96	95	94	93	94	94
3/8"	82	82	86	91	85	78	74	72	88	88
#4	63	50	63	60	64	50	55	48	69	59
#8	43	30	43	38	49	35	39	32	48	36
#16	28	18	30	25	31	20	27	21	35	24
#30	18	11	20	16	23	14	19	14	23	17
#50	11	7	11	9	14	10	14	10	13	11
#100	8	5	7	6	9	8	9	7	8	8
#200	5.9	3.9	5.4	4.5	6.0	6.0	5.9	3.6	5.8	5.6
FAA (%)	45.2	45.3	44.0	44.0	46.1	47.9	48.7	48.5	45	45
AC (%)	5.3	5.3	6.4	6.3	6.0	6.0	5.8	5.8	6.1	6.1
Va (%)	2.5	5.1	3.5	5.1	3.4	5.5	2.0	4.9	3.6	4.5
VMA (%)	13.4	16.1	15.3	16.8	15.1	17.4	15.6	18.4	15.8	16.8
Pbe (%)	4.7	4.8	5.2	5.2	4.9	5.1	5.4	5.5	5.3	5.3

*Notes

IS 2: In-Spec 2; OS 2: Out-Spec 2

5.3 Test Results and Data Analysis

This section presents the volumetric properties (VMA and V_a) and laboratory performance results of all In-Spec and Out-Spec mixtures from 10 airfield mix designs. For upper-limit Mix 1 and Mix 2, shifting the gradation from Out-Spec 2 to In-Spec 2 resulted in a coarser gradation. For the lower limit mix designs (Mix 3, Mix 4, Mix 5, Mix 6, Mix 7, Mix 8, Mix 12, and Mix 13), shifting the gradation from In-Spec 2 to Out-Spec 2 produced a coarser gradation.

5.3.1 Effect of Gradation Change on Volumetric Properties

Figure 31 and **32** present the VMA and V_a values of In-Spec and Out-Spec mixtures across Mix 1 to Mix 13. For all mix designs, Out-Spec 2 exhibited the highest VMA, followed by Out-Spec 1 and In-Spec 1. In-Spec 2 had the lowest VMA. A similar trend was observed for V_a values. These results indicate that moving the gradation curve farther from the MDL, toward either the lower or upper limits, increased the volumetric properties of the asphalt mixture under similar compaction conditions and binder contents.

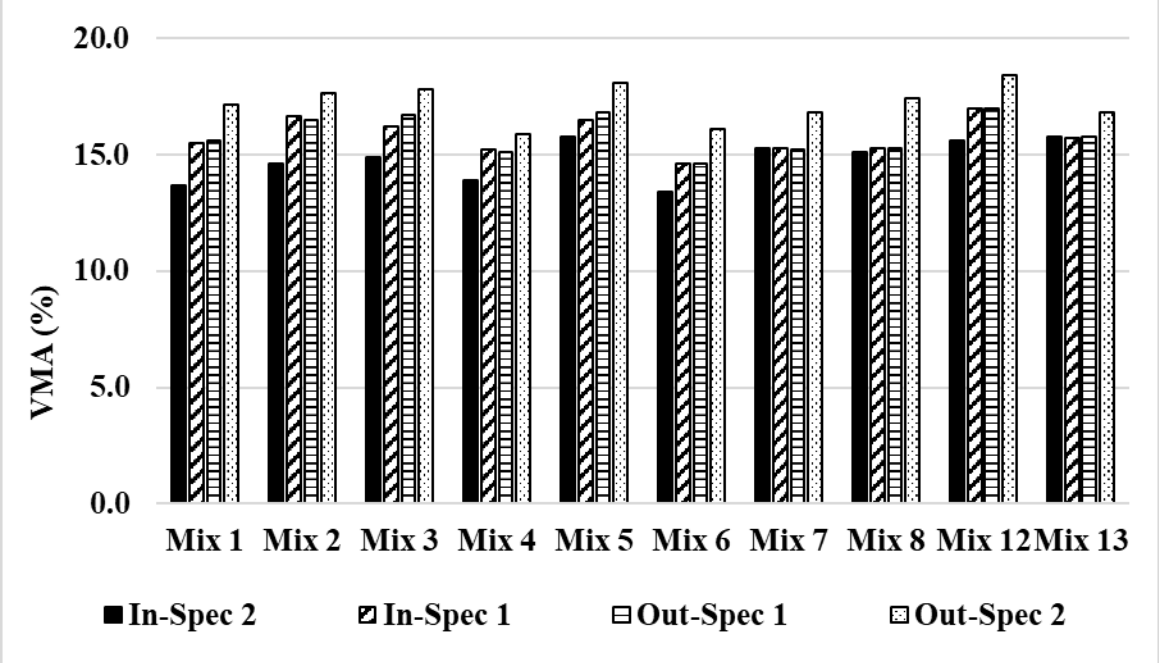

Figure 31. Summary of VMA of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2

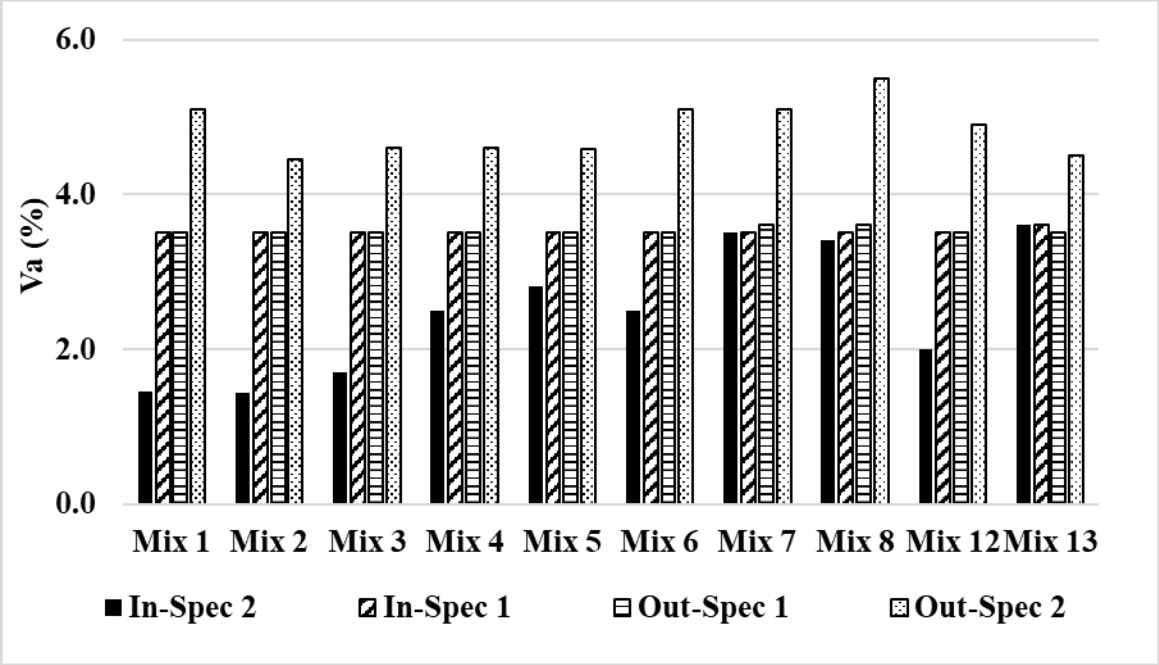

Figure 32. Summary of V_a of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2

5.3.2 Effect of Gradation Change on Laboratory Performance

This section evaluates the effects of systematic gradation shifts on the laboratory performance of the asphalt mixtures, including moisture susceptibility, permeability, durability, rutting resistance, intermediate-temperature cracking resistance, and low-temperature cracking resistance. Performance comparisons were conducted across all four gradations, In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2, to determine whether the volumetric changes documented in the previous section translated into measurable differences in the performance response.

5.3.2.1 Moisture Susceptibility

Figure 33 presents the TSR results for Mix 1 and Mix 3. For both mix designs, the TSR values of all the In-Spec and Out-Spec mixtures were generally equal to or higher than 0.8, ensuring the moisture resistance as specified in P-401. There was an exception for In-Spec 2 of Mix 3, with a TSR value below 0.8. Overall, the results indicate that systematic gradation adjustments within the evaluated range generally ensure moisture resistance. This finding is consistent with the well-

established understanding that moisture resistance in airfield asphalt mixtures is governed primarily by binder properties and aggregate mineralogy rather than by gradation adjustments [211].

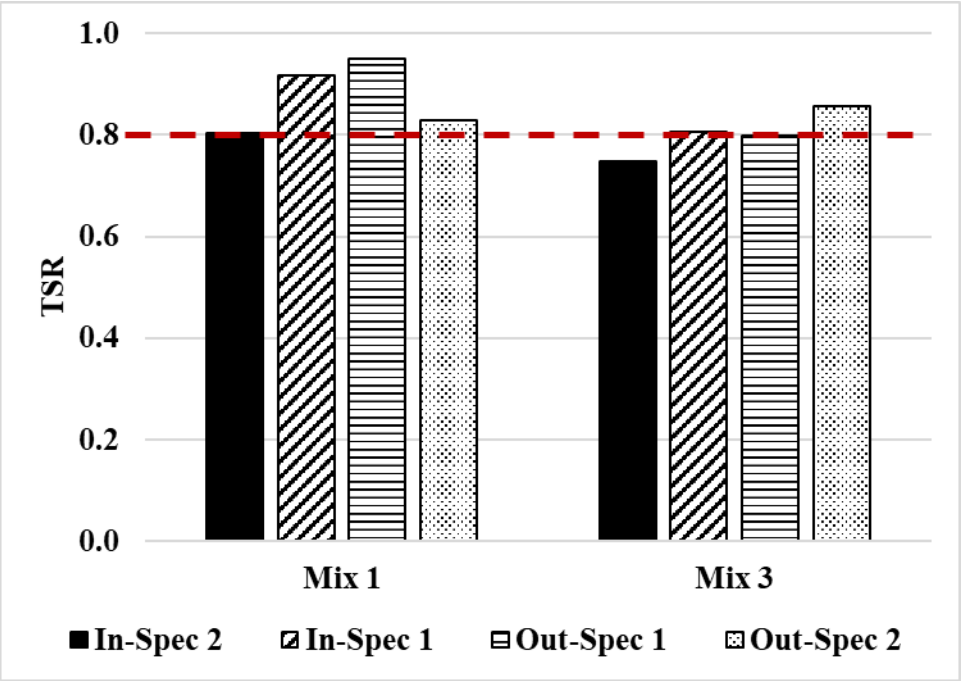

Figure 33. TSR Test Results of Mix 1 and Mix 3

5.3.2.2 Permeability

Figure 34 presents the permeability coefficients and statistical group analysis of In-Spec and Out-Spec gradations of Mix 1 and Mix 3. All measured permeability coefficients were less than 125×10^{-5} cm/s, confirming that the mixtures were effectively impermeable [156, 206]. For Mix 1, group analysis revealed no significant difference between In-Spec 1, Out-Spec 1, and Out-Spec 2; however, all were significantly lower than In-Spec 2. For Mix 3, the permeability was reduced from In-Spec 2 to Out-Spec 2; however, the group analysis found no significant differences among the gradations. In summary, changing gradation from In-Spec 2 to Out-Spec 2 within the evaluated range of volumetric properties maintained the impermeable performance of airfield asphalt mixtures.

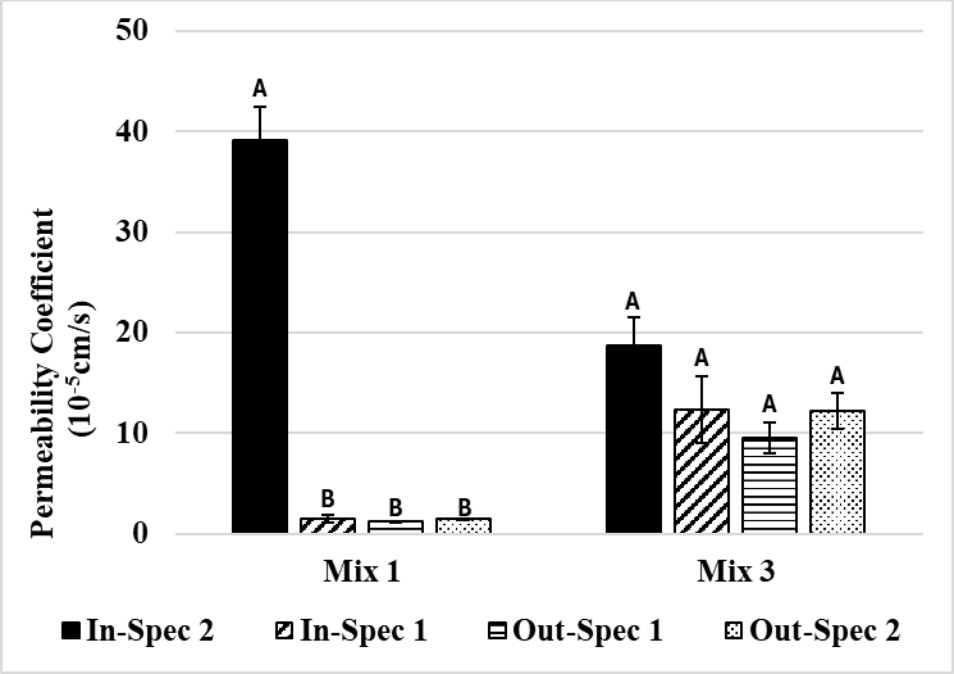

Figure 34. Permeability Test Results of Mix 1 and Mix 3

5.3.2.3 Durability

Figure 35 presents the Cantabro test results and group analysis of In-Spec and Out-Spec gradations of Mix 1 and Mix 3. The Cantabro of all gradations was below 6.3% threshold commonly used for durability evaluation [207], confirming sufficient durability performance. For both mixes, In-Spec 1 exhibited the lowest Cantabro loss, followed by In-Spec 2 and then Out-Spec 2. Out-Spec 1 showed the highest Cantabro loss. However, the group analysis showed no significant differences among gradations for both mix designs. This suggests that changing gradation within the evaluated volumetric range did not significantly impact the durability of airfield asphalt mixtures. This conclusion is consistent with the understanding that durability is primarily influenced by binder content and aggregate-binder adhesion, which were kept similar across four gradations [212].

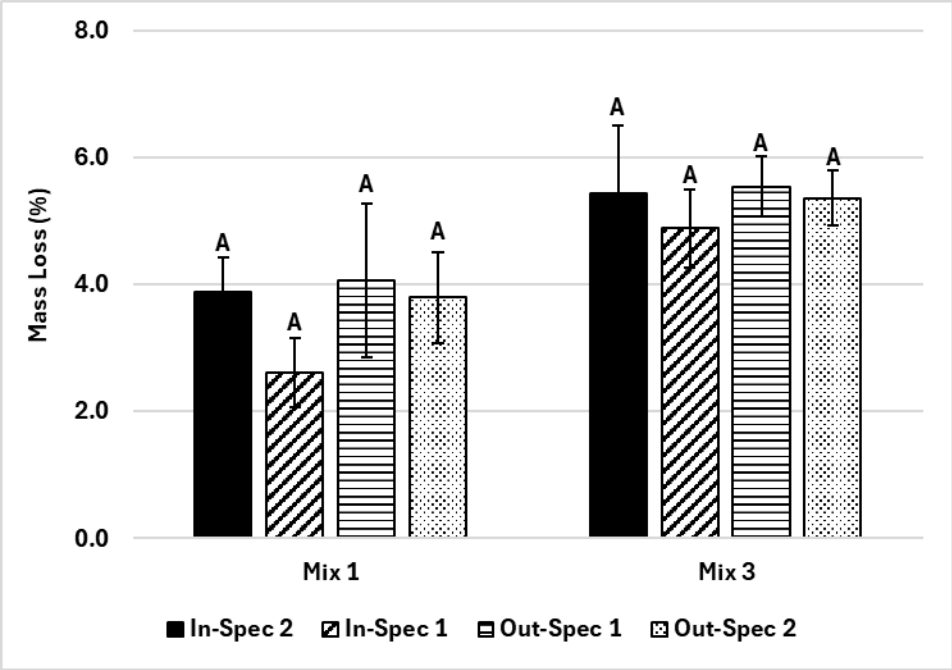

Figure 35. Cantabro Test Results of Mix 1 and Mix 3

5.3.2.4 Rutting Resistance

Figure 36 presents the APA and HWTT test results and group analysis of In-Spec and Out-Spec mixtures from 10 mix designs. The dashed line at 10 mm represents the maximum allowable rut depth threshold specified in the P-401. Mix 6, Mix 8, and Mix 13 were tested using HWTT, while the remaining mix designs were tested using APA.

For the upper limit mix designs, Mix 1 and Mix 2, all gradations had similar rut depths and were below the 10 mm threshold, indicating adequate rutting resistance. Group analysis confirmed no significant difference between the gradations.

For the lower limit mix designs from Mix 3 to Mix 13, the rut depth was generally below the 10 mm threshold, ensuring sufficient rutting resistance. Moreover, the rut depth generally decreased as gradation moved from In-Spec 2 to Out-Spec 2, indicating that a coarser gradation improved rutting resistance. Especially, this improvement was significant for Mix 3, Mix 4, and

Mix 13. There were exceptions for Mix 5 and Mix 6. However, for both mixes, group analysis found no significant differences among the gradations.

In summary, adjusting the mixtures to a coarser gradation either maintained or significantly improved the rutting resistance of airfield asphalt mixtures, which is consistent with the enhanced aggregate interlock and stone-on-stone contact in a coarser aggregate structure [99, 100].

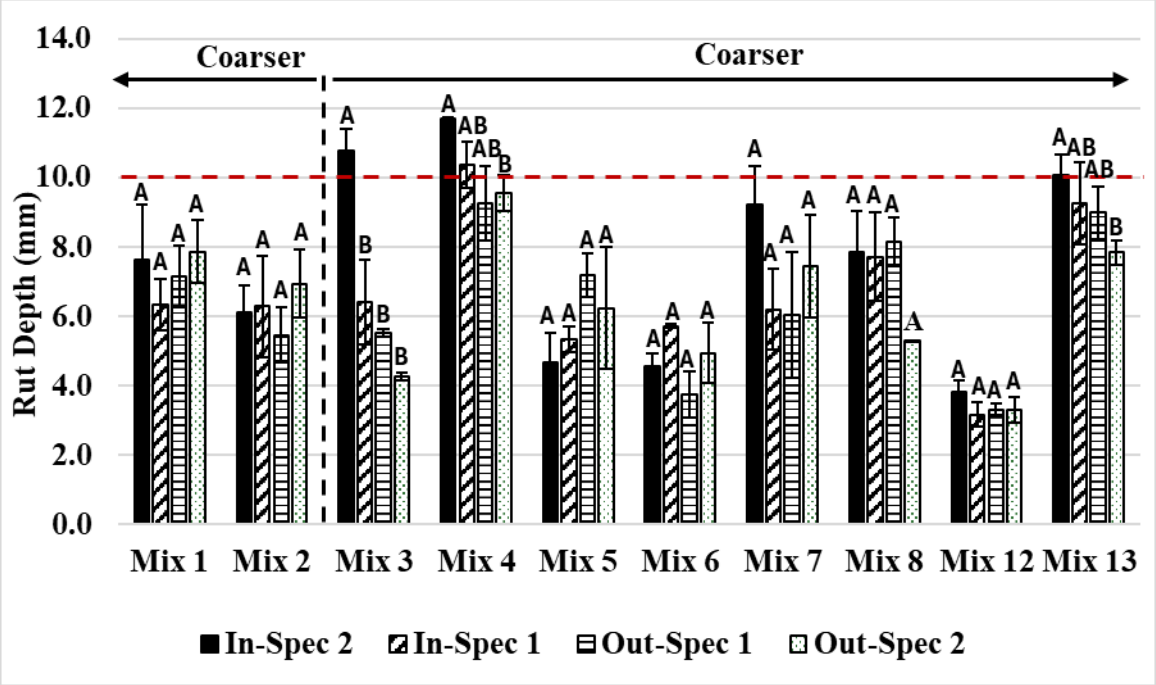

Figure 36. APA/HWTT Results

5.3.2.5 Intermediate- Temperature Cracking Resistance

Intermediate-temperature cracking resistance was evaluated using both CT_{Index} from IDEAL-CT and FI from the I-FIT test. **Figure 37** presents the IDEAL-CT test results and group analysis of In-Spec and Out-Spec gradations from 10 mix designs. For the upper limit mix designs, Mix 1 and Mix 2, adjusting the gradation from Out-Spec 2 to In-Spec 2 increased CT_{Index} values, indicating that a coarser gradation improved intermediate-temperature cracking resistance. These increases were statistically significant based on group analysis.

For several lower limit mix designs, Mix 3, Mix 4, Mix 5, Mix 6, and Mix 8, CT_{Index} values observed a similar trend. Adjusting the gradation coarser from In-Spec 2 to Out-Spec 2 increased CT_{Index} values, with Out-Spec 2 showing a significant improvement over the other gradations, confirmed by group analysis. While other lower limit mix designs, Mix 7, Mix 12, and Mix 13, showed no significant change between gradation adjustments.

Overall, the CT_{Index} results suggested that coarser gradation generally enhanced the intermediate-temperature cracking resistance of airfield asphalt mixtures. In most cases, the improvement was statistically significant.

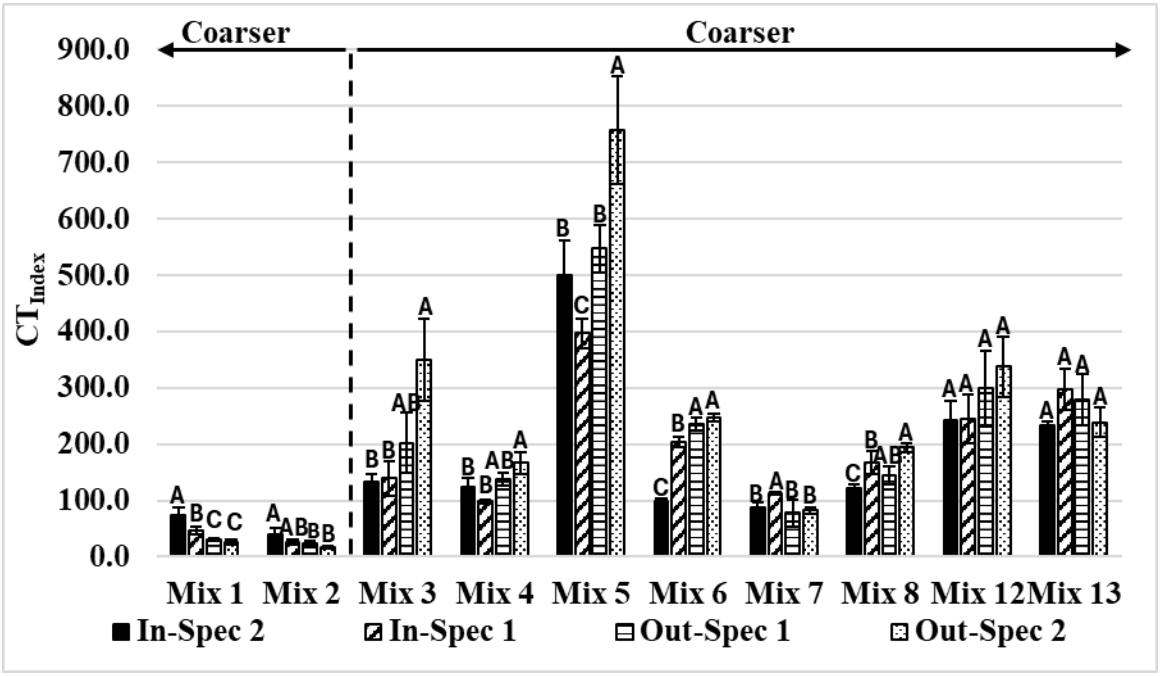

Figure 37. IDEAL-CT Results

Figure 38 presents the I-FIT test results and group analysis of In-Spec and Out-Spec mixtures from 10 mix designs. For the upper limit mix designs, Mix 1 and Mix 2, the FI values increased when the gradation became coarser from Out-Spec 2 to In-Spec 2. The group analysis confirmed that the increases were significant, indicating that coarser gradation significantly improved intermediate-temperature cracking resistance.

For lower limit mix designs, Mix 3, Mix 6, Mix 8, and Mix 12, Out-Spec mixtures had higher FI values than the corresponding In-Spec mixtures, further indicating that coarser gradation improved the intermediate-temperature cracking resistance. Notably, the group analysis revealed the improvement was significant. For the remaining lower limit mixes, Mix 4, Mix 5, Mix 7, and Mix 13, the group analysis showed no significant difference between In-Spec and Out-Spec mixtures.

Overall, both IDEAL-CT and I-FIT tests showed that coarser gradations maintained or significantly enhanced the intermediate-temperature cracking resistance. The trend may be attributed to the effect of gradation on asphalt binder film thickness. Since all four gradations were designed with the same binder content, a coarser gradation having lower fines content and reduced total aggregate surface area results in thicker asphalt binder films around aggregate particles. Thicker binder films are generally associated with improved resistance to crack initiation and propagation [213, 214].

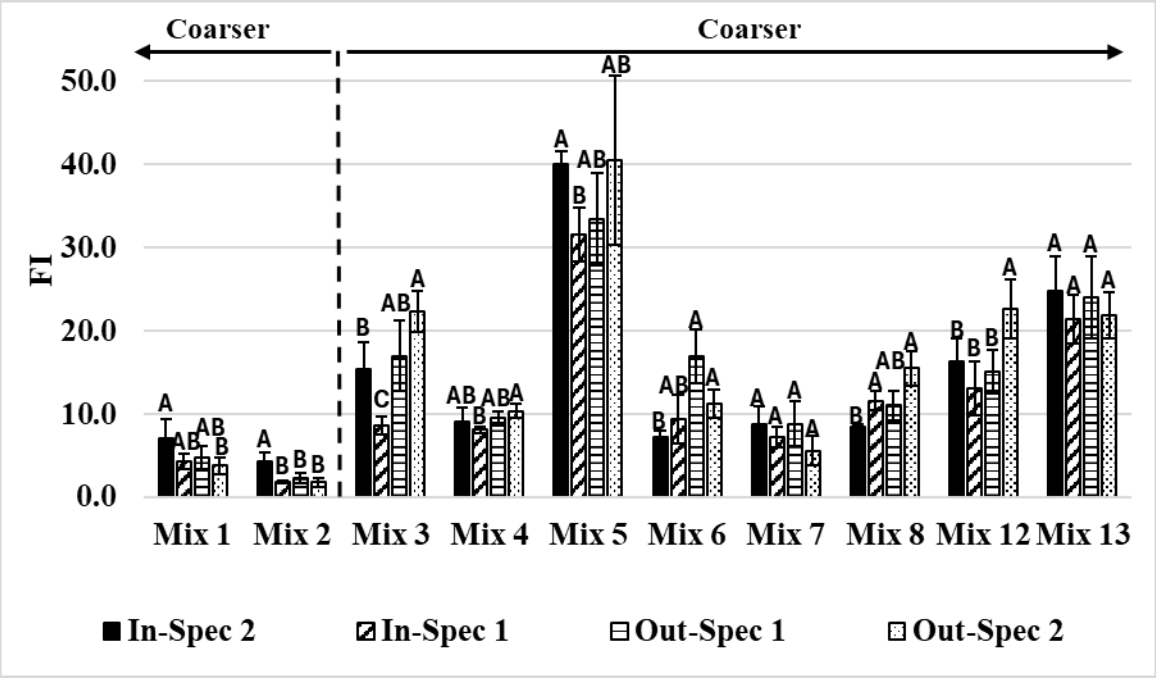

Figure 38. I-FIT Results

5.3.2.6 Low-Temperature Cracking Resistance

Figure 39 presents the fracture energy results obtained from the DCT test and group analysis of In-Spec and Out-Spec mixtures from 10 mix designs. For both upper and lower limit mix designs, adjusting the gradation coarser from In-Spec 2 to Out-Spec 2 generally increased the fracture energy. However, the group analysis showed no significant increase between all mixtures. In summary, increasing gradation coarseness within the evaluated volumetric range did not significantly improve the low-temperature cracking resistance of airfield asphalt mixtures. This outcome is consistent with the observation that low-temperature cracking resistance is primarily a function of asphalt binder low-temperature stiffness and relaxation properties, as well as aggregate type, both of which were held consistent across the four gradation designs [40, 215].

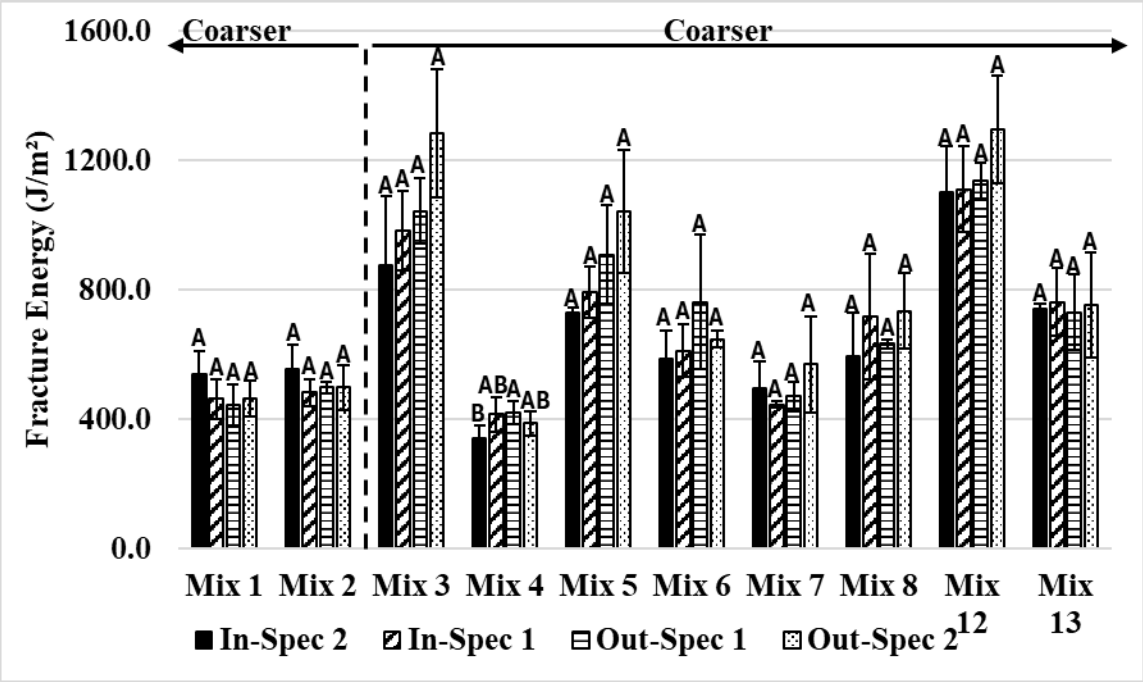

Figure 39. DCT Results

5.4 Summary

This chapter evaluated the effect of gradation changes on the volumetric properties and laboratory performance of airfield asphalt mixtures. In-Spec 2 gradations were adjusted toward the center of the P-401 gradation limits, whereas Out-Spec 2 gradations were positioned further outside the gradation limits. Asphalt binder was maintained within ±0.2% between In-Spec 2 and Out- Spec 2. Note that, unlike the In-Spec 1 and Out-Spec 1, the volumetric properties of the In-Spec 2 and Out-Spec 2 changed as a result of the gradation adjustments. Four gradations (In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2) were compared to quantify how controlled changes in aggregate gradation influence mixture behavior. Based on the test results, the following conclusions were drawn:

- In terms of volumetric properties, moving the gradation farther from the MDL, toward the lower or upper gradation limits, increased the VMA and V_a of airfield asphalt mixtures, given the same compaction efforts and binder contents.

- Regarding moisture resistance and durability, changing the gradation within the evaluated range generally had no significant impact on TSR or Cantabro mass loss, indicating that moisture resistance and durability performance were maintained.
- Regarding permeability performance, the mixture remained impermeable when changing the gradation within the evaluated range.
- Regarding rutting resistance, for mix designs near the lower limits, shifting the gradation toward a coarser aggregate structure generally decreased rut depth and, in several cases, significantly enhanced rutting resistance. For mix designs near the upper limits, rutting resistance remained consistent across different gradation adjustments.
- Regarding intermediate-temperature cracking resistance, IDEAL-CT and I-FIT results demonstrated that coarser gradation maintained or improved the intermediate-temperature cracking resistance of airfield asphalt mixtures. Several mix designs demonstrated statistically significant improvements.
- Regarding low-temperature cracking resistance, DCT results showed that coarser gradation did not statistically significantly improve low-temperature cracking resistance.

The results indicate that the controlled gradation adjustments influence the volumetric structure of asphalt mixtures and have the potential to improve rutting and intermediate-temperature cracking resistance, while leaving durability, permeability, moisture resistance, and low-temperature cracking performance largely unaffected. These observations provide the foundation for the predictive modeling development presented in Chapter 6.

CHAPTER 6 PREDICTION MODELS OF VOLUMETRIC PROPERTIES AND PERFORMANCE

The previous chapter demonstrated that gradation changes can influence the volumetric properties and the rutting and intermediate-temperature cracking resistance of airfield asphalt mixtures. This chapter focuses on developing prediction models to quantitatively characterize these relationships. Four prediction models were developed to characterize the effects of aggregate gradation change on (1) VMA, (2) rut depth, (3) CT_{Index}, and (4) FI. This chapter is organized into six sections. Section 6.1 describes the methodology for model development. Section 6.2 presents two VMA prediction models. Section 6.3 introduces the rut depth prediction model. Section 6.4 presents intermediate-temperature cracking prediction models based on CT_{Index} and FI. Section 6.5 summarizes the performance of the developed models. Finally, Section 6.6 presents the web-based decision tool developed to support the implementation of the developed models in mixture design and specification evaluation.

6.1 Methodology for Model Development

This section presents the detailed methodology for developing and validating prediction models, as shown in **Figure 40**. The modeling procedure includes three sequential steps: (1) database selection, (2) statistical modeling and validation, and (3) back-transformation process.

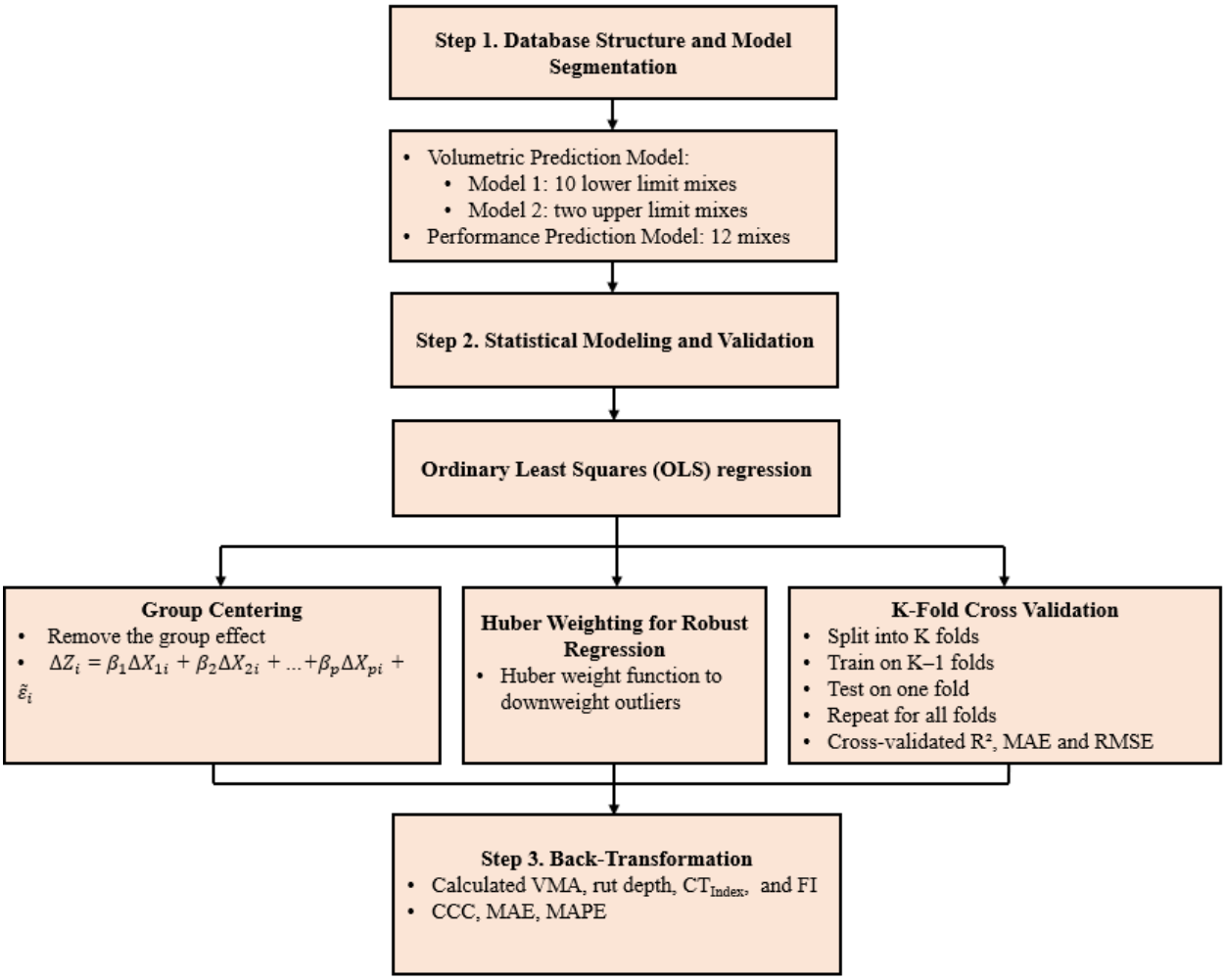

Figure 40. Workflow for the Development of the Prediction Models

6.1.1 Step 1: Data Structure and Model Segmentation

The modeling database was established based on the laboratory results obtained in this study. Among the 13 P-401 mix designs evaluated, 12 were designed with an NMAS of 12.5 mm, while one was designed with an NMAS of 19.0 mm. To maintain consistency in aggregate packing characteristics, gradation limits, and volumetric properties, only the 12 mix designs with a 12.5 mm NMAS were selected for model development. The database includes gradation, aggregate properties (including G_{sb} and FAA), volumetric properties, and laboratory performance test results.

For the volumetric property prediction model, findings from the previous chapter indicate that mixtures designed toward the lower and upper gradation limits exhibited opposite VMA trends. Because a single model could not represent both trends, two separate models were considered.

- The first model was developed using the two upper limit mixes with 18 observations, including In-Spec 1, Out-Spec 1, In-Spec 2, and Out-Spec 2 gradations, along with additional gradations from the adjustment process that failed to meet design requirements but were retained for model development. The model is implemented for the fine-graded mixes with gradation above the MDL, except for No. 200, for which moving the gradation from the upper limit toward the MDL decreases VMA.
- The second model was developed using the 10 lower limit mixes with 52 observations, including the same In-Spec and Out-Spec gradations and additional adjustment phase gradations retained for model development. The model is implemented for S-shaped gradation with percent aggregate passing at fine-fraction sieve sizes below MDL, or for coarse-graded gradation, for which shifting the gradation coarser toward the lower limits increases VMA and

For both volumetric prediction models, the response variable was the change in measured VMA, and the predictor variables included percent aggregate passing at the from sieve 1/2" to No.200, Bailey method parameters (PCS, New CA Ratio, and New FAc Ratio), and aggregate properties (G_{sb} and FAA)

For rutting and intermediate-temperature cracking, the same trend was observed when the gradation was coarser for both the upper and lower limit mixes. Therefore, the performance prediction models used data from all 12 airport mix designs, with 42 observations for rutting prediction models and with 39 observations for cracking prediction models. Predictor variables

were consistent with those used in the VMA models, including sieve sizes and aggregate properties (G_{sb} and FAA). Response variables were the measured rut depth, CT_{Index}, or FI, depending on the model.

6.1.2 Step 2: Model Development and Validation

6.1.2.1 Model Form Selection

Both linear and nonlinear models were evaluated to identify the most appropriate model form. Nonlinear models did not improve prediction accuracy compared to linear regression models and typically require substantially larger datasets to avoid overfitting and to ensure stable parameter estimation. Therefore, with a moderate number of gradations, linear regression models produced more consistent and reliable predictions. For these reasons, Ordinary Least Squares (OLS) regression was selected as the final model for deployment.

The true prediction model for the volumetric and performance of any mixture *i* within group *g* was expressed in **Equation 17**.

$$Z_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots \beta_p X_{pi} + \alpha_{g(i)} + \varepsilon_i \qquad (17)$$

Where:

Z_i = The response variable (VMA, rut depth, CT_{Index}, or FI) for mixture *i*.

X_{ji} = The *j*-th predictor variable of mixture *i* for *j* = 1…, p. Predictors include aggregate gradation (aggregate sieve sizes or Bailey method parameters) and G_{sb} and FAA of the combined aggregate blend.

β_j = Regression coefficients associated with the predictor X_{ji} for *j* = 1…, p

$\alpha_{g(i)}$ = Group-specific baseline effect for P-401 mix design *g*, which is constant for all mixtures from the same P-401 mix design *g*

ε_i = Random error term representing natural laboratory and material variability

6.1.2.2 Group-Centering to Remove Group Effects

Mixtures from the same P-401 mix design exhibit common baseline characteristics, such as aggregate source, binder type, and production conditions. The differences in baseline produce systematic differences in volumetric and performance properties among different P-401 mix designs. If left uncorrected, it could obscure the true relationship between gradation changes and mixture responses. To better isolate the effect of gradation changes, the group-centered method was used to remove design-specific baseline differences.

The group-centered method was implemented following **Equations 18** to **22. Equation 18** presents the group mean form of the true model. Note that the group effect $\alpha_{g(i)}$ was unchanged across all mixtures in the same group *g*, so it is unchanged by averaging.

$$\bar{Z}_g = \beta_1 \bar{X}_{1g} + \beta_2 \bar{X}_{2g} + \cdots + \beta_p \bar{X}_{pg} + \alpha_{g(i)} + \bar{\varepsilon}_g \qquad (18)$$

Where:

$\bar{Z}_g$= Mean response for all mixtures from P-401 mix design *g*

$\bar{X}_{jg}$ = Mean of the predictor variable X_j for all mixtures from P-401 mix design *g*

β_j = Regression coefficients associated with $\bar{X}_{jg}$ for j = 1…, p

$\alpha_{g(i)}$ = Mean of group-specific baseline effect for all mixtures from P-401 mix design *g*,

$\bar{\varepsilon}_g$= Mean of the error terms for group *g*

Equation 19 was obtained by subtracting **Equation** 18 from **Equation 17** to eliminate the group effect $\alpha_{g(i)}$.

$$Z_i - \bar{Z}_g = \left(\beta_1 X_{1i} + \beta_2 X_{2i} + \cdots \beta_p X_{pi} + \alpha_{g(i)} + \varepsilon_i\right) \quad (19)$$

$$- \left(\beta_1 \bar{X}_{1g} + \beta_2 \bar{X}_{2g} + \cdots + \beta_p \bar{X}_{pg} + \alpha_{g(i)} + \bar{\varepsilon}_g\right)$$

$$= \beta_1 (X_{1i} - \bar{X}_{1g}) + \beta_2 (X_{2i} - \bar{X}_{2g}) + \ldots + \beta_p (X_{pi} - \bar{X}_{pg})$$

$$+ (\alpha_{g(i)} - \alpha_{g(i)}) + (\varepsilon_i - \bar{\varepsilon}_g)$$

Equations 20 and **21** present the group-centered response and group-centered predictor variables. These definitions were substituted into **Equation 19** to yield a simplified group-centered equation, as shown in **Equation 22**. This equation is the model form fitted by OLS regression for all prediction models in the study. The fitted coefficients capture the within-group relationship between changes in gradation and changes in the mixture response.

$$\Delta Z_i = Z_i - \bar{Z}_g \quad (20)$$

$$\Delta X_{ji} = X_{ji} - \bar{X}_{jg} \quad (21)$$

$$\Delta Z_i = \beta_1 \Delta X_{1i} + \beta_2 \Delta X_{2i} + \ldots + \beta_p \Delta X_{pi} + \tilde{\varepsilon}_i \quad (22)$$

Where:

ΔZ_i = Group-centered response of mixture *i,* presenting the deviation of mixture *i* response from its group mean.

ΔX_{ji} = Group-centered predictor X_j of mixture i for $j = 1\ldots, p$

$\tilde{\varepsilon}_i = \varepsilon_i - \bar{\varepsilon}_g$ = Group-centered error term

6.1.2.3 Robust Regression Huber Weighting

Following the group-centered process, the next step involved a robust method for handling outliers. OLS is sensitive to outliers because a few extreme residuals can disproportionately influence the fitted coefficients. In the study, the Huber weight function was used to downweight large residuals, preventing outliers from dominating the regression fit. The Huber weighting procedure proceeds

in four steps. Step (a) involved fitting an initial OLS regression model and computing residuals for each mixture using **Equation 23**.

$$r_i = \Delta Z_i - \widehat{\Delta Z}_i \quad (23)$$

Where:

r_i = Residual of mixture *i*

ΔZ_i = Measured group-centered response of mixture *i*

$\widehat{\Delta Z}_i$ = Predicted group-centered response of mixture *i*

Step (b) focuses on standardizing the residuals. First, the median absolute deviation (MAD) of the residuals, which is less sensitive to outliers than the standard deviation, was calculated using **Equation 24**. Next, the MAD was divided by a normalization factor of 0.6745 (the 75th percentile of the standard normal distribution) to be consistent with the standard deviation under normal distribution, using **Equation 25**. Finally, each residual was divided by this estimate to produce standardized residuals using **Equation 26**.

$$\text{MAD} = median(|r_i - median(r_i)|) \quad (24)$$

$$\hat{\sigma} = \frac{MAD}{0.6745} \quad (25)$$

$$u_i = \frac{r_i}{\hat{\sigma}} \quad (26)$$

Step (c) focused on calculating the Huber weight ω_i using the tuning constant c, which was used to identify the outliers, following **Equation 27**. The tuning constant c was set to 1.345, which yielded 95% efficiency when the data were normally distributed, while still protecting against outliers [216, 217]. Observations with $|u_i| \leq 1.345$ (small residuals) receive full weight; observations with $|u_i| > 1.345$ (larger scaled residuals) are downweighted in proportion to their deviation.

$$\omega_i = \begin{cases} 1 \; if \; |u_i| \leq c \\ \frac{c}{|u_i|} \; if \; |u_i| > c \end{cases} \quad (27)$$

Where:

ω_i = Huber weight function of mixture *i*

u_i = Scaled residual of mixture *i*

c = Tuning constant threshold (1.345)

In step (d), the regression model is re-run using the computed weights $(\omega_1, \omega_2, \ldots, \omega_n)$. It produces a new set of residuals, weights, and coefficients. This process repeats until the model coefficients converge.

6.1.2.4 K-Fold Cross-Validation

Finally, to ensure the predictive model performs reliably on new and unseen data, K-fold cross-validation was used. Instead of fitting the model once and checking accuracy on the same database, cross-validation separated the data into K subsets (folds). In each iteration, K–1 subsets were used for training, and one for testing. The process was repeated K times, so that each subset served as the test set once. In the study, the database was built from different airport mix designs, which could serve as different folds. This approach prevented correlated gradation designs from the same airport mix design from appearing in both training and testing folds.

The number of folds matches the number of airports mix designs in each model: two folds for the first volumetric prediction model associated with two upper limit mix designs, 10 folds for the second volumetric prediction model associated with 10 lower limit mix designs, and 12 folds for performance prediction models associated with 12 mix designs.

The performance metrics, including the R^2, RMSE, and MAE as shown in **Equations 28**, **29,** and 30**,** respectively, were used to evaluate the model performance. Model performance was evaluated once across all cross-validated predictions, rather than averaging fold-level metrics, to

ensure an unbiased and statistically consistent assessment. The R^2 value measures the proportion of the total variance in the observed data that the model explains. A higher R^2 indicates stronger agreement between predicted and measured values, implying better model performance. RMSE and MAE provide straightforward interpretations of the average deviation between predicted and measured data. MAE is less sensitive to extreme outliers than RMSE. Lower RMSE and MAE values indicate better model performance.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\Delta Z_i - \widehat{\Delta Z}_i)^2}{\sum_{i=1}^{n}(\Delta Z_i - \overline{\Delta Z})^2} \tag{28}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\Delta Z_i - \widehat{\Delta Z}_i)^2} \tag{29}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | \Delta Z_i - \widehat{\Delta Z}_i | \tag{30}$$

Where:

ΔZ_i = The measured group-centered response of mixture *i*,

$\widehat{\Delta Z}_i$ = Cross-validated predicted group-centered response of mixture *i*,

$\overline{\Delta Z}$ = Mean of all measured group-centered responses

n = Number of mixtures in the database

Note that it was expected that several variables would be highly correlated with each other (multicollinearity) due to the nature of the gradation adjustment. Changing the percentage aggregate passing one sieve resulted in a change in the percentage passing of adjacent sieves. Multicollinearity can inflate the variances of individual coefficients and affect their statistical significance (p-values); therefore, the coefficients in the model should not be interpreted independently. However, when the primary objective is to develop a reliable prediction model

rather than infer individual coefficients, it is not necessary to reduce or eliminate multicollinearity between variables [218]. Therefore, all aggregate sieve sizes were included in the regression to maintain consistency with the sieve structure of the aggregate gradation.

6.1.3 Step 3: Back Transformation

The model was estimated on group-centered data to isolate within-group variation. However, practitioners interpret and use actual calculated VMA and laboratory performance rather than group-centered values. Step 3 in **Figure 40** involves back-transformation to calculate VMA, rut depth, CT_{Index}, and FI from their associated predicted group-centered responses, using **Equation 31**.

$$\hat{Z}_i = \Delta Z_i + \bar{Z}_g \quad (31)$$

Where:

$\hat{Z}_i$= The calculated (back-transformed) response in original scale

ΔZ_i = Group-centered response obtained from the prediction model

$\bar{Z}_g$ = Group mean of the response for group g

Evaluating performance only on the group-centered scale reflects how well the model explains within-group changes, but not how accurately it predicts the original values. Additional performance metrics, including MAE, Concordance Correlation Coefficient (CCC), and Mean Absolute Percentage Error (MAPE), as defined in **Equations 30**, **32, 33**, respectively, were used to better illustrate the agreement between measured and calculated VMA, rut depth, $CT_{Index,}$ and FI. The CCC evaluates both trend and magnitude agreement between measured and calculated values. Since baseline response levels varied across P-401 mix designs, the CCC was calculated separately within each group rather than on the combined dataset to prevent artificial inflation or deflation of agreement caused by between-group mean shifts. A higher average CCC indicates

stronger agreement. MAPE is the average absolute percentage difference between measured and calculated values. Lower MAPE indicates smaller prediction errors and, therefore, better predictive performance.

$$r_c = \frac{2rS_Z S_{\hat{Z}}}{S_Z^2 + S_{\hat{Z}}^2 + (\bar{Z} - \bar{\hat{Z}})^2} \quad (32)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{Z_i - \hat{Z}_i}{Z_i} \right| \quad (33)$$

$$r = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z})(\hat{Z}_i - \bar{\hat{Z}})}{\sqrt{\sum_{i=1}^{n} (Z_i - \bar{Z})^2} \sqrt{\sum_{i=1}^{n} (\hat{Z}_i - \bar{\hat{Z}})^2}} \quad (34)$$

Where:

r_c = Concordance correlation coefficient (CCC)

S_Z and $S_{\hat{Z}}$ = Standard deviation of measured and calculated variables, respectively

$\bar{Z}$ and $\bar{\hat{Z}}$= Mean of measured and calculated variables, respectively

r = Pearson correlation between calculated and measured variables, as shown in **Equation 34**.

Z_i and $\hat{Z}_i$ = The measured and calculated variables, respectively

n = Number of mixtures

6.1.4 Web-Based Tool Development

The developed models were then deployed as a web-based tool, with the workflow presented in **Figure 41**. The tool assumes that these mixtures were designed with the same aggregate source, binder type, and binder content. The process began with the input phase, during which the tool collected information from the existing tested mixtures, including gradation, G_{sb}, and, when available, FAA. Depending on the desired prediction targets, additional inputs, including VMA, rut depth, CT_{Index}, and FI from the same existing mixtures, were incorporated. Afterward, the

gradation, G_{sb}, and optional FAA for the new trial were entered into the tool. During the deployment phase, all inputs were processed through the developed prediction models to predict the group-centered response of the new trial. Back-transformation was then applied to convert the predicted group-centered responses back into the original calculated responses. As a result, the final outputs were expressed as the calculated VMA, rut depth, CT_{Index}, and FI of the new mix design.

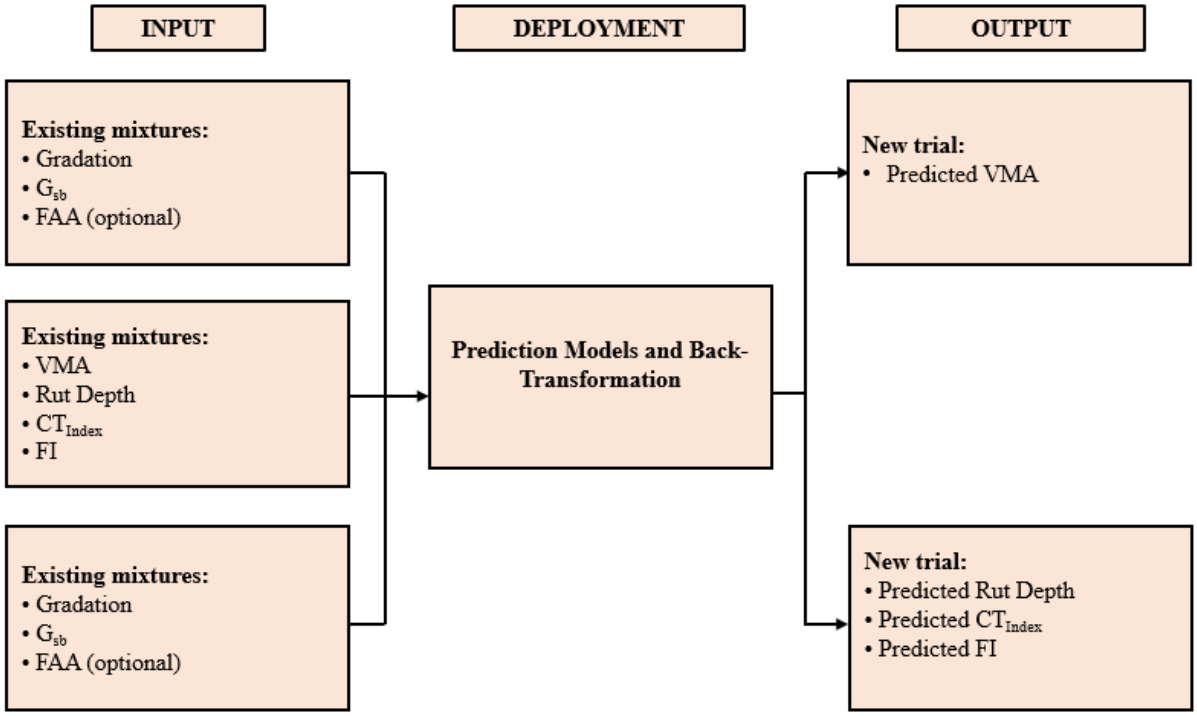

Figure 41. Workflow of the Web-Based Tool.

6.2 Volumetric Properties Prediction Model

Changes in aggregate gradation can influence VMA, with direction depending on whether the gradation falls near the upper or lower P-401 limits. The two sections below describe the two prediction models developed for volumetric properties.

6.2.1 First Volumetric Properties Prediction Model

The first volumetric prediction model was developed for fine-graded gradations with gradation curves located above the MDL. For this model, Bailey method parameters were selected as predictors instead of the percentage passing at individual sieves. This choice was justified for two primary reasons. First, all mixtures included in this model were fine-graded, ensuring that Bailey Method parameters could be defined consistently across all gradation designs. Second, the prediction model was constructed using only two upper limit mixes, resulting in a limited number of observations. Using individual sieve sizes as predictors would have introduced a large number of variables relative to the available sample size, increasing the risk of overfitting and decreasing model stability. The resulting OLS regression equation for the first volumetric prediction model is presented in **Equation 35.**

$$\Delta VMA = 0.2158\Delta PCS + 2.32\Delta New\ CA\ Ratio - 1.55\Delta New\ FA_c\ Ratio \quad (35)$$

Where:

ΔVMA = Group-centered value of VMA

$\Delta No.PCS,\ \Delta New\ CA\ Ratio,\ \Delta New\ FA_c\ Ratio$ = Group-centered value of Bailey method parameters for fine-graded mixtures, including PCS, New CA ratio, and New FA_c ratios (**Table 15**)

Figure 42 presents the distribution of residuals against fitted values of ΔVMA in the first volumetric prediction model. All data points shown represent predictions obtained from the validation sets in the K-fold cross-validation. As shown, the random scatter of residuals against fitted values confirms that a linear model was mathematically appropriate and successfully captured the relationship between the change of gradation and ΔVMA. Furthermore, the vertical spread of residuals remained relatively uniform across the central cluster of fitted values,

suggesting that the critical assumption of homoscedasticity was generally met without evidence of severe heteroscedasticity.

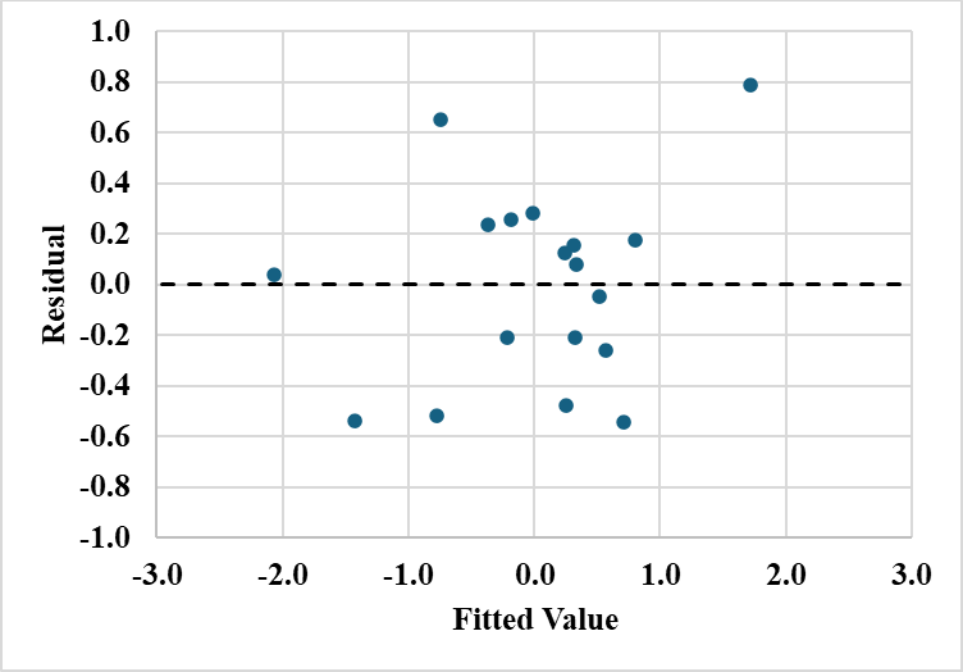

Figure 42. Residuals and Fitted Values of ΔVMA in First Volumetric Prediction Model

R^2, MAE, and RMSE computed based on the validation sets were used to evaluate the model performance, with the results shown in **Table 30**. MAE of 0.31 and RMSE of 0.40 indicate prediction errors of about 0.31–0.40%, which are small relative to the typical VMA range of 14%–18%. The cross-validated R^2 of 86% indicates that the model explained the majority of within-group variation.

Table 30. Statistical Metrics of Predicted and Measured ΔVMA from First Volumetric Prediction Model

MAE	RMSE	K-Fold R^2
0.31	0.40	0.86

Figure 43 presents the measured and calculated VMA values obtained from the back-transformation process. As shown, most of the data points clustered closely around the line of

equality, indicating that the VMA was accurately predicted across two airport mix designs. An average CCC of 0.96 indicated a strong agreement between the measured and calculated VMA values. The MAE of 0.2% VMA indicates that, on average, the calculated VMA differed from the measured value by 0.2%. Despite these promising results, it is crucial to emphasize that the model was developed from only two mix designs, which limits its generalizability. Therefore, the first volumetric prediction model should be applied with caution.

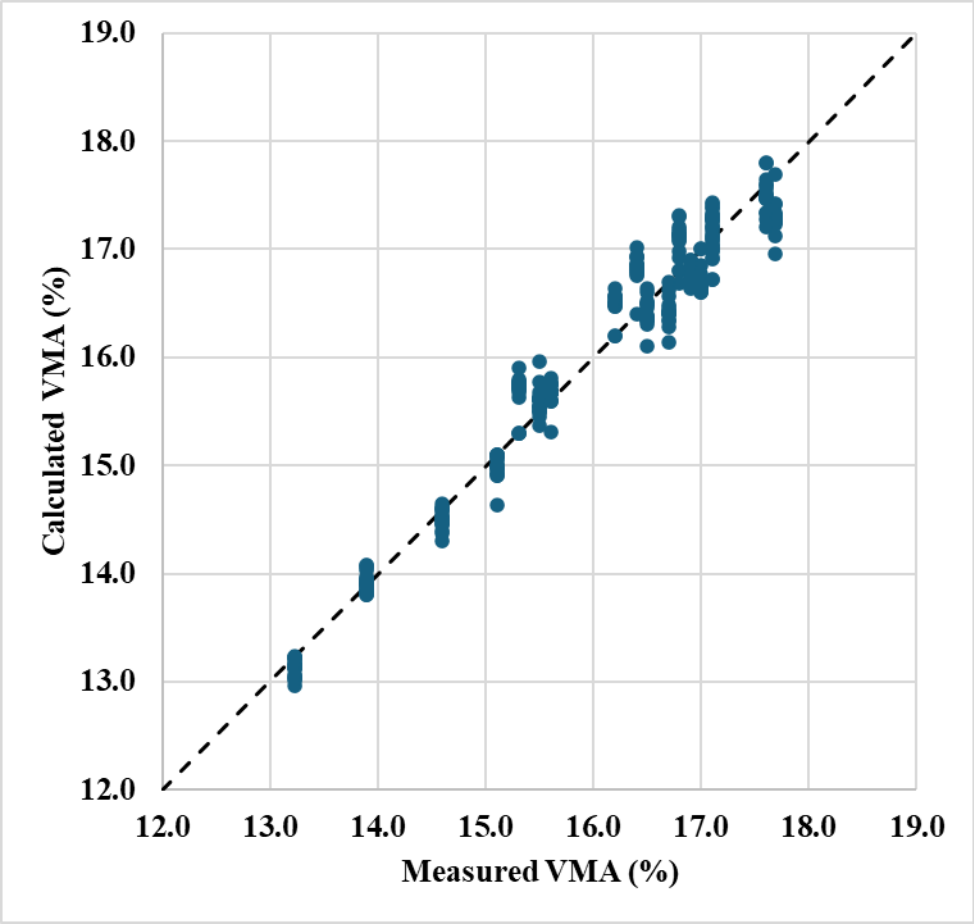

Figure 43. Measured and Calculated VMA from First Volumetric Prediction Model

6.2.2 Second Volumetric Properties Prediction Model

The second model for volumetric prediction was implemented for S-shaped gradation with fine-fraction sieves below MDL and for coarse-graded gradation. The OLS regression equation of the second VMA prediction model is shown in **Equation 36**.

$$\begin{aligned}\Delta VMA = {} & 0.0282 + 0.315\Delta No.\,1/2 - 0.0271\Delta No.\,3/8 - 0.0037\Delta No.\,4 \\ & - 0.044\Delta No.\,8 + 0.099\Delta No.\,16 - 0.324\Delta No.\,30 + 0.970\Delta No.\,50 \\ & - 0.430\Delta No.\,100 - 1.074\Delta No.\,200 + 87.7\Delta G_{sb} + 0.135\Delta FAA\end{aligned} \quad (36)$$

Where:

$\Delta VMA =$ Group-centered value of VMA

$\Delta No.\,1/2, \Delta No.\,3/8 \ldots . \Delta No.\,200 =$ Group-centered values of percentage aggregate passing at No.1/2, No.3/8…No.200 sieve

$\Delta G_{sb} =$ Centered- group value of specific gravity of aggregate blend

$\Delta FAA =$ Centered- group value of FAA of the blend

Figure 44 presents the distribution of residuals against fitted values of ΔVMA in the second volumetric prediction model. All data points shown represent predictions obtained from the validation sets in the K-fold cross-validation. As shown, the random scatter of residuals against fitted values confirms that a linear model was mathematically appropriate and successfully captured the relationship between the change of gradation and ΔVMA. Furthermore, the vertical spread of residuals remained relatively uniform across the central cluster of fitted values, suggesting that the critical assumption of homoscedasticity was generally met without evidence of severe heteroscedasticity.

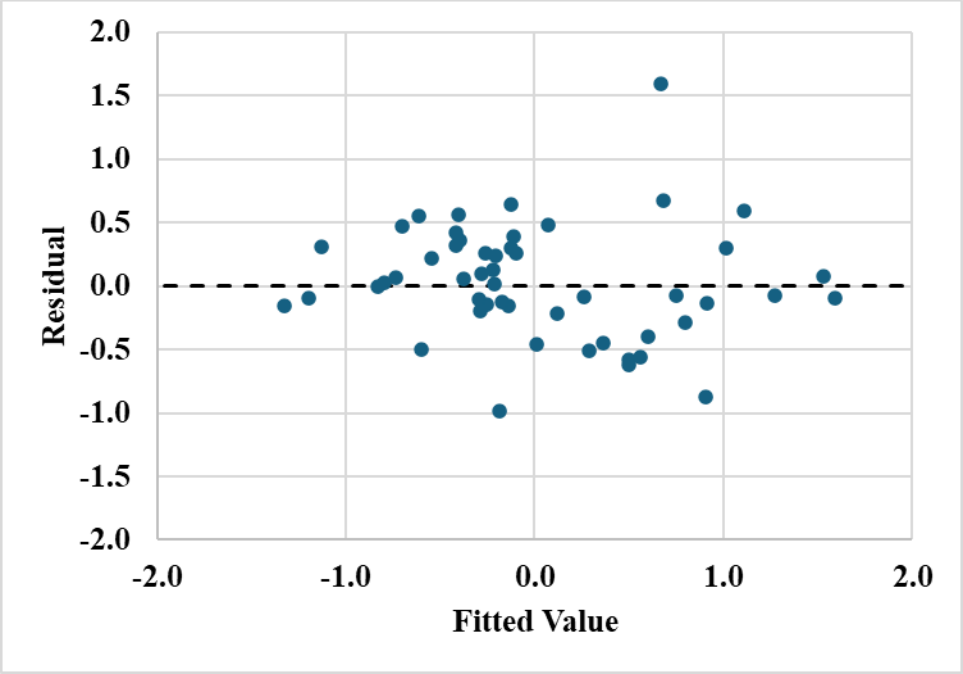

Figure 44. Residuals and Fitted Values of ΔVMA in Second Volumetric Prediction Model

R^2, MAE, and RMSE computed based on the validation sets were used to evaluate the model performance, as shown in **Table 31**. MAE of 0.30 and RMSE of 0.40 indicate prediction errors of about 0.30–0.40%, which are small relative to the typical VMA range of 14%–18%. The cross-validated R^2 of 69% confirmed that the model explained a substantial portion of within-group variation.

Table 31. Statistical Metrics of Predicted and Measured ΔVMA from Second Volumetric Prediction Model

MAE	RMSE	K-Fold R^2
0.30	0.40	0.69

Figure 45 presents the calculated and predicted VMA obtained from the back transformation process. As shown, most of the data points clustered closely around the line of equality, indicating that the VMA was accurately predicted across a wide range of airport mix designs. An average CCC of 0.81 indicated a strong agreement relationship between measured and

calculated VMA. The MAE of 0.4% VMA indicates that, on average, the calculated VMA differed from the measured value by only 0.4%. This falls within the within-laboratory repeatability of VMA measurement of approximately 0.7% [219], indicating that model predictive accuracy is acceptable.

Although several data points still appear as outliers in the plot, this outcome was expected. The Huber method downweighed observations with large residuals but did not remove them from the analysis, leaving these mixtures visually distant from the line of equality even though their influence on the fitted coefficients was substantially reduced.

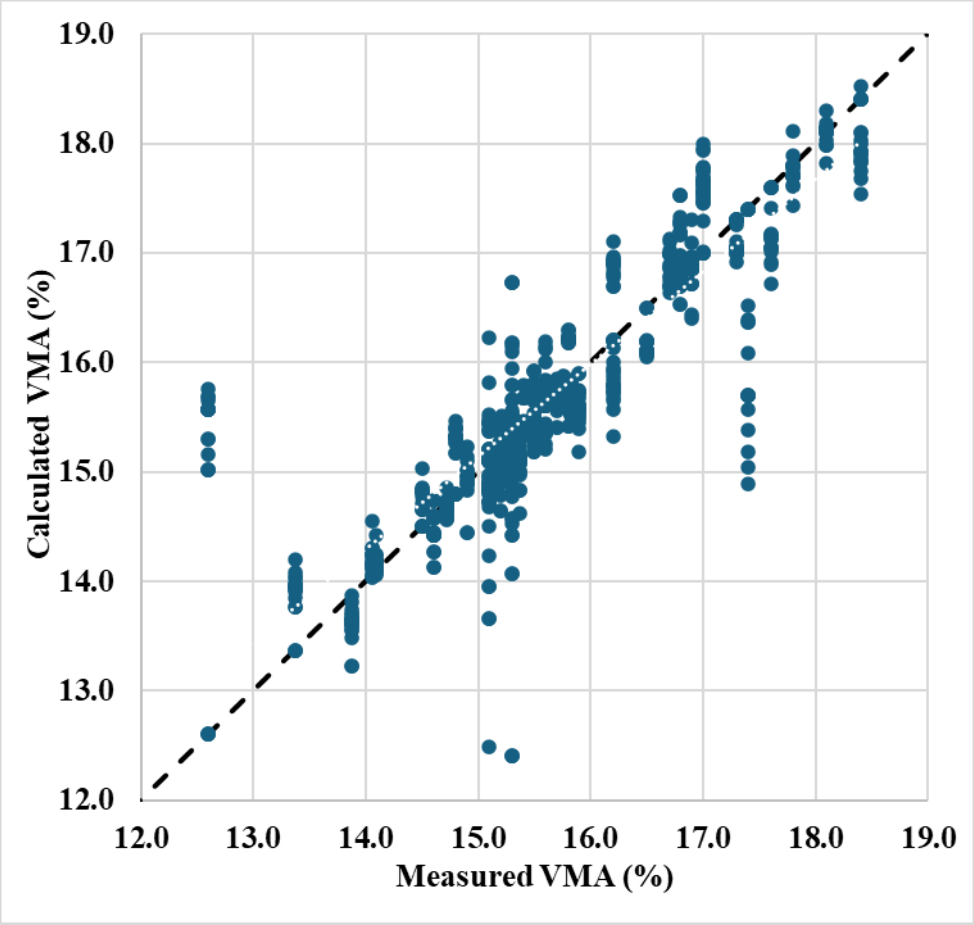

Figure 45. Measured and Calculated VMA from Second Volumetric Prediction Model

In addition, another version of the VMA prediction model was developed without FAA variables, as shown in **Equation 37**. **Table 32** provides the statistical metrics of the model without the FAA variable. The results show that the model without FAA variables performed similarly to the model with FAA variables.

$$\Delta\text{VMA} = 0.0260 + 0.362\Delta\text{No.}\,1/2 - 0.0477\Delta\text{No.}\,3/8 + 0.0018\Delta\text{No.}\,4 - 0.041\Delta\text{No.}\,8 + 0.092\Delta\text{No.}\,16 - 0.308\Delta\text{No.}\,30 + 0.859\Delta\text{No.}\,50 - 0.297\Delta\text{No.}\,100 - 1.109\Delta\text{No.}\,200 + 86\Delta G_{sb} \quad (37)$$

Table 32. Statistical Metrics of Predicted and Measured ΔVMA from Second Volumetric Prediction Model without FAA

MAE	RMSE	K-Fold R^2
0.30	0.40	0.69

6.3 Rut Depth Prediction Model

The OLS regression equation of the rutting prediction model is shown in **Equation 38.**

$$\Delta\text{Rut depth} = -0.0091 + 0.331\Delta\text{No.}\frac{1}{2} - 0.326\Delta\text{No.}\frac{3}{8} + 0.285\Delta\text{No.}\,4 - 0.174\Delta\text{No.}\,8 - 0.734\Delta\text{No.}\,16 + 1.709\Delta\text{No.}\,30 - 0.893\Delta\text{No.}\,50 - 0.212\Delta\text{No.}\,100 - 0.283\Delta\text{No.}\,200 - 52.0\Delta G_{sb} + 0.580\Delta\text{FAA} \quad (38)$$

Where:

$\Delta\text{Rut depth}$ = Centered- group value of rut depth

$\Delta\text{No.}\,1/2, \Delta\text{No.}\,3/8 \ldots \Delta\text{No.}\,200$ = Centered- group value of percentage aggregate passing at No.1/2, No.3/8…No.200 sieves

ΔG_{sb} = Centered- group value of specific gravity of aggregate blend

ΔFAA = Centered- group value of FAA of the blend

Figure 46 presents distribution of residuals against fitted values of ΔRut depth. All data points shown represent predictions obtained from the validation sets in the K-fold cross-validation. As shown, the random scatter of residuals against fitted values confirms that a linear model was mathematically appropriate and successfully captured the relationship between the change of gradation and ΔRut depth. Furthermore, the vertical spread of residuals remained relatively uniform across the central cluster of fitted values, suggesting that the critical assumption of homoscedasticity was generally met without evidence of severe heteroscedasticity.

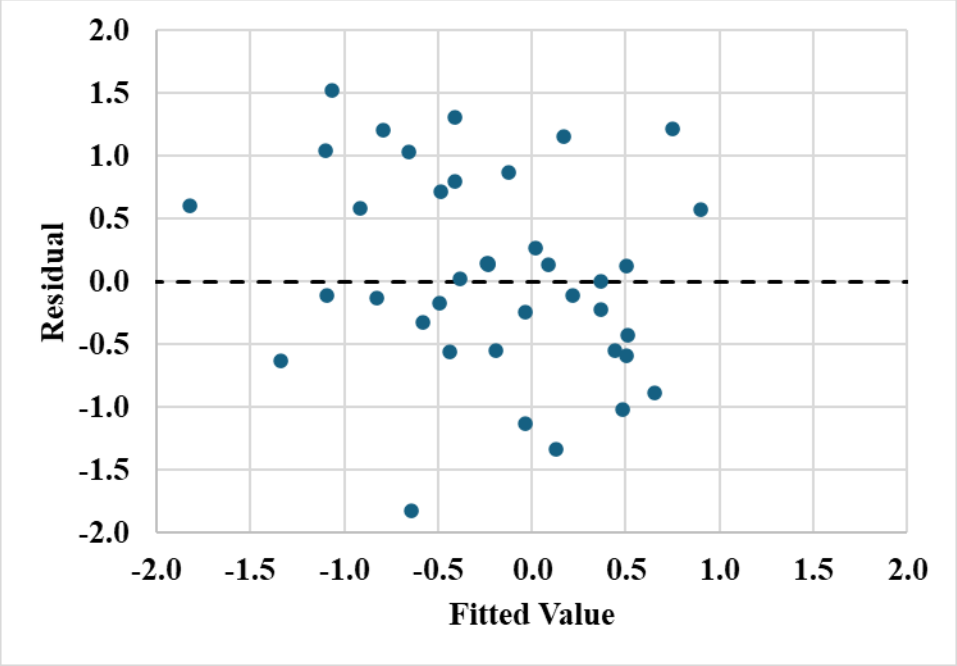

Figure 46. Residuals and Fitted Values of ΔRut Depth

R^2, MAE, and RMSE computed from validation sets were used to evaluate model performance, as shown in **Table 33**. MAE of 0.67 mm and RMSE of 0.81 mm indicate that the typical prediction error was less than 1mm. Given that rut depth values range from about 4 to 10 mm, this level of error is acceptable for screening-level predictions during mix design adjustments. The cross-validated R^2 of 0.39 suggests that about 39% of the within-group variability in rut depth is explained by the changes of gradation and aggregate properties. Although this is lower than the

volumetric models, it is reasonable given the greater experimental variability and mechanical complexity associated with rutting performance.

Table 33. Statistical Metrics of Predicted and Measured ΔRut Depth

MAE	RMSE	K-Fold R^2
0.67	0.81	0.39

Figure 47 presents the measured and calculated rut depth obtained from the back transformation process. As shown, most of the data points clustered closely around the line of equality, indicating that the model accurately predicted rut depth across a wide range of airport mix designs. An average CCC of 0.53 indicated a moderate agreement between the measured and calculated rut depths. The MAE of 0.68 mm indicates that, on average, the calculated rut depth differed from the measured value by only 0.68 mm.

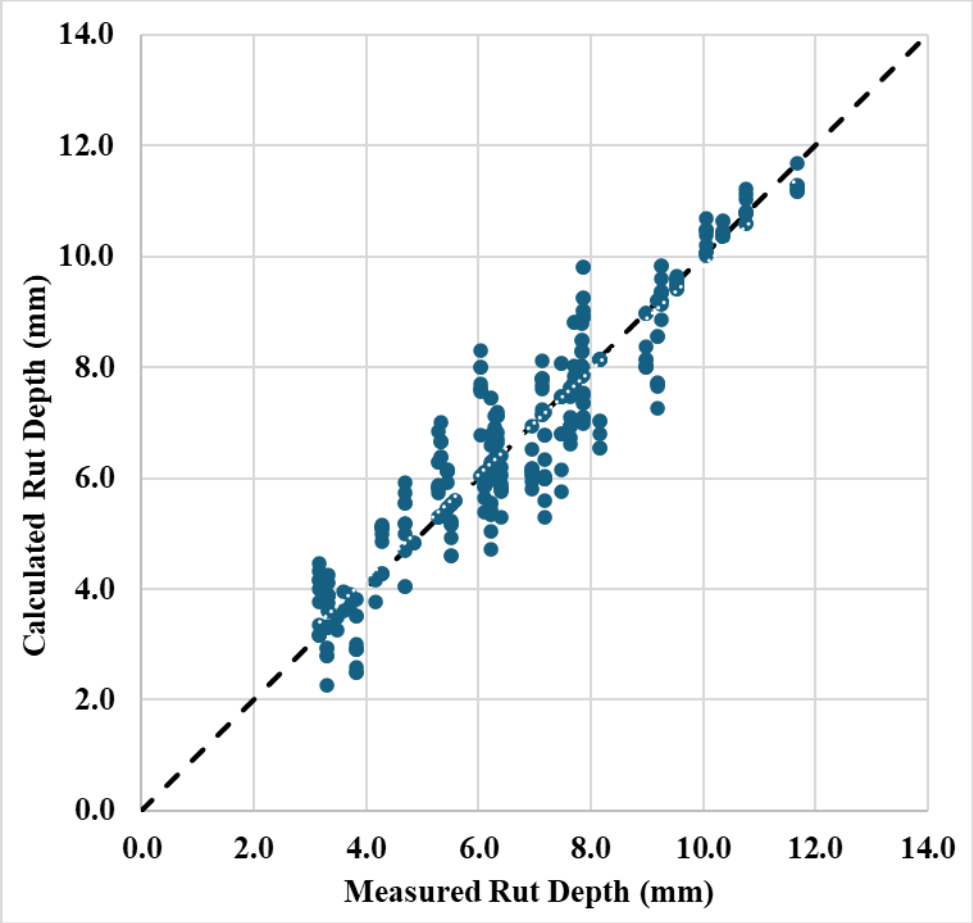

Figure 47. Measured and Calculated Rut Depth

Another version of the rut depth prediction model was developed without FAA variables, as shown in **Equation 39**.

Table 34 provides the statistical metrics of the model without the FAA variable. The results show that the model without the FAA variable provided slightly better performance compared with the model that included the FAA variable.

$$\Delta\text{Rut depth} = -0.0136 + 0.603\Delta\text{No.}\frac{1}{2} - 0.444\Delta\text{No.}\frac{3}{8} + 0.378\Delta\text{No. 4} - 0.192\Delta\text{No. 8} \quad (39)$$
$$- 0.989\Delta\text{No. 16} + 2.113\Delta\text{No. 30} - 1.370\Delta\text{No. 50} + 0.378\Delta\text{No. 100}$$
$$- 0.668\Delta\text{No. 200} - 35.1\Delta G_{sb}$$

Table 34. Statistical Metrics of Predicted and Measured ΔRut Depth from the Prediction Model without FAA

MAE	RMSE	K-Fold R^2
0.60	0.75	0.45

6.4 Intermediate-Temperature Cracking Prediction Model

Intermediate-temperature cracking resistance was assessed using both CT_{Index} and FI, and the following sections present the corresponding prediction models for each.

6.4.1 Intermediate-Temperature Cracking Prediction Model based on CT_{Index}

The OLS regression equation of the prediction model based on CT_{Index} is shown in **Equation 40**.

$$\Delta CT_{Index} = 0.04 + 4.05\Delta No.\frac{1}{2} - 4.66\Delta No.\frac{3}{8} + 27.68\Delta No.4 - 43.12\Delta No.8 + 16.1\Delta No.16 + 16.5\Delta No.30 - 32.2\Delta No.50 + 29.1\Delta No.100 - 17.0\Delta No.200 + 3050\Delta G_{sb} - 15.22\Delta FAA \quad (40)$$

Where:

ΔCT_{Index} = Group-centered CT_{Index}

$\Delta No.1/2, \Delta No.3/8 \ldots. \Delta No.200$ = Group- centered percentage aggregate passing at No.1/2, No.3/8…No.200 sieve

ΔG_{sb}= Group-centered specific gravity of aggregate blend

ΔFAA = Group-centered FAA of blend

Figure 48 presents the distribution of residuals against fitted values of ΔCT_{Index}. All data points shown represent predictions obtained from the validation sets in the K-fold cross-validation. As shown, the random scatter of residuals against fitted values confirms that a linear model was mathematically appropriate and successfully captured the relationship between the change of gradation and ΔCT_{Index}. Furthermore, the vertical spread of residuals remained relatively uniform

across the central cluster of fitted values, suggesting that the critical assumption of homoscedasticity was generally met without evidence of severe heteroscedasticity.

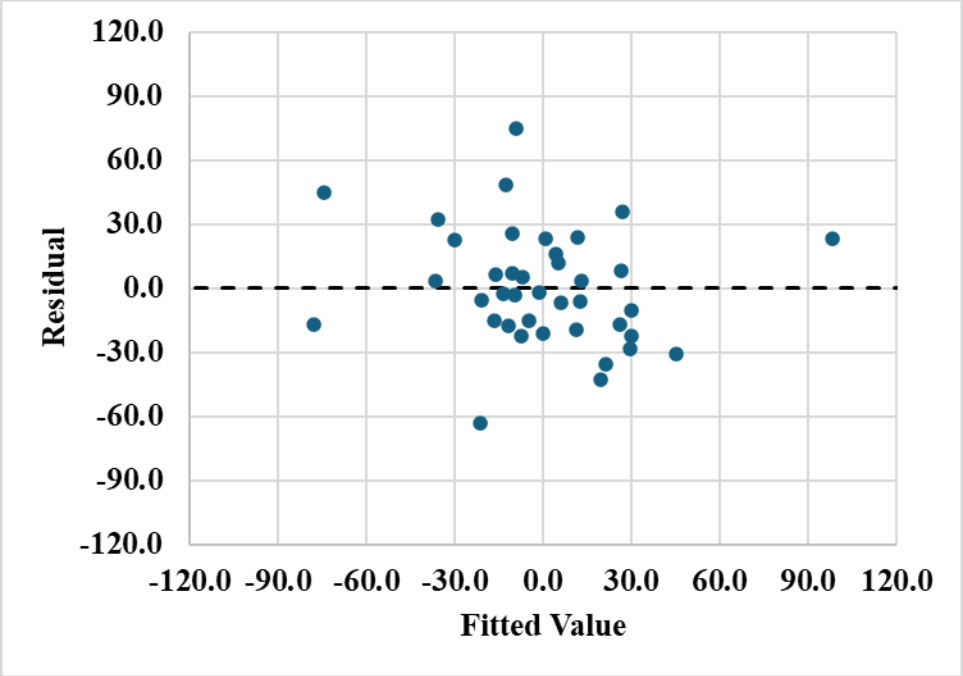

Figure 48. Residuals and Fitted Values of ΔCT_{Index}

R^2, MAE, and RMSE computed from validation sets were used to evaluate the model performance, as shown in **Table 35**. MAE of 16 and an RMSE of 20 indicate a typical prediction error of 16-20. Because CT_{Index} values in this dataset cover a wide range, these absolute errors should be interpreted in the context of inherent test variability. The cross-validated R^2 of 0.62 suggests that the model explains a substantial portion of the variability in unseen data. Given the mechanical complexity and variability in cracking resistance, the predictive performance of a performance prediction model is relatively strong.

Table 35. Statistical Metrics of Predicted and Measured ΔCT_{Index}

MAE	RMSE	K-Fold R^2
16.0	20.0	0.62

Figure 49 presents the measured and calculated CT_{Index} obtained from the back transformation process. As shown, most of the data points clustered closely around the line of equality, indicating that the CT_{Index} was accurately predicted across a wide range of airport mix designs. The average CCC of 0.6 indicated moderate agreement between the measured and calculated CT_{Index}. The MAPE of 17.7% was lower than the typical within-laboratory coefficient of variation (COV) of approximately 20% for the CT_{Index} test [220]. This comparison indicates that the error was lower than the natural experimental variability of the test method, and therefore, the predictive performance can be considered acceptable. In summary, the model performance captured the direction and magnitude of changes across airport mix designs.

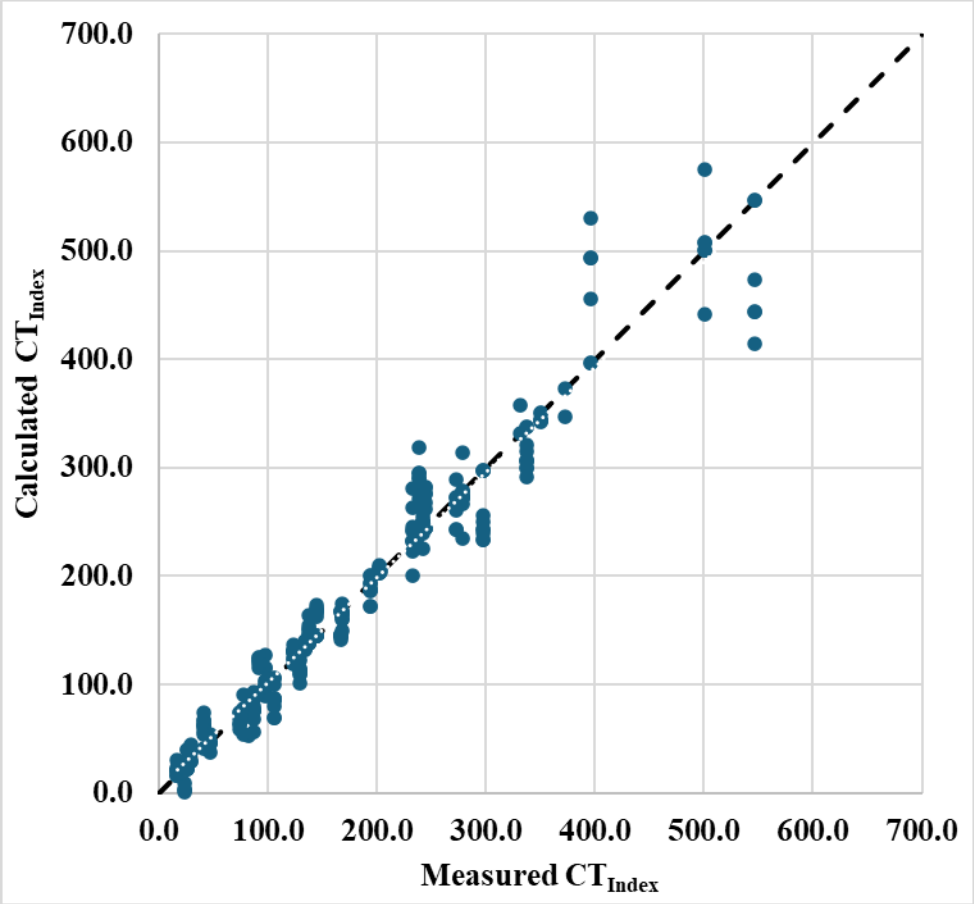

Figure 49. Measured and Calculated CT_{Index}

Another version of the CT_{Index} prediction model was developed without FAA variables, as shown in **Equation 41. Table 36** provides the statistical metrics of the model without the FAA variable. The results show that the model without the FAA variable performed slightly lower than the model that included the FAA variable, indicating that the FAA contributes additional explanatory value for CT_{Index} prediction.

$$\Delta CT_{Index} = -0.13 - 1.90\Delta No.\frac{1}{2} - 3.18\Delta No.\frac{3}{8} + 29.73\Delta No.4 - 47.35\Delta No.8 + 26.1\Delta No.16 + 2.3\Delta No.30 - 16.0\Delta No.50 + 10.3\Delta No.100 - 6.7\Delta No.200 + 2726\Delta G_{sb} \quad (41)$$

Table 36. Statistical Metrics of the CT_{Index} Prediction Model without FAA

MAE	RMSE	K-Fold R^2
18.8	23.5	0.50

6.4.2 Intermediate -Temperature Cracking Prediction Model based on FI

The OLS regression equation of the prediction model based on FI is shown in **Equation 42**.

$$\Delta FI = 0.016 + 0.625\Delta No.\frac{1}{2} - 0.431\Delta No.\frac{3}{8} + 0.044\Delta No.4 - 0.985\Delta No.8 - 0.18\Delta No.16 + 3.22\Delta No.30 - 4.61\Delta No.50 + 4.34\Delta No.100 - 3.01\Delta No.200 + 44.6\Delta G_{sb} + 0.03\Delta FAA \quad (42)$$

Where:

ΔFI = Group-centered FI

$\Delta No.1/2, \Delta No.3/8 \ldots. \Delta No.200$ = Group- centered percentage aggregate passing at No.1/2, No.3/8…No.200 sieve

ΔG_{sb}= Group-centered specific gravity of aggregate blend

ΔFAA = Group-centered FAA of blend

Figure 50 presents the distribution of residuals against fitted values of ΔFI. All data points shown represent predictions obtained from the validation sets in the K-fold cross-validation. As shown, the random scatter of residuals against fitted values confirms that a linear model was mathematically appropriate and successfully captured the relationship between the change of gradation and ΔFI. Furthermore, the vertical spread of residuals remained relatively uniform across the central cluster of fitted values, suggesting that the critical assumption of homoscedasticity was generally met without evidence of severe heteroscedasticity. However, a downward linear trend is visible in the scatter; the model tends to produce positive residuals (underestimation) for lower fitted values and negative residuals (overestimation) for higher fitted values. This pattern suggests that the model may be struggling to capture the full range of extreme values in the dataset.

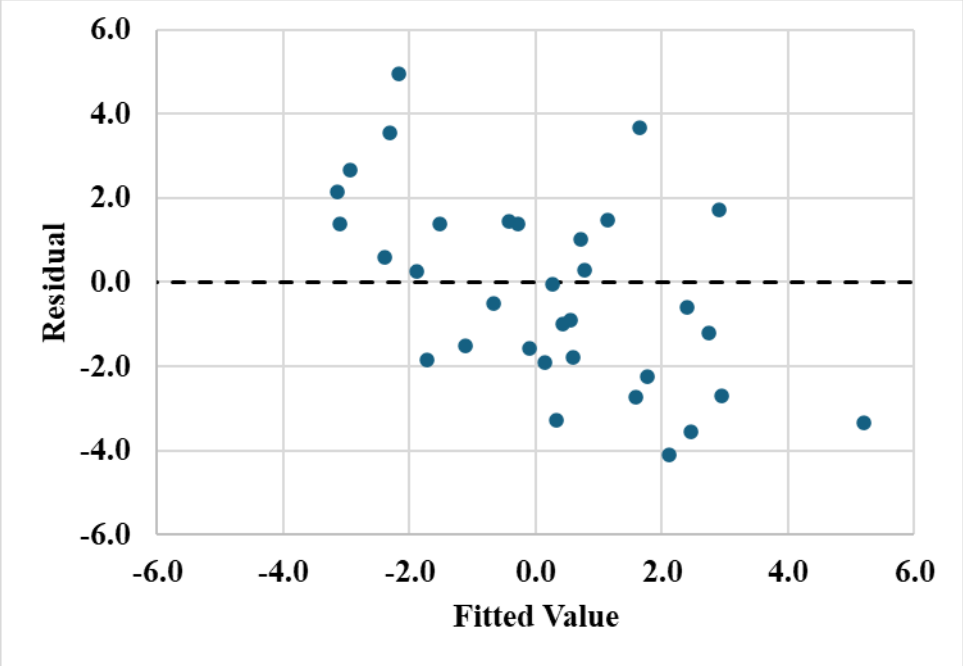

Figure 50. Residuals and Fitted Values of ΔFI

R^2, MAE, and RMSE computed from validation sets were used to evaluate model performance, as shown in **Table 37**. The cross-validated R^2 value close to zero indicates minimal within-group variation in FI after accounting for group means. In other words, once airport-specific

baseline effects were removed, the resulting ΔFI values were minimal and largely influenced by experimental noise.

Table 37. Statistical Metrics of Predicted and Measured ΔFI

MAE	RMSE	K-Fold R^2
1.9	2.4	0.0

Figure 51 presents the measured and calculated FI obtained from the back-transformation process. After back-transformation, the large between-group differences in FI were restored. A CCC of 0.45 reflects poor to moderate concordance between measured and predicted values, indicating the presence of meaningful systematic bias in absolute terms. However, it confirmed a moderate positive linear relationship, suggesting that the model captures the overall trend in FI despite limited absolute agreement. The MAPE of 17.0% was lower than the typical within-laboratory coefficient of variation for FI testing, which is approximately 25% [138], indicating that the model error was lower than the natural experimental variability of the test method. The model is thus best interpreted as a screening and comparative tool, capable of identifying trends of FI across gradations rather than precisely estimating the FI value.

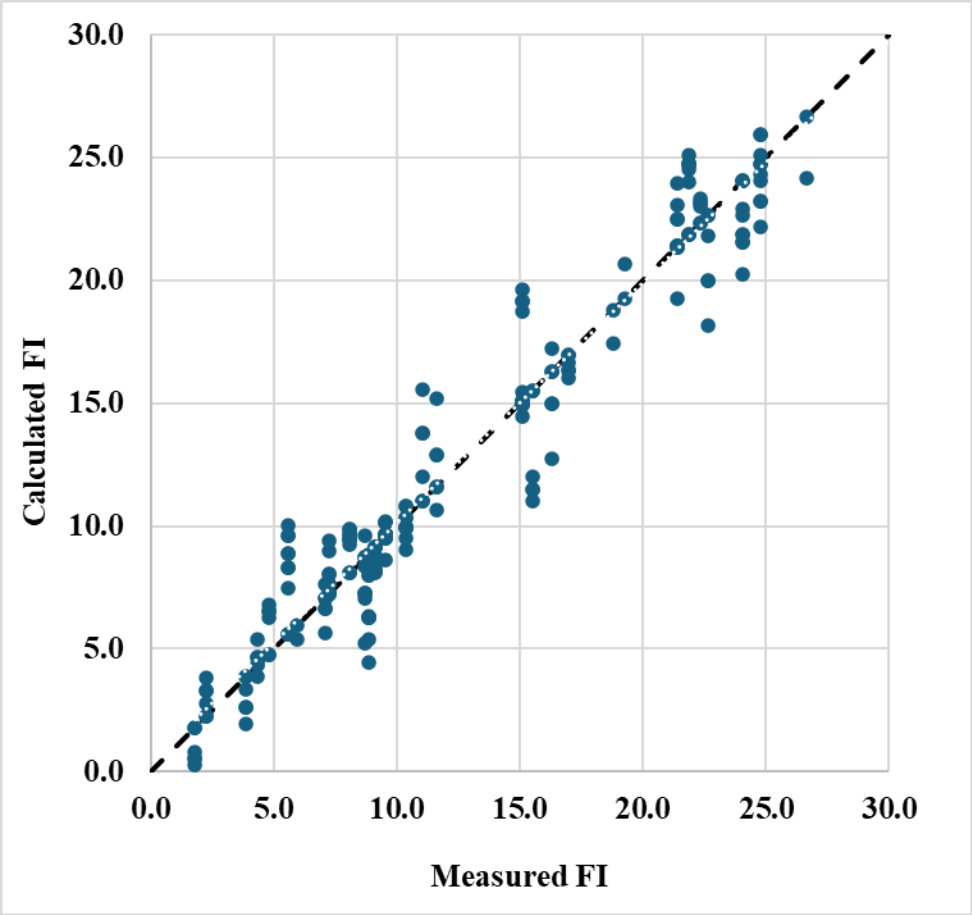

Figure 51. Measured and Calculated FI

Another version of the FI prediction model was developed without FAA variables, as shown in **Equation 43**.

Table 38 provides the statistical metrics of the model without the FAA variable. The results show that the model without the FAA variable performed slightly better than the model that included it.

$$\Delta FI = 0.017 + 0.637\Delta No.\frac{1}{2} - 0.434\Delta No.\frac{3}{8} + 0.049\Delta No.4 - 0.986\Delta No.8 \quad (43)$$
$$- 0.20\Delta No.16 + 3.25\Delta No.30 - 4.64\Delta No.50 + 4.38\Delta No.100$$
$$- 3.04\Delta No.200 + 45.5\Delta G_{sb}$$

Table 38. Statistical Metrics of Predicted and Measured ΔFI from Prediction Model without FAA

MAE	RMSE	K-Fold R^2
1.6	2.0	0.0

6.5 Model Performance Summary

Table 39 presents the cross-validated performance metrics (K-fold R^2, RMSE, and MAE) for all prediction models on the group-centered scale, including with and without FAA, for each model. Among the four prediction models, the VMA models demonstrated the strongest predictive performance, followed by the CT_{Index} model, then the rut depth model, while the FI model showed the weakest group-centered performance. Regarding the effect of FAA inclusion, the CT_{Index} model benefited most noticeably from including FAA, with R^2 improving from 0.50 to 0.62 and MAE decreasing from 18.8 to 16.0, confirming that FAA is a meaningful predictor of intermediate-temperature cracking resistance. For the rut depth model, excluding FAA slightly improved performance, suggesting that FAA contributed noise rather than a predictive signal for rutting. For the second VMA model and the FI model, FAA had a minimal effect on model accuracy.

Table 39. Cross-Validated Performance Metrics of Prediction Models (Group-Centered Scale)

Model	Number of Mix Design	With FAA			Without FAA		
		K-Fold R^2	RMSE	MAE	K-Fold R^2	RMSE	MAE
VMA First Model	2	0.86	0.40%	0.31%	N/A	N/A	N/A
VMA Second Model	10	0.69	0.40%	0.30%	0.69	0.40%	0.30%
Rut Depth	12	0.39	0.81 mm	0.67 mm	0.45	0.75 mm	0.60 mm
CT_{Index}	12	0.62	20	16	0.5	23.5	18.8
FI	12	0.0	2.4	1.9	0.0	2.0	1.6

Table 40 summarizes the back-transformed prediction accuracy of the four models that include FAA, using the CCC and the MAE or MAPE computed on the actual response scale after group means are restored. Among the prediction models, the VMA models achieved the strongest agreement between measured and calculated values. The CT_{Index} model demonstrated moderate agreement, followed closely by the rut depth model, while the FI model showed the weakest agreement. The MAE and MAPE of all models remained within or below the natural experimental variability of their test methods.

Table 40. Back-Transformed Prediction Accuracy of Models with FAA

Model	Number of Mix Design	CCC	MAE/MAPE
VMA First Model	2	0.96	MAE = 0.20%
VMA Second Model	10	0.81	MAE = 0.40%
Rut Depth	12	0.53	MAE = 0.68 mm
CT_{Index}	12	0.60	MAPE = 17.7%
FI	12	0.45	MAPE = 17.0%

6.6 Web-Based Decision Tool

Figure 52 illustrates the interface of the Volumetric and Performance Prediction tool, a web-based application developed based on the prediction models and back-transformation methods presented in this chapter. The tool can be accessed at: https://asphalt-parameter-predictor.vercel.app/.

The application enables users to estimate volumetric properties and laboratory performance for a new trial using gradation, aggregate characteristics, and laboratory data from existing tested mixtures. It is designed for cases where both the existing and new trial mixtures use the same aggregate source, binder type, and binder content. The reliability of predictions may decrease if different materials or binder contents are used in the new trial.

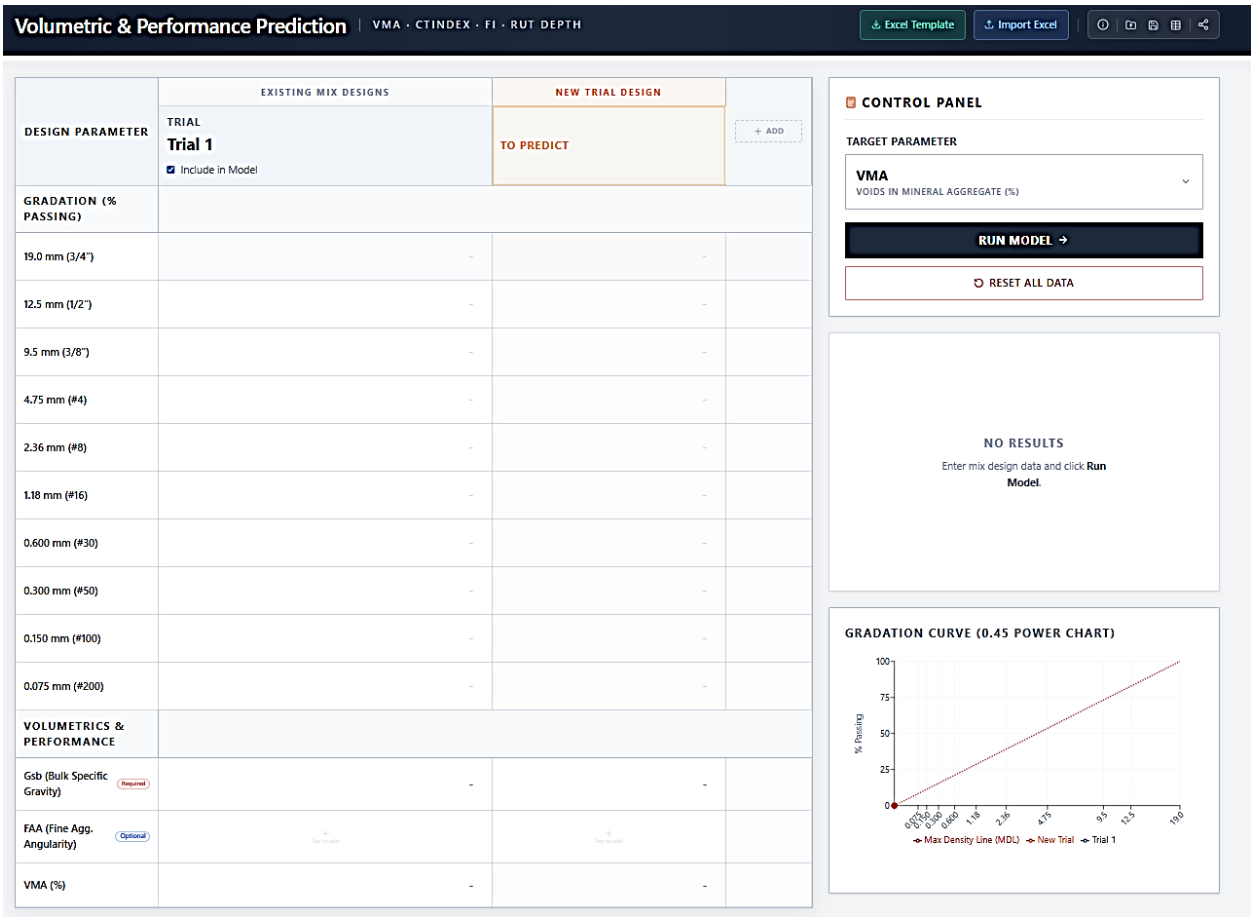

Figure 52. A Screenshot of the Web-Based Decision Tool for Volumetric Property and Laboratory Performance Prediction

The following steps describe how to execute the tool.

Step 1: Select the Target Parameters

Open the tool and locate the Control Panel on the right side of the interface. In the navigation menu, select the parameters that need to be predicted. Options include VMA, rut depth, CT_{Index}, and FI. Note that each target parameter requires a different set of inputs.

Step 2: Enter Input Data for Both the Existing and the New Trial.

The tool requires two types of input: data from existing trial mix(es) that have already been laboratory tested, and data from the new trial mix for which predictions are needed. For the existing trial mix, go to the Design Parameters, Volumetric, and Performance table, and under the Existing

Mix Design column, enter the aggregate gradation, G_{sb}, and FAA (if available) of the combined aggregate blend. Depending on the target parameter selected in Step 1, additional input will be required: measured VMA values if volumetric prediction is selected, or laboratory test results (rut depth, CT_{Index}, and FI) if laboratory performance prediction is selected.

The tool accepts multiple existing trial mixes to improve prediction stability. To add a second or subsequent existing trial, click the Add Trial button ("+") in the table. Each added trial appears as a new row. To include or exclude a specific trial from the prediction calculation, use the checkbox underneath the trial name. At least one existing trial must be checked for the model to run.

For the new trial, the design for which predictions will be generated, enter the aggregate gradation, G_{sb}, and FAA (if available) of the combined aggregate blend in the New Trial Design column.

Step 3: Run the Prediction Model

Once all data is entered, click the “Run Model” button to run the prediction. The tool will execute the appropriate model and display the results for the selected parameters. Note that if the FAA is not provided, the tool will use the model version that does not include the FAA. For VMA prediction, the tool will automatically choose the first or second VMA prediction model based on the gradation of the trial mix.

Step 4: Other functions

The tool also provides several functions for managing project files:

- Excel Template: Used to prepare input design data
- Import Excel: Import the input design data
- Load Project: Open a previously saved project.

- Save Project: Save the current project for future use.
- Export Data: Download project data and prediction results as an Excel file.
- Share Project: Generate a shareable link to allow collaborators to access the project.

Figure 53 presents an example of using the tools to predict the CT_{Index} of a new trial based on the mix design information from Trial 1. The first column includes the gradation, $G_{sb,}$ and CT_{Index} of Trial 1. The second column presents the gradation and G_{sb} of the new trial. Since the FAA was unavailable, the tool applied the model without it. Based on the inputs, the predicted CT_{Index} for the new trial is 294.5.

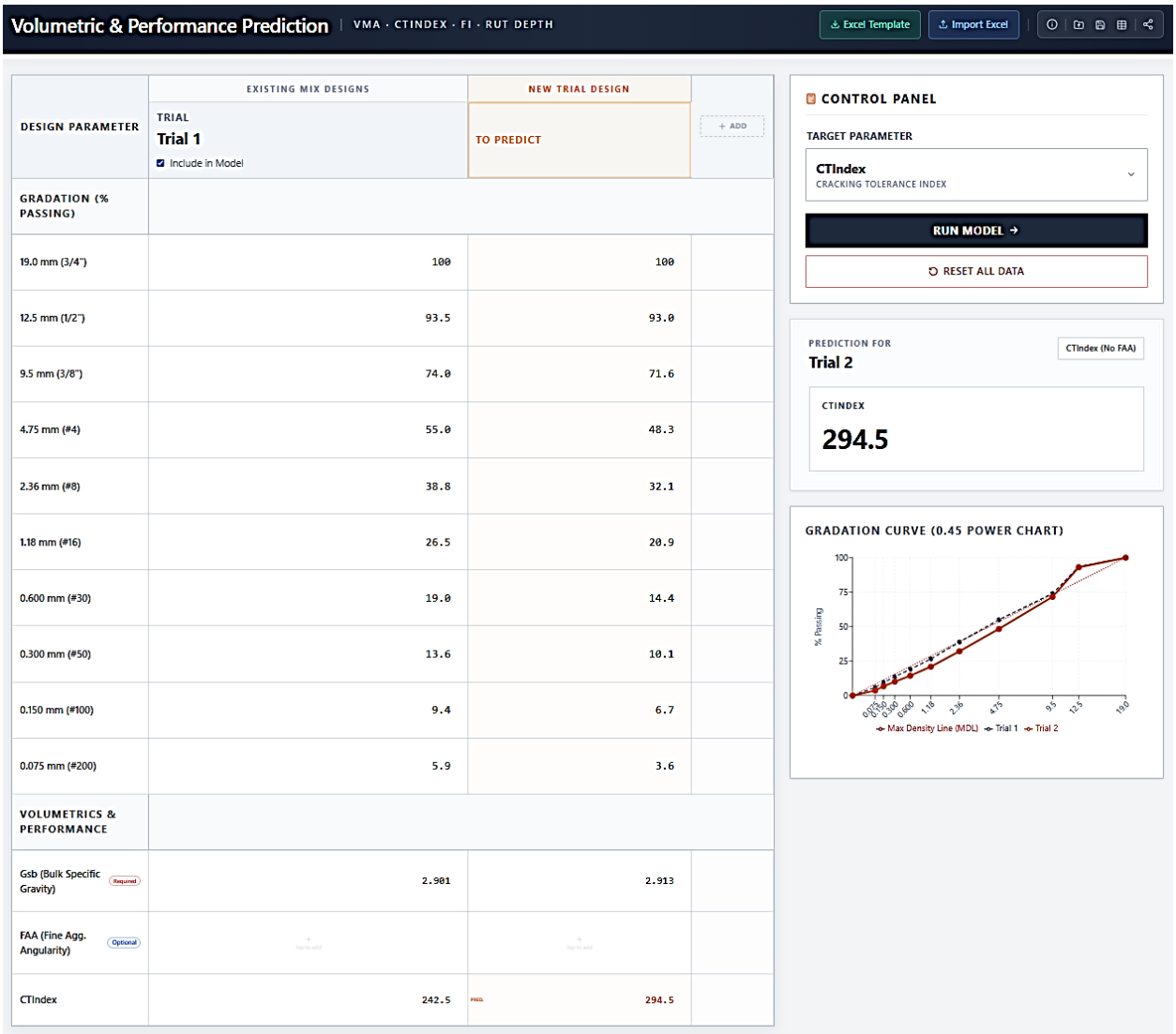

Figure 53. Example of Prediction Tool for CT_{Index}

CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS

This study focused on two objectives: (1) evaluating the effect of gradation outside the P-401 gradation limits while maintaining the volumetric properties and binder content, and using findings to revise the P-401 gradation limits, and (2) evaluating the effect of gradation changes on volumetric properties and laboratory performance of airfield asphalt mixtures and developing prediction models to quantify these relationships. A total of 13 P-401 mix designs from airports across four LTPP climate zones were selected. Four gradation designs, including two inside the P-401 gradation limits (In-Spec 1 and In-Spec 2) and two outside the P-401 gradation limits (Out-Spec 1 and Out-Spec 2), were evaluated through comprehensive laboratory performance testing. The conclusions and recommendations were summarized in this chapter.

7.1 Proposed Revisions of Current P-401 Gradation Limits

The first objective of this study was to assess whether asphalt mixtures with gradations outside the current P-401 gradation limits could deliver performance comparable to mixtures within those limits, provided that volumetric properties and asphalt binder content were kept consistent. Using 13 P-401 mix designs collected from different airports, mixtures with gradations outside the existing gradation limits (Out-Spec 1) were compared against those within the limits (In-Spec 1). The main findings are summarized below:

- Moisture susceptibility was not adversely affected by adjusting gradation outside the P-401 gradation limits, while maintaining volumetric and binder content.
- Permeability remained within acceptable limits for all mixtures, indicating that the evaluated gradation changes did not compromise the impermeability requirements of fine-graded airfield asphalt pavements.

- Durability, assessed using the Cantabro mass loss test, was similar for both In-Spec 1 and Out-Spec 1 mixtures, suggesting comparable resistance to raveling.
- Rutting resistance, measured using the HWTT and APA tests, was not significantly influenced by shifting the gradation outside the current gradation limits.
- Intermediate-temperature cracking resistance, evaluated using CT_{Index}, was generally maintained or improved when gradation shifted below the lower limits, but it tended to decrease when gradation exceeded the upper limits.
- Intermediate-temperature cracking resistance assessed using FI was generally maintained or improved when gradation was adjusted outside the gradation limits.
- Low-temperature cracking resistance, evaluated using the DCT test, showed no statistically significant differences between In-Spec 1 and Out-Spec 1 mixtures.

Based on these findings, it was concluded that adjusting the aggregate gradation outside the P-401 gradation limits is feasible without compromising laboratory performance, provided that volumetric properties and binder content are maintained. It is recommended to reduce the lower limit of 12.5 mm NMAS mixtures by 2% at the No. 16, No. 30, and No. 50 sieves. These adjustments may also provide mixture designers with improved flexibility in controlling volumetric properties and enhancing aggregate packing, without compromising performance requirements.

7.2 Effect of Gradation Change on Mixture Volumetric and Laboratory Performance Properties

The second objective was to evaluate the effects of systematic gradation adjustments on the volumetric properties and laboratory performance of asphalt mixtures. This analysis compared the mixture behaviors of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2 from 10 P-401 mix designs.

The gradation was coarser when moving from Out-Spec 2 to In-Spec 2 for upper limit mixes and from In-Spec 2 to Out-Spec 2 for lower limit mixes. The key findings and conclusions are summarized as follows:

- Volumetric properties: Shifting the gradation away from the MDL toward either the lower or upper gradation limits consistently increased VMA and air voids, when binder content and compaction effort were held constant.
- Moisture susceptibility: Gradation adjustments within the evaluated range generally did not practically affect moisture resistance.
- Durability: Gradation adjustments within the evaluated range did not significantly affect durability performance
- Permeability: Gradation adjustments within the evaluated range did not compromise impermeability requirements for airfield asphalt mixtures.
- Rutting resistance: Rutting resistance was maintained or significantly improved as gradations became coarser.
- Intermediate-temperature cracking resistance: Both IDEAL-CT and I-FIT results consistently demonstrated that coarser gradations enhanced intermediate-temperature cracking resistance.
- Low-temperature cracking resistance: Coarser gradations did not significantly enhance low-temperature cracking resistance.

In summary, adjusting gradation away from the MDL improves VMA, and coarser gradation can improve rutting and intermediate-temperature cracking resistance, while maintaining durability, permeability, moisture resistance, and low-temperature cracking performance. The findings further support the proposed revisions to P-401 gradation limits and provide the quantitative foundation for the prediction models.

7.3 Models for Predicting Changes in Volumetric and Performance Properties Due to Gradation Changes

Prediction models were developed to quantitatively characterize the effects of aggregate gradation change on (1) VMA, (2) rut depth, (3) CT_{Index}, and (4) FI. Overall, the results indicate that variations in mixture performance can be predicted with reasonable reliability based on changes in aggregate gradation and aggregate properties, provided that binder type, binder content, and aggregate source remain constant. To support practical application, these models were integrated into a web-based tool that enables engineers to estimate VMA and laboratory performance for new trial mix designs using mix design parameters and performance data from previously tested mixtures.

7.4 Recommendations for Future Research

The recommendations for future research related to the revision of P-401 gradation limits and the development of prediction models are summarized as follows:

- Expansion of gradation limits revision to additional NMAS: In this study, gradation limit adjustments were exclusively focused on 12.5 mm NMAS mixtures. Future studies should collect and analyze a broader range of airport asphalt mix designs, including those with 19 mm and 9.5 mm NMAS.
- Advanced machine learning models as the database expands: OLS regression was selected in this study due to its relatively stable performance with a limited number of observations. As more airport mix design data become available and the database grows, future research should explore the use of nonparametric and machine-learning models, which may better capture nonlinear relationships and complex interactions among variables.

- Develop prediction models incorporating binder properties: The current modeling framework predicts changes in mixture behavior as a function of gradation variations. Future research should expand this approach by developing models that also account for changes in binder content and binder properties, enabling the prediction of their effects on mixture performance.

REFERENCES

1. FAA, *150/5370-10H: Standards for Specifying Construction of Airports.* Office of Airport Safety & Standards, Airport Engineering Division, Federal Aviation Administration, Washington, DC, 2018.
2. AASHTO, *AASHTO M323: Standard Specification for Superpave Volumetric Mix Design.* AASHTO, Washington, D.C, 2022.
3. Warren, F.J., *Pavement (U.S. Patent No. 727,505).* U.S. Patent and Trademark Office, Washington, DC, 1903.
4. Richardson, C., *The Modern Asphalt Pavement*. 1912: John Wiley & Sons, New York, Second Edition.
5. Huber, G., *History of Asphalt Mix Design in North America, Part 1.* Asphalt Magazine, Asphalt Institute, 2696 Research Park Drive, Lexington, KY 40511-8480, 2013.
6. Nijboer, L.W., *Plasticity as a Factor in the Design of Dense Bituminous Road Carpets.* Elsevier Publishing Co., Inc., New York, NY, 1948.
7. McFadden, G. and W. Ricketts. *Design and Control of Asphalt Paving Mixtures for Military Installations*. in *Proceedings, The Association of Asphalt Paving Technologists, Volume 17, pp. 93-113*. 1948.
8. Regan, G.L., *A Laboratory Study of Asphalt Concrete Mix Designs for High-Contact Pressure Aircraft Traffic.* USAEWES, Geotechnical Laboratory, PO Box 631, Vicksburg, MS, 1987.
9. Newman, J. and R.B. Freeman, *Evaluation of strategic highway research program (SHRP) products for application to airport pavements*. 2000, United States. Federal Aviation Administration.
10. Cooley, L., et al., *Implementation of Superpave Mix design for Airfield Pavements.* FAA Airport Technology Transfer Conference, Washington, DC, 2007.
11. Cooley Jr, L.A., R. Ahlrich, and R.S. James. *Design of Hot Mix Asphalt for Airfield Pavements using the Superpave Gyratory Compactor*. in *2010 FAA Worldwide Airport Technology Transfer Conference, Atlantic City, New Jersey, USA* 2010.
12. FAA, *AC 150/5370-10 Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 1974.
13. FAA, *AC 150/5370-10 CH1: Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 1977.
14. FAA, *AC 150/5370-10A Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 1989.
15. FAA, *AC 150/5370-10A includes changes thru CHG 12.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC:, 1999.
16. FAA, *AC 150/5370-10D Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 2008.
17. FAA, *AC 150/5370-10E Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 2009.
18. FAA, *AC 150/5370-10F Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 2011.
19. FAA, *AC 150/5370-10G Standard Specifications for Construction of Airports.* U.S. Department of Transportation, Federal Highway Administration, Washington, DC, 2014.

20. Crane, J., *Latest FAA P-401 Specification Takes Off.* CMS, California Asphalt Magazine, 2020 Quality Issue, 2020.
21. Coree, B. and W.P. Hislop, *A Laboratory Investigation into the Effects of Aggregate-Related Factors of Critical VMA in Asphalt Paving Mixtures*. 2000, Center for Transportation Research and Education, Iowa State University, IA
22. Hislop, W.P. and B.J. Coree, *VMA as a Design Parameter in Hot-Mix Asphalt.* Mid-continent Transportation Symposium Proceedings, 2000.
23. Aschenbrener, T. and C. MacKean, *Factors that Affect the Voids in the Mineral Aggregate in Hot Mix Asphalt.* Report. No. CDOT-DTD-R-92-13, Colorado Department of Transportation, 4201 E Arkansas Avenue, Denver, CO 1992.
24. Dae-Wook Park, A.C., Joe Button *Effects of Aggregate Gradation and Angularity on VMA and Rutting Resistance.* Research Report ICAR –201-3F, Texas Transportation Institute, The Texas A&M University System, College Station, Texas 2001.
25. Purcell, E.M. and S.A. Cross, *Effects of Aggregate Angularity on VMA and Rutting of KDOT Superpave Level 1 Mixes.* No. K-TRAN: KU-98-5, Kansas Dept. of Transportation, Topeka, 2001.
26. Aschenbrener, T., *How to Increase Voids in Mineral Aggregate: Guidelines to Increase VMA of Superpave Mixtures.* 1996.
27. Cheung, L.W. and A.R. Dawson, *Effects of Particle and Mix Characteristics on Performance of Some Granular Materials.* Transportation Research Record: Journal of the Transportation Research Board, 1787(1), 90-98., 2002.
28. Bessa, I.S., et al., *Aggregate Shape Properties and their Influence on the Behavior of Hot-Mix Asphalt.* Journal of Materials in Civil Engineering, 2015. **27**(7): p. 04014212.
29. Institute, A., *MS-2 Asphalt Mix Design Methods.* 2014.
30. McGennis, R.B., et al., *Background of SuperPave Asphalt Mixture Design and Analysis*. 1995, Office of Technology Applications, Federal Highway Administration. United States.
31. Leon, L.P. and R. Charles, *Aggregate Angularity on the Permanent Deformation Zones of Hot Mix Asphalt.* Global Journal of Research in Engineering, 2015. **15**(3): p. 25-29.
32. Souza, L.T., et al., *Experimental Testing and Finite-Element Modeling to Evaluate the Effects of Aggregate Angularity on Bituminous Mixture Performance.* Journal of Materials in Civil Engineering, 2012. **24**(3): p. 249-258.
33. Pan, T., E. Tutumluer, and S.H. Carpenter, *Effect of Coarse Aggregate Morphology on Permanent Deformation Behavior of Hot Mix Asphalt.* Journal of Transportation Engineering, 2006. **132**(7): p. 580-589.
34. Renteria, D., et al., *Discrete Element Analysis of SCB Variability: Asphalt Mixtures.* [Master's thesis, The University of Texas Rio Grande Valley]. ScholarWorks @ UTRGV. https://scholarworks.utrgv.edu/etd/343, 2017.
35. Saadeh, S. and M. El Asmar, *Sensitivity Analysis of the IDEAL CT Test using the Distinct Element Method.* Mineta Transportation Institute College of Business, San José State University, San José, CA 95192-0219, 2023.
36. Park, D.W. and H.S. Lee, *Test Methods for Fine Aggregate Angularity Considering Resistance of Rutting.* KSCE Journal of Civil Engineering, 2002. **6**(4): p. 421-427.
37. Topal, A. and B. Sengoz, *Evaluation of Compacted Aggregate Resistant Test Compared With the Fine Aggregate Angularity Standards.* Construction and Building Materials, 2008. **22**(5): p. 993-998.

38. Ramli, I., et al., *Fine Aggregate Angularity Effects on Rutting Resistance of Asphalt Mixture.* Jurnal Teknologi, 2013. **65**(3).
39. Iuele, T., F.G. Praticò, and R. Vaiana, *Fine Aggregate Properties vs Asphalt Mechanical Behavior: An Experimental Investigation*, in *Pavement and Asset Management*. 2019, CRC Press. p. 309-315.
40. Johnson, E., et al., *Investigation of Superpave Fine Aggregate Angularity Criterion for Asphalt Concrete.* Transportation Research Record: Journal of the Transportation Research Board, 2007. **1998**(1): p. 75-81.
41. Chowdhury, A., et al., *Evaluation of Superpave Fine Aggregate Angularity Specification.* Report No. ICAR – 201-1, Texas Transportation Institute The Texas A&M University System College Station, Texas, 2001.
42. Leomar Fernandes, J., et al., *Uncompacted Void Content of Fine Aggregate as Quality Indicator of Materials Used In Superpave Mixtures.* Transportation Research Record, 2000. **1723**(1): p. 37-44.
43. Bahia, H. and A. Stakston, *The Effect off Fine Aggregate Angularity, Asphalt Content and Performance Graded Asphalts on Hot Mix Asphalt Performance.* Wisconsin Highway Research Program, Wisconsin. Department of Transportation, 2003.
44. Lee, C.-J., T.D. White, and T.R. West, *Effect of Fine Aggregate Angularity on Asphalt Mixture Performance.* Publication FHWA/IN/JTRP-98/20. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2000. https://doi.org/10.5703/1288284313227, 2000.
45. Roque, R., et al., *Evaluation of Superpave (Trademark) Criteria for VMA And Fine Aggregate Angularity. Volume 2 of 2: Fine Aggregate Angularity (FAA)*. 2002: Florida Department of Transportation, 605 Suwannee St, Tallahassee, FL 32399.
46. Kandhal, P.S. and F. Parker, *Aggregate Tests Related to Asphalt Concrete Performance in Pavements.* Transportation Research Board, 500 Fifth Street, NW, Washington, DC United States 20001, 1998. **405**.
47. Aho, B.D., W.R. Vavrik, and S.H. Carpenter, *Effect of Flat and Elongated Coarse Aggregate on Field Compaction of Hot-Mix Asphalt.* Transportation Research Board, 500 Fifth Street, NW, Washington, DC, United States 20001, 2001. **1761**(1): p. 26-31.
48. Buchanan, M.S., *Evaluation of the Effect of Flat and Elongated Particles on the Performance of Hot Mix Asphalt Mixtures*. 2000, NCAT Report 00-03, National Center for Asphalt Technology, Auburn, AL.
49. Wang, S., Y. Miao, and L. Wang, *Investigation of the Force Evolution in Aggregate Blend Compaction Process and the Effect of Elongated and Flat Particles using DEM.* Construction Building Materials, 2020. **258**: p. 119674.
50. Chen, J.S., M. Chang, and K. Lin, *Influence of Coarse Aggregate Shape on the Strength of Asphalt Concrete Mixtures.* Journal of The Eastern Asia Society for Transportation Studies, 2005. **6**: p. 1062-1075.
51. Mukhopadhyay, A.K., et al., *Treatments for Clays In Aggregates Used to Produce Cement Concrete, Bituminous Materials, and Chip Seals: Technical Report*. 2013, Report No. FHWA/TX-13/0-6444-1, Texas Department of Transportation Research and Technology Implementation, Austin, Texas
52. Dong, Q., et al., *Reduction of Moisture Susceptibility of Cold Asphalt Mixture with Portland Cement and Bentonite Nanoclay Additives.* Journal of Cleaner Production, 2018. **176**: p. 320-328.

53. Yin, F., et al., *Novel Method for Moisture Susceptibility and Rutting Evaluation Using Hamburg Wheel Tracking Test.* Transportation Research Record: Journal of the Transportation Research Board, Washington, D.C., 2014, pp. 1–7. DOI: 10.3141/2446-01, 2014(1).
54. Afsha, S., et al., *Impact of Clay Contamination on Rutting Performance of Asphalt Mixtures.* Transportation Research Record, 2024. **2678**(10): p. 356-366.
55. Dahir, S., *A Review of Aggregate Selection Criteria for Improved Wear Resistance and Skid Resistance of Bituminous Surfaces.* Journal of Testing Evaluation, 1979. **7**(5): p. 245-253.
56. Wu, Y., F. Parker, and P.S. Kandhal, *Aggregate Toughness/Abrasion Resistance and Durability/Soundness Tests Related to Asphalt Concrete Performance in Pavements.* Transportation Research Record: Journal of the Transportation Research Board, 1998. **1638**(1): p. 85-93.
57. Amirkhanian, S.N., D. Kaczmarek, and J. Burati Jr, *Effects of Los Angeles Abrasion Test Values on the Strengths of Laboratory Prepared Marshall Specimens* Transportation Research Record: Journal of the Transportation Research Board, 1991(1301).
58. Masad, E., et al., *Correlation of Fine Aggregate Imaging Shape Indices with Asphalt Mixture Performance.* Transportation Research Record: Journal of the Transportation Research Board, 2001. **1757**(1): p. 148-156.
59. Hassan, H.M.Z., et al., *Study on the Influence of Aggregate Strength and Shape on the Performance of Asphalt Mixture.* Construction Building Materials, 2021. **294**: p. 123599.
60. Hislop, W.P., *The Difficult Nature of Minimum VMA: A Historical Perspective.* Transportation Research Record: Journal of the Transportation Research Board, 1681(1), 148-156, 1998.
61. Hudson, S. and R. Davis. *Relationship of aggregate voidage to gradation*. in *Assoc Asphalt Paving Technol Proc*. 1965.
62. Kandhal, P.S. and R.B. Mallick, *Design of New-Generation Open-Graded Friction Courses (Revised)*. 1999, Report No: 99-3, National Center for Asphalt Technology, Auburn, AL.
63. NAPA, *Designing and Constructing SMA Mixtures—State-of-the-practice.* National Asphalt Pavement Association, 5100 Forbes Boulevard Lanham, Maryland 2002.
64. Interactive, P. *Pavement types/mix types. Pavement Interactive.* [cited 2025 July 27,].
65. Koch, J., *Effects of Aggregate Blend Gradations on Hot Mix Asphalt Volumetrics* Department of Civil Engineering and Construction Bradley, University Peoria, IL 2021.
66. Williams, S.G. and J.K. Foreman, *Effects of various HMA material properties on pavement performance*. 2006, Arkansas State Highway and Transportation Department, Little Rock, AR.
67. Chadbourn, B.A., et al., *The Effect of Voids in Mineral Aggregate (VMA) on Hot Mix Asphalt Pavements.* Report No. MN/RC - 2000-13, University of Minnesota, Department of Civil Engineering, Minneapolis, MN 55455-0116 1999.
68. Bessette, L., *Evaluation of the Effects of Aggregate Gradation And Compaction Effort on the Voids in Mineral Aggregate in Asphalt Concrete*. 2013, West Virginia University: Graduate Theses, Dissertations, and Problem Reports. 7301, West Virginia University, https://researchrepository.wvu.edu/etd/7301.
69. Huber, G. and T. Shuler. *Providing Sufficient Void Space for Asphalt Cement: Relationship of Mineral Aggregate Voids and Aggregate Gradation*. in *Effects of aggregates and mineral fillers on asphalt mixture performance*. 1992. ASTM International.

70. Kandhal, P.S., K.Y. Foo, and J.A. D'Angelo, *Field Management of Hot Mix Asphalt Volumetric properties.* NCAT Report 95-04, National Center for Asphalt Technology, Auburn, AL, 1996. **1299**.
71. Xie, H., L.A. Cooley, and M.H. Huner, *4.75 mm NMAS stone matrix asphalt (SMA) mixtures.* NCAT Report 12-10, National Center for Asphalt Technology, Auburn, AL, 2003: p. 03-05.
72. Leiva, F. and R. West, *Using the Primary Control Sieve Index to Define Gradation Type and as a Factor Related to Asphalt Mixture Properties* 2021: NCAT Report 21-01, National Center for Asphalt Technology Auburn University, Auburn, Alabama
73. Zaniewski, J.P. and C. Mason, *An Evaluation of The Bailey Method to Predict Voids in The Mineral Aggregate.* Asphalt Technology Program Department of Civil and Environmental Engineering Morgantown, West Virginia 2006.
74. Liu, H., P. Hao, and J. Xu, *Effects of Nominal Maximum Aggregate Size on the Performance of Stone Matrix Asphalt.* Applied Sciences, 2017. **7**(2): p. 126.
75. USDOT, *Superpave Mix Design and Gyratory Compaction Levels.* Technical Brief, FHWA-HIF-11-031, Federal Highway Administration, USDOT, 2010.
76. Shen, S. and H. Yu, *Analysis of Aggregate Gradation and Packing for Easy Estimation of Hot-Mix-Asphalt Voids in Mineral Aggregate.* Journal of Materials in Civil Engineering, 2011. **23**(5): p. 664-672.
77. Pouranian, M., *Aggregate Packing Characteristics of Asphalt Mixtures*. 2019, Purdue University.
78. Zhang, H., et al., *Establishment and application of theoretical-empirical prediction model of VMA in hot mix asphalt mixture.* Case Studies in Construction Materials, 2023. **18**: p. e01986.
79. Khalifa, M.O.L., *The Behavior of Asphaltic Concrete Constructed with Large-Sized Aggregate*. 1969: University of Illinois at Urbana-Champaign.
80. Roberts, F.L., et al., *Hot Mix Asphalt Materials, Mixture Design and Construction.* 1996.
81. Graziani, A., et al., *An Application to the European Practice of the Bailey Method for HMA Aggregate Grading Design.* Procedia-Social Behavioral Sciences, 2012. **53**: p. 990-999.
82. Golalipour, A., et al., *Effect of Aggregate Gradation on Rutting of Asphalt Pavements.* Procedia-Social Behavioral Sciences, 2012. **53**: p. 440-449.
83. Scherocman, J., *Stone Mastic Asphalt Reduces Rutting.* Better Roads, 1991. **61**(11).
84. Scherocman, J. *Construction of SMA test sites in the US*. in *AAPT meeting*. 1992.
85. Asi, I.M., *Laboratory Comparison Study for the Use of Stone Matrix Asphalt in Hot Weather Climates.* Construction and Building Materials, 2006. **20**(10): p. 982-989.
86. Zhao, Z., et al., *Factors Affecting the Rutting Resistance of Asphalt Pavement based on the Field Cores using Multi-Sequenced Repeated Loading Test.* Construction and Building Materials, 2020. **253**: p. 118902.
87. Sangsefidi, E., H. Ziari, and M. Sangsefidi, *The Effect of Aggregate Gradation Limits Consideration on Performance Properties And Mixture Design Parameters of Hot Mix Asphalt.* KSCE Journal of Civil Engineering, 2016. **20**(1): p. 385-392.
88. Musselman, J.A., et al., *Superpave field implementation: Florida's early experience.* Transportation Research Record, 1998. **1609**(1): p. 51-60.
89. Kandhal, P.S. and L. Cooley Jr, *The restricted zone in the Superpave aggregate gradation specification.* NCHRP Report 464, Transportation Research Board, National Research Council National Academy Press, Washington, D.C., 2001(Project 9-14 FY'98).

90. Kim, Y.-R., et al., *Effects of Aggregate Structure on Hot-Mix Asphalt Rutting Performance in Low Traffic Volume Local Pavements.* Construction and Building Materials, 2009. **23**: p. 2177-2182.
91. Cross, S.A., et al. *Effects of Gradation on Performance of Asphalt Mixtures*. in *78th Annual Meeting of the Transportation Research Board, Washington, DC*. 1999.
92. Garba, R., *Permanent deformation properties of asphalt concrete mixtures.* 2002.
93. El-Basyouny, M. and M. Mamlouk. *Effect of aggregate gradation on rutting potential of Superpave mixes*. in *In the 8th Annual Meeting of the Transportation Research Board. Washington, DC: Transportation Research Board*. 1999.
94. Hasan, M.M., *Evaluation of Fine and Coarse Graded Asphalt Mixes*. 2021, University of New Mexico.
95. Kandhal, P.S. and R.B. Mallick, *Effect of Mix Gradation on Rutting Potential of Dense-Graded Asphalt Mixtures.* Transportation Research Record, 2001. **1767**(1): p. 146-151.
96. Chowdhury, A., J.W. Button, and J.D. Grau, *Effects of Superpave Restricted Zone on Permanent Deformation.* Report No. 201-2, Texas Transportation Institute, The Texas A&M University System, College Station, Texas 2001.
97. Hand, A.J., et al., *Gradation Effects on Hot-Mix Asphalt Performance.* Transportation Research Record, Washington, D.C, 2001. **1767**(1): p. 152-157.
98. West, R., et al., *Phase IV Ncat Pavement Test Track Findings.* NCAT Report 12-10, National Center for Asphalt Technology, Auburn University, Auburn, Alabama, 2012.
99. Hasan, M.M., et al., *Laboratory Performance Evaluation of Fine And Coarse-Graded Asphalt Concrete Mix.* Journal of Materials in Civil Engineering, 2019. **31**(11): p. 04019259.
100. Pan, Y., et al., *A Laboratory Evaluation of Factors Affecting Rutting Resistance of Asphalt Mixtures using Wheel Tracking Test.* Case Studies in Construction Materials, 2023. **18**: p. e02148.
101. Huang, W., P. Yu, and X. Huang. *Study on Grading Envelope of AC-16 Asphalt Mixture based on Rutting Resistance Performance*. in *2011 IEEE International Conference on Automation and Logistics (ICAL)*. 2011. IEEE.
102. Gao, Y., et al., *Variability Evaluation of Gradation for Asphalt Mixture in Asphalt Pavement Construction.* Automation in Construction, 2021. **128**: p. 103742.
103. Chen, J.-S. and M.-C. Liao, *Evaluation of internal resistance in hot-mix asphalt (HMA) concrete.* Construction and Building Materials, 2002. **16**(6): p. 313-319.
104. Lv, Q., et al., *Influence of Gradation on Asphalt Mix Rutting Resistance Measured by Hamburg Wheel Tracking Test.* Construction and Building Materials, 2020. **238**: p. 117674.
105. Zhang, Y., et al., *Mechanical Evaluation of Aggregate Gradation to Characterize Load Carrying Capacity and Rutting Resistance of Asphalt Mixtures.* Construction Building Materials, 2019. **205**: p. 499-510.
106. Teklu, W., *Effect of Gradation of Aggregates on the Rutting Performance of Hot Mix Asphalt.* Addis Ababa University, 2015.
107. Thompson, G., *Investigation of the Bailey Method for the design and analysis of dense-graded MHAC using Oregon aggregates*. 2007, Oregon. Dept. of Transportation. Research Unit.
108. Aurilio, V., W.J. Pine, and P. Lum. *The Bailey Method Achieving Volumetrics and HMA Compactability*. in *Proceedings of The Annual Conference-Canadian Technical Asphalt Association*. 2005. Victoria British Columbia, Canada.

109. Vavrik, W.R., et al., *The Bailey Method of Gradation Evaluation: The Influence of Aggregate Gradation and Packing Characteristics on Voids in The Mineral Aggregate* Association of Asphalt Paving Technologists, White Bear Lake, MN, 2001. **70**.
110. Manjunath, K. and P.D. NB, *Design of Hot Mix Asphalt using Bailey Method of Gradation.* International Journal of Research in Engineering Technology, 2014. **3**(06): p. 386-393.
111. Wen-jiang, L., et al., *Effect of CA ratio on asphalt mixture property based on Bailey method.* Journal of Chang'an University (Natural Science Edition), 2005. **25**(04): p. 5-8.
112. Gierhart, D., *Analysis of Oklahoma Mix Designs for the National Center for Asphalt Technology Test Track using the Bailey Method.* Transportation Research Record, Washington, D.C, 2007. **33**.
113. Bhasin, A., et al., *Analysis of South Texas Aggregates for Use in Hot Mix Asphalt.* FHWA/TX-050-4203-4, Texas Transportation Institute, The Texas A&M University System, College Station, Texas 2005. **4**.
114. Christensen, D.W., et al., *Refinement of Current WisDOT HMA Mixture Application Guidelines Related to NMAS and Aggregate Characteristics*. 2013: Wisconsin Highway Research Program, University of Wisconsin-Madison.
115. Hafeez, I., et al., *Investigating the Effects of Maximum Size of Aggregate on Rutting Potential of Stone Mastic Asphalt.* Journal of Engineering Applied Sciences, 2012.
116. del Pilar Vivar, E. and J.E. Haddock, *HMA Pavement Performance and Durability.* Report No. FHWA/IN/JTRP-2005/14. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2006. https://doi.org/10.5703/1288284313391, 2006.
117. Ma, H., et al., *Influence of Fine Aggregate Content on Low-Temperature Cracking of Asphalt Pavements*. 2016: ASTM International.
118. Tan, Y., et al. *Effect of Aggregate Grading on Low Temperature Cracking Resistance in Asphalt Mixtures Base on Mathematical Statistic*. in *Proceedings of Eighth International RILEM conference on cracking in pavements*. 2008.
119. Nian, T., et al., *Improved Discrete Element Numerical Simulation and Experiment on Low-Temperature Anti-Cracking Performance of Asphalt Mixture based on PFC2D.* Construction and Building Materials, 2021. **283**: p. 122792.
120. Nejad, F.M., E. Aflaki, and M. Mohammadi, *Fatigue behavior of SMA and HMA mixtures.* Construction and Building Materials, 2010. **24**(7): p. 1158-1165.
121. Gui-ying, L., M. Zhong-chuan, and G. Ze-yu, *Experimental Study on Aggregate Gradation Influence on the Performance-related Properties of Structural-layer Asphalt Mixtures.* Highway, 2019. **5**: p. 219-225.
122. Lv, S., et al., *Research on Strength and Fatigue Properties of Asphalt Mixture with Different Gradation Curves.* Construction Building Materials, 2023. **364**: p. 129872.
123. Saleh, M.F., *Effect of Aggregate Gradation and Percentage of Air Voids on the Fatigue Behaviour of Hot Mix Asphalts.* Road Transport Research: A Journal of Australian New Zealand Research Practice, 2012. **21**(3): p. 53-63.
124. Sousa, J.B., et al., *Effect of Aggregate Gradation on Fatigue Life of Asphalt Concrete Mixes.* Transportation Research Record, 1998. **1630**(1): p. 62-68.
125. Hasan, M.M., M.A. Hasan, and R.A. Tarefder. *Evaluation of Fatigue Performance of Coarse and Fine Graded Asphalt Concrete Mix Employing Viscoelastic Continuum Damage (VECD) Model*. in *International Airfield and Highway Pavements Conference 2019*. 2019. American Society of Civil Engineers Reston, VA.

126. Jiangcai, C., et al., *Investigation on Temperature Shrinkage Characteristics of the Combined Structure in Asphalt Pavement.* Frontiers in Materials, 2022. **9**: p. 1055641.
127. Moon, K.H., A.C. Falchetto, and W.S. Yeom, *Low-temperature performance of asphalt mixtures under static and oscillatory loading.* Arabian Journal for Science and Engineering, 2014. **39**: p. 7577-7590.
128. Christensen, D.W. and R.F. Bonaquist, *Evaluation of indirect tensile test (IDT) procedures for low-temperature performance of hot mix asphalt*. Vol. 530. 2004: Transportation Research Board, Washington, DC.
129. Garcia-Gil, L., R. Miró, and F.E. Pérez-Jiménez, *Evaluating the role of aggregate gradation on cracking performance of asphalt concrete for thin overlays.* Applied Sciences, 2019. **9**(4): p. 628.
130. Hainin, M.R., et al., *Fatigue Life of Malaysian Hot Mix Asphalt Mixtures.* Malaysian Journal of Civil Engineering, 2013. **25**(1).
131. Lou, K., et al., *Comprehensive Study about Effect of Basalt Fiber, Gradation, Nominal Maximum Aggregate Size and Asphalt on The Anti-Cracking Ability of Asphalt Mixtures. Applied Sciences,* 2021. **11**(5): p. 2289.
132. Shen, S. and S. Carpenter, *Development of an Asphalt Fatigue Model Based on Energy Principles.* Asphalt Paving Technology-Proceedings, 2007. **76**: p. 525.
133. Abo-Qudais, S. and I. Shatnawi, *Prediction of Bituminous Mixture Fatigue Life Based on Accumulated Strain.* Construction and Building Materials, 2007. **21**(6): p. 1370-1376.
134. Kim, S.S., J.J. Yang, and R.A. Etheridge, *Effects of Mix Design Variables on Flexibility Index of Asphalt Concrete Mixtures.* International Journal of Pavement Engineering, 2020. **21**(10): p. 1275-1280.
135. Etheridge, R.A., et al., *Evaluation of Fatigue Cracking Resistance of Asphalt Mixtures using Apparent Damage Capacity.* Journal of Materials in Civil Engineering, 2019. **31**(11): p. 04019257.
136. Christensen, D.W. and R.F. Bonaquist, *Volumetric Requirements for Superpave Mix Design*. Vol. 567. 2006: Transportation Research Board.
137. Sreedhar, S. and E. Coleri, *Effects of Binder Content, Density, Gradation, and Polymer Modification on Cracking and Rutting Resistance of Asphalt Mixtures Used in Oregon.* Journal of Materials in Civil Engineering, 2018. **30**(11): p. 04018298.
138. Abbas, A.R., et al., *Crack Resistance and Durability of Ohio DOT Asphalt Mixtures using I-FIT & IDEAL-CT: Phase 2*. 2021, Ohio. Dept. of Transportation. Office of Statewide Planning and Research.
139. Garcia, V.M., et al., *Effect of Aggregate Gradation on Performance of Asphalt Concrete Mixtures.* Journal of Materials in Civil Engineering, 2020. **32**(5): p. 04020102.
140. Haryanto, I. and O. Takahashi, *Crack Initiation Assessment of Wearing Course Asphalt Mixtures using Aggregate Gradation Characteristic.* Journal of Engineering Technological Sciences, 2007. **39**(1): p. 28-42.
141. Lu, Q., P. Fu, and J. Harvey, *Laboratory Evaluation of the Noise and Durability Properties of Asphalt Surface Mixes* Report No. UCPRC-RR-2009-07, Office of Roadway Research, Division of Research and Innovation, California Department of Transportation,, 2009.
142. Shiva Kumar, G. and S. Suresha, *State of the Art Review on Mix Design and Mechanical Properties of Warm Mix Asphalt.* Road Materials Pavement Design, 2019. **20**(7): p. 1501-1524.

143. Ji, T., et al., *Durability Behavior of Asphalt Mixtures in Regard to Material Properties and Gradation Type.* Frontiers in Materials, 2023. **10**: p. 1151479.
144. Watson, D., N.A. Qureshi, and Z. Xie, *Improvements of Open-Graded Friction Course in Alabama ALDOT 930-790.* NCAT Report 20-07, National Center for Asphalt Technology Auburn University, Auburn, Alabama 2020.
145. Xie, Z., et al., *Five-Year Performance of Improved Open-Graded Friction Course on the NCAT Pavement Test Track.* Transportation Research Record, 2019. **2673**: p. 544 - 551.
146. Tran, T., *Open-Graded Friction Courses Suitable for Suburban Environments*. 2023: Auburn University, Auburn, Alabama.
147. Bennert, T. and L.A. Cooley, *Evaluate the Contribution of The Mixture Components on the Longevity and Performance of FC-5.* FDOT Final Report No. BDS15 977-01, 2014.
148. Kanitpong, K., N. Charoentham, and S. Likitlersuang, *Investigation on the Effects of Gradation And Aggregate Type to Moisture Damage of Warm Mix Asphalt Modified with Sasobit.* International Journal of pavement engineering, 2012. **13**(5): p. 451-458.
149. Khan, M.K., A. Hussain, and S. Khanzada, *Effect of Gradation on Moisture Susceptibility of Asphalt Paving Mixtures*, in *International Conference on Sustainable Development in Civil Engineering*. 2017: MUET, Pakistan.
150. AASHTO, *AASHTO T283: Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage.* AASHTO, Washington, D.C, 2022.
151. Aodah, H.H., Y.N.A. Kareem, and S. Chandra, *Effect of Aggregate Gradation on Moisture Susceptibility and Creep in HMA.* World Academy of Science and Technology, 2012. **6**: p. 12-28.
152. Nekkanti, H., B.J. Putman, and B. Danish, *Influence of Aggregate Gradation and Nominal Maximum Aggregate Size on the Performance Properties of OGFC Mixtures.* Transportation Research Record, Washington, DC, 2019. **2673**(1): p. 240-245.
153. Yan, Y., J.P. Zaniewski, and D. Hernando, *Development of a Predictive Model to Estimate Permeability of Dense-Graded Asphalt Mixture based on Volumetrics.* Construction and Building Materials, 2016. **126**: p. 426-433.
154. Cooley Jr, L., E.R. Brown, and S. Maghsoodloo, *Developing Critical Field Permeability and Pavement Density Values for Coarse-Graded Superpave Pavements.* Transportation research record, 2001. **1761**(1): p. 41-49.
155. Gogula, A.K., M. Hossain, and S.A. Romanoschi, *A study of Factors Affecting the Permeability of Superpave Mixes in Kansas*. 2004, Kansas. Dept. of Transportation.
156. Brown, E.R., et al., *Relationships of HMA In-Place Air Voids, Lift Thickness, and Permeability.* NCHRP Report 531, Transportation Research Board. National Research Council, Washington DC, 2004. **1**.
157. Westerman, J.R. *AHTD's Experience with Superpave Pavement Permeability*. in *Arkansas Superpave Symposium*. 1998.
158. Choubane, B., G.C. Page, and J.A. Musselman, *Investigation of Water Permeability of Coarse Graded Superpave Pavements.* Journal of the Association of Asphalt Paving Technologists, 1998. **67**.
159. Cooley, L.A., B.D. Prowell, and E. Brown, *Issues Pertaining to the Permeability Characteristics of Coarse Graded Superpave Mixes.* 2002.
160. Mallick, R.B., et al., *An Evaluation of Factors Affecting Permeability of Superpave Designed Pavements.* Report No 03-02, National Center for Asphalt Technology, Auburn University, Alabama, 2003.

161. Chen, J., et al., *Permeability Loss of Open-Graded Friction Course Mixtures due to Deformation-Related and Particle-Related Clogging: Understanding from a Laboratory Investigation.* Journal of Materials in Civil Engineering, 2015. **27**(11): p. 04015023.
162. Hardiman, C., *The Improvement of Water Drainage Function and Abrasion Loss of Conventional Porous Asphalt.* Proceedings of the Eastern Asia Society for Transportation Studies, 2005. **5**: p. 671-678.
163. Momm, L. and L. Bariani Bernucci. *Study of the Aggregate for the Pervious Asphalt Concrete*. in *Performance Testing and Evaluation of Bituminous Materials ptebm'03. proceedings of the 6th international rilem symposium held zurich, switzerland, 14-16 april 2003*. 2003.
164. Hasan, M.A., R.A. Tarefder, and Z.H. Khan, *Laboratory Investigation of Open Graded Friction Courses with Different Binder Grades and Aggregate Gradations.* Transportation Research Record, 2023: p. 03611981231172749.
165. Khosla, N.P. and S. Sadasivam, *Determination of Optimum Gradation for Resistance to Permeability, Rutting and Fatigue Cracking*. 2005: No. FHWA/NC/2004-12, Transportation Research Board, Washington, D.C.
166. Hainin, M.R., L.A. Cooley Jr, and B.D. Prowell, *An Investigation of Factors Influencing Permeability of Superpave Mixes.* International Journal of Pavements, 2003. **2**(2): p. 41-52.
167. Zhang, H., *Asphalt Mixture Volume Index and Anti-Rutting Performance Unified Prediction Model and its Application.* Master's thesis, School of Civil Engineering, Shandong Univ, 2009.
168. Liu, H.Y., *Research on Asphalt Mixture Volume Parameter Characteristics and Gradation Design Method.* Master's thesis, Highway School, Chang'an Univ., 2007.
169. Li, B., *Experimental Study on Critical Mineral Gap Ratio of Asphalt Mixture.* Master's thesis, School of Traffic & Transportation Engineering, Changsha Univ. of Science and Technology., 2006.
170. Inoue, T., Y. Gunji, and H. Akagi. *Rational Design Method of Hot Mix Asphalt based on Calculated VMA*. in *Proceedings of The 3rd Eurasphalt and Eurobitume Congress Held Vienna, May 2004*. 2004.
171. Yu, A. and N. Standish, *A Study of the Packing of Particles with a Mixture Size Distribution.* Powder Technology, 1993. **76**(2): p. 113-124.
172. Pouranian, M.R. and J.E. Haddock, *Determination of Voids in the Mineral Aggregate and Aggregate Skeleton Characteristics of Asphalt Mixtures using a Linear-Mixture Packing Model.* Construction Building Materials, 2018. **188**: p. 292-304.
173. Li, X. and Z. He, *Prediction Model for Connected Voids Ratio of the Porous Asphalt Mixture.* Advances in Civil Engineering, 2020. **2020**(1): p. 3760435.
174. Othman, K., *Prediction of the Hot Asphalt Mix Properties using Deep Neural Networks.* Beni-Suef University Journal of Basic Applied Sciences, 2022. **11**(1): p. 40.
175. Shaikh, S., et al., *Data-Driven Predictive Models for Evaluating Optimum Binder Content and Volumetric Properties of Bituminous Mixtures using Design Variables.* Measurement, 2025. **253**: p. 117411.
176. Ezzat, E.N. and A.H. Abed, *Prediction of Modified Asphalt Concrete Rutting Depth using Statistical Model.* Materials Today: Proceedings 2021. **42**: p. 1784-1790.

177. Majidifard, H., et al., *Developing a Prediction Model for Rutting Depth of Asphalt Mixtures using Gene Expression Programming.* Construction Building Materials, 2021. **267**: p. 120543.
178. Ma, W., *A Method for Preliminary Prediction of Asphalt Mixture Permanent Deformation Test Results*, in *Airfield and Highway Pavements 2025*. p. 238-245.
179. Lang, H., I.L. Al-Qadi, and U. Mohamed Ali, *Designing Asphalt Concrete Mixtures Based on Balanced Mix Design Requirements Utilizing a Two-Step Approach.* Transportation Research Record, Washington, D.C, 2025. **2679**(10): p. 870-887.
180. Erten, K.M. and R. Gürfidan, *Regression-Based Performance Prediction in Asphalt Mixture Design and Input Analysis with SHAP.* Applied Sciences 2025. **15**(19): p. 10779.
181. Tong, B., et al., *Machine Learning–Based Prediction and Optimization of Balanced Mixture Design Performance Indices.* Transportation Research Record, Washington, D.C, 2025: p. 03611981251324209.
182. Cooper Jr, S.B., et al., *Development of a Predictive Model based on an Artificial Neural Network for the Semicircular Bend Test.* Transportation Research Record, Washington, D.C, 2016. **2576**(1): p. 83-90.
183. Nguyen, L.N., et al., *Machine Learning Approaches for Predicting Cracking Tolerance Index (CTIndex) of Asphalt Concrete Containing Reclaimed Asphalt Pavement.* Plos one, 2023. **18**(10): p. e0287255.
184. Rad, S.M., et al., *Evaluation of Asphalt Mixture Cracking Resistance and Development of a Machine Learning-Based Application for Cracking Tolerance Index Prediction.* Construction Building Materials, 2025. **490**: p. 142519.
185. Mirzaiyanrajeh, D., et al., *Developing a Prediction Model for Low-Temperature Fracture Energy of Asphalt Mixtures using Machine Learning Approach.* International Journal of Pavement Engineering, 2023. **24**(2): p. 2024185.
186. FAA, *Superpave Mixture Design Guide.* Federal Highway Administration, US Department of Transportation, Washington, DC, 2001.
187. Anirudh, N., K. Mallesh, and M.I. Anjum, *Influence of Particle Index of Coarse Aggregate and its Influence on Properties of Asphalt Concrete Mixtures.* International Journal of Research in Engineering Technology, 2014. **3**(6): p. 304.
188. Pine, W.J. and J. Wielinski, *The Bailey Method.* Asphalt Magazine, Asphalt Institute, 2696 Research Park Drive, Lexington, KY 40511-8480, 2023.
189. William, R.V., J.P. William, and H.C. Samuel, *Bailey Method for Gradation Selection in HMA Mixture Design.* Transportation Research Board, Washington, DC 2002.
190. AASHTO, *AASHTO T19: Standard Method of Test for Bulk Density (Unit Weight) and Voids in Aggregate.* AASHTO, Washington, D.C, 2025.
191. Hajj, E.Y. and N.G. Elias, *Balanced Mix Design: Rutting Performance Tests* National Asphalt Pavement Association (NAPA), 6406 Ivy Lane, Suite 350, Greenbelt, MD 2025.
192. ASTM, *ASTM D4867: Standard Test Method for Effect of Moisture on Asphalt Mixtures.* ASTM, West Conshohocken, PA, 2017, 2024.
193. AASHTO, *AASHTO T340: Standard Method of Test for Determining Rutting Susceptibility of Asphalt Mixtures Using the Asphalt Pavement Analyzer (APA).* AASHTO, Washington, D.C, 2023.
194. AASHTO, *AASHTO T 324: Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures.* AASHTO, Washington, D.C, 2019.

195. AASHTO, *AASHTO T401: Standard Method of Cantabro Abrasion Loss of Asphalt Mixture Specimens.* AASHTO, Washington, D.C, 2022.
196. FDOT, *FM 5-565: Measurement of Water Permeability of Compacted Asphalt Paving Mixtures.* State Materials Office, Gainesville, Florida, 2015., 2015.
197. AASHTO, *ASTM D8225-19: Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature.* AASHTO, Washington, D.C, 2019.
198. AASHTO, *AASHTO T393: Standard Method of Test for Determining the Fracture Potential of Asphalt Mixtures Using the Illinois Flexibility Index Test (I-FIT).* AASHTO, Washington, D.C, 2022.
199. ASTM, *ASTM D7313: Standard Test Method for Determining Fracture Energy of Asphalt Mixtures Using the Disk-Shaped Compact Tension Geometry.* ASTM, West Conshohocken, PA, 2017, 2020.
200. AASHTO, *AASHTO T 304: Uncompacted Void Content of Fine Aggregate.* AASHTO, Washington, D.C, 2022.
201. AASHTO, *AASHTO T27: Standard Method of Test for Sieve Analysis of Fine and Coarse Aggregates.* AASHTO, Washington, D.C, 2024.
202. ASTM, *ASTM C670: Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials.* ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA, 2025.
203. Azari, H., *Precision Estimates of AASHTO T283: Resistance of Compacted Hot Mix Asphalt to Moisture-Induced Damage.* NCHRP Web-Only Document 166, Washington, D.C, 2010.
204. Kim, S. and B.J. Coree, *Evaluation of Hot Mix Asphalt Moisture Sensitivity using the Nottingham Asphalt Test Equipment*. 2005, Center for Transportation Research and Education, Iowa State University.
205. Moraes, R., R. Velasquez, and H.U. Bahia, *Measuring the Effect of Moisture on Asphalt–Aggregate Bond with the Bitumen Bond Strength Test.* Transportation Research Record, Washington, D.C, 2011. **2209**(1): p. 70-81.
206. West, R.C., et al., *Laboratory Refinement and Field Validation of 4.75 mm Superpave Designed Asphalt Mixtures.* NCAT Report 11-01, National Center for Asphalt Technology, Auburn University, Alabama, 2011.
207. Habbouche, J., et al., *Mechanistic-Based Evaluation of Performance Thresholds for Balanced Mix Design Asphalt Surface Mixtures.* Report VTRC 26-R08, Virginia Transportation Research Council, 530 Edgemont Road Charlottesville, VA 22903 2025.
208. Habbouche, J., I. Boz, and S.D. Diefenderfer, *Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II.* Report No. FHWA/VTRC 23-R3, Virginia Transportation Research Council, 530 Edgemont Road Charlottesville, VA 22903 2022.
209. Tran, N., *Benchmarking Cracking Resistance of Asphalt Mixtures in Alabama.* 67th Annual Alabama Transportation Conference, Alabama, 2024.
210. Hajj, R., et al., *Updates to Mechanistic-Empirical Design Inputs for Illinois Flexible Pavements.* FHWA-ICT-24-008, Illinois Center for Transportation, University of Illinois at Urbana-Champaign, Urbana, IL 61801, 2024.
211. Hicks, R.G., *Moisture Damage in Asphalt Concrete*. Vol. 175. 1991: Transportation Research Board, Washington, D.C.

212. Alvarez, A.E., A.E. Martin, and C. Estakhri, *A Review of Mix Design and Evaluation Research for Permeable Friction Course Mixtures.* Construction Building Materials, 2011. **25**(3): p. 1159-1166.
213. Oliver, J.W., *The Effect of Binder Film Thickness on Asphalt Cracking and Ravelling.* Road Transport Research: A Journal of Australian New Zealand Research Practice, 2011. **20**(3): p. 3-13.
214. Elseifi, M.A., et al., *Validity of Asphalt Binder Film Thickness Concept in Hot-Mix Asphalt.* Transportation Research Record, Washington, D.C, 2008. **2057**(1): p. 37-45.
215. Jung, D. and T.S. Vinson, *Low-Temperature Cracking: Test Selection.* Strategic Highway Research Program National Research Council, 2101 Constitution Avenue N.W. Washington, DC 20418, 1994(SHRP-A-400).
216. Heritier, S., et al., *Robust Methods in Biostatistics*. 2009: John Wiley & Sons.
217. Huber, P.J., *Robust Estimation of a Location Parameter*, in *Breakthroughs in Statistics: Methodology and Distribution*. 1992, Springer. p. 492-518.
218. Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. , *Applied Linear Statistical Models.* 5th Edition, McGraw-Hill, Irwin, New York., 2005.
219. Hand, A. and A. Epps, *Effects of Test Variability on Mixture Volumetrics and Mix Design Verification.* Asphalt Paving Technology, 2000. **69**: p. 635-674.
220. Zhou, F., *Development of an IDEAL Cracking Test for Asphalt Mix Design, Quality Control and Quality Assurance* NCHRP IDEA Project 195, IDEA Programs Transportation Research Board, 500 Fifth Street, NW Washington, DC 2019.

APPENDIX A STEP-BY-STEP PROCEDURE FOR PREDICTION MODEL DEVELOPMENT

This appendix provides a complete, software-specific, and step-by-step procedure for replicating the prediction model development described in Chapter 6. The study used Microsoft Excel and Minitab Statistical Software to support the development process. Microsoft Excel was used for calculating group-centered responses and predictor variables. Minitab software was used to run Huber weighted regression, retrieve residuals, and perform K-fold cross-validation. The examples shown in the appendix are based on the second VMA prediction model, since the same development process was applied to the other prediction models.

Step 1. Data Preparation and Group Centering

Group-centering removes between-group (airport-to-airport) variation so that the regression fits only the within-group relationship between gradation changes and mixture responses. All centering is performed in Excel using **Equations 20** and **21**. **Table 41** presents the input data, where each row corresponds to an asphalt mixture observation. The table includes the following columns:

- Group ID: Identifies the P-401 airport mix designs. Mixtures that shared the same group ID were from the same P-401 mix design. In the example, the second VMA prediction model included a total of 10 group IDs.
- Predictor variables: the group-centered values of percentage aggregate passing at No.1/2, No.3/8…No.200 sieve ($\Delta\text{No.}\,1/2$,$\Delta\text{No.}\,3/8$….$\Delta\text{No.}\,200$) group-centered value of the specific gravity of the aggregate blend ΔG_{sb}, group-centered value of fine aggregate angularity of the blend (ΔFAA).
- Response variables: the group-centered value of VMA (ΔVMA).

Table 41. Input Data with Group-Centered Variables for Second VMA Prediction Model

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Group I	Delta No. 1/2	Delta No.3/8	Delta No.4	Delta No.8	Delta No.16	Delta No.30	Delta No.50	Delta No.100	Delta No.200	Delta Gsb	Delta FAA	Delta VMA
1	1	-0.3	-0.8	0.4	1.8	4.1	5.0	3.1	-0.0	-0.5	-0.0	-1.4	-1.3
2	1	1.7	4.9	3.0	1.9	2.9	3.2	2.3	0.9	0.6	-0.0	-0.9	-0.8
3	1	1.7	4.9	3.0	1.3	0.4	-0.1	-0.2	-0.1	-0.1	0.0	0.2	0.0
4	1	1.7	4.9	3.0	1.4	0.9	0.6	0.6	0.8	0.7	-0.0	0.1	0.0
5	1	-2.5	-6.8	-0.8	1.9	-1.0	-2.5	-1.9	-0.3	-0.1	0.0	0.9	0.5
6	1	-2.5	-7.0	-8.5	-8.3	-7.3	-6.2	-3.9	-1.2	-0.6	0.0	1.1	1.6
7	2	-0.1	-0.1	0.2	0.1	1.9	2.2	0.8	0.4	0.4	-0.0	-0.3	-1.2
8	2	0.4	-0.3	5.3	5.7	4.4	3.4	2.5	1.8	1.3	-0.0	0.1	-0.7
9	2	-0.1	-0.1	0.0	-0.4	0.0	0.0	-0.1	-0.1	0.1	-0.0	0.2	-0.5
10	2	-0.1	-0.1	0.0	-0.3	1.3	1.4	-0.1	-0.4	-0.1	-0.0	-0.2	-0.4
11	2	0.1	0.4	-2.4	-3.3	-3.0	-2.2	-0.5	0.1	0.1	0.0	0.5	0.2
12	2	-0.2	-0.4	-2.5	-2.6	-2.8	-2.3	-0.8	-0.3	-0.2	0.0	-0.1	0.6
13	2	-0.1	-0.0	2.4	4.4	2.3	1.1	0.2	-0.4	-0.6	0.0	-0.5	0.7
14	2	0.2	0.5	-3.0	-3.6	-4.1	-3.6	-1.9	-1.3	-1.0	0.0	0.2	1.4
15	3	1.1	2.1	2.6	1.2	0.7	0.8	1.2	1.1	0.9	-0.0	-1.1	-0.8
16	3	1.1	2.1	4.1	2.8	1.9	1.7	1.8	1.6	1.3	-0.0	0.8	-0.8
17	3	-0.4	-0.8	1.3	1.9	1.4	1.0	0.5	0.3	0.3	-0.0	0.3	-0.1
18	3	-0.4	-0.8	-1.9	-1.5	-1.1	-0.9	-0.8	-0.7	-0.6	0.0	0.6	0.2
19	3	-1.3	-2.6	-6.1	-4.3	-2.9	-2.6	-2.7	-2.4	-2.0	0.0	-0.5	1.5
20	4	-0.1	-0.5	-2.2	-1.2	-0.0	0.4	0.5	0.3	0.2	-0.0	-0.4	-1.1
21	4	0.2	0.0	-1.0	0.6	1.4	1.4	1.2	0.6	0.3	-0.0	-0.4	-0.4
22	4	0.9	2.0	4.2	4.4	3.4	2.4	1.6	0.7	0.3	-0.0	-0.4	-0.1
23	4	-1.5	-2.4	-2.7	-2.6	-2.1	-1.7	-1.1	-0.6	-0.4	0.0	0.6	-0.1
24	4	0.4	1.0	1.8	-1.1	-2.5	-2.5	-2.0	-0.9	-0.6	0.0	0.6	1.7
25	5	-1.3	-2.0	-1.8	-2.1	-1.3	-0.7	-0.3	-0.0	0.0	-0.0	0.0	-0.3
26	5	0.2	0.8	-0.3	2.1	1.4	0.8	0.4	0.2	0.1	0.0	0.0	-0.1
27	5	0.2	0.8	3.0	2.1	1.4	0.8	0.4	0.2	0.1	0.0	0.0	0.2
28	5	0.8	0.5	-0.9	-2.2	-1.4	-0.9	-0.6	-0.3	-0.2	-0.0	0.0	0.3
29	6	-1.2	-2.9	-1.3	-2.0	-2.3	-1.6	-0.1	0.2	0.3	-0.0	0.0	-0.4
30	6	0.8	1.8	-1.1	-1.7	-1.3	-0.7	0.1	0.5	0.7	0.0	0.0	-0.3
31	6	-0.4	-0.4	3.3	3.2	2.2	1.2	0.4	0.2	0.2	0.0	0.0	-0.3
32	6	-0.3	-0.5	1.3	2.3	2.5	1.7	0.2	-0.2	-0.3	-0.0	0.0	-0.2
33	6	-0.7	-2.1	-2.1	0.0	1.3	1.3	0.2	0.0	-0.0	-0.0	0.0	0.0
34	6	1.8	4.2	0.2	-1.9	-2.3	-1.9	-1.0	-0.7	-0.7	0.0	0.0	1.2
35	7	-2.2	-3.2	-1.5	-0.6	-0.6	-0.5	-0.6	-0.4	-0.3	0.0	-0.5	-0.1
36	7	4.4	6.9	3.1	1.7	1.4	1.1	1.0	0.8	0.6	-0.0	0.9	0.0
37	7	-2.2	-3.7	-1.7	-1.1	-0.7	-0.6	-0.3	-0.4	-0.3	0.0	-0.5	0.0
38	8	1.0	6.0	3.8	0.1	0.8	0.3	-0.1	-0.5	-0.4	-0.0	-1.2	-2.5
39	8	1.3	4.9	7.9	7.8	5.2	3.9	1.7	0.4	-0.0	-0.0	-0.2	-0.0
40	8	-0.7	-2.9	-2.8	-0.5	2.9	2.5	1.5	0.2	0.3	-0.0	-1.2	0.2
41	8	-1.4	-5.9	-3.1	-1.8	-3.1	-2.4	-1.0	0.3	0.3	0.0	1.0	0.2
42	8	-0.4	-2.0	-5.8	-5.7	-5.9	-4.5	-2.1	-0.3	-0.0	0.0	1.6	2.3
43	9	-0.4	-0.6	2.2	1.9	1.1	0.8	0.6	0.4	0.4	-0.0	0.0	-0.3
44	9	-0.7	-1.4	-2.4	-3.3	-3.3	-1.8	-0.2	0.2	0.2	0.0	0.0	-0.2
45	9	0.4	0.9	4.8	6.5	6.1	3.5	0.5	-0.1	-0.1	-0.0	0.0	-0.2
46	9	0.6	0.9	-4.6	-5.0	-4.1	-2.5	-1.1	-0.6	-0.3	0.0	0.0	0.8
47	10	-0.3	-1.0	0.0	0.5	0.7	0.8	0.9	1.1	1.3	-0.0	-0.1	-1.5
48	10	1.1	3.4	1.4	1.3	1.4	1.2	0.6	0.5	0.4	-0.0	0.1	-0.2
49	10	-0.3	-1.0	0.1	0.3	0.1	0.0	-0.1	-0.3	-0.5	-0.0	-0.5	-0.1
50	10	-1.0	-3.6	-1.0	-0.3	-0.8	-1.0	-0.6	-0.2	-0.1	0.0	0.8	-0.1
51	10	1.4	5.5	6.1	4.5	3.5	2.7	1.7	0.5	-0.2	-0.0	-0.4	0.5
52	10	-0.9	-3.3	-6.7	-6.3	-4.9	-3.8	-2.6	-1.5	-0.9	0.0	0.0	1.3

Step 2: Run Initial OLS Regression in Minitab

This first regression uses equal weights of 1 (standard OLS) to obtain an initial set of residuals.

1. Open Minitab and import the Excel file from Step 1 using File → Open → Excel.
2. Navigate to Stat → Regression → Regression → Fit Regression Model.
3. In the regression dialog box, as shown in **Figure 54**:

- Responses: Set to ΔVMA

- Continuous predictors: Set to $\Delta No.1/2, \Delta No.3/8 \ldots . \Delta No.200$, ΔG_{sb}, ΔFAA (If the model includes fine aggregate angularity)

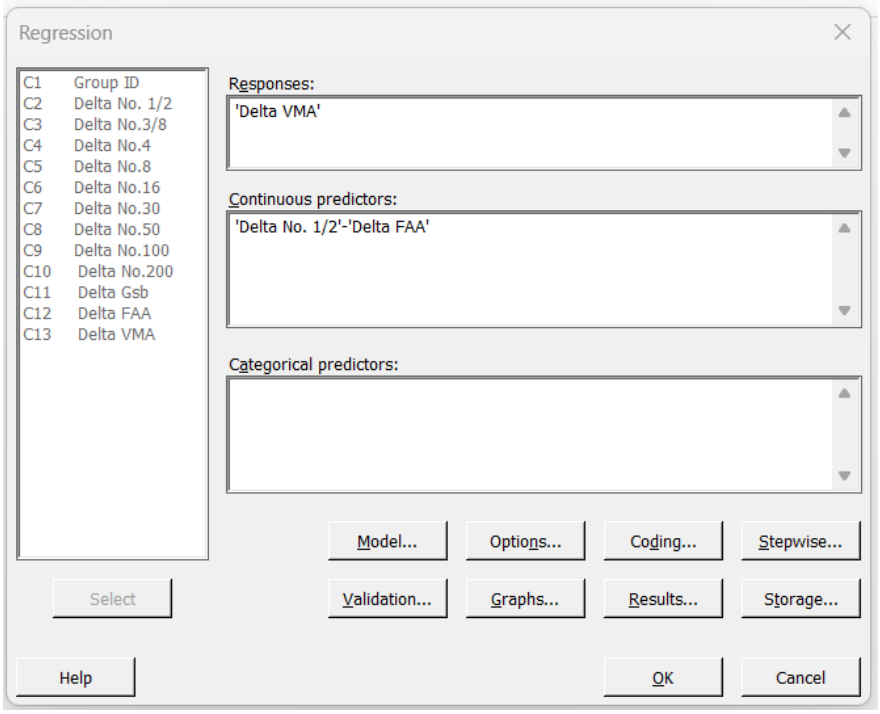

Figure 54. Minitab Regression Dialog Box

4. Click the Storage in the regression dialog box (**Figure 54**) to open Regression: Storage dialog box (**Figure 55**), check the Residuals box and click OK. This will create a new column labeled as RESI in the Minitab worksheet when you run the regression.

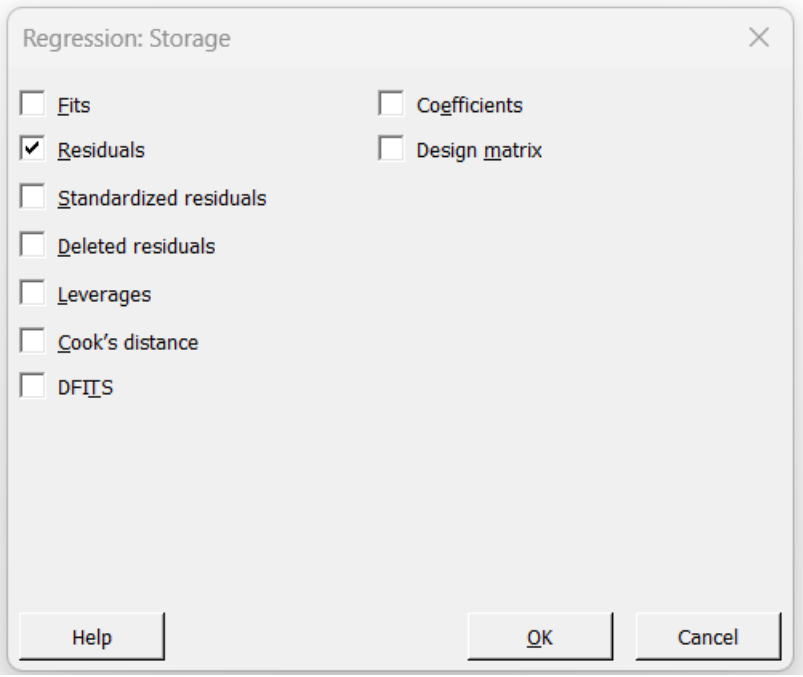

Figure 55. Minitab Regression: Storage Options

5. Click OK in the main regression dialog box (**Figure 54**) to run the regression. Minitab will produce the initial regression output and store residuals in column RESI (C14), as shown in Table 42. The residuals represent the difference between the measured and fitted group-centered response values.

Table 42. Minitab Initial Regression Output with Residual in Column C14

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Group I	Delta No. 1/	Delta No.3/	Delta No.	Delta No.	Delta No.1	Delta No.3	Delta No.5	Delta No.10	Delta No.2	Delta Gs	Delta FA	Delta VM	RESI
1	1	-0.3	-0.8	0.4	1.8	4.1	5.0	3.1	-0.0	-0.5	-0.0	-1.4	-1.3	-0.20982
2	1	1.7	4.9	3.0	1.9	2.9	3.2	2.3	0.9	0.6	-0.0	-0.9	-0.8	0.35725
3	1	1.7	4.9	3.0	1.3	0.4	-0.1	-0.2	-0.1	-0.1	0.0	0.2	0.0	-0.39081
4	1	1.7	4.9	3.0	1.4	0.9	0.6	0.6	0.8	0.7	-0.0	0.1	0.0	0.54327
5	1	-2.5	-6.8	-0.8	1.9	-1.0	-2.5	-1.9	-0.3	-0.1	0.0	0.9	0.5	-0.28027
6	1	-2.5	-7.0	-8.5	-8.3	-7.3	-6.2	-3.9	-1.2	-0.6	0.0	1.1	1.6	-0.01962
7	2	-0.1	-0.1	0.2	0.1	1.9	2.2	0.8	0.4	0.4	-0.0	-0.3	-1.2	-0.48799
8	2	0.4	-0.3	5.3	5.7	4.4	3.4	2.5	1.8	1.3	-0.0	0.1	-0.7	0.06016
9	2	-0.1	-0.1	0.0	-0.4	0.0	0.0	-0.1	-0.1	0.1	-0.0	0.2	-0.5	-0.04840
10	2	-0.1	-0.1	0.0	-0.3	1.3	1.4	-0.1	-0.4	-0.1	-0.0	-0.2	-0.4	0.07762
11	2	0.1	0.4	-2.4	-3.3	-3.0	-2.2	-0.5	0.1	0.1	0.0	0.5	0.2	0.01634
12	2	-0.2	-0.4	-2.5	-2.6	-2.8	-2.3	-0.8	-0.3	-0.2	0.0	-0.1	0.6	0.26403
13	2	-0.1	-0.0	2.4	4.4	2.3	1.1	0.2	-0.4	-0.6	0.0	-0.5	0.7	-0.25470
14	2	0.2	0.5	-3.0	-3.6	-4.1	-3.6	-1.9	-1.3	-1.0	0.0	0.2	1.4	0.37294
15	3	1.1	2.1	2.6	1.2	0.7	0.8	1.2	1.1	0.9	-0.0	-1.1	-0.8	0.26874
16	3	1.1	2.1	4.1	2.8	1.9	1.7	1.8	1.6	1.3	-0.0	0.8	-0.8	-0.10826
17	3	-0.4	-0.8	1.3	1.9	1.4	1.0	0.5	0.3	0.3	-0.0	0.3	-0.1	0.28057
18	3	-0.4	-0.8	-1.9	-1.5	-1.1	-0.9	-0.8	-0.7	-0.6	0.0	0.6	0.2	-0.49721
19	3	-1.3	-2.6	-6.1	-4.3	-2.9	-2.6	-2.7	-2.4	-2.0	0.0	-0.5	1.5	0.05616
20	4	-0.1	-0.5	-2.2	-1.2	-0.0	0.4	0.5	0.3	0.2	-0.0	-0.4	-1.1	-0.23423
21	4	0.2	0.0	-1.0	0.6	1.4	1.4	1.2	0.6	0.3	-0.0	-0.4	-0.4	-0.16578
22	4	0.9	2.0	4.2	4.4	3.4	2.4	1.6	0.7	0.3	-0.0	-0.4	-0.1	0.17220
23	4	-1.5	-2.4	-2.7	-2.6	-2.1	-1.7	-1.1	-0.6	-0.4	0.0	0.6	-0.1	-0.15641
24	4	0.4	1.0	1.8	-1.1	-2.5	-2.5	-2.0	-0.9	-0.6	0.0	0.6	1.7	0.38423
25	5	-1.3	-2.0	-1.8	-2.1	-1.3	-0.7	-0.3	-0.0	0.0	-0.0	0.0	-0.3	0.28141
26	5	0.2	0.8	-0.3	2.1	1.4	0.8	0.4	0.2	0.1	0.0	0.0	-0.1	-0.68615
27	5	0.2	0.8	3.0	2.1	1.4	0.8	0.4	0.2	0.1	0.0	0.0	0.2	0.03624
28	5	0.8	0.5	-0.9	-2.2	-1.4	-0.9	-0.6	-0.3	-0.2	-0.0	0.0	0.3	0.36851
29	6	-1.2	-2.9	-1.3	-2.0	-2.3	-1.6	-0.1	0.2	0.3	-0.0	0.0	-0.4	0.03617
30	6	0.8	1.8	-1.1	-1.7	-1.3	-0.7	0.1	0.5	0.7	0.0	0.0	-0.3	-0.17853
31	6	-0.4	-0.4	3.3	3.2	2.2	1.2	0.4	0.2	0.2	0.0	0.0	-0.3	0.06497
32	6	-0.3	-0.5	1.3	2.3	2.5	1.7	0.2	-0.2	-0.3	-0.0	0.0	-0.2	0.02752
33	6	-0.7	-2.1	-2.1	0.0	1.3	1.3	0.2	0.0	-0.0	-0.0	0.0	0.0	-0.05410
34	6	1.8	4.2	0.2	-1.9	-2.3	-1.9	-1.0	-0.7	-0.7	0.0	0.0	1.2	0.10396
35	7	-2.2	-3.2	-1.5	-0.6	-0.6	-0.5	-0.6	-0.4	-0.3	0.0	-0.5	-0.1	0.37781
36	7	4.4	6.9	3.1	1.7	1.4	1.1	1.0	0.8	0.6	-0.0	0.9	0.0	-0.47966
37	7	-2.2	-3.7	-1.7	-1.1	-0.7	-0.6	-0.3	-0.4	-0.3	0.0	-0.5	0.0	0.10184
38	8	1.0	6.0	3.8	0.1	0.8	0.3	-0.1	-0.5	-0.4	-0.0	-1.2	-2.5	-1.38765
39	8	1.3	4.9	7.9	7.8	5.2	3.9	1.7	0.4	-0.0	-0.0	-0.2	-0.0	0.34808
40	8	-0.7	-2.9	-2.8	-0.5	2.9	2.5	1.5	0.2	0.3	-0.0	-1.2	0.2	0.60968
41	8	-1.4	-5.9	-3.1	-1.8	-3.1	-2.4	-1.0	0.3	0.3	0.0	1.0	0.2	-0.47477
42	8	-0.4	-2.0	-5.8	-5.7	-5.9	-4.5	-2.1	-0.3	-0.0	0.0	1.6	2.3	0.90465
43	9	-0.4	-0.6	2.2	1.9	1.1	0.8	0.6	0.4	0.4	-0.0	0.0	-0.3	0.14360
44	9	-0.7	-1.4	-2.4	-3.3	-3.3	-1.8	-0.2	0.2	0.2	0.0	0.0	-0.2	-0.41429
45	9	0.4	0.9	4.8	6.5	6.1	3.5	0.5	-0.1	-0.1	-0.0	0.0	-0.2	0.42834
46	9	0.6	0.9	-4.6	-5.0	-4.1	-2.5	-1.1	-0.6	-0.3	0.0	0.0	0.8	-0.15764
47	10	-0.3	-1.0	0.0	0.5	0.7	0.8	0.9	1.1	1.3	-0.0	-0.1	-1.5	-0.04985
48	10	1.1	3.4	1.4	1.3	1.4	1.2	0.6	0.5	0.4	-0.0	0.1	-0.2	0.04462
49	10	-0.3	-1.0	0.1	0.3	0.1	0.0	-0.1	-0.3	-0.5	-0.0	-0.5	-0.1	-0.36571
50	10	-1.0	-3.6	-1.0	-0.3	-0.8	-1.0	-0.6	-0.2	-0.1	0.0	0.8	-0.1	-0.71113
51	10	1.4	5.5	6.1	4.5	3.5	2.7	1.7	0.5	-0.2	-0.0	-0.4	0.5	0.72973
52	10	-0.9	-3.3	-6.7	-6.3	-4.9	-3.8	-2.6	-1.5	-0.9	0.0	0.0	1.3	0.35234

Step 3: Calculate Huber Weight

Huber weights are computed from the residuals. Observations with large residuals receive reduced weights to limit the influence of outliers on the regression.

1. Navigate to View → Command Line/History. A Command Line window will appear. Add a code which was used to calculate Huber weights, built based on **Equations 24**, **25,** and **26**, to

the Command Line window, as shown in **Figure 56**. Note that only the first line needs to be changed in each iteration.

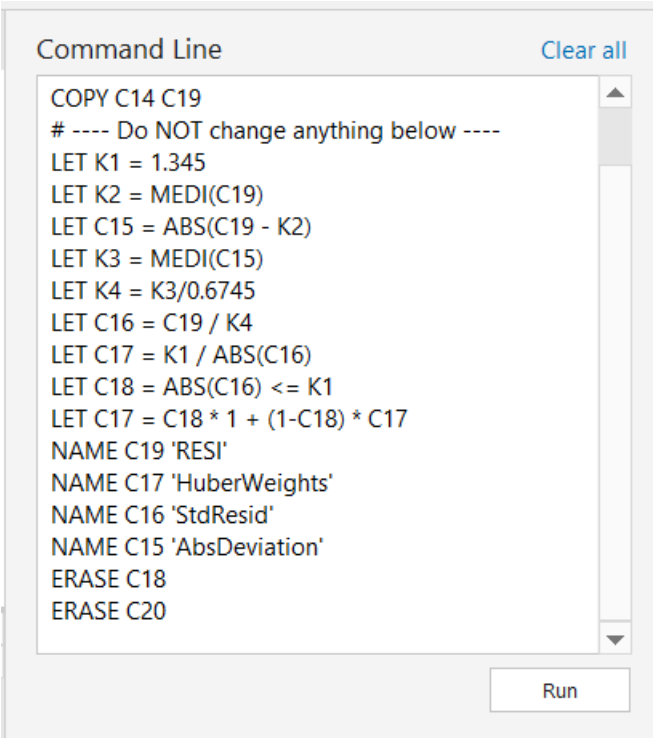

Figure 56. Command Line with Huber Weight Calculation Code

2. Click Run in the Command Line window (**Figure 56**). The Huber weights will be calculated and stored in column C17 (HuberWeights).

Step 4. Run Huber Weighted Regression in Minitab

In this step, the Huber weights computed in Step 3 are incorporated into the regression model to reduce the influence of outliers.

1. Repeat Step 2.2 to open the regression dialog box **(Figure 54)** and keep all predictors and response settings identical to Step 2.
2. Click Options in the main regression dialog box (**Figure 54**) to open the Regression: Options dialog box (**Figure 57**). Locate the Weights field, select C17 HuberWeights, and click Select. Then click OK to close the dialog box.

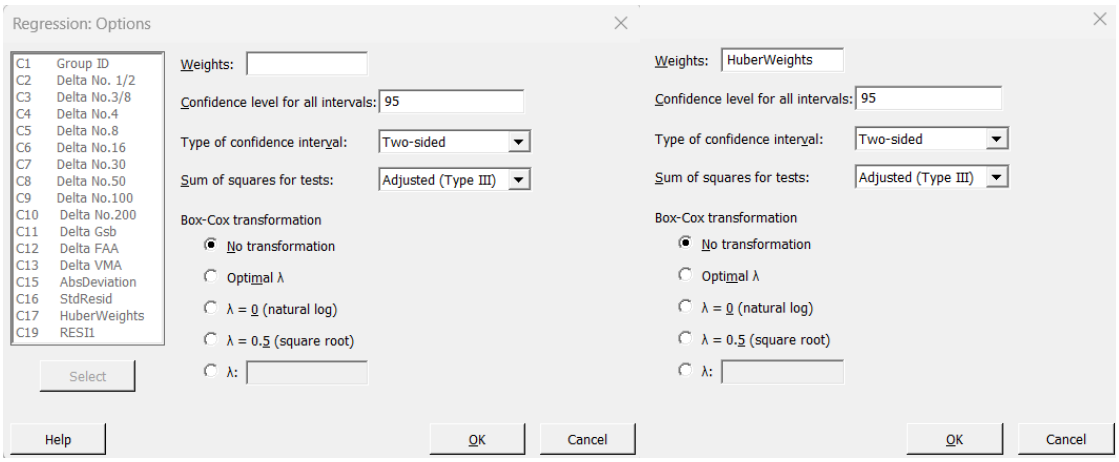

Figure 57. Minitab Regression: Options- Selecting Huber Weights

3. Click OK in the main regression dialog box (**Figure 54**) to run the weighted regression. A residual that now accounts for Huber weights was stored in Column C20.

Step 5. Recalculate Huber Weights and Iterate

1. Return to the Command Line window (View → Command Line/History). In the code from Step 3, update the first line only, change COPY C14 C19 to COPY C20 C19 (or whichever column contains the most recent residuals). Click Run. Updated Huber weights will be stored in column C17.
2. Repeat steps 4 and 5.1 until the regression coefficients stabilize across iterations.

Step 6. K-fold Cross-Validation

When the regression coefficients stabilize, K-fold cross-validation is conducted to evaluate whether the model can reliably predict unseen data.

1. Navigate to Stat → Regression → Regression → Fit Regression Model to open the main regression dialog box (**Figure 54**).
2. Click Validation to open the Regression: Validation dialog box. Under Validation method, select K-fold cross-validation. Select “Assign of each fold by ID column” option. In the ID

Column, locate the C1 Group ID from the column list, and click Select. Click OK to close the dialog box. The process is shown in **Figure 58**.

3. Click OK in the main regression dialog box (**Figure 54**) to run the cross-validation.

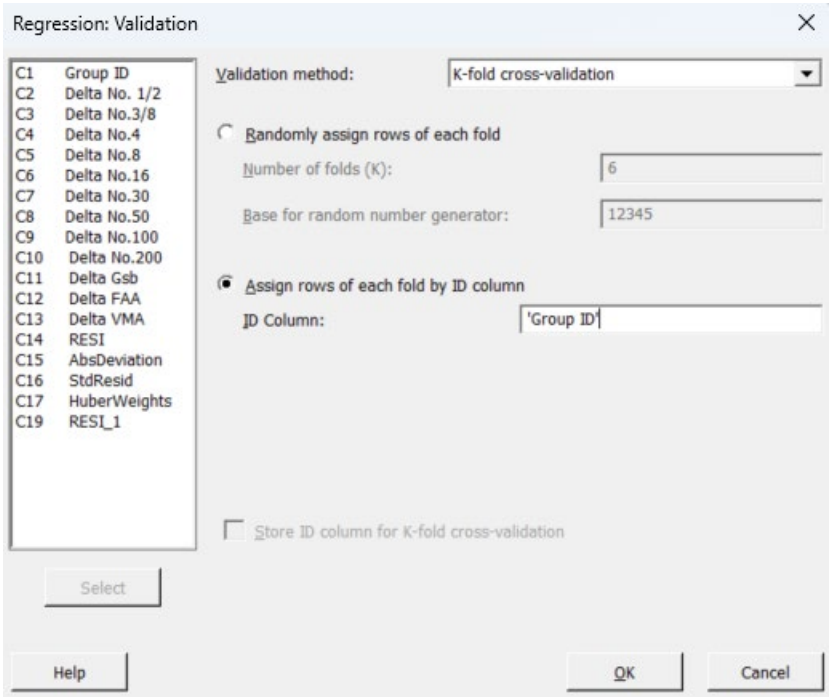

Figure 58. K-Fold Cross- Validation using Group ID

The results are shown in **Figure 59**, including the regression equation and model performance metrics.

Method

Weights	HuberWeights
Cross-validation ID	Group ID
Cross-validation	10-fold

Regression Equation

Delta VMA = 0.0282 + 0.314 Delta No. 1/2 - 0.0271 Delta No.3/8 - 0.0036 Delta No.4
- 0.044 Delta No.8 + 0.099 Delta No.16 - 0.324 Delta No.30 + 0.969 Delta No.50
- 0.430 Delta No.100 - 1.074 Delta No.200 + 87.7 Delta Gsb + 0.135 Delta FAA

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
0.300079	86.26%	82.48%	77.61%	0.394390	69.14%

Figure 59. Minitab Results of Second VMA Prediction Model