Airport Asphalt Pavement Technology Program

Revisions to Aggregate Gradation Bands for P-401 Mixtures

Final Report

November 2026



Credit: AAPTP

Airport Asphalt Pavement Technology Program

PREPARED BY:

National Center for Asphalt Technology at Auburn University
Nam H. Tran, Trung Tran, Chen Chen, Fan Yin

Heritage Research Group
Bill Pine, Gerry Huber, Katie DeCarlo, Faustina Keuliyan Rodriguez

University of Nevada, Reno
Adam Hand, Elie Y. Hajj, George Aad, Naveen Parameswaran

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

Acknowledgments

The authors gratefully acknowledge project co-chairs Brett Williams and Kristen Morris and panel members Matthew Hoyle, Brian Wood, Jeremy Barron, Jim Scherocman, Hadi Rashidi, and Richard Willis for their guidance, project oversight, and review of deliverables. The authors also thank Chris Campbell and Marissa Childs for their support in materials preparation and testing, Josue Garita for assistance in developing the web-based tool, and Tianhao Yan and Bertram Zinner for reviewing the model development procedure and results.



Credit: AAPTP

Table of Contents

Executive Summary **1**

Chapter 1. Introduction **3**

1.1 Problem Statement 3

1.2 Research Objectives 5

1.3 Organization of the Document 5

Chapter 2. Literature Review **6**

2.1 Evolution of Aggregate Gradation Limits for Asphalt Mixtures 6

2.1.1 Sieve Sizes and Gradations 6

2.1.2 Marshall Method and the Origin of P-401 Gradation Limits 7

2.1.3 The Superpave Framework and Current AASHTO Gradation Control 11

2.1.4 Evaluation of Superpave Gradation for Airfields 12

2.1.5 Summary 14

2.2 Impact of Aggregate Properties on Volumetric Properties 15

2.3 Impact of Aggregate Consensus and Source Properties on Mixture Performance 16

2.3.1 Impact of Consensus Properties of Aggregates on Mixture Performance 16

2.3.2 Impact of Sand Equivalent Value on Mixture Performance 19

2.3.3 Impact of Source Properties of Aggregate on Mixture Performance 20

2.3.4 Summary 22

2.4 Impact of Aggregate Gradation on Volumetric Properties of Asphalt Mixtures 23

2.5 Impact of Aggregate Gradation on Performance of Asphalt Mixtures 25

2.5.1 Rutting Resistance 25

2.5.2 Cracking Resistance 26

2.5.3 Durability 28

2.5.4 Permeability 30

2.5.5 Summary 31

2.6 Bailey Method 33

2.6.1 An Approach to Understanding Aggregate Packing 33

2.6.2 NMAS and PCS in the Bailey Method 33

2.6.3 Establishing CA Volume to Determine Mix Type 34

2.6.4 Aggregate Packing 36

2.6.5 Summary 46

2.7 Prediction Models for Volumetric Properties and Performance of Asphalt Mixtures 46

2.7.1 Prediction Models for Volumetric Properties 46
2.7.2 Prediction Models for Rut Depth 48
2.7.3 Prediction Models for Cracking Resistance 50
2.7.4 Summary 52
Chapter 3. Research Methodology 54
3.1 Overview 54
3.2 Experimental Design 54
3.3 Selection of Existing P-401 Mix Designs and Materials 56
3.4 Adjustment of Aggregate Gradation in P-401 Mix Designs 58
3.5 Laboratory Performance and Aggregate Blend Evaluation 60
3.5.1 Moisture Susceptibility 61
3.5.2 Rutting Resistance 62
3.5.3 Durability 64
3.5.4 Permeability 64
3.5.5 Intermediate-Temperature Cracking Resistance 66
3.5.6 Low-Temperature Cracking Resistance 68
3.5.7 Fine Aggregate Angularity Test 69
3.6 Analysis of Test Results 71
Chapter 4. Development of Proposed Revisions to P-401 Gradation Limits 72
4.1 Execution of Part 1 of Experimental Plan 72
4.1.1 Step 1: Design of In-Spec 1 72
4.1.2 Step 2: Design of Out-Spec 1 74
4.1.3 Step 3: Laboratory Performance Evaluation 74
4.1.4 Step 4: Evaluation for Gradation Limits Revision 74
4.2 In-Spec 1 and Out-Spec 1 Design Summary 75
4.3 Test Results and Data Analysis 80
4.3.1 Moisture Susceptibility Evaluation 80
4.3.2 Permeability Evaluation 82
4.3.3 Durability (Cantabro Mass Loss) 83
4.3.4 Rutting Resistance 84
4.3.5 Intermediate-Temperature Cracking Resistance 86
4.3.6 Low-Temperature Cracking Resistance 89
4.3.7 Summary 91

4.4 Proposed Revisions to P-401 Gradation Limits 91
4.4.1 Upper-Limit Evaluation 91
4.4.2 Lower-Limit Evaluation 92
4.4.3 Engineering Rationale for Proposed Adjustments to Lower Limits 93
Chapter 5. Effect of Gradation Changes on Volumetric and Performance Properties .. 95
5.1 Execution of Part 2 of Experimental Design 95
5.2 Mix Design Summary 97
5.3 Test Results and Data Analysis 101
5.3.1 Effects of Gradation Changes on Volumetric Properties 101
5.3.2 Effects of Gradation Changes on Laboratory Performance 102
5.4 Summary 109
Chapter 6. Prediction Models for Volumetric and Performance Properties 111
6.1 Methodology for Model Development 111
6.1.1 Step 1: Database Structure and Model Segmentation 112
6.1.2 Step 2: Statistical Modeling and Validation 114
6.1.3 Step 3: Back-Transformation to Original Response Scale 119
6.1.4 Web-Based Deployment 120
6.2 Prediction Model for Volumetric Properties 121
6.2.1 First Prediction Model for Volumetric Properties 121
6.2.2 Second Prediction Model for Volumetric Properties 124
6.3 Rut Depth Prediction Model 128
6.3.1 Model Formulation 128
6.3.2 Model Evaluation 129
6.3.3 Cross-Validation 129
6.3.4 Back-Transformation and Comparison to Measured Rut Depth 130
6.3.5 Rutting Prediction Model Without Fine Aggregate Angularity 131
6.4 Intermediate-Temperature Cracking Prediction Models 132
6.4.1 Intermediate-Temperature Cracking Prediction Model Based on CT_{Index} 132
6.4.2 Intermediate-Temperature Cracking Prediction Model Based on FI 136
6.5 Model Performance Summary 139
6.6 Web-Based Decision Tool 140
Chapter 7. Conclusions and Recommendations 144
7.1 Proposed Revisions to Current P-401 Gradation Limits 144
7.2 Effects of Gradation Changes on Mixture Volumetric and Performance Properties 146

7.3 Models for Predicting Changes in Volumetric and Performance Properties Due to Gradation Changes 147
7.4 Recommendations for Future Research 147
References 149
Appendix A: Step-by-Step Procedure for Prediction Model Development 164

List of Figures

Figure 1. Gradation Limits for 3/4-Inch Maximum Size Asphaltic Concrete Mixture [7] 10
Figure 2. Illustration of Current Superpave Gradation Limits 12
Figure 3. An Aggregate Blend Plotted Within P-401 Gradation Limits and Superpave Control Points [13] 13
Figure 4. Illustration of the Bailey Method Applied to FG Mixtures, Showing Redefinition of Control Sieves and Aggregate Fractions Using the Combined Gradation Passing the PCS [102] 41
Figure 5. Fine Fraction (Passing the Original PCS) Treated as a New Blend Gradation for Evaluating FG Mixtures [102] 44
Figure 6. Conceptual Relationship Between the "Old" FA_c Ratio and VMA for FG Mixtures [102] 45
Figure 7. Experimental Plan 55
Figure 8. Geographic Locations of the Selected P-401 Mix Designs 58
Figure 9. Illustration of Adjustments of Gradation Close to Upper Limits 59
Figure 10. Illustration of Adjustments of Gradation Close to Lower Limits 60
Figure 11. Tensile Strength Ratio (TSR) 61
Figure 12. Asphalt Pavement Analyzer (APA) 63
Figure 13. Hamburg Wheel-Tracking Test (HWTT) 63
Figure 14. Cantabro Test 64
Figure 15. Florida Permeability Test 65
Figure 16. Indirect Tensile Asphalt Cracking Test (IDEAL-CT) 67
Figure 17. Illinois Flexibility Index Test (I-FIT) 68
Figure 18. Disk-Shaped Compact Tension (DCT) Test 68
Figure 19. DCT Sample 69
Figure 20. Fine Aggregate Angularity (FAA) Test 71
Figure 21. Experimental Design (Part 1) for Revising P-401 Gradation Limits 73
Figure 22. Asphalt Binder Contents of In-Spec 1 and Out-Spec 1 78
Figure 23. VMA of In-Spec 1 and Out-Spec 1 79
Figure 24. V_{be} of In-Spec 1 and Out-Spec 1 80
Figure 25. TSR Test Results of In-Spec 1 and Out-Spec 1 81
Figure 26. Permeability Test Results of In-Spec 1 and Out-Spec 1 83
Figure 27. Cantabro Mass Loss of In-Spec 1 and Out-Spec 1 84
Figure 28. APA/HWTT Test Results of In-Spec 1 and Out-Spec 1 85
Figure 29. CT_{Index} Results for In-Spec 1 and Out-Spec 1 87
Figure 30. FI Values of In-Spec 1 and Out-Spec 1 88
Figure 31. Fracture Energy from DCT of In-Spec 1 and Out-Spec 1 90
Figure 32. Experimental Design (Part 2) for Evaluating the Effect of Gradation Changes on Volumetric and Performance Properties 97
Figure 33. Comparison of VMA of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2 102

Figure 34. Comparison of AV of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2..... 102
Figure 35. TSR Test Results of Mix 1 and Mix 3 103
Figure 36. Permeability Test Results of Mix 1 and Mix 3 104
Figure 37. Cantabro Test Results for Mix 1 and Mix 3 105
Figure 38. APA/HWTT Test Results 106
Figure 39. IDEAL-CT Results..... 107
Figure 40. I-FIT Results 108
Figure 41. DCT Results 109
Figure 42. Workflow for the Development of the Prediction Models 112
Figure 43. Workflow of the Prediction Tool..... 121
Figure 44. Residuals and Fitted Values of ΔVMA in the First Volumetric Prediction Model 123
Figure 45. Measured and Calculated VMA from First Volumetric Prediction Model 124
Figure 46. Residuals and Fitted Values of ΔVMA in Second Volumetric Prediction Model 126
Figure 47. Measured and Calculated VMA from Second Volumetric Prediction Model 127
Figure 48. Residuals and Fitted Values of ΔRut Depth 129
Figure 49. Measured and Calculated Rut Depth..... 131
Figure 50. Residuals and Fitted Values of ΔCT_{Index}..... 133
Figure 51. Measured and Calculated CT_{Index} 135
Figure 52. Residuals and Fitted Values of ΔFI..... 137
Figure 53. Measured and Calculated FI 138
Figure 54. Screenshot of the Web-Based Tool for Estimating VMA, Rutting, and Cracking Properties..... 141
Figure 55. Example of Prediction Tool for CT_{Index} 143
Figure 56. Minitab Regression Dialog Box 166
Figure 57. Minitab Regression: Storage Options 167
Figure 58. Command Line with Huber Weight Calculation Code 169
Figure 59. Minitab Regression: Options—Selecting Huber Weights 170
Figure 60. K-fold Cross-Validation Using Group ID..... 171
Figure 61. Minitab Results of Second VMA Prediction Model..... 171

List of Tables

Table 1. Gradation and VMA Requirements in P-401 and AASHTO M 323 Superpave Specifications [1, 2] 4
Table 2. 1948 Gradation Limits for Airfield Asphalt Mixtures [8] 9
Table 3. Adaptation of Superpave Gradation Limits for Airfield Asphalt Mixtures [8] 14
Table 4. Summary of Findings on Impact of Aggregate Properties on VMA 15
Table 5. Summary of Findings on the Impact of CAA and FAA on Mixture Performance 17
Table 6. Summary of Findings on the Impact of FE Particles on Mixture Performance 18
Table 7. Summary of Findings on the Impact of SE Value and Clay Content on Mixture Performance 19
Table 8. Summary of Findings on the Impact of Aggregate Toughness on Mixture Performance 21
Table 9. Effect of Aggregate Consensus and Source Properties on Mixture Performance 22
Table 10. Effect of Gradation on VMA 24
Table 11. Effect of Aggregate Gradation on Rutting Resistance of Asphalt Mixtures 25
Table 12. Effect of Aggregate Gradation on Cracking Resistance of Asphalt Mixtures 27
Table 13. Effect of Aggregate Gradation on Durability of Asphalt Mixtures 29
Table 14. Effect of Aggregate Gradation on Permeability of Asphalt Mixtures 31
Table 15. Impact of Gradation on Asphalt Mixture Performance 32
Table 16. Primary Control Sieves Used in the Bailey Method [159, 160] 34
Table 17. Suggested CA CUW Values for Each Mix Type [102] 35
Table 18. Half Sieve and PCS for Each NMAS [102] 37
Table 19. Suggested CA Ratio Ranges for CG and SMA Mixtures [102] 38
Table 20. Suggested Ranges for CA, FA_c, and FA_f Ratios for CG Mixtures [102] 40
Table 21. Suggested Ranges for CA, FA_c, and FA_f Ratios for SMA Mixtures [102] 40
Table 22. Suggested Ranges for Recalculated CA, FA_c, and FA_f Ratios for FG Mixtures Based on the Controlling Blend Passing the PCS [102] 42
Table 23. Summary of VMA Prediction Models 48
Table 24. Summary of Rut Depth Prediction Models 50
Table 25. Summary of Cracking Performance Prediction Models 52
Table 26. P-401 Mix Design Summary 57
Table 27. Laboratory Performance Tests 61
Table 28. Design Summary of In-Spec 1 and Out-Spec 1 76
Table 29. TSR Difference Results 82
Table 30. P-Values of Permeability Test Results 83
Table 31. P-Values of Cantabro Mass Loss 84
Table 32. P-Values of Rut Depths 86
Table 33. D2s and P-Values of CT_{Index} for In-Spec 1 and Out-Spec 1 87
Table 34. D2s and P-Values of FI of In-Spec 1 and Out-Spec 1 89
Table 35. D2S and P-Values of DCT Test Results of In-Spec 1 and Out-Spec 1 90

Table 36. Number of Mixes and Percentage Deviations from Upper Limits by Sieve Size 92
Table 37. Number of Mixes and Percentage Deviations from Lower Limits by Sieve Size 92
Table 38. Current and Revised P-401 Gradation Limits for 12.5-mm NMAS Mix Designs 93
Table 39. Design Summary of In-Spec 2 and Out-Spec 2 99
Table 40. Statistical Metrics of Predicted and Measured ΔVMA for First Volumetric Prediction Model 123
Table 41. Statistical Metrics of Predicted and Measured ΔVMA for Second Volumetric Prediction Model 126
Table 42. Statistical Metrics of Predicted and Measured ΔVMA for Second Volumetric Prediction Model without FAA 128
Table 43. Statistical Metrics of Predicted and Measured ΔRut Depth 129
Table 44. Statistical Metrics of Predicted and Measured ΔRut Depth from the Prediction Model Without FAA 132
Table 45. Statistical Metrics of Predicted and Measured ΔCT_{Index} 134
Table 46. Statistical Metrics of the CT_{Index} Prediction Model Without FAA 135
Table 47. Statistical Metrics of Predicted and Measured ΔFI 137
Table 48. Statistical Metrics of Predicted and Measured ΔFI from Prediction Model Without FAA 139
Table 49. Cross-Validated Performance Metrics of Prediction Models (Group-Centered Scale) 139
Table 50. Back-Transformed Prediction Accuracy of Models with FAA 139
Table 51. Input Data with Group-Centered Variables for Second VMA Prediction Model .. 165
Table 52. Minitab Initial Regression Output with Residual in Column C14 168

Acronyms and Abbreviations

AAPTP	Airport Asphalt Pavement Technology Program
AASHTO	American Association of State Highway and Transportation Officials
ANN	Artificial neural network
APA	Asphalt Pavement Analyzer
ARZ	Above the restriction zone
AV	Air void(s)
BMD	Balanced Mix Design
BRZ	Below the restriction zone
CA	Coarse aggregate
CA CUW	Coarse aggregate chosen unit weight
CAA	Coarse aggregate angularity
CCC	Concordance correlation coefficient
CMHB	Coarse matrix high binder
CG	Coarse-graded
CMOD	Crack mouth opening displacement
CUW	Chosen unit weight
DCT	Disk-shaped compact tension
DEM	Discrete-element modeling
FA	Fine aggregate
FAA	Fine aggregate angularity
FA_c	Fine aggregate coarse ratio
FA_f	Fine aggregate fine ratio
FE	Flat and elongated
FI	Flexibility Index
FG	Fine-graded
G_{sb}	Bulk specific gravity
GEP	Gene expression programming
G_f	Fracture energy
G_{sb}	Bulk specific gravity
HMA	Hot mix asphalt
HWTT	Hamburg Wheel-Tracking Test
IDEAL-CT	Indirect Tensile Asphalt Cracking Test
IDOT	Illinois Department of Transportation
I-FIT	Illinois Flexibility Index Test
ITS	Indirect tensile strength
JMF	Job mix formula
LA	Los Angeles (abrasion)
LTOA	Long-term oven-aged
LTPP	Long-Term Pavement Performance

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
MEPDG	Mechanistic-Empirical Pavement Design Guide
ML	Machine learning
NAPA	National Asphalt Pavement Association
NCAT	National Center for Asphalt Technology
NMAS	Nominal maximum aggregate size
OGFC	Open-graded friction course
OLS	Ordinary least squares
PCS	Primary control sieve
PG	Performance grade
R^2	Coefficient of determination
RAP	Reclaimed asphalt pavement
RF	Random Forest
RMSE	Root mean squared error
RRI	Rutting Resistance Index
RUW	Rodded unit weight
SCB	Semi-circular bend
SCS	Secondary control sieve
SE	Sand equivalent
SHAP	Shapley Additive exPlanations
SHRP	Strategic Highway Research Program
SMA	Stone matrix asphalt
SPR	Simplified Practice Recommendation
SVR	Support vector regression
TCS	Tertiary control sieve
TRZ	Through the restriction zone
TSR	Tensile strength ratio
USACE	U.S. Army Corps of Engineers
UW	Unit weight
VCA	Voids in coarse aggregate
VECD	Viscoelastic continuum damage
VMA	Voids in mineral aggregate
WMA	Warm mix asphalt

Executive Summary

The P-401 specification establishes aggregate gradation limits for asphalt mixtures used in airport pavements. These gradation limits are intended to produce durable mixtures capable of supporting heavy aircraft loads. In practice, however, mixture designers often encounter difficulty meeting both the gradation limits and the required voids in mineral aggregate (VMA) during mix design. Adjustments made to satisfy gradation limits at one sieve may result in failure at other sieves or in insufficient VMA. Solutions such as reprocessing fine aggregates or sourcing materials from distant locations can address these problems, but they increase costs and limit the use of locally available aggregates. These challenges raise questions about whether the current gradation limits can be adjusted while still achieving satisfactory mixture performance.

This study evaluated the feasibility of adjusting the P-401 gradation limits while maintaining mixture volumetric properties and performance. The research addressed two primary objectives: (1) evaluate the laboratory performance of mixtures with gradations outside the current P-401 gradation limits while maintaining comparable volumetric properties and asphalt binder content (including effective binder content), and (2) assess how systematic gradation changes influence mixture volumetric characteristics and laboratory performance properties.

A laboratory program was conducted using 13 P-401 mix designs representing different aggregate sources and binder grades used across the United States. For each mix design, mixtures within the gradation limits were compared with those outside them, while maintaining similar volumetric properties and binder content. The mixtures were evaluated using laboratory tests for moisture susceptibility, rutting resistance, durability, permeability, and cracking resistance.

The results indicate that mixtures with gradations outside the current P-401 gradation limits can achieve laboratory performance comparable to that of mixtures designed within the gradation limits, provided volumetric properties and binder content are maintained. Moisture resistance, durability, permeability, rutting resistance, and low-temperature cracking resistance were generally unaffected by moderate gradation shifts, while intermediate-temperature cracking resistance showed limited sensitivity to gradation changes.

Based on these findings, the study recommends reducing the lower limits at the No. 16, No. 30, and No. 50 sieves by 2 percent for Gradation 2 mixtures with a 12.5-mm nominal maximum aggregate size. No changes are recommended to the current P-401 volumetric requirements, and the existing upper limits are retained at all sieves.

The 2-percent value reflects the extent to which gradations could be shifted below the current limits while keeping the mixtures fine-graded, holding binder content and

volumetric properties approximately constant, and working within the aggregate fractions available in the selected mix designs. It should therefore be understood as a conservative revision supported by the mixtures evaluated, rather than an optimal or maximum allowable adjustment. Because the recommendation reflects the range of gradations evaluated in this study, agencies adopting the revised limits should monitor mix designs and field performance during initial implementation.

The revision provides greater flexibility in aggregate blending and improved control of mixture volumetric properties while maintaining expected performance. In addition, prediction models and a web-based tool were developed to help mix designers evaluate the effects of gradation adjustments on mixture volumetrics and performance test results. Overall, the findings support the proposed changes to the P-401 gradation limits while maintaining satisfactory performance of airport asphalt mixtures.

Chapter 1. Introduction

1.1 Problem Statement

Airport runways, taxiways, and aprons are constructed using asphalt mixtures designed in accordance with Item P-401 of Federal Aviation Administration Advisory Circular 150/5370-10H, *Standard Specifications for Construction of Airports* [1]. Under the P-401 specification, asphalt mixtures are designed based on volumetric criteria, including a design air void (AV) content of 3.5 percent and a minimum voids in mineral aggregate (VMA) requirement that varies with the nominal maximum aggregate size (NMAS). Laboratory specimens can be compacted using either the Marshall hammer or Superpave gyratory method. Additionally, the mix design is evaluated for moisture resistance through the tensile strength ratio (TSR) test.

For airfield asphalt pavements serving aircraft with a gross weight exceeding 60,000 lb, P-401 further requires evaluation of rutting resistance using either the Asphalt Pavement Analyzer (APA) or the Hamburg Wheel-Tracking Test (HWTT). Aggregate requirements under P-401 specify the use of crushed coarse aggregate (stone, gravel, and/or slag), fine aggregate (screenings and/or natural sand), and mineral filler as needed. These materials must meet the requirements for angularity, shape, toughness, soundness, and cleanliness. Although natural (uncrushed) sand may be used to adjust gradation or improve workability, its use is limited to a maximum of 15 percent of the total aggregate to mitigate rutting potential under heavy aircraft loading.

The P-401 specification differs from the Superpave volumetric specification, American Association of State Highway and Transportation Officials (AASHTO) M 323 [2], in several important respects, particularly regarding aggregate gradation limits, as illustrated in Table 1. P-401 mix designs typically employ fine-graded gradations to reduce permeability and require minimum VMA values that are 1-percent higher than those specified under Superpave to accommodate increased asphalt binder content for improved durability. In addition, P-401 imposes full gradation limits, with both upper and lower limits specified at all sieve sizes. In contrast, the Superpave specification controls gradation at only four sieve sizes. Although full gradation limits help achieve durability in airfield asphalt mixtures, they have also introduced significant challenges during the mix design process. Practitioners report difficulty meeting both gradation and VMA requirements simultaneously due to the specification's restrictive limits. Adjustments made to correct a gradation violation at one sieve often result in violations at other sieves or cause VMA deficiencies. To comply with the specification, mix designers may resort to reprocessing fine aggregates or sourcing materials from distant locations, which can increase costs and delay construction.

Table 1. Gradation and VMA Requirements in P-401 and AASHTO M 323 Superpave Specifications [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						
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**
Min VMA (%)	14.0	13.0	15.0	14.0	16.0	15.0

The P-401 specification also imposes stricter limits on the No. 200 sieve than the Superpave specification, creating additional challenges during design and production. Contractors may attempt to address these limits by using washed screenings, natural sand, or waste dust; however, the resulting blends frequently remain near the limits for the fine aggregate fractions. These issues highlight concerns regarding whether the current gradation controls are overly restrictive and whether some flexibility could be introduced without compromising mixture performance.

Furthermore, full gradation limits are applied during laboratory mix design, whereas construction quality control and acceptance focus on a limited number of sieves. Acceptance of P-401 mixtures is based primarily on laboratory AV, field density, pavement grade, and surface roughness, rather than strict compliance with the full gradation limits. This distinction raises questions about the necessity of full gradation limits during the design phase.

Collectively, these challenges have prompted airport pavement designers and practitioners to question whether the P-401 gradation limits could be adjusted while maintaining required volumetric properties and performance. Guidance is needed to assess how gradation adjustments influence volumetric characteristics and laboratory performance test results during the mix design process.

1.2 Research Objectives

To address the challenges, this study aimed to achieve the following two objectives:

- Evaluate the effect of gradation outside the P-401 specification limits while maintaining similar volumetric properties and asphalt binder content (including effective binder content) for each P-401 mix design. This objective will guide adjustments to the P-401 gradation limits without compromising mixture performance.
- Evaluate the effect of gradation changes on volumetric properties and laboratory performance of P-401 mix designs. This objective supports the development of predictive models to quantify the effects of gradation adjustments on the volumetric properties and laboratory performance indicators of airfield asphalt mixtures.

1.3 Organization of the Document

This document is organized into seven chapters. Chapter 1 presents the problem statement and research objectives. Chapter 2 presents a review of the literature on aggregate gradation limits, the impact of aggregate properties and gradation on volumetrics and performance, the Bailey Method for adjusting gradation, and predictive models utilized to quantify these effects. Chapter 3 describes the research methodology and laboratory testing program. Chapter 4 focuses on the development of proposed revisions to the P-401 gradation limits. Chapter 5 evaluates the effects of gradation changes on volumetric properties and laboratory performance. Chapter 6 presents predictive models that correlate gradation changes to volumetric and performance responses. Chapter 7 summarizes the study conclusions and provides recommendations for future implementation.

Chapter 2. Literature Review

This chapter provides a literature review on the development of aggregate gradation in asphalt mixtures, with a specific focus on airfield applications. It examines how the properties and gradation of aggregates influence the volumetric and performance characteristics of these mixtures. The chapter summarizes gradation selection approaches, such as the Bailey Method, and emphasizes their importance in assessing aggregate packing and its effect on VMA for both hot mix asphalt (HMA) and warm mix asphalt (WMA). Additionally, it reviews predictive modeling techniques used in asphalt mixture design. This information lays a foundation for the experimental work and analyses presented in subsequent chapters.

2.1 Evolution of Aggregate Gradation Limits for Asphalt Mixtures

Aggregate gradation has long been recognized as a governing factor in the performance of asphalt mixtures, influencing aggregate packing, internal stability, void structure, and the balance between aggregate interlock and the contribution of the asphalt binder. Although current asphalt specifications treat gradation as a primary design variable, this role emerged gradually, progressing from empirical practice to scientific investigation and, ultimately, institutional standardization [3, 4]. This section discusses the evolution of aggregate gradation concepts as they relate to asphalt mixture design, with emphasis on the technical developments that culminated in P-401 requirements and current AASHTO gradation standards.

2.1.1 Sieve Sizes and Gradations

In 1921, the U.S. Department of Commerce initiated national standardization efforts to simplify industrial practices. Under the direction of Herbert Hoover, the Division of Simplified Practice was established within the National Bureau of Standards (now the National Institute of Standards and Technology [NIST]). As part of these efforts, the Department issued simplified practice recommendations across multiple industries. In June 1936, with input from the National Sand and Gravel Association, the National Crushed Stone Association, and the National Slag Association, *Simplified Practice Recommendation (SPR) R 163-36* was developed for the aggregate industry. This recommendation was later updated and reissued in 1948 as SPR R 163-48 [5].

SPR R 163 established a standardized system of sieves with square openings based on a logarithmic progression, in which each sieve size is approximately one-half the size of the preceding sieve. The standard series includes 3-inch, 1-1/2-inch, 3/4-inch, 3/8-inch, No. 4, No. 8, No. 16, No. 30, No. 50, No. 100, and No. 200 sieves. Additional intermediate sizes, including 4 inch, 2 inch, 1 inch, and 1/2 inch, were later introduced to improve resolution where gaps between standard sizes were considered too large.

The SPR framework was subsequently adopted into national standards. In 1937, ASTM published SPR R 163-36 as a tentative specification, ASTM D448-37T, *Standard Classification of Aggregate for Road and Bridge Construction*. The specification became permanent in 1947 and was updated in 1948 to align with SPR R 163-48. ASTM continues to maintain this standard, currently designated as ASTM D448-12 (reapproved 2017) [5].

Similarly, the American Association of State Highway Officials (AASHO) adopted SPR R 163-36 in 1942 as Specification M 43-42. This specification evolved into the current AASHTO M 43 standard, *Standard Specification for Sizes of Aggregate for Road and Bridge Construction* [6].

This standardized sieve system provides the foundation for defining aggregate gradation in asphalt mixtures. Gradation curves derived from these sieve sizes describe the particle size distribution and play a central role in controlling aggregate packing, void structure, and mixture performance.

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

The origin of airfield asphalt mixture gradation limits is based on research conducted by the U.S. Army Corps of Engineers (USACE) during and following World War II. Rapid increases in aircraft weight, tire pressure, and traffic repetition exposed the inadequacy of existing pavement design and mixture proportioning practices, which had been developed primarily for highway applications. Airfield pavements were required to withstand substantially higher contact stresses under slow-moving or stationary loads while maintaining durability under fuel exposure and limited maintenance opportunities. These demands necessitated a more rigorous, performance-oriented approach to asphalt mixture design, with particular emphasis on aggregate structure and gradation control.

2.1.2.1 USACE Research on Airfield Pavements

In response to increasing demands from military aviation, USACE initiated a coordinated research program to evaluate the performance of asphalt mixtures for airfield pavements. In 1943, the Tulsa District completed a study assessing the suitability of existing mix design methods, including the Hubbard–Field and Hveem Stabilometer methods, in comparison with the newly developed Marshall method. The Marshall method was selected as the primary design approach due to its practicality and repeatability in evaluating mixture stability under load [7].

In 1944, full-scale test sections were constructed at the Waterways Experiment Station in Vicksburg, Mississippi. The test sections included a strong aggregate base to isolate the structural response of the asphalt layers. Fifteen asphalt mixtures with varying gradations and asphalt contents were designed using Marshall procedures and placed for evaluation.

The test sections were subjected to wheel loads representative of military aircraft, including 15,000, 37,000, and 60,000 lb. For the 15,000-lb load, the asphalt thickness was

1.5 inches, with a design Marshall stability of 150 lb. For the 37,000-lb load, the thickness increased to 3 inches, with a stability of 350 lb. For the 60,000-lb load, the pavement consisted of a 5-inch asphalt structure placed in two lifts (3-inch binder course and 2-inch surface course), with a design Marshall stability of 550 lb.

For each surface mixture, three asphalt binder contents were evaluated: optimum, optimum minus 10 percent, and optimum minus 20 percent. The surface mixtures had an NMAS of 3/4 inch, with approximately 50 percent passing the No. 10 (2.0 mm) sieve. The gradation was described as well-graded and generally following Fuller's curve. The mineral filler content (passing the No. 200 sieve) was 3.5, 6.0, and 8.0 percent for mixtures designed at 150, 350, and 550-lb Marshall stability, respectively.

Two aggregate types were evaluated for the surface mixtures: crushed limestone with natural sand and uncrushed gravel with natural sand. The asphalt binder was a 120–150 penetration grade, which corresponds approximately to a performance grade (PG) 58-28 binder under current classification.

The binder course mixtures used the same aggregate sources and binder grade but were designed with a coarser gradation. Both binder and surface mixtures had an NMAS of 3/4 inch; however, the binder course contained approximately 40 percent passing the No. 10 sieve, compared to 50 percent for the surface course.

This experimental program enabled a systematic evaluation of the effects of aggregate gradation and asphalt content on deformation resistance, stability, and structural performance under aircraft loading [7]. The results demonstrated that mixture performance is highly sensitive to aggregate structure. These findings established that airfield pavements require distinct mixture design considerations and that aggregate gradation plays a critical role, beyond asphalt content, in controlling stability under high contact stresses.

2.1.2.2 Definition of Upper and Lower Gradation Limits

A key outcome of the USACE research was the explicit linkage between asphalt mixture Marshall stability and the coarse aggregate fraction (defined as the percent retained on the No. 10 sieve). For the mixtures evaluated, a maximum aggregate size of 3/4 inch (19.0 mm) coarse aggregate was defined as material retained on the No. 10 (2.0 mm) sieve. By systematically varying the proportion of coarse aggregate while holding other variables constant, researchers identified clear trends between aggregate structure and Marshall stability.

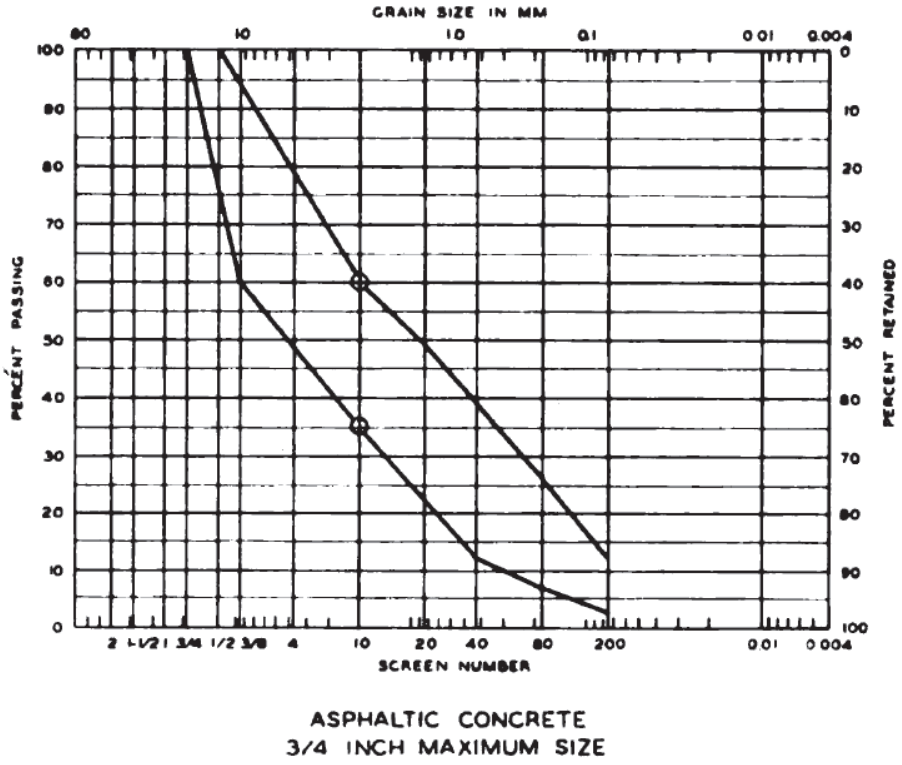
Test results showed that Marshall stability increased as the coarse aggregate fraction increased from approximately 40 percent to about 65 percent of the aggregate blend. Within this range, increased stone-on-stone contact enhanced internal friction and load-carrying capacity. However, when the coarse aggregate fraction exceeded approximately

65 percent, Marshall stability declined. Excessive coarse aggregate content disrupted the continuity of the aggregate skeleton, reduced effective interlock, and limited the ability of the finer fraction to stabilize the structure. Conversely, mixtures containing less than approximately 40 percent coarse aggregate relied excessively on the fine fraction and asphalt binder for stability, resulting in reduced resistance to deformation [7].

Based on these observations, USACE defined upper and lower gradation limits for dense-graded asphalt mixtures (asphaltic concrete) by bounding the allowable coarse aggregate fraction, as listed in Table 2 and shown in Figure 1. These envelopes were translated into sieve-based limits by specifying corresponding ranges for percent passing the No. 10 sieve [7].

Table 2. 1948 Gradation Limits for Airfield Asphalt Mixtures [8]

Sieve Size	Gradation Limits Percent Passing		
	Asphaltic Concrete	Stone Filled Sand Asphalt	Sand Asphalt
3/4 inch	100	100	–
1/2 inch	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



Source: Gayle McFadden and Walter Ricketts

Figure 1. Gradation Limits for 3/4-Inch Maximum Size Asphaltic Concrete Mixture [7]

Source: Reproduced from "Design and Control of Asphalt Paving Mixtures for Military Installations," by G. McFadden and W. Ricketts, 1948, in *Proceedings of the Association of Asphalt Paving Technologists*, *17*, p. 93–113. Work of the U.S. Government in the public domain.

The resulting gradation limits formed the technical basis for airfield asphalt mixture specifications adopted after World War II. These limits were incorporated into military guidance and later adopted in civil aviation specifications, where they evolved into the gradation limits now codified in the P-401 specifications. Importantly, the defining feature of these specifications is not adherence to a specific gradation shape but rather control of aggregate structure that reflects the unique performance demands imposed by aircraft loading.

The USACE research established a durable framework for airfield asphalt mixture design by directly linking aggregate gradation to performance. Although compaction methods, materials, and testing technologies have evolved, the fundamental concept of bounding aggregate structure through upper and lower gradation limits has remained central to airfield specifications. As a result, modern P-401 gradation limits represent the direct continuation of a design philosophy rooted in controlled experimentation and field validation.

2.1.3 The Superpave Framework and Current AASHTO Gradation Control

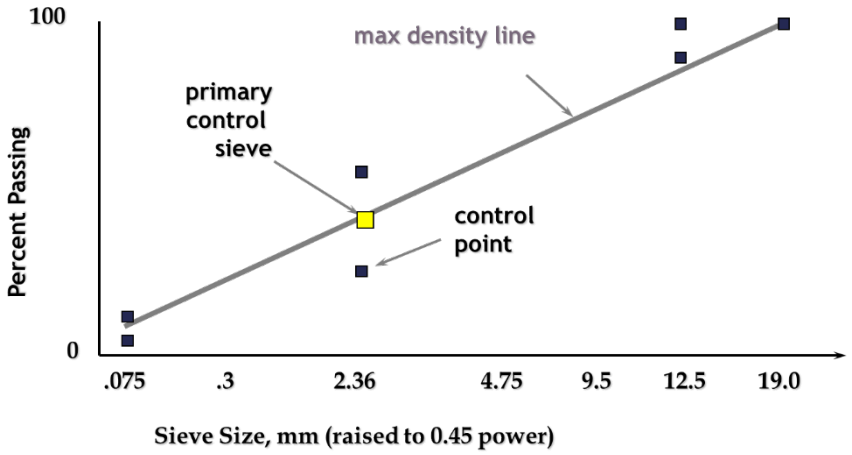
Current AASHTO gradation control for asphalt mixtures was formalized through the Strategic Highway Research Program (SHRP), which translated earlier analytical concepts into an operational gradation specification for highway pavements designed with the Superpave mixture design system. Building on Nijboer's theoretical work, Superpave adopted the 0.45-power chart as a standardized tool for visualizing aggregate structure and comparing gradation shapes across mixtures with different maximum aggregate sizes. By plotting cumulative percent passing against sieve size raised to the 0.45 power, gradations could be evaluated relative to a reference maximum density line, providing insight into aggregate packing and the balance between coarse and fine fractions [9, 10].

To ensure consistent application of maximum density lines drawn on a 0.45-power chart, Superpave introduced standardized definitions for maximum aggregate size and NMAS. NMAS was defined as one sieve size larger than the first sieve to retain more than 10 percent of the aggregate, while maximum aggregate size was defined as one sieve size larger than the NMAS. These definitions resolved longstanding inconsistencies in gradation description and ensured that the maximum density line was drawn consistently across mixtures and agencies [11]. This clarification represented a key advance in making analytical gradation evaluation reproducible and providing the basis for a specification.

The Superpave framework further classified gradations using a limited set of control points rather than prescribing detailed gradation envelopes. Central to this approach was the primary control sieve (PCS), defined as the sieve size that separates coarse and fine aggregate. The PCS changes for different NMAS mixtures. Gradations passing below the PCS are classified as coarse-graded (CG), while those passing above are classified as fine-graded (FG). This binary classification provided agencies with the flexibility to select gradation types appropriate to traffic levels, climate, and material availability, while maintaining a consistent analytical framework for evaluating aggregate structure [2].

Early versions of the Superpave specification also included a restricted zone defined around the maximum density line between the 0.3-mm and 2.36-mm sieves. The restricted zone was intended to discourage gradations containing excessive amounts of natural sand, which are known to produce mixtures with rutting susceptibility. Subsequent research demonstrated that fine aggregate shape and texture measured with fine aggregate angularity (FAA) prevented inclusion of excessive natural sand, making the restricted zone unnecessary. As a result, the restricted zone was removed from the Superpave specification, reflecting a shift away from prescriptive gradation constraints toward performance-related material properties [12].

Collectively, these elements define the current AASHTO approach to gradation control: standardized aggregate size definitions, analytical evaluation using the 0.45-power chart, and limited control points as shown in Figure 2.



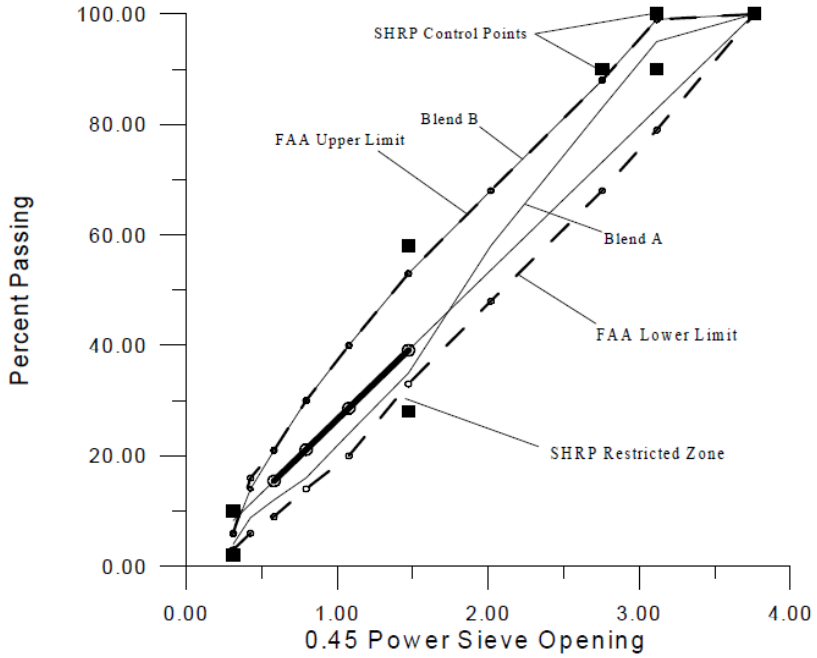
Source: Adapted by AAPTP from Indiana Department of Transportation Certified Asphalt Technician (ICAT) Program Manual 2026.

Figure 2. Illustration of Current Superpave Gradation Limits

2.1.4 Evaluation of Superpave Gradation for Airfields

Following implementation of the Superpave system for highway pavements, the Federal Aviation Administration undertook a formal evaluation to determine whether Superpave mix design concepts—including gradation control and compaction procedures—were applicable to airfield pavements. This assessment was motivated by the potential benefits of harmonizing specifications across transportation sectors while maintaining the performance requirements unique to airfield operations. For example, Figure 3 shows a design gradation that meets both the P-401 and Superpave gradation limits. However, the Federal Aviation Administration evaluation concluded that although certain Superpave elements were adaptable, applying highway gradation and compaction criteria directly to airfield mixtures was inappropriate without modification [13].

A primary concern identified in the assessment was whether Superpave gradation control, when combined with gyratory compaction levels developed for highways, could provide adequate VMA for airfield mixtures. Airfield pavements require higher effective asphalt contents than highway pavements to ensure durability under fuel exposure, environmental aging, and limited maintenance access. The assessment found that Superpave gradations compacted at standard highway gyratory levels often produced low VMA, thereby constraining asphalt content and increasing susceptibility to permeability- and durability-related distress [13].



FAA = Federal Aviation Administration.

Source: Federal Aviation Administration

Figure 3. An Aggregate Blend Plotted Within P-401 Gradation Limits and Superpave Control Points [13]

Source: Reproduced from *Evaluation of Strategic Highway Research Program (SHRP) Products for Application to Airport Pavements* (Report No. DOT/FAA/AR-99/13) by J. K. Newman and R. B. Freeman, Federal Aviation Administration, 2000. Public Domain.

Superpave gradation control points permit coarser gradations than those historically allowed in Federal Aviation Administration P-401 specifications. While such gradations may be acceptable for highway pavements, the increased permeability associated with coarser gradations could compromise durability in airfield environments, particularly under fuel exposure, and pose a risk of foreign object damage to engines. As a result, the research recommended tighter control of aggregate structure for airfield mixtures, including modest adjustments to lower control points when Superpave concepts were evaluated for potential adaptation, as shown in Table 3 [8].

Table 3. Adaptation of Superpave Gradation Limits for Airfield Asphalt Mixtures [8]

Sieve Size U.S. (mm)	Percentage by Weight Passing Sieves			
	1-1/2" max	1" max	3/4" max	1/2" max
1-1/2 inches (37.5)	100	–	–	–
1 inch (25.0)	86–98	100	–	–
3/4 inch (19.0)	68–93	76–97	100	–
1/2 inch (12.5)	57–81	67–87	77–98	100
3/8 inch (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

Collectively, research by the Federal Aviation Administration and the Airport Asphalt Pavement Technology Program found that Superpave gradation and compaction criteria cannot be applied directly to airfield pavements [8, 13]. The unique loading conditions, durability demands, and risk tolerance associated with airfield pavements require strict controls on permeability and long-term performance. Cooley et al. concluded that no changes to the airfield gradation requirements were necessary once the lower control points for the PCS in the highway Superpave specifications were aligned with those in the P-401 specifications [8]. Neither evaluation examined the lower limits at the No. 16, No. 30, and No. 50 sieves, which are the subject of the revisions proposed in this report.

2.1.5 Summary

The evolution of aggregate gradation limits reflects a progression from empirical practice to performance-driven specification. For airfield pavements, this evolution culminated in explicit gradation limits derived from controlled experimentation under aircraft loading, forming the basis of Federal Aviation Administration P-401. These requirements embody an enduring balance among density, durability, and constructability achieved through deliberate control of aggregate structure.

Current highway specifications emphasize analytical flexibility and volumetric balance, whereas airfield specifications retain tighter control of gradation to address unique loading and durability demands. The historical record demonstrates that P-401 gradation limits represent a distinct, performance-based lineage rather than a legacy. Accordingly, current research should focus on validating and refining these gradation limits using modern analytical tools and performance tests while preserving the proven structural principles that have governed airfield asphalt pavement performance for decades.

2.2 Impact of Aggregate Properties on Volumetric Properties

Studies have examined how the properties of aggregates—especially the Superpave consensus properties—affect the volumetric characteristics of asphalt mixtures. The Superpave specification (AASHTO M 323) establishes minimum VMA requirements based on the NMAS. However, at each NMAS level, additional factors, such as the percentage of crushed coarse aggregate and the FAA, can also affect VMA [14, 15].

Among consensus properties, aggregate angularity has been shown to correlate strongly with VMA. Angular aggregates tend to increase interparticle friction and resist dense packing, leading to higher VMA values than rounded aggregates [16]. Several experimental studies have confirmed this relationship under controlled laboratory conditions. The key findings from these studies are summarized in Table 4.

Table 4. Summary of Findings on Impact of Aggregate Properties on VMA

Aggregate Property	Ref.(s)	Study Approach	Volumetrics Evaluated	Key Findings
Crushed Aggregate Content, FAA	[14, 15]	Review and comparative analyses	VMA adequacy	Aggregate angularity and crushed content were more reliable indicators of VMA than NMAS
FAA	[17]	Laboratory mixtures with four aggregate types	VMA	Mixtures with crushed aggregates exhibited higher VMA than those with rounded natural sand
FAA	[17]	Fine aggregate replacement study	VMA	Using 30% rounded natural sand caused a sharp reduction in VMA (R^2 = 0.70 between FAA and VMA)
FAA	[18]	Constant gradation with progressive sand replacement	VMA	Lower FAA mixtures consistently produced lower VMA
Aggregate Shape	[19]	Comparative particle shape evaluation	VMA	Cubical aggregates resisted compaction and produced higher VMA than flat or elongated particles

R^2 = coefficient of determination.

The literature consistently indicates that aggregate angularity and shape significantly impact VMA, often exceeding the effects based on NMAS [14-18]. Higher FAA increases interparticle friction and reduces aggregate packing efficiency, leading to increased VMA, whereas rounded fine aggregates promote denser packing and lower VMA [17, 18].

Aggregate shape also affects volumetric properties. Cubical particles resist reorientation during compaction and tend to form a more open aggregate structure, resulting in higher VMA. In contrast, flat and elongated particles align more readily during compaction, allowing tighter packing and lower VMA values [19].

In summary, aggregate properties—particularly angularity and shape—play a critical role in controlling volumetric characteristics of asphalt mixtures. These findings suggest that reliance on NMAS alone is insufficient to ensure appropriate VMA levels and underscore

the importance of explicitly considering aggregate properties during mixture design, especially for specifications with stringent volumetric requirements.

2.3 Impact of Aggregate Consensus and Source Properties on Mixture Performance

2.3.1 Impact of Consensus Properties of Aggregates on Mixture Performance

2.3.1.1 Impact of Coarse Aggregate Angularity and Fine Aggregate Angularity on Mixture Performance

Aggregate angularity, including coarse aggregate angularity (CAA) and FAA, is widely recognized as a key factor influencing the rutting resistance of asphalt mixtures [20, 21]. Increased angularity enhances internal friction and aggregate interlock, thereby improving resistance to shear deformation under repeated loading. For this reason, minimum CAA and FAA requirements are commonly specified in asphalt mixture design [22, 23].

Several studies have examined the influence of CAA and FAA on rutting and cracking performance using laboratory testing, mechanistic analysis, and numerical modeling. The key findings from these studies are summarized in Table 5, followed by a synthesis of overall trends.

Overall, the literature consistently indicates that CAA plays a dominant role in controlling rutting resistance by strengthening the aggregate skeleton and improving shear deformation resistance [20-26]. This effect is most noticeable in mixtures with softer binders, while very stiff binders can diminish the impact of aggregate angularity [17]. Similar benefits of increased CAA have also been reported for WMA mixtures [17].

The influence of CAA on cracking performance is more complex. Increased angularity can promote stress concentrations that accelerate crack initiation, yet it can also increase binder demand and enhance energy dissipation [25]. As a result, reported trends vary with mixture type. For example, mixtures with high binder content, such as stone matrix asphalt (SMA) mixtures, benefited from higher angularity, whereas Superpave mixtures showed mixed or opposite trends depending on aggregate strength and angularity level [27, 28].

Table 5. Summary of Findings on the Impact of CAA and FAA on Mixture Performance

Aggregate Property	Ref.(s)	Mixture Type / Test Method	Performance Evaluated	Key Findings
CAA	[20-26]	HMA; rutting, creep, APA, shear tests	Rutting resistance	Higher CAA improved aggregate interlock and significantly reduced permanent deformation
CAA	[25]	HMA; laboratory tests and finite element modeling	Rutting resistance	Increased CAA enhanced shear resistance and reduced lateral flow
CAA	[17]	HMA with PG 64-22, 76-22, 82-22	Rutting resistance	CAA effects were significant for softer binders but diminished for very stiff binders
CAA	[17]	WMA; laboratory rutting tests	Rutting resistance	Strong linear relationship between increased CAA and improved rutting resistance
CAA	[25, 27, 28]	SCB, IDEAL-CT; Superpave and CMHB mixes	Cracking resistance	Effect depended on mix type; higher angularity improved peak load in CMHB but not consistently in Superpave
FAA	[29-33]	Wheel tracking, Marshall, APA, dynamic modulus	Rutting resistance	Higher FAA generally improved mixture stability and reduced rut depth
FAA	[34-36]	Shear-related laboratory tests	Rutting resistance	FAA alone was not a sufficient indicator of shear strength or rutting resistance
FAA	[18, 37]	Mixtures with varying gradation and binder content	Rutting resistance	Rutting behavior depended on gradation, binder content, and aggregate–binder compatibility
FAA	[29]	Fatigue testing with different sands	Cracking resistance	Higher FAA and rougher particles improved fatigue resistance

SCB = semi-circular bend; IDEAL-CT = indirect tensile asphalt cracking test; CMHB = coarse matrix high binder

FAA has also been shown to influence mixture performance, particularly rutting resistance. Several studies reported improved rutting performance for mixtures containing crushed fine aggregates with higher FAA compared with mixtures containing rounded natural sand [29-33]. However, other studies demonstrated that FAA alone is not a sufficient predictor of shear strength or rutting resistance and must be considered together with gradation, binder content, and aggregate–binder compatibility [34-37]. Research has also shown that mixture gradation can outweigh FAA effects, with FG mixtures exhibiting better rutting resistance regardless of FAA level [18].

Regarding cracking resistance, limited studies suggest that higher FAA, combined with rougher, stronger fine aggregates, can improve fatigue performance [29]. Nonetheless, the overall effect of FAA on cracking remains dependent on interactions among multiple mixture variables.

In summary, CAA and FAA are critical contributors to asphalt mixture performance, particularly rutting resistance. However, their effects cannot be evaluated in isolation. Binder properties, gradation structure, and mixture type strongly influence the extent to

which aggregate angularity improves performance, underscoring the need for an integrated mixture design approach.

2.3.1.2 Impact of Flat and Elongated Particles on Mixture Performance

Flat and elongated (FE) particles are coarse aggregate particles characterized by a high aspect ratio, typically defined as a maximum-to-minimum dimension ratio exceeding 5:1, although alternative limits are specified by some agencies. FE particles are generally undesirable because they reduce aggregate packing efficiency, weaken load transfer, and increase susceptibility to particle breakage during compaction and traffic loading.

Under the Superpave system developed through SHRP, FE particles with a 5:1 ratio are limited to a maximum of 10 percent by weight for mixtures subjected to traffic levels exceeding 3 million equivalent single axle loads. This requirement reflects concerns regarding the influence of particle shape on mixture stability and durability.

Several studies have investigated the relationship between FE content and asphalt mixture performance. The key findings from these studies are summarized in Table 6.

Table 6. Summary of Findings on the Impact of FE Particles on Mixture Performance

FE Aspect Ratio / Content	Ref.(s)	Mixture Type / Test Method	Performance Evaluated	Key Findings
FE > 5:1 (< 10%)	[26]	HMA; permanent deformation tests	Rutting resistance	When FE content was well below 10%, FE particles had no significant effect on rutting
FE > 2:1	[38]	HMA; laboratory performance evaluation	Rutting and fatigue resistance	Stricter FE limits improved resistance to permanent deformation and fatigue cracking
Increased FE content	[39, 40]	HMA; construction and traffic simulations	Rutting resistance	Higher FE content led to increased aggregate breakdown and greater rutting susceptibility
High FE content	[41]	Asphalt pavements; load transfer analysis	Structural performance	FE particles reduced compaction efficiency and load transfer capability
FE vs. cubical particles	[42]	Asphalt mixtures; rutting tests	Rutting resistance	Mixtures with FE particles exhibited lower rutting resistance than mixtures with cubical aggregates

The literature indicates that moderate FE contents within current gradation limits may not significantly degrade rutting performance, provided that FE particles remain well below threshold levels [26]. However, higher FE contents increase the likelihood of aggregate breakage, disrupt aggregate packing, and reduce load transfer efficiency, leading to inferior rutting performance [39-42].

Research also suggests that stricter FE limits than those currently specified may offer performance benefits. Kandhal and Parker [38] reported improved resistance to both rutting and fatigue cracking when a more conservative FE ratio of 2:1 was adopted,

compared with the commonly used 5:1 limit in the United States. These findings indicate that FE particle limits influence not only rutting resistance but also longer-term cracking performance.

In summary, FE particles affect asphalt mixture performance primarily by altering aggregate structure and durability. While limited FE content may be acceptable, excessive FE particles reduce mixture stability and should be controlled to ensure adequate rutting and fatigue resistance, particularly for heavily loaded pavements.

2.3.2 Impact of Sand Equivalent Value on Mixture Performance

Clay content in aggregates, commonly quantified using the sand equivalent (SE) value, represents the proportion of fine clay-like materials present in the fraction of aggregate passing the No. 4 sieve. Elevated clay content is undesirable because clay minerals tend to absorb moisture and swell, thereby weakening the adhesive bond between the asphalt binder and aggregate particles [43-45]. This degradation of the aggregate–binder interface reduces mixture durability and increases its susceptibility to distress.

Previous studies have shown that clay contamination can adversely affect both rutting resistance and moisture damage resistance of asphalt mixtures. The primary findings from the literature addressing the influence of SE value and clay contamination on mixture performance are summarized in Table 7.

Table 7. Summary of Findings on the Impact of SE Value and Clay Content on Mixture Performance

Aggregate Property	Ref.(s)	Study Approach	Performance Evaluated	Key Findings
Clay content (SE value)	[43-45]	Mechanistic and laboratory studies	Adhesion and durability	Clay swelling weakened aggregate–binder bonding and reduced mixture durability
Clay contamination (MBV-based)	[46]	HWTT on mixtures with active and inactive clays	Rutting and moisture resistance	Clay contamination with MBV > 6 mg/g significantly reduced rutting resistance and moisture damage resistance
Active vs. inactive clays	[46]	Laboratory mixtures with controlled clay types	Rutting and moisture resistance	Active clays caused more severe performance deterioration than inactive clays

MBV = methylene blue value

Overall, the literature indicates that higher clay content adversely affects asphalt mixture performance by weakening the aggregate–binder bond and increasing moisture susceptibility [43-46]. Clay swelling upon moisture exposure disrupts the internal structure of the mixture, increasing the potential for rutting, stripping, and premature failure.

Experimental evidence further suggests that the type and activity of clay minerals are critical factors. Active clays, such as bentonite and natural clays identified through methylene blue value (MBV) testing, cause substantially greater performance degradation than inactive clays, particularly when MBV exceeds threshold levels [46]. These findings

highlight the importance of controlling clay contamination in aggregates and support the use of SE and MBV tests as screening tools for aggregate quality.

In summary, SE value serves as an important indicator of aggregate cleanliness and mixture durability. Adequate control of clay content is essential to ensure satisfactory rutting resistance and moisture damage performance, especially for asphalt mixtures subjected to heavy loading or moisture exposure.

2.3.3 Impact of Source Properties of Aggregate on Mixture Performance

2.3.3.1 Impact of Toughness on Mixture Performance

Aggregate toughness reflects the ability of coarse aggregates to resist abrasion and mechanical breakdown during handling, construction, and in-service loading. It is commonly evaluated using the Los Angeles (LA) abrasion test, which measures the percentage of mass loss from a standard aggregate sample—typically retained on the No. 8 sieve—subjected to tumbling, impact, and grinding with steel balls. Reported loss values generally range from 30 to 45 percent, with higher LA values indicating more friable aggregates and lower values indicating greater durability [22].

Aggregate toughness influences asphalt mixture performance primarily by affecting particle integrity, surface texture retention, and resistance to polishing and wear. Aggregates with lower LA abrasion loss are less prone to degradation during production and compaction and tend to provide improved resistance to surface wear and skid loss under traffic [22]. Dahir [47] demonstrated that aggregate hardness, toughness, and mineralogy strongly influence wear resistance and skid resistance of bituminous surfaces. These effects are particularly important for mixtures such as open-graded friction course (OGFC) and SMA, which lack the cushioning effect of fine aggregates and therefore subject coarse aggregate particles to high contact stresses [48].

Several studies have examined the relationship between aggregate toughness and asphalt mixture strength and performance. The key findings from these studies are summarized in Table 8.

Overall, the literature suggests that aggregate toughness contributes to mixture durability and strength, particularly by limiting aggregate breakdown during production and service [22, 47-49]. Lower LA abrasion values are generally associated with higher tensile strength and improved resistance to surface wear, especially in surface mixtures subjected to high contact stresses.

Table 8. Summary of Findings on the Impact of Aggregate Toughness on Mixture Performance

Aggregate Property	Ref.(s)	Study Approach	Performance Evaluated	Key Findings
LA abrasion value	[22]	Standard abrasion testing	Durability and wear resistance	Lower LA values indicated more durable aggregates with better resistance to mechanical degradation
Aggregate toughness and mineralogy	[47]	Laboratory evaluation of surface wear	Skid and wear resistance	Toughness and mineralogy strongly influenced wear resistance, especially for surface mixtures
Aggregate toughness in OGFC and SMA	[48]	Mechanistic assessment	Durability under contact stress	Coarse aggregates in gap-graded mixtures experienced higher contact stresses due to lack of fines
LA abrasion value (28–55%)	[49]	Marshall ITS (dry and wet)	Tensile strength	Mixtures with lower LA values exhibited significantly higher ITS
Aggregate texture vs. toughness	[50]	Rutting performance analysis	Rutting resistance	Aggregate texture showed stronger correlation with rutting than toughness or hardness
LA abrasion value	[51]	Performance correlation study	Overall mixture performance	Weak linear correlation between LA value and mixture performance

ITS = indirect tensile strength.

However, multiple studies indicate that aggregate toughness alone is not a reliable predictor of rutting resistance or overall mixture performance. Research has shown that surface texture and particle shape often correlate more strongly with performance metrics than LA abrasion loss [50, 51]. These findings highlight that toughness should be considered one component of aggregate quality, evaluated in combination with other properties such as angularity, texture, and shape.

In summary, aggregate toughness plays an important role in maintaining particle integrity and surface durability, but its influence on mixture performance is secondary to other aggregate characteristics. An integrated evaluation of aggregate properties is therefore necessary to ensure satisfactory asphalt mixture performance.

2.3.3.2 Impact of Soundness on Mixture Performance

Beyond toughness and abrasion resistance, aggregate performance is also governed by soundness and hardness, which describe the ability of aggregate particles to resist degradation caused by environmental actions such as wetting–drying and freeze–thaw cycles. When the asphalt binder coating remains intact, these environmental stresses typically have a limited effect on mixture performance. However, aggregates containing weak or soft particles are more susceptible to breakdown during compaction and service, which facilitates moisture intrusion and accelerates distress mechanisms such as raveling, stripping, and, in severe cases, rutting.

Previous studies have shown that aggregate degradation associated with poor soundness compromises the integrity of the aggregate skeleton and weakens the binder–aggregate bond, thereby reducing the durability of asphalt mixtures. As fractured or porous aggregate

particles provide pathways for moisture ingress, damage accumulates more rapidly under traffic loading. Consequently, selecting sound, durable aggregates is essential for maintaining long-term pavement performance, particularly in environments subject to moisture exposure and temperature cycling [48].

2.3.4 Summary

Table 9 summarizes the reported effects of aggregate consensus and source properties on asphalt mixture performance. The table highlights the primary performance aspects influenced by each property and the general trends identified in prior studies.

Table 9. Effect of Aggregate Consensus and Source Properties on Mixture Performance

Aggregate Property	Study Approach	Performance Evaluated	Key Findings
CAA	Laboratory performance testing	Rutting resistance	Higher CAA improved aggregate interlock and mechanical stability, leading to higher rutting resistance [20, 21, 24-26]
CAA	Cracking-related laboratory tests	Cracking resistance	Effect on cracking was mixed and dependent on mixture type and binder characteristics [25, 27, 28]
FAA	Rutting and stability tests (APA, wheel tracking, Marshall)	Rutting resistance	Higher FAA increased internal friction and improved rutting resistance [23, 29-33]
FAA	Fatigue testing with different fine aggregates	Fatigue cracking resistance	Higher FAA combined with strong, rough particles improved fatigue resistance [29]
FE	Laboratory rutting and structural evaluation	Rutting resistance	Higher FE content reduced load transfer efficiency and rutting resistance [39, 40, 42]
SE value (clay content)	Moisture damage and rutting tests (HWTT, MBV)	Rutting resistance	Higher clay content (lower SE) increased rutting susceptibility and stripping potential [43-46]
SE value (clay content)	Adhesion and durability assessment	Moisture damage	Higher clay content weakened binder–aggregate bonding and reduced moisture resistance [43-45]
Toughness (LA abrasion)	LA abrasion testing and mixture strength evaluation	Rutting resistance	Lower LA loss values were associated with higher mixture strength and load resistance [22]
Toughness (LA abrasion)	Wear and skid resistance studies	Durability	Lower LA loss indicated more durable aggregates with better skid and wear resistance [22, 47]
Soundness	Environmental durability assessment	Durability and rutting resistance	Sound and durable aggregates improved long-term durability and resistance to rutting under moisture and temperature cycling [48]

2.4 Impact of Aggregate Gradation on Volumetric Properties of Asphalt Mixtures

The gradation of aggregates is the primary factor influencing the volumetric properties of asphalt mixtures, especially VMA. Early research conducted by McLeod [11, 52] demonstrated that gradation affects how aggregates pack together and the structure of the voids within the mixture. Later analytical and experimental studies confirmed that the impact of gradation on VMA results from particle arrangement, particle size distribution, and interactions between adjacent sieve fractions [53].

Different types of mixtures yield varying VMA ranges based on their gradation design strategy. OGFC mixtures generally exhibit the highest VMA due to their low fine aggregate content and a well-connected, gap-graded void structure [54]. SMA mixtures, which utilize a gap-graded gradation with fines filling the gaps between particles, achieve intermediate levels of VMA; these levels are generally higher than those of dense-graded mixtures but lower than OGFC [55]. Dense-graded mixtures, designed with a continuous aggregate size distribution, achieve tighter packing and therefore lower VMA [2, 56]. Experimental comparisons using the same aggregate sources have further confirmed that SMA mixtures have higher VMA than dense-graded mixtures, even when both have similar AV contents. This difference is primarily attributed to the gap-graded structure of SMA mixtures and their greater binder and filler contents [57].

The distance between the gradation curve and the maximum density line (MDL) has been shown to correlate strongly with VMA. Gradations located near the MDL on the 0.45-power chart produce the tightest packing and the lowest VMA, while deviations from the MDL—either above (FG) or below (CG)—lead to an increase in VMA [58, 59]. Laboratory investigations confirmed that increasing this distance results in higher VMA for both CG and FG mixtures [60]. Furthermore, VMA is particularly sensitive to gradation spacing between the No. 8 and No. 200 sieves. However, establishing a single statistical relationship between gradation and VMA is challenging due to the interactions among various gradation variables [16, 61].

The percentage of aggregate passing specific sieves also affects VMA. Kandhal et al. [62] showed that the percent passing the No. 4 sieve and smaller sieves are among the best predictors of VMA. For OGFC mixtures, increasing the percent passing the No. 4 sieve reduced VMA until stone-on-stone contact was lost and the aggregate skeleton collapsed [54]. Additionally, the amount of material passing the No. 200 sieve was shown to significantly influence VMA, with lower dust contents increasing VMA and higher contents reducing it, particularly for gradations on the fine side of the MDL [16, 55, 63]. Gradation near the PCS was found to be especially influential, with opposite trends observed for CG, FG, and SMA mixtures as described in the Bailey Method and related studies [64, 65].

NMAS further influences VMA. Laboratory studies showed that mixtures with smaller NMAS generally exhibit higher VMA, while those with larger NMAS tend to have lower VMA [66, 67]. This trend is reflected in the Superpave minimum VMA requirements, which increase as NMAS decreases [68]. However, additional research demonstrated that NMAS alone is insufficient to predict critical VMA levels. Factors such as gradation shape, fineness modulus, and aggregate texture were shown to be stronger indicators of VMA than NMAS alone, suggesting that current minimum VMA criteria may exclude mixtures capable of satisfactory field performance [15].

Given the well-established relationship between gradation and VMA, several modeling and simulation approaches have been developed to predict VMA directly from gradation characteristics. Discrete-element modeling studies quantified aggregate packing behavior and produced gradation-based weighting factors that reasonably predict VMA [69]. Other analytical and empirical models introduced concepts such as central particle size and skeleton porosity, demonstrating strong correlations with measured VMA and enabling rapid screening of gradation alternatives without extensive laboratory testing [70, 71].

The combined effects of gradation type, distance from the MDL, sieve-level distribution, PCS-region gradation, and NMAS on VMA are summarized in Table 10. This table consolidates findings across the literature and provides a concise reference for understanding how aggregate gradation impacts mixture volumetrics. These relationships provide the foundation for the next section, which discusses how gradation affects the performance of asphalt mixtures.

Table 10. Effect of Gradation on VMA

Gradation Factor	Impact on VMA
Gradation Type	VMA is highest for OGFC, followed by SMA, reflecting progressively tighter aggregate packing [2, 54-57]
Distance from MDL	Increased distance from MDL increases VMA due to reduced packing efficiency [58-61]
Percent Passing Specific Sieves	Percent passing the No. 4, No. 8, No. 16, No. 30, No. 50, No. 100, and No. 200 sieves are key predictors of VMA [62]
	Lower No. 200 content increases VMA, particularly for fine-side gradations [16, 55, 63]
	For SMA and OGFC, lower No. 4 content increases VMA by maintaining stone-on-stone contact [54, 55]
	For CG and SMA mixtures, lower PCS passing increases VMA [64, 65]
	For FG mixtures, higher PCS passing increases VMA [64, 65]
NMAS	Smaller NMAS increases VMA [66, 67]

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 [72, 73].

2.5.1 Rutting Resistance

The literature shows that rutting resistance is primarily governed by aggregate gradation structure, which impacts aggregate interlock, stone-on-stone contact, and load transfer [74-78]. The key findings from these studies are summarized in Table 11.

Table 11. Effect of Aggregate Gradation on Rutting Resistance of Asphalt Mixtures

Gradation Factor	Ref.(s)	Study Approach	Performance Measure	Key Findings
Overall gradation structure and mixture type	[74-78]	Laboratory and field performance studies	Rut depth, wheel tracking	Rutting resistance is governed by gradation structure; SMA generally exhibits superior rutting resistance due to a stable coarse aggregate skeleton
Gradation type (coarse vs. fine graded)	[12, 79, 80]	Laboratory mixture evaluations	Rutting resistance indicators	Coarser gradations, particularly near the lower envelope limit, improve rutting resistance by increasing aggregate interlock and internal friction
Restricted zone and gradation classification	[12, 80-88]	Laboratory testing and NCAT Test Track	Rut depth, field performance	Restricted zone compliance is not a reliable predictor of rutting resistance; satisfactory performance was observed for mixtures not meeting the restricted zone requirements
PCS-based gradation and coarse aggregate structure	[89-91]	Wheel-tracking and gradation classification studies	Rut depth, stripping resistance	Reducing PCS passing enhances stone-on-stone contact and improves resistance to rutting and stripping
Sieve-level distribution and Bailey parameters	[92-99] [100, 101]	Sieve fraction sensitivity and Bailey Method studies	Rut depth	Rutting resistance is sensitive to No. 4 and intermediate sieve fractions; Bailey CA ratio strongly correlates with rutting resistance by reflecting coarse aggregate skeleton quality

CA = coarse aggregate; NCAT = National Center for Asphalt Technology.

Mixtures with inadequate gradation may exhibit poor resistance to permanent deformation even when high-quality aggregates and binders are used [74]. Among common mixture types, SMA generally provides superior rutting resistance due to the development of a stable coarse aggregate skeleton through stone-on stone contact [75-78].

Rutting resistance is strongly influenced by the position of the gradation curve within the gradation envelope. Coarser gradations, particularly those near the lower bound of the

envelope, typically exhibit improved rutting resistance by increasing aggregate contact points, increasing internal friction, and reducing shear deformation [12, 79, 80].

Early Superpave specifications introduced a restricted zone to control rutting; however, subsequent studies yielded inconsistent results. While some research reported improved rutting resistance for mixtures below the restricted zone [12, 80], other studies and full-scale NCAT Test Track evaluations showed that mixtures violating restricted zone criteria can still achieve satisfactory rutting performance [81-88]. The findings led to the removal of the restricted zone from AASHTO specifications and to a transition to gradation classification based on PCS and NMAS [89].

Rutting resistance is also influenced by the percentage of material passing through the sieve sizes near the PCS. Reducing the percentage of aggregate passing the PCS enhances stone-on-stone contact and improves resistance to rutting and stripping [90, 91]. Similarly, maintaining an optimal range of material passing the No. 4 sieve improves aggregate interlock and limits disruption of the aggregate skeleton [92-95]. Intermediate sieve fractions further contribute to rutting resistance by providing stiffness and structural support [96, 98]. Bailey Method parameters, particularly the coarse aggregate (CA) ratio, can be used to quantify the effectiveness of the coarse aggregate skeleton and are strongly correlated with rut depth [99-102].

Overall, the findings indicate that rutting resistance is more reliably explained by gradation structure than by gradation limits alone. Rutting performance is maximized when gradation promotes a stable CA skeleton, effective stone-on-stone contact, and a balanced distribution across critical sieve sizes, underscoring the need to use performance tests to evaluate rutting resistance rather than relying on gradation limits.

2.5.2 Cracking Resistance

The influence of aggregate gradation on cracking resistance is governed by its effects on effective binder, mixture stiffness, and aggregate structure, with impacts that vary by cracking mechanism [103-106] , as summarized in Table 12.

For low-temperature cracking, dust content and effective asphalt film thickness are the dominant factors. Without changing binder content, increasing dust content reduces the effective binder available for thermal stress relaxation, thereby reducing cracking resistance [104]. As a result, coarser gradations with thicker asphalt films generally improve resistance to thermal cracking [90, 105, 107]. Although several studies reported improved low-temperature performance with increasing NMAS [108, 109], other findings show that larger NMAS can reduce cracking resistance under certain conditions, and NMAS effects are typically secondary to dust content and effective binder content [106].

At intermediate temperatures, cracking resistance is primarily influenced by gradation types. Coarser gradations with lower fine contents, likely having a higher effective binder

content or film thickness, show increased IDEAL-CT cracking tolerance indices when the aggregate source and binder grade are constant [28].

Table 12. Effect of Aggregate Gradation on Cracking Resistance of Asphalt Mixtures

Cracking Mechanism	Gradation Factor	Ref.(s)	Evaluation Method	Key Findings
Low-temperature cracking	Fine content, asphalt film thickness, NMAS	[90, 104-106, 108-112]	Low-temperature cracking tests and thermal cracking indices	Higher fine content reduces free binder and cracking resistance; coarser gradations and thicker asphalt films generally improve low-temperature cracking resistance. NMAS effects are secondary to fine content
Intermediate-temperature cracking	Gradation structure and fines content	[28]	IDEAL-CT	Coarser gradations with reduced fines improve intermediate-temperature cracking tolerance independent of aggregate type
Fatigue cracking	Gradation structure, NMAS, and mix stiffness	[103, 110-119]	Fatigue testing, VECD analysis, cracking indices	Finer gradations and smaller NMAS often improve fatigue resistance through increased flexibility, while optimized coarse gradations can also perform well when stiffness is controlled
Reflective cracking	NMAS and compaction	[103, 120, 121]	Laboratory and field-oriented cracking evaluations	Smaller NMAS mixtures are associated with improved reflective cracking resistance and better compactability
Stiffness-controlled cracking (cross-mechanism)	Sieve-level distribution and dust content	[122-126]	Modulus testing, predictive modeling, IDEAL-CT, I-FIT	Dust and intermediate sieve fractions control stiffness and asphalt film thickness, strongly influencing cracking susceptibility across mechanisms

VECD = viscoelastic continuum damage; I-FIT = Illinois Flexibility Index Test.

The influence of gradation on fatigue cracking is more variable and reflects competing effects of stiffness and flexibility. Studies reported improved fatigue resistance with finer gradations and smaller NMAS mixtures, attributed to increased flexibility, delayed crack initiation, and improved compaction [103, 110-113, 115-119, 127]. However, other investigations demonstrated that CG mixtures can also achieve good fatigue performance when the binder content and aggregate structure are optimized, as indicated by dynamic modulus and viscoelastic continuum damage (VECD)-based analyses [114]. These results indicate that fatigue performance depends on how gradation affects stiffness and damage evolution rather than on gradation type (coarse vs. fine) alone.

For reflective cracking, smaller NMAS mixtures are generally associated with improved resistance, largely due to enhanced compactability and reduced stress concentrations at joints and cracks [103, 120, 121].

Across cracking mechanisms, the percentages of materials passing through intermediate sieves and the dust content can influence cracking behavior by affecting the mixture's stiffness and the thickness of the asphalt film. Higher dust content reduces film thickness

and increases stiffness, leading to lower cracking indices and greater fatigue susceptibility. This relationship has been demonstrated through modulus-based models and cracking tests such as IDEAL-CT and the Illinois Flexibility Index Test (I-FIT) [122-126].

In summary, the findings suggest that the effect of gradation on cracking resistance depends on the specific mechanism involved. Low-temperature cracking is affected by fine content and binder availability, intermediate-temperature cracking by gradation type, and fatigue and reflective cracking by stiffness, compactability, and NMAS-related effects [28, 90, 103-128]. Although several studies have reported improved low-temperature performance, these findings support using cracking tests to assess cracking resistance in asphalt mixture design rather than relying solely on gradation limits [103, 125, 128].

2.5.3 Durability

Aggregate gradation affects the durability of asphalt mixtures by influencing raveling resistance and moisture susceptibility. The literature indicates that NMAS, gradation type, and percent passing on specific sieve sizes are the primary factors governing durability performance. The key findings from these studies are summarized in Table 13.

Reducing NMAS from larger sizes (e.g., 9.5 or 12.5 mm) to smaller sizes (e.g., 4.75 mm) has been shown to significantly improve resistance to raveling [129]. Similar trends have been reported for SMA mixtures, in which smaller NMAS consistently led to improved raveling resistance [67]. These findings indicate that smaller aggregate sizes promote improved particle interlock and reduce aggregate loss under abrasive loading.

NMAS also influences moisture resistance, with smaller NMAS generally associated with lower moisture sensitivity [130]. Ji et al. [129] showed that dense-graded mixtures with 12.5 mm-NMAS exhibited lower stripping numbers in the HWTT than mixtures with 19 mm NMAS. Similar trends were observed for SMA mixtures, in which reductions in NMAS decreased moisture sensitivity [67]. For OGFC mixtures, Watson et al. [131] and Xie et al. [132] reported that 9.5-mm NMAS mixtures retained higher tensile strength after moisture and freeze–thaw conditioning than 12.5-mm NMAS mixtures.

In addition to NMAS, the gradation type also affects durability. Finer gradations generally improve raveling resistance by increasing particle contact points and cohesive bonding. Cantabro testing demonstrated that SMA mixtures with finer gradation exhibited higher raveling resistance, while similar improvements were reported for OGFC mixtures with finer gradations [131, 133, 134].

Table 13. Effect of Aggregate Gradation on Durability of Asphalt Mixtures

Gradation Factor	Ref.(s)	Study Approach	Performance Measure	Key Findings
NMAS	[67, 129-132, 135]	Laboratory durability testing, HWTT, moisture conditioning	Cantabro loss, stripping ratio, tensile strength	Smaller NMAS improves durability by enhancing raveling resistance and reducing moisture sensitivity across dense-graded, SMA, and OGFC mixtures
Gradation type (coarse- vs. fine-graded)	[130, 132–136]	Cantabro and moisture damage	Cantabro loss, TSR	Finer gradations improve raveling resistance and moisture durability by increasing particle contact and reducing void connectivity
Void structure and permeability	[136-138]	Mechanistic interpretation and laboratory testing	Moisture damage indicators	CG mixtures exhibit greater moisture susceptibility due to interconnected voids, increased pore pressure, and aggregate breakage during compaction
Percent passing on No. 4, No. 8, and smaller	[59, 139]	Gradation sensitivity and durability testing	Cantabro loss, moisture resistance	Higher percentages passing the No. 4 and finer sieves improve durability by filling voids, increasing contact points, and enhancing cohesive bonding

Finer gradations also provide greater moisture resistance than coarser gradations. CG mixtures contain more interconnected voids that facilitate water ingress but restrict drainage, resulting in increased pore pressure under traffic loading and greater susceptibility to moisture damage [136]. Modified Lottman testing showed that fine gradations achieved higher tensile strength ratios than middle or coarse gradations [137], and a similar trend was also reported by Aodah et al. [138]. In addition, larger aggregates in CG mixtures are more prone to breakage during compaction, producing uncoated surfaces that absorb water and contribute to stripping [138].

Additionally, durability is influenced by gradation at specific sieve sizes. Chadbourn et al. [59] reported that gradation deviations, particularly at the No. 8 sieve and smaller sizes, significantly affected VMA and moisture resistance. For mixtures with the same NMAS, higher percentages passing the No. 4 sieve resulted in lower Cantabro loss and improved raveling resistance, as smaller particles may fill voids, increase aggregate contact, and enhance cohesive bonding [139].

Overall, the literature indicates that durability improves with smaller NMAS, finer gradation, and balanced gradation, which collectively enhance raveling resistance and reduce moisture damage. These findings suggest that durability performance is governed primarily by gradation type rather than NMAS or gradation limits alone and should be evaluated in performance tests.

2.5.4 Permeability

The literature indicates that permeability is governed by gradation type and NMAS, through their combined effects on AV connectivity and pore structure [140-150]. As summarized in Table 14, CG mixtures exhibit higher permeability than FG mixtures at similar AV contents because CA structures promote interconnected void networks that facilitate water flow [140-144].

Permeability also increases significantly once AV contents exceed approximately 6 percent, reflecting the transition from isolated voids to continuous flow paths [151, 152]. This behavior demonstrates that AV content alone is insufficient to characterize permeability without accounting for gradation structure.

NMAS also affects permeability at a given AV level. Larger NMAS mixtures generally exhibit higher permeability due to larger pore sizes and more continuous void channels, a trend observed consistently for dense-graded, OGFC, and SMA mixtures [67, 141, 145-150]. Reducing NMAS changes pore structure and lowers permeability, although the magnitude of this effect varies by mixture type.

Permeability is also sensitive to gradation at some sieve sizes, particularly the No. 4 and intermediate sieves. Lower percentages passing the No. 4 sieve significantly increase permeability by enlarging void spaces, whereas increasing the proportions of No. 8 and No. 16 aggregates disrupts void continuity and reduces permeability [54, 153]. Hainin et al. developed a model, shown in Equation (1), showing that permeability reflects the combined influence of gradation and construction-related factors, including CA ratio, in-place AV, and lift thickness [154].

$$\begin{aligned}\mathrm{Ln}\,(\mathrm{k}) \;=\; & -19.2 \;+\; 5.96\mathrm{Ln}(\mathrm{CL}) \;+\; 1.47(\mathrm{CA\ Ratio}) \;+\; 0.078(\mathrm{P12.5}) \\ & +\; 0.0485(\mathrm{P1.18}) \;+\; 0.00928(\mathrm{Ndes}) \;-\; 0.0124(\mathrm{Ave.\ Thickness})\end{aligned} \tag{1}$$

Where:

Ln (k) = natural log of permeability
Ln (CL) = natural log of AV 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
Ndes = design compaction effort of mix design
Ave. Thickness= average thickness of a given core (lift thickness at location of core)

Overall, these findings show that permeability is more reliably explained by gradation structure and NMAS than by AV content alone, underscoring the need for gradation-sensitive evaluation of permeability when durability and moisture resistance are critical performance considerations [143, 144, 151, 152, 154].

Table 14. Effect of Aggregate Gradation on Permeability of Asphalt Mixtures

Gradation Factor	Ref.(s)	Study Approach	Permeability Measure	Key Findings
Gradation type (coarse vs. fine)	[140-144]	Laboratory testing, field measurements, surveys	Laboratory and field permeability	CG mixtures exhibited higher permeability than FG mixtures at similar AV contents due to increased void connectivity
AV connectivity threshold	[151, 152]	Field and laboratory evaluation of Superpave mixtures	In-place permeability	CG Superpave mixtures became permeable once AV exceeded approximately 6%
NMAS (dense-graded mixtures)	[141, 145, 146]	Laboratory permeability testing at controlled AV	Hydraulic conductivity	Permeability increased substantially with increasing NMAS at the same AV content
NMAS (OGFC and SMA mixtures)	[67, 147-150]	OGFC and SMA mixture comparisons	Hydraulic conductivity	Increasing NMAS increased permeability in OGFC mixtures, while decreasing NMAS reduced permeability in SMA mixtures
Percent passing (No. 4 and intermediate sieves)	[54, 153]	Falling-head permeameter testing; gradation optimization studies	Laboratory permeability	Lower No. 4 passing increased permeability; increasing No. 8 and No. 16 fractions disrupted void continuity and reduced permeability
Combined gradation and construction effects	[154]	Statistical modeling of field data	Predicted permeability	Permeability increased with higher CA ratio, greater passing of selected sieves, and higher AV content

2.5.5 Summary

Aggregate gradation significantly influences asphalt mixture performance by affecting the aggregate skeleton, binder distribution, AV connectivity, and stiffness. The effects of gradation vary depending on performance requirements and often involve tradeoffs among rutting resistance, cracking resistance, durability, and permeability; there may not be a single optimal gradation condition. Table 15 summarizes the key findings from Sections 2.5.1 to 2.5.4, emphasizing how gradation parameters influence different performance mechanisms.

Overall, mixtures with coarser gradations and larger NMAS tend to improve rutting resistance by promoting stone-on-stone contact and load transfer but may reduce cracking resistance and increase permeability. In contrast, mixtures with finer gradations and smaller NMAS generally enhance fatigue life, low-temperature cracking resistance, and durability by increasing effective asphalt film thickness, flexibility, and moisture resistance, while reducing interconnected voids. The percentage of materials passing specific sieve sizes—particularly in relation to the PCS and dust content—affects performance by influencing aggregate interlock, stiffness, and effective binder content.

The literature also shows that gradation limits alone are insufficient predictors of performance. Mixed findings associated with the restricted zone and NMAS demonstrate

that acceptable performance can be achieved across a range of gradations when the aggregate skeleton is stable. Collectively, these findings reinforce the need to evaluate gradation effects within a performance-based mixture design framework, in which rutting, cracking, durability, and permeability are assessed directly using appropriate performance tests rather than inferred solely from gradation criteria.

Table 15. Impact of Gradation on Asphalt Mixture Performance

Gradation Parameter	Performance Type	Impact on Performance
Gradation Type	Rutting resistance	SMA > Dense-graded [75-78] CG > FG [90, 91] Lower bound > Mid-bound > Upper bound [79]
	Intermediate-temp cracking	Coarser gradation → Increase cracking resistance[28]
	Fatigue cracking	Finer gradation → Increase fatigue life [110-113] Coarse gradation→ Increase fatigue cracking resistance [114]
	Low-temp cracking	Coarse gradation → Increase low-temp cracking resistance (thicker asphalt film) [90, 104, 105]
	Durability	Finer gradations → Increase raveling and moisture resistance [133, 134, 136-138]
	Permeability	Coarse gradations → Increase permeability [140, 142-144, 151, 152]
Restricted Zone (ARZ, TRZ, BRZ)	Rutting resistance	Mixed findings [12, 80-84, 86-88]
Percent Passing Certain Sieve	Rutting resistance	No. 4, No. 8, No. 16, No. 30, No. 50 sieves greatly affect rutting [92-95, 97, 98]
	Cracking resistance	Significant influence on fatigue and reflective cracking performance [122, 125] High dust content (No. 200) → Decreasing fatigue resistance [124]
	Permeability	Lower passing No. 4 sieve → Higher permeability [54] Lower passing No. 4 size aggregates and higher passing of No. 8 and No. 16 → Lower permeability [153]
Bailey Method Parameters	Rutting resistance	Strong correlation with rut depth [94, 99, 100] Higher FA_c → Increase rutting resistance [94, 98]
	Cracking resistance	Higher CA and FA_c ratios → Reduce crack resistance [126]
NMAS	Rutting resistance	Larger NMAS → Increase rut resistance [67, 155-157]
	Low-temp cracking	Mixed findings [67, 108, 109][158]
	Fatigue cracking	Smaller NMAS → Increase fatigue life [103, 115-119]
	Reflective cracking	Smaller NMAS → Increase the reflective cracking resistance [120]
	Durability (raveling)	Smaller NMAS → Increase raveling resistance [67, 129]
	Moisture damage	Smaller NMAS → Increase moisture resistance [67, 131, 132, 135]
	Permeability	Larger NMAS → Increase permeability [141, 145] [147, 148]
CA Ratio	Permeability	Higher CA ratio → Increase permeability [154]

FA_c = FA course ratio; ARZ = above restriction zone; TRZ = through restriction zone; BRZ = below restriction zone.

2.6 Bailey Method

2.6.1 An Approach to Understanding Aggregate Packing

The Bailey Method is a systematic approach to improve the understanding of aggregate packing in asphalt mixtures. Robert D. Bailey developed the method in the early 1980s while serving as a civil engineer with District 5 of the Illinois Department of Transportation (IDOT). It was based on his extensive experience in the design, production, and placement of Portland cement concrete mixtures [159].

During that period, premature rutting was the primary distress observed in early HMA overlays placed on Portland cement concrete pavements on interstate routes in Illinois and other regions. Asphalt mixtures were designed using 50 Marshall blows, with the optimum asphalt content selected at 2.0 percent AV. Later practice increased compaction to 75 blows and selected the optimum asphalt content at 4.0 percent AV. Although Marshall stability and flow requirements were commonly specified, stability values did not consistently correlate with field rutting performance. The Bailey Method was introduced to address this limitation and improve rutting resistance [160].

At that time, manufactured sand was not commonly used in asphalt mixtures in Illinois. To address the limited availability of fine aggregates with what is now known as FAA, early mixtures designed using the Bailey Method focused on achieving a coarse gradation. This approach aimed to enhance stone-on-stone contact, relying on the angularity of crushed coarse aggregates to improve internal friction and resistance to rutting [160].

Over time, the Bailey Method evolved through application and training across the United States and internationally. Adaptation to local specifications and materials expanded its applicability to a range of mixture types and nominal sizes. The method is now commonly applied to FG, CG, and SMA mixtures [102].

2.6.2 NMAS and PCS in the Bailey Method

For many years, the Bailey Method defined NMAS as the first sieve larger than the first sieve retaining more than 10 percent of the aggregate. After review of production data, this threshold was revised to 15 percent in late 2006. For most CG mixtures, this change does not affect NMAS classification. The revision most often affects mixtures initially appearing as FG 19.0-mm mixtures, which may instead be classified as CG 12.5-mm mixtures [102].

Two features distinguish the Bailey Method from other combined gradation analysis approaches. First, it defines a PCS that separates coarse and fine aggregate fractions. Second, it establishes the aggregate blend by volume before converting to mass. The PCS is calculated as 0.22 times the NMAS, rounded to the nearest standard sieve size. This sieve forms the basis for evaluating aggregate packing characteristics within the combined gradation [159, 160]. Table 16 lists the PCS used in the Bailey Method.

Table 16. Primary Control Sieves Used in the Bailey Method [159, 160]

NMAS	NMAS × 0.22	Closest Standard Sieve
37.5 mm (1-1/2")	8.250 mm	9.5 mm (3/8")
25.0 mm (1")	5.500 mm	4.75 mm (No. 4)
19.0 mm (3/4")	4.180 mm	4.75 mm (No. 4)
12.5 mm (1/2")	2.750 mm	2.36 mm (No. 8)
9.5 mm (3/8")	2.090 mm	2.36 mm (No. 8)
4.75 mm (No. 4)	1.045 mm	1.18 mm (No. 16)

2.6.3 Establishing CA Volume to Determine Mix Type

The Bailey Method uses loose and rodded unit weight (UW) measurements to establish CA volume and define mix type. UW tests are conducted on each individual virgin aggregate used in a mix design, with results expressed in pounds per cubic foot or kilograms per cubic meter. Void content is determined for both loose and rodded conditions [160].

UW testing is performed on aggregates in the as-received condition, using the full stockpile gradation, in accordance with AASHTO T 19. While AASHTO T 19 provides general testing requirements, the Bailey Method recommends UW container sizes based on NMAS: 0.5 ft^3 for 37.5-mm (1.5-in.) stockpiles, 0.25 ft^3 for stockpiles with NMAS ranging from 25.0 mm (1 in.) to 9.5 mm (3/8 in.), and 1/30 ft^3 (4-inch-diameter Proctor mold) for stockpiles with a 4.75-mm (No. 4) or smaller NMAS [160]. Although UW testing is not performed on mineral fillers, baghouse fines, hydrated lime, or recycled materials, the Bailey Method provides procedures for incorporating these materials while maintaining the target CA volume.

For dense-graded mixtures, including FG and CG, the loose unit weight (LUW) of each CA is used to establish CA volume and mix type. The LUW condition represents the onset of aggregate interlock and particle-to-particle contact within the coarse fraction (material retained on the PCS). CG mixtures typically target CA volumes near 100 percent of the CA LUW condition, whereas FG mixtures typically target 80 percent or less of the CA LUW condition to promote separation of the coarse fraction and support by the fine fraction [160]. The selected percentage of CA LUW is referred to as the CA chosen unit weight (CUW) and serves as the primary design parameter used to distinguish between FG and CG mixtures.

Within a given aggregate blend, CAs are virgin aggregates with 49.9 percent or less passing the PCS of the blend, while fine aggregates (FAs) are those with 50.0 percent or more passing the PCS of the blend. When multiple virgin CA sources are used, the designer selects a CA blend by volume totaling 100 percent. This selection influences both the calculated CA volume and the gradation of the coarse fraction (retained on the PCS). The voids within the coarse fraction then establish the volume available for the fine fraction. The combined CA and FA volumes are treated as a unit volume, typically 1.0 cubic foot or 1.0 cubic meter [160].

In dense-graded mixtures, some (CG mixtures) or most (FG mixtures) of the fine fraction contributes to the strength of the aggregate structure. Accordingly, 100 percent of the rodded unit weight (RUW) of the FA is used to fill voids within the coarse fraction. When multiple FA sources are used, the designer selects an FA blend by volume totaling 100 percent, which directly affects the gradation and packing of the fine fraction (passing the PCS) [160].

For SMA mixtures, aggregate structure is dominated by the coarse fraction, and a high degree of stone-on-stone contact is required. In this case, the fine fraction primarily serves to fill voids within the coarse fraction, as opposed to significantly contributing to the strength of the aggregate structure. Additionally, the RUW of each CA is used to establish CA volume corresponding to the selected CA CUW, while 100 percent of the LUW of each FA is used to fill voids within the coarse aggregate skeleton. Multiple CA and FA sources are handled using the same volumetric blending approach applied to dense-graded mixtures [102].

Table 17 summarizes suggested CA CUW ranges for each mix type. These ranges reflect reported applications of the Bailey Method and guidance developed through training and field experience [102].

Table 17. Suggested CA CUW Values for Each Mix Type [102]

Mix Type	Min % CA LUW	Max % CA LUW	Min % CA RUW	Max % CA RUW
FG	60	80	–	–
CG	95	105	–	–
SMA	–	–	110	125

For aggregate blends with a combined aggregate bulk specific gravity (G_{sb}) between approximately 2.60 and 2.70, a CA CUW of about 90 percent of CA LUW has been observed to correspond to approximately 44 percent passing the PCS, which is near the intersection with the MDL on a standard 0.45-power curve [159]. A 10-percent change in CA CUW often corresponds to a shift of about 4 to 5 percent in the amount passing the PCS. During production, if the percent passing the PCS fluctuates by ±2 to 3 percent relative to the design target, the effective CA CUW may vary from approximately 85 to 95 percent of CA LUW. Such variation can result in a mixture exhibiting characteristics that transition between FG and CG behavior.

When notable differences exist between the coarse and fine fractions in particle shape, surface texture, or strength, transitions between FG and CG characteristics can complicate field compaction. To reduce this risk, a gap is intentionally maintained between the upper recommended CA LUW limit for FG mixtures (80 percent) and the lower recommended CA LUW limit for CG mixtures (95 percent), as shown in Table 17 [102].

2.6.4 Aggregate Packing

Aggregate packing is governed by (a) gradation (e.g., continuously graded vs. gap-graded), (b) type and amount of compaction effort (e.g., 50 Marshall blows vs. 75 gyrations), (c) particle shape (e.g., cubical vs. flat and/or elongated particles), (d) particle texture, which can impact consolidation, and (e) particle strength, which can affect the potential for degradation during mix production and compaction. Together, these factors control aggregate structure and directly influence VMA. Variations in material properties, production processes, and plant configuration can alter these factors, even for the same mixture design, leading to differences in packing behavior and compactability [160].

Within the Bailey Method, aggregate packing is addressed through four interrelated principles: (a) CA volume, (b) CA ratio, (c) FA course ratio (FA_c), and (d) FA fine ratio (FA_f). These principles provide a structured framework for understanding how changes in aggregate gradation and packing affect mix type, VMA, and compactability [102].

2.6.4.1 Principle 1: Coarse Aggregate Volume

Principle 1 establishes that CA volume is the primary factor determining mix type. CA volume is expressed as a percentage of CA LUW for FG and CG mixtures or as a percentage of CA RUW for SMA mixtures. By specifying the CA volume, the Bailey Method determines the coarse fraction volume and the corresponding volume available to the fine fraction.

The Bailey Method is not intended to force CA interlock in all mixtures. Instead, it provides a structured approach to help designers and production personnel ensure that the intended mix type is both designed and produced, and to help them understand how changes in CA volume influence VMA and mixture compactability [160].

For mixtures with CA interlock (i.e., CG and SMA), increasing CA volume reduces the percent passing the PCS and increases VMA by strengthening aggregate interlock. Although the design volume of the fine fraction decreases, increased interlock restricts compaction of the fine fraction within the CA skeleton. As a result, the fine fraction achieves a lower density, and VMA increases despite the reduced fine fraction design volume [160].

For mixtures without CA interlock (i.e., FG), decreasing CA volume increases the percent passing the PCS and increases VMA through two mechanisms. First, floating coarse particles are replaced by an equivalent mass of fine aggregate particles that contain internal voids. Second, the larger fine fraction volume contains more particles that must rearrange under the same compaction effort, limiting achievable density and increasing VMA [102].

Across FG, CG, and SMA mixtures, VMA is controlled primarily by the properties of the fine fraction and the degree to which it is compacted. The Bailey Method treats each mix type as a unit volume: once the CA volume is defined, a corresponding volume of voids is

created to be filled by the fine fraction. Any remaining voids in the aggregate structure, defined as VMA, reside within the overall fine fraction [160].

2.6.4.2 Principle 2: The Coarse Fraction

Principle 2 characterizes the particle size distribution of the coarse fraction (material retained on the PCS) using the CA ratio. Within this fraction, the smallest particles are too large to fit into the voids created by the largest particles. As a result, the gradation of the coarse fraction influences its packing characteristics, the compaction achieved in the fine fraction, and the mixture's susceptibility to segregation, particularly for mixtures with CA interlock (i.e., CG and SMA) [160].

To evaluate the coarse fraction, the *half sieve* is defined as one-half the NMAS. The half sieve and PCS together bound the particle sizes used to compute the CA ratio. Table 18 lists the corresponding half sieve and PCS for common NMAS values [102].

Table 18. Half Sieve and PCS for Each NMAS [102]

NMAS	Half Sieve	PCS
37.5 mm (1-1/2")	19.0 mm (3/4")	9.5 mm (3/8")
25.0 mm (1")	12.5 mm (1/2")	4.75 mm (No. 4)
19.0 mm (3/4")	9.5 mm (3/8")	4.75 mm (No. 4)
12.5 mm (1/2")	**6.35 mm (1/4")**	2.36 mm (No. 8)
9.5 mm (3/8")	4.75 mm (No. 4)	2.36 mm (No. 8)
4.75 mm (No. 4)	2.36 mm (No. 8)	1.18 mm (No. 16)

The CA ratio is determined using Equation (2):

$$\text{CA ratio} = \frac{\%\ \textit{Passing Half Sieve} - \%\ \textit{Passing PCS}}{100\% - \%\ \textit{Passing Half Sieve}} \quad (2)$$

The numerator of the equation represents particles commonly referred to as *interceptors*. These particles are too large to fit into voids formed by larger coarse particles and therefore disrupt packing, reduce compactability of the coarse fraction, and limit compaction achieved in the fine fraction. Higher CA ratios are associated with higher VMA and mixtures that are more difficult to compact but less susceptible to segregation [160].

The denominator represents particles referred to as *pluggers*. Increasing the proportion of pluggers improves compactability of the coarse fraction and enhances compaction of the fine fraction. Lower CA ratios, therefore, tend to reduce VMA and improve field compactability, but at the cost of increased segregation susceptibility [160].

For CG and SMA mixtures, VMA generally increases as the CA ratio increases and decreases as the CA ratio decreases. Low CA ratios are associated with greater segregation potential, while high CA ratios reduce field compactability [160]. For FG mixtures, this parameter is referred to as the *Old CA ratio*. In FG mixtures, the coarse fraction floats within

the fine fraction and has little influence on VMA; however, the Old CA ratio remains an important indicator of segregation susceptibility [102].

Table 19 provides suggested CA ratio ranges for CG and SMA mixtures, along with Old CA ratio ranges for FG mixtures, as starting guidance. These ranges reflect reported applications of the Bailey Method and should be refined based on aggregate characteristics and local performance experience [102].

Table 19. Suggested CA Ratio Ranges for CG and SMA Mixtures [102]

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

For CG mixtures with a larger number of coarse fraction particle sizes, the CA ratio can be closer to 1.0. For CG mixtures with fewer coarse fraction particle sizes, the CA ratio needs to be lower to achieve a compactible coarse fraction. As an example, a 37.5-mm (1-1/2-inch) NMAS mix has five coarse fraction particle sizes (e.g., 50.0–37.5, 37.5–25.0, 25.0–19.0, 19.0–12.5, and 12.5–9.5), whereas a 4.75-mm (No. 4) NMAS mix only has three (e.g., 9.5–4.75, 4.75–2.36, and 2.36–1.18).

In SMA mixtures, with a significantly higher CA volume (to achieve the degree of CA interlock desired), the overall coarse fraction must be more compactible than in a CG mix. Therefore, an SMA mix is gap-graded, resulting in a lower CA ratio. As shown in Table 19, the recommended CA ratio range for an SMA, relative to the recommended range of an equal NMAS CG mix, is considerably lower. Reducing the CA ratio (fewer interceptors, more pluggers) in an SMA creates a coarse fraction that is more compactible in the lab and the field. However, the lower CA ratio in an SMA mix typically does not result in segregation susceptibility because it is generally offset by the use of higher effective asphalt binder contents and modified asphalt binders.

In FG mixtures, the Old CA ratio still relates to segregation susceptibility. Since the coarse fraction floats within the overall fine fraction, it is possible to use a higher Old CA ratio with little to no impact on VMA. Table 19 outlines the suggested maximum limit for the Old CA ratio for the various FG NMAS sizes. Exceeding this limit can result in a reduction of the mix NMAS, turning an FG mix into a CG mix of the next-smaller NMAS size if the PCS changes as well. This is what often occurs with a mix typically viewed as being a 19.0-mm (3/4-inch) FG mix that is actually a 12.5-mm (1/2-inch) CG mix. Understanding mix size and type is crucial to understanding aggregate packing.

2.6.4.3 Principle 3: The Overall Fine Fraction

The fine fraction (material passing the PCS) in CG, SMA, and FG mixtures contains sufficient particle sizes that can be evaluated as a blend. If the gradation of the fine fraction is examined independently after removal of the coarse fraction (material retained on the PCS), the percent passing the PCS is 100 percent, and the percent retained on the next-smaller sieve typically exceeds 15 percent. Under these conditions, the PCS serves as both the maximum sieve and the NMAS for the fine fraction [160].

Because the fine fraction contains multiple particle sizes, a second break between coarse and fine particles can be defined within this fraction. This break is referred to as the secondary control sieve (SCS), which separates the coarse and fine portions of the fine fraction. The SCS is determined as 0.22 times the PCS, rounded to the nearest standard sieve size, consistent with the Bailey Method framework [102]. Principle 3 uses the FA_c ratio to characterize the gradation of the fine fraction. The FA_c ratio is defined in Equation (3):

$$\mathbf{FA}_c \textbf{ ratio} = \frac{\%\ \boldsymbol{Passing\ SCS}}{\%\ \boldsymbol{Passing\ PCS}} \quad \textbf{(3)}$$

Based on Equation (3), as the FA_c ratio decreases, the fine fraction becomes coarser and packing efficiency decreases, which increases VMA. When the FA_c ratio increases, the fine fraction becomes finer, and packing initially improves, leading to a decrease in VMA. This trend continues only up to a point. Increasing the FA_c ratio has an effect similar to increasing the percent passing the PCS in a CG mixture. As a CG mixture becomes finer, VMA decreases until the mixture transitions to FG behavior. Beyond this transition, further increases in fineness lead to higher VMA. The same response occurs within the fine fraction. Once the fine fraction transitions from CG-like to FG–like behavior, further increases in the FA_c ratio reduce packing efficiency and increase VMA. This response is referred to as the FA_c dip, a change in packing mechanism within the fine fraction [102].

2.6.4.4 Principle 4: The Fine Part of the Fine Fraction

The fine part of the fine fraction (material passing the SCS) in CG, SMA, and FG mixtures also contains sufficient particle sizes that can be evaluated as a blend. If this portion of the gradation is examined independently after removal of the coarser portion (material retained on the SCS), the percent passing the SCS is 100 percent, and the percent retained on the next-smaller sieve typically exceeds 15 percent. Under these conditions, the SCS serves as both the maximum sieve and the NMAS for the fine part of the fine fraction.

Because multiple particle sizes exist below the SCS, an additional break between coarse and fine particles can be defined within this fraction. This break, referred to as the tertiary control sieve (TCS), separates the coarse and fine portions of the fine part of the fine fraction. The TCS is determined as 0.22 times the SCS, rounded to the nearest standard sieve size, consistent with the Bailey Method framework.

Principle 4 uses the FA_f ratio to characterize the gradation of the fine part of the fine fraction (minus SCS). The FA_f ratio is defined in Equation (4):

$$\mathbf{FA_f\ ratio} = \frac{\%\ \boldsymbol{Passing\ TCS}}{\%\ \boldsymbol{Passing\ SCS}} \quad (4)$$

Based on Equation (4), as the FA_f ratio decreases, the finest portion of the aggregate structure becomes coarser and packing efficiency decreases, which increases VMA. As the FA_f ratio increases, the fine fraction becomes finer, and packing efficiency initially improves, leading to a decrease in VMA. This trend continues only up to a point. Increasing the FA_f ratio has an effect similar to increasing fineness in a CG mixture. As fineness increases, VMA decreases until the mixture transitions to FG–like behavior. Beyond this transition, further fining reduces packing efficiency and increases VMA. Within the fine part of the fine fraction, this characteristic is referred to as the FA_f DIP. The FA_f DIP reflects a change in packing mechanism within the finest portion of the aggregate structure and provides guidance for selecting FA_f ratio ranges that balance compactability and volumetric stability.

2.6.4.5 Suggested Thresholds for CG and SMA Mixtures

Table 20 and Table 21 provide suggested ranges for the CA ratio, FA_c ratio, and FA_f ratio for CG and SMA mixtures. These ranges are based on documented applications of the Bailey Method and are intended as initial guidance for mixture design and evaluation [160].

Table 20. Suggested Ranges for CA, FA_c, and FA_f Ratios for CG Mixtures [102]

NMAS	37.5 mm (1-1/2")	25.0 mm (1")	19.0 mm (3/4")	12.5 mm (1/2")	9.5 mm (3/8")	4.75 mm (No. 4)
CA ratio	0.80-0.95	0.70-0.85	0.60-0.75	0.50-0.65	0.40-0.55	0.30-0.45
FA_c ratio	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*
FA_f ratio	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*	0.35-0.50*

* Avoid values less than 0.40

Table 21. Suggested Ranges for CA, FA_c, and FA_f Ratios for SMA Mixtures [102]

NMAS	37.5 mm (1-1/2")	25.0 mm (1")	19.0 mm (3/4")	12.5 mm (1/2")	9.5 mm (3/8")	4.75 mm (No. 4)
CA ratio	–	0.45-0.60	0.35-0.50	0.25-0.40	0.15-0.30	0.05-0.20
FA_c ratio	–	0.60-0.85	0.60-0.85	0.60-0.85	0.60-0.85	0.60-0.85
FA_f ratio	–	0.65-0.90	0.65-0.90	0.65-0.90	0.65-0.90	0.65-0.90

For CG mixtures, the recommended ranges reflect the need to balance CA packing, fine fraction compactability, and segregation resistance across different NMAS. For SMA mixtures, the recommended ranges reflect the higher CA volume required to achieve stone-on-stone contact and the corresponding need for a more compactible coarse fraction.

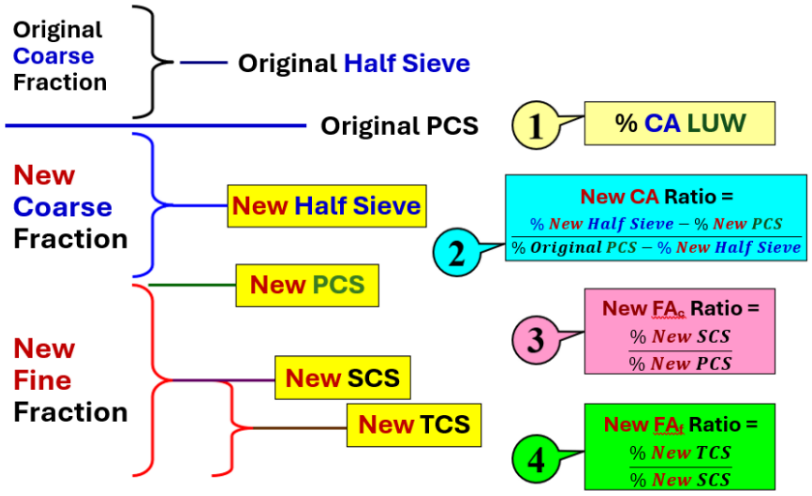
The suggested ranges should be refined as needed based on aggregate characteristics, mixture type, and local performance experience. Guidance for FG mixtures is addressed separately in the following section.

2.6.4.6 Suggested Thresholds for FG Mixtures

What makes an FG mix fine-graded is having a volume of fine fraction that exceeds the volume of voids in the coarse fraction. Therefore, the volume of the coarse fraction must be less than the CA LUW condition. In an FG mixture, the characteristics of the aggregate blend are governed by the amount and properties of the fine fraction.

For FG mixtures, the Bailey Method evaluates aggregate packing by focusing on the portion of the combined gradation passing the PCS. In this approach, the original PCS of the combined gradation is treated as the new maximum sieve and NMAS of the controlling blend (i.e., the fine fraction). The Bailey Method relationships are then reapplied to this controlling blend to define a new half sieve, new PCS, new SCS, and new TCS, where applicable.

Using these new control sieves, a new CA ratio, a new FA_c ratio, and a new FA_f ratio are calculated for the FG mixture. As with the Bailey Method framework applied to the original blend, the half sieve is defined as 0.5 times the NMAS of the blend being evaluated, and each control sieve is defined as 0.22 times the NMAS of that blend. This process allows the fine fraction to be evaluated using the same aggregate packing principles applied to CG and SMA mixtures. Figure 4 provides a schematic illustration to demonstrate this process for FG mixtures. Suggested thresholds for FG mixtures corresponding to typical NMAS values are summarized in Table 22.



© 2004 Heritage Research Group

Figure 4. Illustration of the Bailey Method Applied to FG Mixtures, Showing Redefinition of Control Sieves and Aggregate Fractions Using the Combined Gradation Passing the PCS [102]

Source: Reproduced from *The Bailey Method: Achieving Volumetrics and HMA Compactability* by W. J. Pine, Heritage Research Group. Reprinted with permission.

Table 22. Suggested Ranges for Recalculated CA, FA_c, and FA_f Ratios for FG Mixtures Based on the Controlling Blend Passing the PCS [102]

NMAS	37.5 mm (1-1/2")	25.0 mm (1")	19.0 mm (3/4")	12.5 mm (1/2")	9.5 mm (3/8")	4.75 mm (No. 4)
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 percent passing for the blend below the original PCS.
* For these mixes, only the New CA and FA_c Ratios can be determined.

2.6.4.6.1 Special Cases for FG Mixtures

For FG mixtures, Principle 1 generally indicates that increasing VMA requires making the mix finer. In some cases, however, FG mixtures respond in the opposite manner, where making the mix finer (as measured by percent passing the original PCS) decreases VMA. These responses reflect situations in which an FG mixture begins to behave more like a CG mixture, making the characterization of aggregate packing more difficult.

The following indicators suggest that an FG mixture may be behaving as a CG mixture:

1. Incorrect NMAS determination
2. CA CUW greater than 80-percent CA LUW
3. "Old" CA ratio exceeding the suggested maximum limit for an FG mixture
4. "Old" FA_c ratio less than 0.450
5. Voids in the CA of the mix (VCA_{mix}) achieved less than the voids corresponding to 90-percent CA LUW

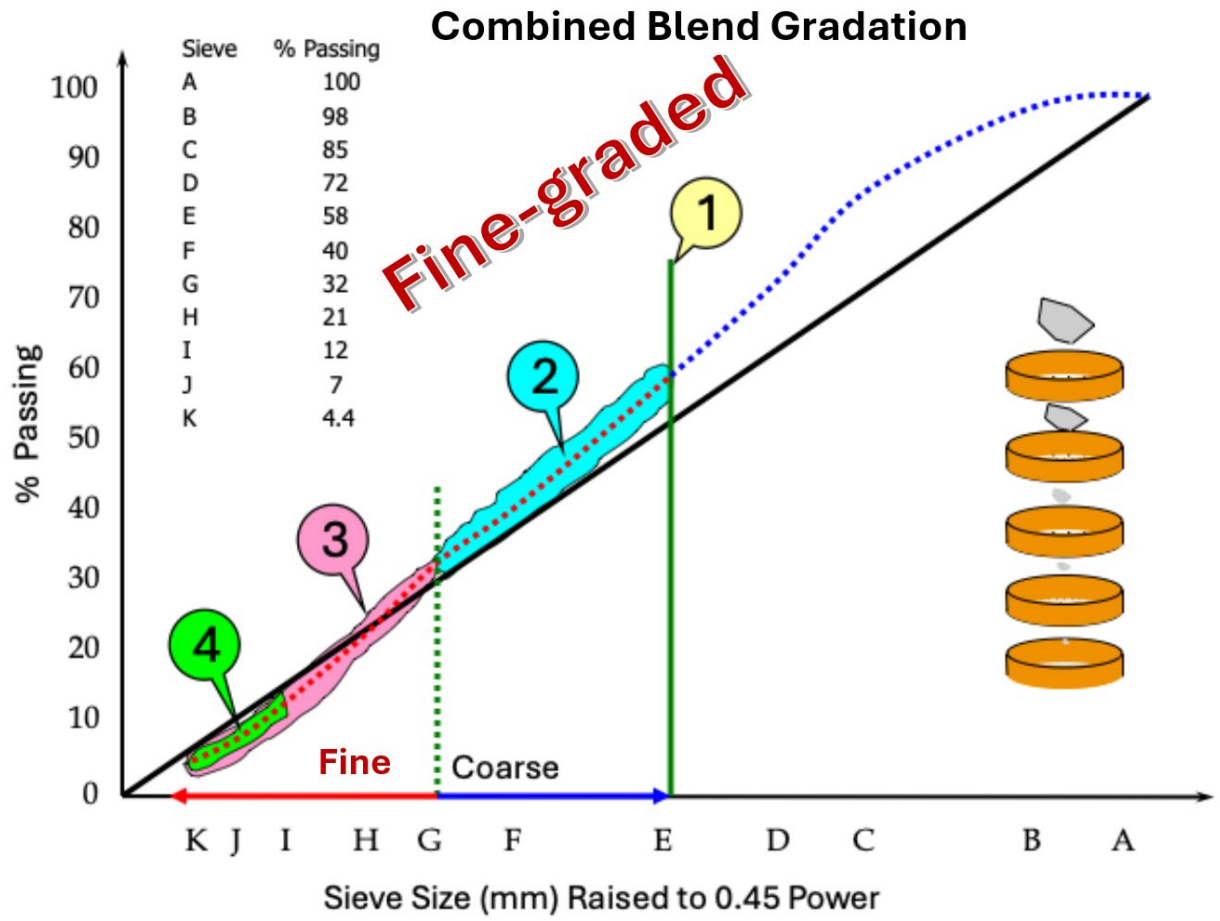
The first indicator (i.e., incorrect NMAS determination) was a primary reason for revising the NMAS definition in the Bailey Method. A common example is a mixture perceived as a 19.0-mm (3/4-inch) FG mix that has 85 percent or more passing the 12.5-mm (1/2-inch) sieve. From a Bailey Method perspective, such a mixture is more accurately classified as a 12.5-

mm NMAS. This reclassification shifts the PCS to the 2.36-mm (No. 8) sieve and often shows a CA CUW exceeding 95 percent of CA LUW, indicating CG behavior.

The second indicator (i.e., CA CUW greater than 80 percent of CA LUW) reflects the fact that virgin fine aggregate in an asphalt mixture often compacts to a density greater than its RUW. This occurs because the asphalt binder acts as a lubricant and applied compaction effort (e.g., 75 gyrations) densifies the fine fraction more effectively than dry rodding alone. As the fine fraction volume decreases, CA particles move closer together. The net result is that the coarse fraction makes up a greater percentage of the reduced total volume than originally designed. In some cases, particle-to-particle contact within the material retained on the PCS occurs, causing the mixture to respond as CG when gradation changes are made to adjust VMA.

The third indicator (i.e., an "Old" CA ratio exceeding the suggested maximum for FG mixtures) indicates an excess of interceptors and insufficient pluggers within the coarse fraction. This condition reduces the void size between floating CA particles and limits the packing efficiency of the fine fraction (material passing the original PCS). This indicator often coincides with an incorrectly defined NMAS (i.e., the first indicator). When the PCS shifts to a smaller sieve, the mixture may properly be classified as CG.

By definition, an FG mixture has a fine fraction (material passing the original PCS) volume that exceeds the void volume of the coarse fraction. In this condition, the fine fraction controls the characteristics of the whole aggregate structure, making it essential to determine whether the fine fraction itself behaves as FG or CG. As illustrated in Figure 5, the fine fraction (i.e., material passing the original PCS) is treated as a "new" blend gradation for evaluating aggregate packing in FG mixtures.



© 2004 Heritage Research Group

Figure 5. Fine Fraction (Passing the Original PCS) Treated as a New Blend Gradation for Evaluating FG Mixtures [102]

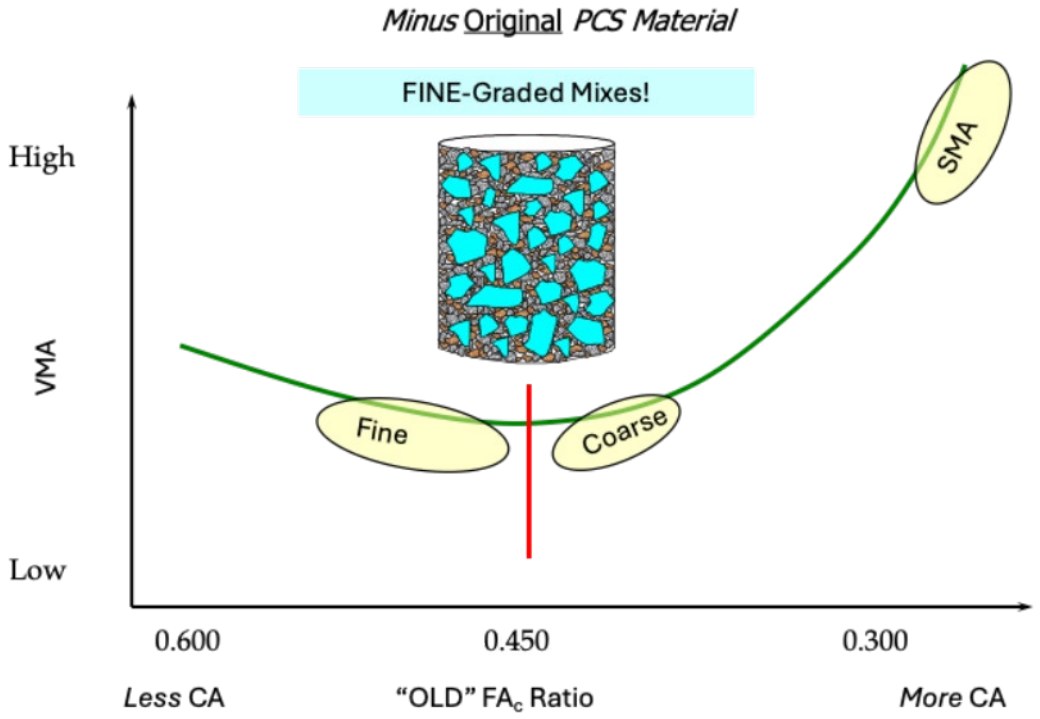
Source: Reproduced from *The Bailey Method: Achieving Volumetrics and HMA Compactability* by W. J. Pine, Heritage Research Group. Reprinted with permission.

The fourth indicator (i.e., "Old" FA_c ratio less than 0.450) is used to characterize the fine fraction (minus original PCS) of an FG mixture as either being FG or CG. The "Old" FA_c ratio is determined using Equation (5):

$$\text{"Old" } FA_c \text{ ratio} = \frac{\%\ \textit{Passing Original SCS}}{\%\ \textit{Passing Original PCS}} \quad (5)$$

The lowest, or dip, in VMA response (bottom of the curve in Figure 6) has been observed at an Old FA_c ratio near 0.450. Values greater than this threshold (left of the dip) indicate an FG fine fraction, while lower values (right of the dip) indicate a CG fine fraction. When the fine fraction is CG, making an FG mixture coarser in terms of percent passing the original PCS may increase VMA rather than decrease it. This occurs because reduced packing efficiency in the fine fraction limits compaction as floating coarse particles move closer

together (i.e., a coarser FG mix), resulting in a smaller void between the floating coarse fraction particles in which to orient the CG fine fraction.



© 2004 Heritage Research Group

Figure 6. Conceptual Relationship Between the "Old" FA_c Ratio and VMA for FG Mixtures [102]

Source: Reproduced from *The Bailey Method: Achieving Volumetrics and HMA Compactability* by W. J. Pine, Heritage Research Group. Reprinted with permission.

The fifth indicator (i.e., VCA_{mix} less than the voids corresponding to 90-percent CA LUW) involves comparing the measured VCA_{mix} with the theoretical voids associated with the designed CA CUW and CA blend by volume. In FG mixtures, the fine fraction often compacts to a density exceeding its RUW because compaction effort is focused solely on the fine fraction. As a result, the actual CA CUW achieved may exceed the design value. In these cases, the effective or actual CA CUW corresponding to the measured VCA_{mix} should be determined to assess whether the mixture has reached or approached coarse aggregate interlock. If so, the mixture should be treated as CG when making gradation adjustments to change VMA.

These indicators are directly applicable to many P-401 mixtures, which are commonly designed as FG due to specification gradation limits. The Bailey Method provides a framework to better understand the complex behavior of FG mixtures by identifying which portion of the aggregate blend controls the aggregate structure. This understanding supports the characterization of volumetric properties, performance test results, and field compaction.

2.6.5 Summary

The Bailey Method includes practical rules of thumb that describe how changes in each principle influence VMA, typically expressed as the gradation change required to produce a 1-percent change in VMA. Understanding mix type and the role of each principle allows designers and production quality control personnel to better interpret VMA results from mix design trials and production samples, and to anticipate the direction and magnitude of VMA changes before adjusting the combined gradation.

Changes in AV generally follow changes in VMA, although they are also influenced by variations in effective asphalt binder content. Estimating VMA and AV is a detailed process that is addressed through Bailey Method training, where participants work through example datasets and apply provided guidelines and spreadsheet tools during implementation. The Bailey Method can be used to:

- Evaluate existing mixtures to accurately identify mix type and NMAS and to better understand why some mixtures performed as intended while others did not, including refinement of acceptable ranges for the four principles based on local materials.
- Develop new mixtures to meet specification requirements, where feasible, while considering field compaction and segregation susceptibility.
- Understand differences in VMA and AV between samples as a function of gradation changes during mix design trials and production.
- Estimate changes in VMA and AV prior to making gradation adjustments during mix design and production.
- Improve understanding of aggregate structure and its influence on performance test results.
- Reduce trial-and-error, saving time and lowering production risk.

2.7 Prediction Models for Volumetric Properties and Performance of Asphalt Mixtures

Researchers have developed predictive models to quantify how aggregate characteristics and gradation influence volumetric properties and performance metrics. These models vary from traditional regression methods to advanced machine learning techniques that capture nonlinear relationships among different variables. The subsequent sections present these models, organized by prediction response: volumetric properties, rutting resistance, and cracking resistance.

2.7.1 Prediction Models for Volumetric Properties

The relationship between aggregate gradation, aggregate characteristics, and VMA is well established. As a result, numerous predictive models have been developed to estimate VMA using gradation parameters, aggregate properties, and packing-related indices. These

models generally fall into three categories: (1) parametric regression models, (2) mechanistic or packing-based models, and (3) data-driven machine learning models.

Early VMA prediction efforts relied primarily on parametric regression techniques. Zang [161] developed a linear regression model using percentages passing 12 sieves ranging from 26.5 mm to 0.075 mm, along with aggregate specific gravity parameters. Subsequent studies incorporated gradation descriptors derived from aggregate packing concepts. Vavrik et al. [160] proposed a polynomial regression model using Bailey Method parameters, including the CA ratio, FA_c ratio, and FA_f ratio. Liu [162] simplified this formulation by using only the CA and FA_c ratios, while Li [163] expanded the model to include interaction terms among the three Bailey ratios. These regression-based approaches demonstrated that gradation ratios grounded in packing theory provide greater explanatory power than sieve-based variables alone.

Mechanistic models based on particle packing theory were also developed to explicitly describe the mechanisms of void formation. Hudson and Davis [53] and Shen and Yu [69] introduced models that quantify changes in void volume as successive aggregate size groups are added to a base aggregate skeleton. Inoue et al. [164] proposed a two-aggregate system model incorporating particle shape factors and void ratios. Yu and Standish [165], as applied by Pouranian and Haddock [166], used specific volume concepts to account for particle interactions across size ranges. Li and He [167] further combined packing principles with gradation parameters by relating VMA to voids in CA and density ratios. While these mechanistic models offer physical insight into aggregate packing behavior, they often require detailed material characterization and increased computational effort.

More recently, data-driven machine learning (ML) approaches have been applied to predict VMA with higher accuracy and flexibility. Othman [168] employed deep artificial neural networks using aggregate gradation as the sole input and demonstrated moderate predictive capability. Shaikh et al. [169] developed ML-based VMA prediction models using a large dataset and multiple input variables, including aggregate source, NMAS, surface area, binder type, and compaction level. Among the evaluated algorithms, ensemble methods—particularly Gradient Boosting Regression—provided the highest accuracy, significantly outperforming traditional linear regression. These ML-based approaches highlight the potential of data-driven techniques to capture nonlinear relationships among mixture variables, though their applicability depends on the availability of sufficiently large, representative datasets. Table 23 summarizes representative regression and mechanistic models proposed for VMA prediction.

Table 23. Summary of VMA Prediction Models

Study	Models	Primary Inputs	Key Characteristics / Notes
Zang [161]	Linear regression	Percent passing multiple sieves (26.5 mm to 0.075 mm); aggregate specific gravity	Sieve-based model; relatively simple but limited physical interpretation
Vavrik et al. [160]	Polynomial regression	Bailey Method ratios (CA, FA_c, FA_f)	Packing-based variables; directly linked to aggregate structure
Liu [162]	Regression	CA ratio, FA_c ratio	Simplified Bailey-based model with reduced inputs
Li [163]	Regression with interaction terms	CA, FA_c, FA_f ratios and interactions	Improved explanatory power; increased model complexity
Hudson and Davis [53]	Mechanistic (packing-based)	Interference factors between aggregate size groups	Incremental packing approach; physically based but data-intensive
Shen and Yu [69]	Mechanistic (packing-based)	Aggregate volume fractions; interference factors	Explicit void evolution with size additions
Inoue et al. [164]	Mechanistic	Particle shape factors; void ratios	Two-aggregate system; incorporates shape effects
Yu and Standish [165]; Pouranian & Haddock [166]	Packing-based (specific volume)	Specific volume of aggregate size groups	Captures particle interaction effects across gradation
Li and He [167]	Hybrid (packing + regression)	Voids in CA; density ratios; gradation parameters	Links VMA to VCA concepts and density relationships
Othman [168]	Neural network	Gradation (percent passing 3/8" to No. 200)	Early ML application; moderate prediction accuracy
Shaikh et al. [169]	ML (ensemble methods)	Aggregate source, NMAS, surface area, binder type, compaction	High accuracy; captures nonlinear interactions but data-dependent

2.7.2 Prediction Models for Rut Depth

Researchers have developed a wide range of models to predict rut depth in asphalt mixtures, spanning traditional regression-based formulations, evolutionary algorithms, and data-driven ML approaches. These models differ in complexity, data requirements, and predictive capability, reflecting tradeoffs between interpretability and accuracy.

2.7.2.1 Regression and Evolutionary Models

Early rut-depth prediction efforts relied on regression models to relate permanent deformation to mixture properties and test conditions. Ezzat and Abed [170] developed a multiple linear regression model for polymer-modified asphalt mixtures using four variables—loading cycles, aggregate gradation (percent passing the No. 8 sieve), binder viscosity, and test temperature—to estimate rut depth measured from wheel-tracking tests. The model achieved moderate predictive performance (R^2 = 0.69; standard error of the regression [SER] = 0.35) and demonstrated acceptable statistical validity as indicated by residual analysis.

To better capture nonlinear behavior, Majidifard et al. [171] applied gene expression programming (GEP) to HWTT results. Using 13 input variables representing mixture composition, binder properties, recycled material content, gradation characteristics, and test conditions, the GEP model achieved improved prediction accuracy relative to linear regression, with a correlation coefficient of 0.75, root mean square error (RMSE) of 1.81, and a mean absolute error (MAE) of 1.27.

2.7.2.2 Machine Learning Models

ML approaches, particularly ensemble tree-based methods, have shown strong potential for rut depth prediction due to their ability to model nonlinear relationships and interactions among multiple variables. Ma [172] developed a Random Forest (RF) model to predict the Rutting Resistance Index (RRI) using more than 4,300 observations from 25 mixture types. Key inputs included mixture type, aggregate properties, binder grade, asphalt content, additives, reclaimed asphalt pavement (RAP) content, and NMAS. The RF model achieved an R^2 of 0.64 and outperformed a convolutional neural network evaluated in the same study.

Lang et al. [173] applied an RF model to a large IDOT database comprising more than 8,000 HWTT samples from nearly 3,800 mix designs. Using a broad set of mixtures, binders, recycled materials, and volumetric variables, the model achieved strong predictive performance (R^2 = 0.79) without requiring local trend adjustment, indicating good generalization across materials, mixture types, and production conditions.

2.7.2.3 Comparative Algorithm Studies

Several studies have directly compared multiple ML algorithms to identify optimal modeling approaches. Erten and Remzi [174] evaluated a suite of algorithms for predicting Marshall stability and flow, with the Extra Trees algorithm providing the best overall performance. Feature importance analysis indicated that volumetric parameters, fine aggregate fractions, and binder properties were dominant predictors.

Tong et al. [175] compared linear regression, RF, XGBoost, and Support Vector Regression (SVR) models for predicting APA rut depth using 648 specimens. SVR provided the best predictive performance (R^2 = 0.77), followed by XGBoost and RF, while linear regression performed poorly. These findings highlight the importance of nonlinear modeling approaches when predicting rutting behavior.

Table 24 summarizes representative rut-depth prediction models, highlighting model class, primary input variables, and key characteristics relevant to mixture design and performance evaluation.

Table 24. Summary of Rut Depth Prediction Models

Study	Rutting Parameter	Model Class	Primary Inputs	Key Characteristics / Notes
Ezzat & Abed [170]	Rut depth (wheel-tracking test)	Linear regression	Loading cycles, gradation (% passing No. 8), binder viscosity, temperature	Simple, interpretable model; moderate accuracy for polymer-modified mixes
Majidifard et al. [171]	Rut depth (HWTT)	Evolutionary (GEP)	Mix composition, binder grade, recycled content, gradation, test conditions	Captures nonlinear effects better than linear regression
Ma [172]	RRI	ML (RF)	Mix type, aggregate properties, binder PG, asphalt content, RAP, NMAS	Robust across mixture types; outperformed CNN model
Lang et al. [173]	Rut depth (HWTT)	ML (RF)	Mix characteristics, binder properties, recycled content, volumetrics	High accuracy using large DOT database; strong generalization
Erten & Remzi [174]	Marshall flow	ML (Extra Trees)	Gradation, binder properties, volumetric parameters	Highlights importance of VMA and fine aggregate fractions
Tong et al. [175]	Rut depth (APA test)	ML (SVR / XGB / RF)	Volumetrics, RAP, gradation, binder viscosity	Nonlinear ML models outperform regression approaches

CNN = convolutional neural network; DOT = department of transportation.

2.7.3 Prediction Models for Cracking Resistance

A range of models has been developed to predict cracking performance in asphalt mixtures, reflecting the diversity of cracking mechanisms, test methods, and performance indices used in practice. These models span empirical regression, ML, and hybrid approaches, with increasing emphasis on data-driven methods to capture nonlinear relationships between mixture characteristics and cracking resistance.

2.7.3.1 Models for Intermediate-Temperature Cracking

Several studies have focused on predicting intermediate-temperature cracking resistance using fracture-based indices. Cooper et al. [176] developed a nonlinear artificial neural network (ANN) model to predict the critical strain energy release rate (Jc) from the semi-circular bend (SCB) test. Following screening of 24 candidate variables to remove multicollinearity and non-significant inputs, the final model used four parameters—percent passing the No. 200 sieve, dust-to-asphalt ratio, asphalt film thickness, and polymer content. The ANN model achieved strong predictive performance (R^2 = 0.93 for training and R^2 = 0.81 for validation), indicating suitability as a screening-level tool during mixture design.

Lang et al. [173] developed a two-step ML framework to predict the Flexibility Index (FI) from I-FIT using a large IDOT database. An RF model incorporating 32 mixture, aggregate, binder, and volumetric variables achieved an R^2 of 0.80, with a SHAP (Shapley Additive exPlanations) value analysis identifying asphalt content and aggregate fractions near the NMAS as the most influential predictors.

Multiple studies have developed prediction models for the cracking tolerance index (CT_{Index}) using different datasets and ML algorithms. Nguyen et al. [177] evaluated multivariate adaptive regression splines, RF, and k-nearest neighbors models using 107 laboratory samples. The RF model performed best, achieving near-perfect prediction accuracy (R^2 = 0.995), with SHAP analysis indicating that percent passing the No. 4 sieve, RAP content, and binder content were dominant predictors.

Rad et al. [178] focused on long-term oven-aged (LTOA) mixtures, developing CT_{Index} prediction models using 34 plant-produced mixes from Oklahoma. Among the algorithms evaluated, XGBoost provided the best performance (R^2 = 0.81), with residual analysis confirming stable prediction across mixture types.

Tong et al. [175] compared linear regression, RF, XGBoost, and SVR models using 648 specimens from Virginia. XGBoost achieved the highest accuracy (R^2 = 0.92), followed closely by RF and SVR, while linear regression performed comparatively poorly. These results highlight the importance of nonlinear ML approaches for CT_{Index} prediction.

2.7.3.2 Models for Low-Temperature Cracking

Mirzaiyanrajeh et al. [179] developed prediction models for low-temperature fracture energy (G_f) using an extensive dataset of 852 disk-shaped compact tension (DCT) test results from 71 mixtures. Among three modeling approaches evaluated, Self-Validated Ensemble Modeling provided the most efficient balance between accuracy and computational cost, achieving R^2 values of 0.93 for training and 0.87 for testing. Sensitivity analysis identified design traffic level, low-temperature binder grade, percent passing the 3/8-inch sieve, and VMA as the most influential variables.

Table 25 summarizes representative cracking prediction models, highlighting cracking mode, model class, primary input variables, and key characteristics relevant to mixture design and performance evaluation.

Table 25. Summary of Cracking Performance Prediction Models

Study	Cracking Mode / Index	Model Class	Primary Inputs	Key Takeaways
Cooper et al. [176]	Intermediate-temperature cracking (SCB Jc)	ANN	Fines content, dust-to-asphalt ratio, film thickness, polymer content	Strong screening tool; highlights importance of fines and binder film
Lang et al. [173]	Intermediate-temperature cracking (I-FIT FI)	Random Forest	Asphalt content, aggregate fractions near NMAS, volumetrics	High accuracy using large DOT database; interpretable via SHAP
Nguyen et al. [177]	Intermediate-temperature cracking (CT_{Index})	Random Forest	Gradation (% passing No. 4), RAP, binder content	Excellent predictive capability; gradation and RAP dominate
Rad et al. [178]	Cracking after aging (CT_{Index}, LTOA)	XGBoost	Gradation, volumetrics, RAP, binder grade	Effective for aged mixtures; suitable for durability assessment
Tong et al. [175]	Intermediate-temperature cracking (CT_{Index})	XGBoost / RF / SVR	Volumetrics, RAP, gradation, binder viscosity	Nonlinear ML models outperform regression approaches
Mirzaiyanrajeh et al. [179]	Low-temperature cracking (DCT Gf)	Ensemble modeling	Binder PG, traffic level, gradation, VMA	Traffic level and binder grade dominate low-temperature cracking

2.7.4 Summary

This section discusses models for predicting volumetric properties, rut depths, and cracking resistance of asphalt mixtures. Prediction models for volumetric properties rely largely on gradation, aggregate specific gravity, asphalt content, and packing variables, reflecting the central role of aggregate packing and binder distribution in defining mixture structure. These same variables—often supplemented by binder rheological properties, aging condition, and recycled material content—also appear as dominant inputs in rutting and cracking models.

Rutting prediction models extend volumetric variables by incorporating mixture stiffness, binder grade, and test conditions, and ML approaches demonstrate better predictive performance by capturing nonlinear interactions among these factors. Cracking prediction models, by contrast, place greater emphasis on fine fraction characteristics, binder film thickness, aging effects, and volumetric balance—particularly VMA and AV—which govern fracture resistance at intermediate and low temperatures. ML methods consistently outperform traditional regression while retaining interpretability.

Collectively, these findings underscore the interdependence of mixture volumetrics, rutting resistance, and cracking performance. Accurate characterization of aggregate structure and volumetric balance not only improves prediction of individual performance metrics but also supports an integrated, performance-based framework for mixture design. This

synthesis reinforces the importance of gradation control and volumetric consistency in Federal Aviation Administration P-401 mixtures and highlights the value of data-driven modeling approaches for evaluating tradeoffs among durability, rutting resistance, and cracking performance within a mix design.

Chapter 3. Research Methodology

3.1 Overview

This chapter describes the experimental plan for evaluating the feasibility of adjusting the aggregate gradation limits for P-401 mix designs. These adjustments could provide greater flexibility in using locally available aggregates to design gradations that meet the current volumetric criteria while ensuring optimal performance.

The research plan was structured in two complementary parts. The first part evaluated the performance of mixtures with aggregate gradations intentionally positioned outside the current P-401 gradation limits. These mixtures were compared directly with corresponding mixtures designed within the existing gradation limits while maintaining similar volumetric properties and asphalt binder contents (including effective binder contents). This controlled comparison isolates the effect of gradation on performance and establishes a technical basis for assessing potential modifications to the P-401 gradation limits.

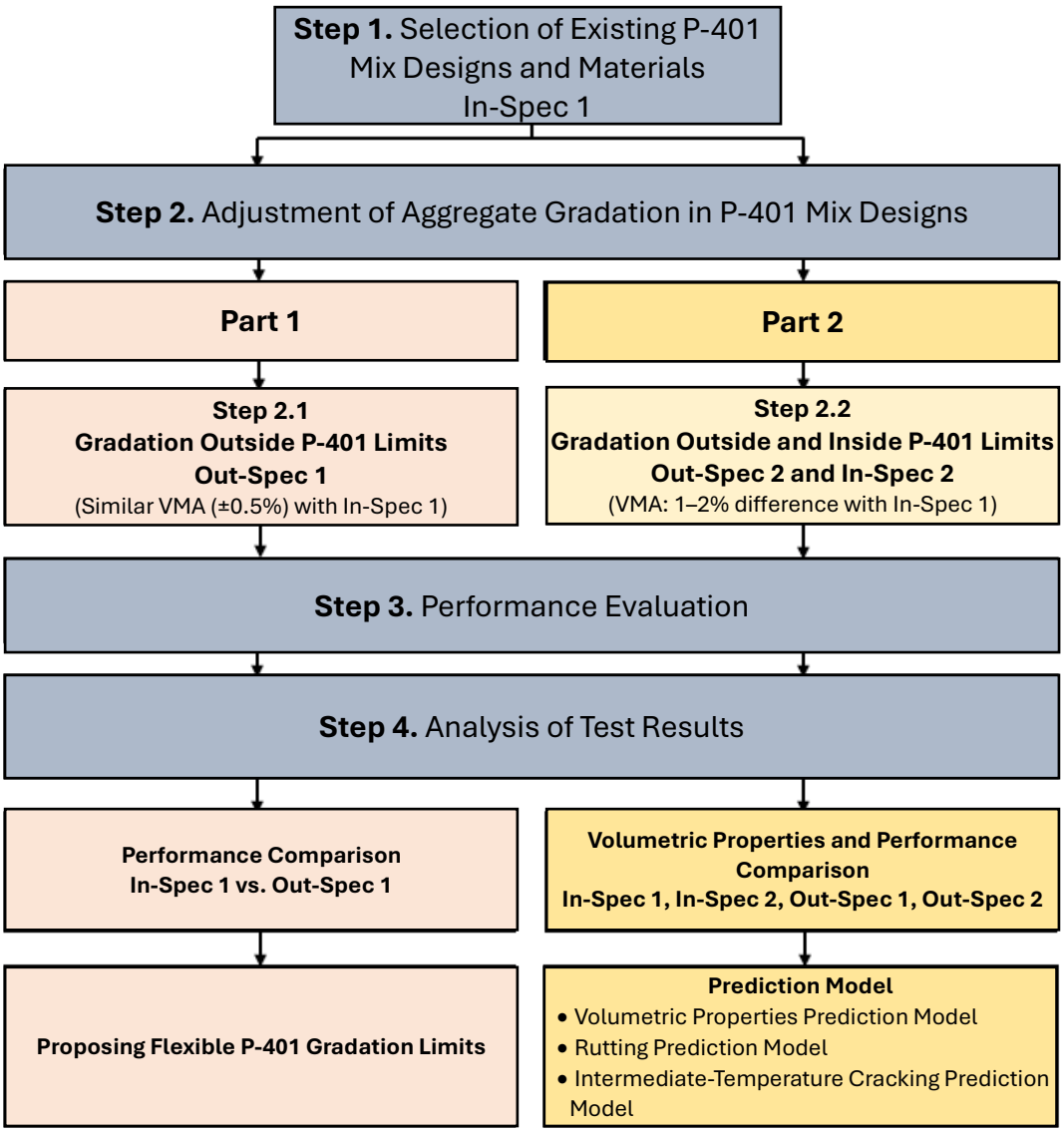
The second part examined the effects of changes in gradation on volumetric properties and laboratory performance. Adjustments were made to shift gradations further toward the inside and outside of the gradation limits while maintaining consistent binder content. The resulting data were used to develop predictive relationships quantifying how changes in gradation affect VMA and key performance indicators, including rutting and cracking resistance.

Together, these two components provide a structured framework for evaluating potential modifications to P-401 gradation limits that can increase aggregate sourcing flexibility without compromising mixture performance.

3.2 Experimental Design

The two-part experimental plan consisted of four steps, as illustrated in Figure 7. First, P-401 mix designs used for airport asphalt pavements were selected to represent a variety of aggregate sources, binder grades, and climatic conditions. The focus was on mix designs that had difficulty meeting both the current gradation limits and the VMA requirements.

The second step differed between Parts 1 and 2 of the experimental plan. For Part 1, two gradations were developed using the component materials from each selected P-401 mix design: In-Spec 1 and Out-Spec 1. The In-Spec 1 gradation met current gradation limits, whereas the Out-Spec 1 gradation was positioned outside the gradation limits. Both In-Spec 1 and Out-Spec 1 were designed to have similar VMA (±0.5 percent) and asphalt binder content (±0.2 percent), determined at a design AV of 3.5 percent, thereby isolating the influence of gradation changes on laboratory performance.



Source: AAPTP

Figure 7. Experimental Plan

For Part 2, two additional gradations were developed: In-Spec 2 and Out-Spec 2. The In-Spec 2 gradation was shifted to a more central position within the gradation limits while maintaining the same binder content as In-Spec 1, resulting in a VMA change of approximately 1 to 2 percentage points. The Out-Spec 2 gradation was positioned further outside the gradation limits, also with the same binder content, resulting in a VMA increase of approximately 1 to 2 percentage points. Maintaining consistent binder content across all four gradations (In-Spec 1, Out-Spec 1, In-Spec 2, and Out-Spec 2) enabled the isolation of the effects of gradation change on volumetric properties and laboratory performance.

Third, the mixtures were evaluated through a comprehensive laboratory testing program that included standard P-401 performance tests, as well as additional tests assessing cracking resistance, durability, and permeability.

Laboratory results were analyzed to compare the behavior of mixtures between the two gradations in Part 1 and across the four gradations in Part 2, with the aim of developing quantitative relationships linking changes in gradation to volumetric properties and performance indicators.

This integrated experimental framework provides a structured basis for evaluating the implications of gradation adjustments on airfield asphalt mixture behavior and supports data-driven consideration of potential updates to Federal Aviation Administration P-401 gradation limits. The following sections describe each step in detail.

3.3 Selection of Existing P-401 Mix Designs and Materials

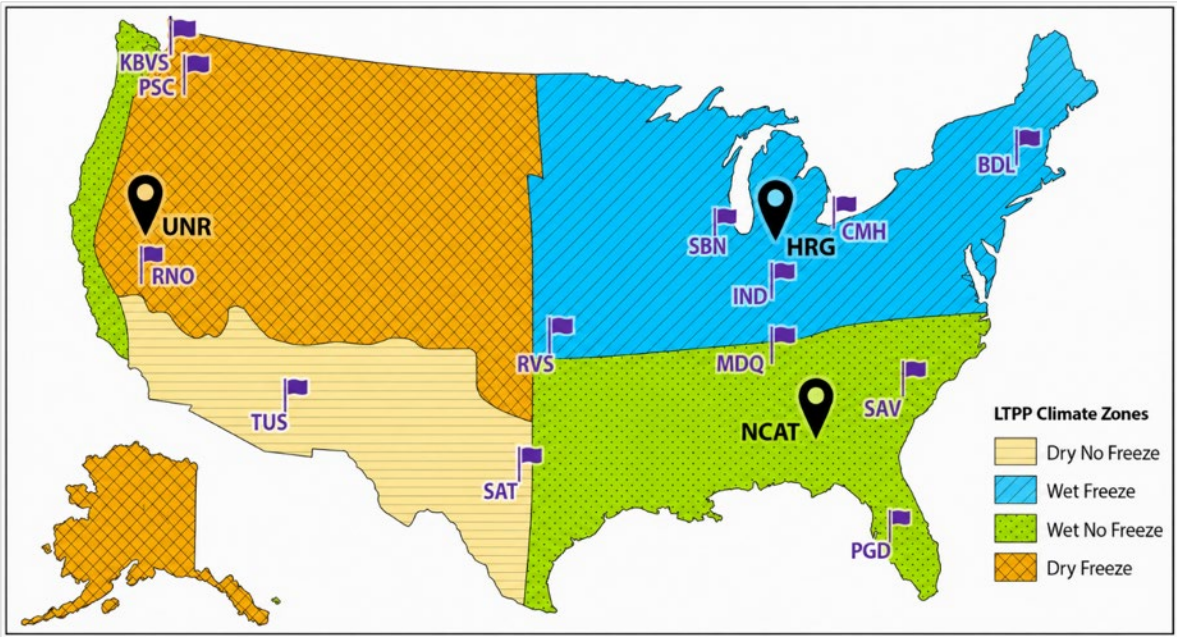
Step 1 involved selecting existing airfield P-401 mix designs for gradation adjustment and laboratory evaluation based on the following considerations:

- The selected mix designs presented challenges in meeting current P-401 gradation limits and VMA requirements. In particular, emphasis was placed on mixtures with gradations positioned near specification boundaries or with the potential to shift toward or beyond gradation limits at specific sieve sizes.
- The selected materials represented a range of aggregate sources and asphalt binder grades across the four Long-Term Pavement Performance (LTPP) program climatic zones: dry/no-freeze, dry/freeze, wet/no-freeze, and wet/freeze. This distribution ensured that the evaluation captured regional variations in materials and environmental conditions.
- Current P-401 specifications require 75 blows or gyrations for airports serving aircraft heavier than 60,000 lb and 50 blows or gyrations for airports serving lighter aircraft. To reflect more demanding design conditions, this study prioritized mix designs using 75 blows or gyrations.

Table 26 summarizes the selected mix designs, including airport location, aggregate type, gradation classification, and binder grade. Figure 8 presents the geographic distribution of the selected airports. A total of 13 P-401 mix designs were included in the evaluation. These consisted of three mix designs (KBVS, PSC, and RNO) from the dry/freeze climate zone; two mix designs (TUS and SAT) from the dry/no-freeze zone; four mix designs (RVS, MDQ, PGD, and SAV) from the wet/no-freeze zone; and four mix designs (SBN, IND, BDL, and CMH) from the wet/freeze zone. All mixtures were designed with a 12.5 mm-NMAS, except for TUS, which used a 19 mm NMAS.

Table 26. P-401 Mix Design Summary

Airport	State	Aggregate Type	P-401 Gradation Classification	Binder
San Antonio International Airport (SAT)	Texas	Basalt, limestone, and natural sand	12.5 mm NMAS-FG	PG 76-22
Huntsville Executive Airport (MDQ)	Alabama	Limestone and Couch sand	12.5 mm NMAS-FG	PG 67-22
Punta Gorda Airport (PGD)	Florida	Limestone and sand	12.5 mm NMAS-FG	PG 76-22
Tulsa Riverside Airport (RVS)	Oklahoma	Limestone	12.5 mm NMAS-FG	PG 70-28
Savannah/Hilton Head International Airport (SAV)	Georgia	Granite	12.5 mm NMAS-FG	PG 76-22
Indianapolis International Airport (IND)	Indiana	Regular limestone	12.5 mm NMAS-FG	PG 64-22
South Bend International Airport (SBN)	Indiana	Dolomitic limestone	12.5 mm NMAS-FG	PG 76-28
John Glenn Columbus International Airport (CMH)	Ohio	Limestone and natural sand	12.5 mm NMAS-FG	PG 76-22
Tucson International Airport (TUS)	Arizona	Limestone and quartzite	19 mm NMAS-FG	PG 76-22
Reno-Tahoe International Airport (RNO)	Nevada	Basalt	12.5 mm NMAS-FG	PG 64-28
Skagit Regional Airport (KBVS)	Washington	Limestone	12.5 mm NMAS-FG	PG 70-22
Tri-Cities Airport (PSC)	Washington	Limestone	12.5 mm NMAS-FG	PG 64-28NV
Bradley International Airport (BDL)	Connecticut	Traprock	12.5 mm NMAS-FG	PG 76-22



UNR = University of Nevada, Reno; HRG = Heritage Research Group; NCAT = National Center for Asphalt Technology.

Source: AAPTP

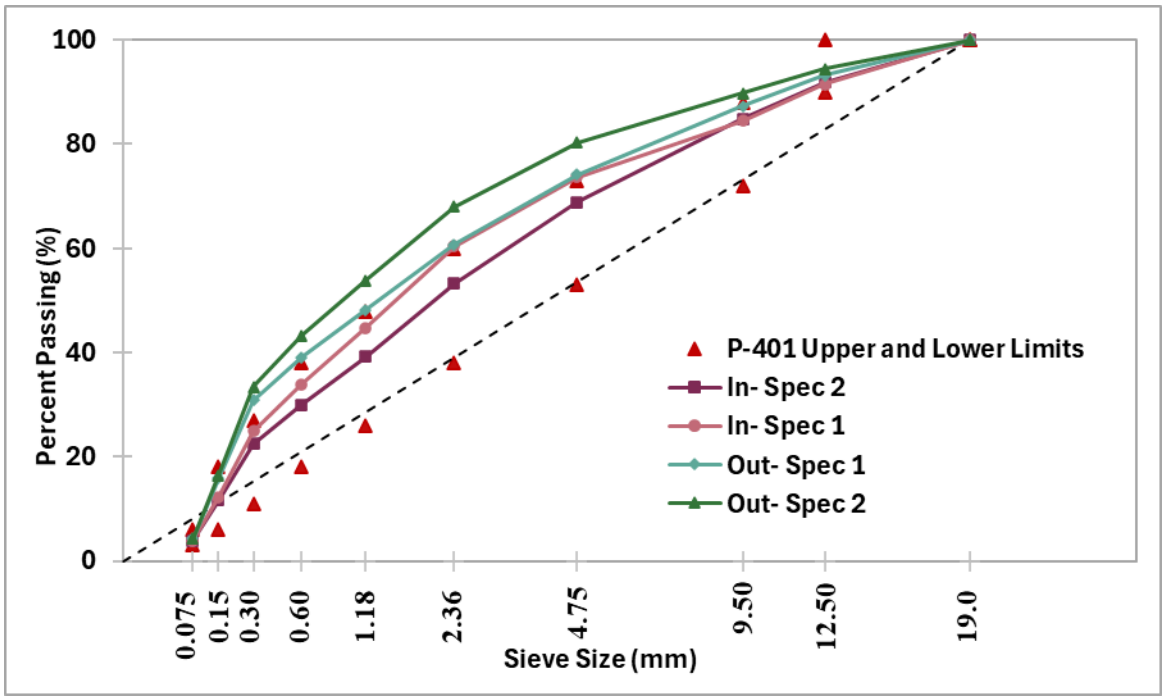
Figure 8. Geographic Locations of the Selected P-401 Mix Designs

3.4 Adjustment of Aggregate Gradation in P-401 Mix Designs

The gradation of In-Spec 1 was established as the control gradation for each selected P-401 mix design. In Part 1 of the experimental plan, the Bailey Method was used to adjust the In-Spec 1 gradation, resulting in a gradation referred to as Out-Spec 1 that fell outside the current P-401 gradation limits while maintaining similar volumetric properties. The target VMA for Out-Spec 1 was controlled within ±0.5 percentage points of In-Spec 1, and the asphalt binder content was kept within ±0.2 percent, using the same binder type. These controls were implemented to isolate the impact of gradation change on mixture performance.

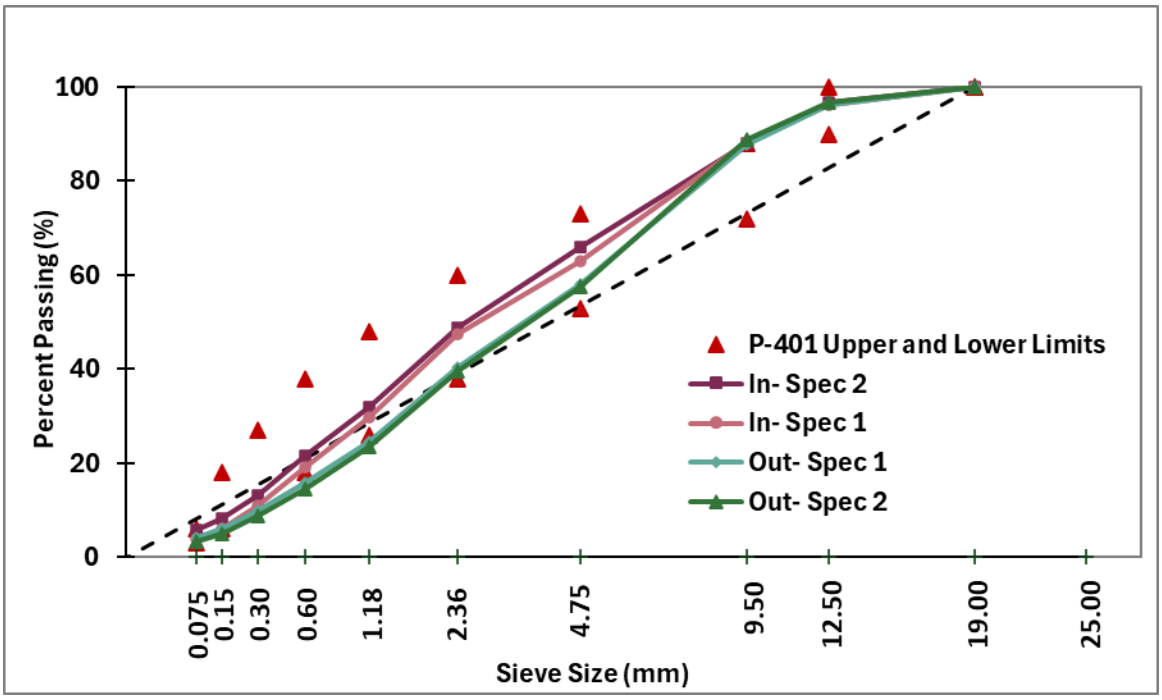
For Part 2, two additional gradations were developed. One, designated as In-Spec 2, was adjusted to a more central position within the gradation limits. The second, designated as Out-Spec 2, was shifted further outside the gradation limits than Out-Spec 1. In-Spec 2 and Out-Spec 2 were designed with the same binder content as In-Spec 1 and Out-Spec 1, respectively, to maintain consistency in binder-related effects. Relative to In-Spec 1, In-Spec 2 was developed to reduce VMA by approximately 1 to 2 percentage points, whereas Out-Spec 2 was developed to increase VMA by approximately 1 to 2 percentage points. As a result, each selected P-401 mix design generated four gradation conditions: two within the gradation limits (In-Spec 1 and In-Spec 2) and two outside the gradation limits (Out-Spec 1 and Out-Spec 2). Maintaining similar binder contents across all four gradations enabled a systematic evaluation of how gradation position and associated volumetric shifts influence performance.

Figure 9 and Figure 10 illustrate representative gradation adjustments when the In-Spec 1 gradation is positioned near the upper or lower P-401 limits. In both scenarios, shifting the gradation outside the gradation limits increased the separation between the gradation curve and the MDL. When the In-Spec 1 gradation was near the upper limit, progressive movement from In-Spec 2 to Out-Spec 2 resulted in a finer gradation. Conversely, when the In-Spec 1 gradation was near the lower limit, movement from In-Spec 2 to Out-Spec 2 resulted in a coarser gradation.



Source: AAPTP

Figure 9. Illustration of Adjustments of Gradation Close to Upper Limits



Source: AAPTP

Figure 10. Illustration of Adjustments of Gradation Close to Lower Limits

3.5 Laboratory Performance and Aggregate Blend Evaluation

A comprehensive laboratory testing program was conducted to evaluate the performance of the In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2 mixtures. The program included performance tests required by the P-401 specification, as well as additional tests to provide a broader assessment of mixture behavior.

Rutting resistance and moisture susceptibility were evaluated using the APA, HWTT, and TSR, in accordance with P-401 requirements. To further characterize cracking performance, durability, and permeability, additional laboratory tests were performed. Low-temperature cracking resistance was evaluated using the DCT test. Intermediate-temperature cracking resistance was assessed using the IDEAL-CT and I-FIT. Durability was evaluated using the Cantabro test, and permeability was measured using the Florida permeability test.

Table 27 summarizes the laboratory performance tests used in this study, including the applicable test standards, the primary performance parameters, and the target AV contents for the test specimens. Unless otherwise noted, rutting, moisture susceptibility, durability, and permeability tests were conducted at 7.0±0.5 percent AV to represent a conservative condition. Cracking tests (IDEAL-CT, I-FIT, and DCT) were performed at 5.0±0.5 percent AV to align with Airport Asphalt Pavement Technology Program (AAPTP) studies [180], representing the in-place density target and enhancing repeatability and consistency in cracking resistance characterization.

Table 27. Laboratory Performance Tests

Performance	Mixture Test	Test Standard	Test Parameter	AV (%)
Rutting	APA	AASHTO T 340 at 250 psi hose pressure	Rut depth (mm)	7.0±0.5
	HWTT	AASHTO T 324		7.0±0.5
Moisture	TSR	ASTM D4867	TSR ratio	7.0±0.5
Intermediate-Temperature Cracking	I-FIT	AASHTO T 393	FI	5.0±0.5
	IDEAL-CT	ASTM D8225	CT_{Index}	5.0±0.5
Low-Temperature Cracking	DCT	ASTM D7313	Fracture energy (J/m^2)	5.0±0.5
Durability	Cantabro	AASHTO T 401	Mass loss (%)	7.0±0.5
Permeability	Florida Permeability Test	FM 5-565	Permeability coefficient (10^{-5}cm/s)	7.0±0.5

3.5.1 Moisture Susceptibility

Moisture susceptibility was evaluated using the TSR test in accordance with ASTM D4867 [181], as shown in Figure 11. Six cylindrical specimens, 150 mm in diameter and 95±5 mm in height, were compacted to a target AV content of 7.0±0.5 percent. The specimens were divided into two groups: conditioned (wet) and unconditioned (control).



Source: AAPTP

Figure 11. Tensile Strength Ratio (TSR)

Unconditioned specimens were stored at room temperature. Prior to testing, they were immersed in a 25±1 °C water bath for 20 min. Conditioned specimens were vacuum-saturated to achieve a degree of saturation between 70 and 80 percent. For mixtures

representing freeze–thaw climates, freeze–thaw conditioning was applied. This consisted of freezing the saturated specimens at −18 °C for 16 hr, followed by thawing in a 60 °C water bath for 24 hr. After thawing, specimens were conditioned in a 25±1 °C water bath for 1 hr prior to testing.

Both conditioned and unconditioned specimens were tested for indirect tensile strength by applying a diametral compressive load at a constant deformation rate of 50 mm/min until failure. The maximum load was recorded, and the indirect tensile strength was calculated using Equation (**6**). The TSR was calculated using Equation (7). A TSR value of 80 percent was used as the minimum acceptable threshold in accordance with P-401 requirements.

$$\mathbf{S_t = \frac{2P}{\pi t D}} \tag{6}$$

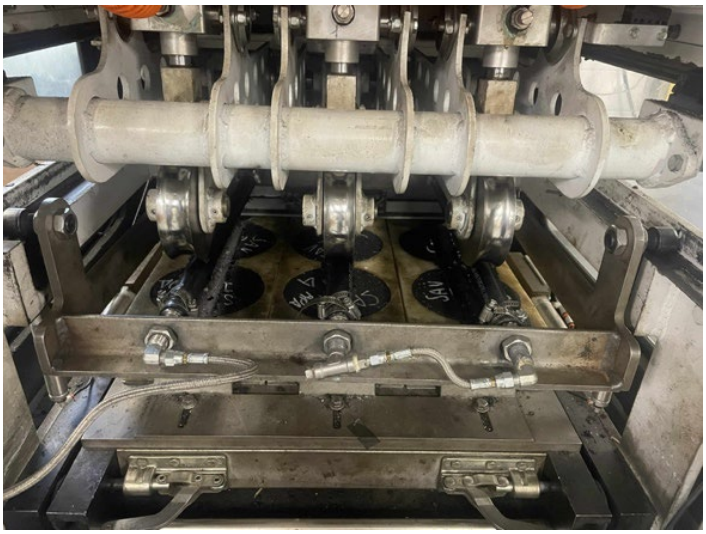
Where:

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

$$\boldsymbol{TSR = \frac{\text{Average Strength of Conditioned Specimens}}{\text{Average Strength of Unconditioned Specimens}} \times 100} \tag{7}$$

3.5.2 Rutting Resistance

Rutting resistance was evaluated using either the APA, in accordance with AASHTO T 340 [182], or the HWTT, following AASHTO T 324 [183]. For the APA test, six cylindrical specimens, 150 mm in diameter and 75±2 mm in height, were prepared and tested using a three-wheel APA device (Figure 12). Specimens were compacted to a target AV content of 7.0±0.5 percent, consistent with the protocol followed in the AAPTP Balanced Mix Design (BMD) rutting study, which represents a conservative condition. Prior to testing, specimens were conditioned at 64 °C for a minimum of 6 hr and a maximum of 24 hr. During testing, a 250 lb wheel load with a hose pressure of 250 psi was applied. The test was continued until 8,000 passes were completed or the rut depth reached 14 mm. Rut depth progression was monitored using a linear variable differential transducer. Lower rut depths indicate improved rutting resistance. In accordance with P-401 specifications, the maximum allowable rut depth is 10 mm after 4,000 passes.



Source: AAPTP

Figure 12. Asphalt Pavement Analyzer (APA)

For the HWTT (Figure 13), cylindrical specimens were also compacted to 7.0±0.5 percent AV and then cut horizontally to produce two test specimens from each compacted sample. The trimmed specimens were prepared in accordance with AASHTO T 324 [183]. For each gradation design, two sets of specimens were conditioned in a 50 °C water bath for 45 min prior to testing. A steel wheel weighing 158±1.0 lb reciprocated across the submerged specimens at a rate of 52±2 passes per minute. Testing continued until 20,000 passes were completed. Rut depth was continuously recorded during testing. In accordance with P-401 requirements, the maximum allowable rut depth is 10 mm after 20,000 passes.



Source: AAPTP

Figure 13. Hamburg Wheel-Tracking Test (HWTT)

3.5.3 Durability

Durability and resistance to raveling were evaluated using the Cantabro test in accordance with AASHTO T 401 [184], as shown in Figure 14. Cylindrical specimens 150 mm in diameter and 115±1 mm in height were compacted to a target AV content of 7.0±0.5 percent. A minimum of three replicate specimens were prepared for each gradation design.



Source: AAPTP

Figure 14. Cantabro Test

Prior to testing, specimens were conditioned in an environmental chamber at 25 °C for a minimum of 4 hr. Each specimen was then placed in the LA abrasion drum without a steel charge and rotated for 300 revolutions at a speed of 30 to 33 rpm. After completion of the rotations, the specimen was carefully removed, and the loose particles were collected.

Cantabro loss was calculated as the percentage mass loss relative to the initial specimen mass, as shown in Equation (8). Lower Cantabro loss values indicate improved resistance to raveling and enhanced mixture durability.

$$\textbf{Cantabro Loss} = \frac{\mathbf{M_{initial}} - \mathbf{M_{final}}}{\mathbf{M_{initial}}} \cdot \mathbf{100} \quad (8)$$

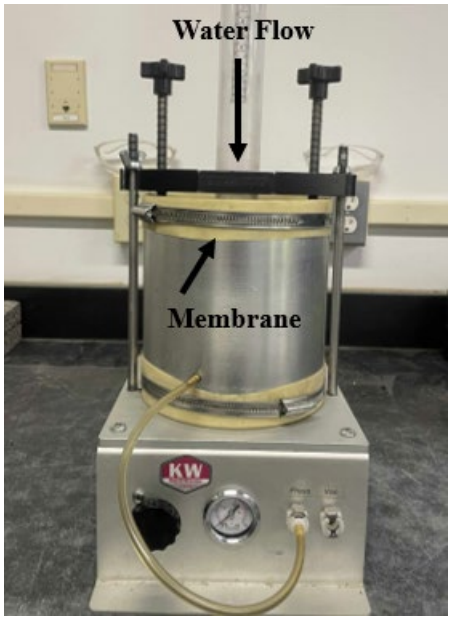
Where:

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

3.5.4 Permeability

Permeability was evaluated using the Florida permeability test in accordance with FM 5-565 [185], employing a falling-head permeability apparatus with 150-mm (6-inch) diameter cylindrical specimens, as shown in Figure 15. Each specimen was prepared by trimming

approximately 25 mm (1 inch) from both the top and bottom surfaces of the compacted sample to remove surface irregularities and ensure uniform flow conditions.



Source: AAPTP

Figure 15. Florida Permeability Test

For each gradation design, three replicate specimens were tested at a target AV content of 7.0±0.5 percent, representing a conservative condition for permeability assessment. Prior to testing, specimens were fully submerged in water for a minimum of 1 hr at ambient temperature to achieve saturation.

Following saturation, the specimen was placed on the pedestal plate and assembled with the upper chamber, which included a graduated cylinder and a sealing membrane. The latex membrane was inflated to seal the specimen perimeter, ensuring that water flow was restricted to the vertical direction. Water was then introduced into the graduated cylinder to a level above the upper timing mark and allowed to flow through the specimen. The time required for the water level to drop between two designated head levels was recorded.

The coefficient of permeability, k, was calculated using Darcy's law for a falling-head test, as shown in Equation (9). A reference temperature of 20 °C (68 °F) was used for viscosity correction. Higher values of k indicate greater mixture permeability.

$$\mathbf{k} = \frac{\mathbf{aL}}{\mathbf{At}} \cdot \mathbf{ln}\frac{\mathbf{h_1}}{\mathbf{h_2}} \cdot \mathbf{t_c} \quad (9)$$

Where:

k = coefficient of permeability (cm/s)
a = inside cross-sectional area of the buret (cm^2)
L = average thickness of the specimen (cm)
A = cross-sectional area of the specimen (cm^2)
t = elapsed time between head measurements (s)
h_1 = initial head (cm)
h_2 = final head (cm)
t_c = temperature correction factor for water viscosity

3.5.5 Intermediate-Temperature Cracking Resistance

Intermediate-temperature cracking resistance was evaluated using two tests: the IDEAL-CT in accordance with ASTM D8225 [186], as shown in Figure 16, and the I-FIT following AASHTO T 393 [187], as shown in Figure 17.

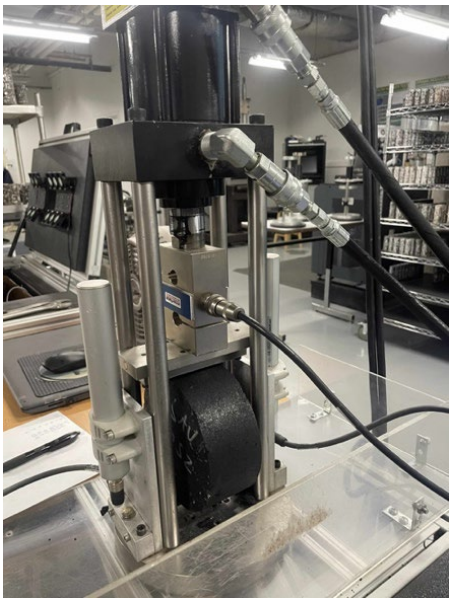
The IDEAL-CT test characterizes cracking resistance by calculating the CT_{Index}, which is derived from the load-displacement response under monotonic loading. Cylindrical specimens 150 mm in diameter and 62±2 mm in height were compacted to a target AV content of 5.0±0.5 percent. This AV level was selected to improve test repeatability and reduce variability, consistent with findings from the AAPTP BMD cracking study. A minimum of four replicate specimens were prepared for each gradation design.

Prior to testing, specimens were conditioned at 25±1 °C. A diametral compressive load was applied at a constant displacement rate of 50 mm/min until complete fracture occurred. Load and displacement were recorded continuously during the test. The CT_{Index} was calculated using Equation (10). Higher CT_{Index} values indicate improved resistance to intermediate-temperature cracking.

$$\mathbf{CT_{Index}} = \frac{\boldsymbol{t}}{\mathbf{62}} * \frac{\boldsymbol{l_{75}}}{\mathbf{D}} * \frac{\mathbf{G_f}}{|\boldsymbol{m_{75}}|} * \mathbf{10^6} \quad (10)$$

Where:

$\mathrm{CT_{Index}}$= Cracking Tolerance Index
$|G_f|$ = fracture energy (J/m^2)
$|m_{75}|$ = absolute value of the post-peak slope m_{75} (N/m)
l_{75} = displacement at 75% of the peak load after the peak (mm)
D = specimen diameter (mm)
t = specimen thickness (mm)



Source: AAPTP

Figure 16. Indirect Tensile Asphalt Cracking Test (IDEAL-CT)

For the I-FIT procedure, a cylindrical specimen 150 mm in diameter and 160±1 mm in height was cut horizontally to produce two disks approximately 50±1 mm thick. Each disk was then cut vertically along the diameter to create two SCB specimens. A notch 15±0.5 mm deep was introduced at the midpoint of the flat edge of each SCB specimen to promote controlled crack initiation.

Similar to the IDEAL-CT specimens, SCB specimens were compacted to a target AV content of 5.0±0.5 percent to improve repeatability. A minimum of four replicate specimens were tested for each gradation design. Prior to testing, specimens were conditioned at 25±1 °C for at least 2 hr.

Each notched SCB specimen 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 surface directly above the notch at a constant displacement rate of 50 mm/min until fracture occurred. Load-displacement data were recorded continuously. FI was calculated using Equation (11). Higher FI values indicate greater flexibility and improved resistance to crack propagation at intermediate temperatures.

$$\mathbf{FI} = \frac{\mathbf{G_f} * \mathbf{A}}{|\mathbf{m}|} \tag{11}$$

Where:

G_f = fracture energy (J/mm^2)
$|m|$ = absolute value of the post-peak slope (N/mm)
A = empirically derived calibration constant

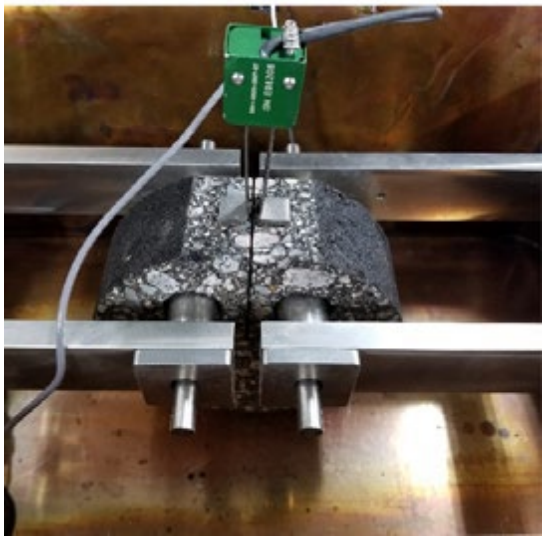


Source: AAPTP

Figure 17. Illinois Flexibility Index Test (I-FIT)

3.5.6 Low-Temperature Cracking Resistance

Low-temperature cracking resistance was evaluated using the DCT test in accordance with ASTM D7313 [188], as shown in Figure 18. The DCT test quantifies fracture energy under monotonic tensile loading and provides an indication of resistance to thermal cracking.



Source: AAPTP

Figure 18. Disk-Shaped Compact Tension (DCT) Test

A cylindrical specimen 150 mm in diameter and 160±1 mm in height was cut horizontally to produce two disk-shaped specimens approximately 50±1 mm thick. A notch was introduced at the center of one edge to control crack initiation, and two loading holes were drilled near the top of the specimen to allow load application through steel pins (Figure 19). A minimum of four replicate specimens were tested for each gradation design.



Source: AAPTP

Figure 19. DCT Sample

Consistent with the intermediate-temperature cracking tests, specimens were compacted to a target AV content of 5.0±0.5 percent to improve repeatability and reduce variability. Prior to testing, specimens were conditioned for 2 to 8 hr at a temperature 10 °C above the low-temperature PG of the asphalt binder.

During testing, the specimen was mounted in the DCT fixture, and a monotonic tensile load was applied through steel loading pins at a constant displacement rate of 0.017 mm/s. Load and crack mouth opening displacement (CMOD) were recorded continuously. Fracture energy was calculated by dividing the total work of fracture (area under the load–CMOD curve) by the ligament area (area of the uncracked section ahead of the notch), as shown in Equation (12). Higher fracture energy values indicate improved resistance to low-temperature cracking.

$$G_f = \frac{AREA}{B(W - a)} \quad (12)$$

Where:

G_f = fracture energy (J/m^2)
AREA = area under load–CMOD$_{fit}$ curve
B = specimen thickness (m)
W – *a* = initial ligament length (m)

3.5.7 Fine Aggregate Angularity Test

FAA was evaluated in accordance with AASHTO T 304-22 [189], Method A, to determine the uncompacted void content of the fine aggregate fraction of each blend. Method A was selected because it measures void content under standardized conditions that reflect particle shape, angularity, and surface texture.

The aggregate blend was first batched according to the design stockpile proportions and washed over the No. 200 sieve to remove dust and fines. The washed material was oven-dried and sieved into individual size fractions in accordance with AASHTO T 27 [190]. Four size fractions were collected and proportioned as follows:

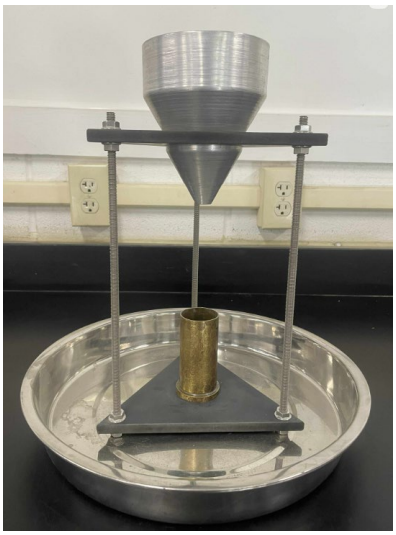
- 2.36 mm (No. 8) to 1.18 mm (No. 16): 44 g
- 1.18 mm (No. 16) to 600 µm (No. 30): 57 g
- 600 µm (No. 30) to 300 µm (No. 50): 72 g
- 300 µm (No. 50) to 150 µm (No. 100): 17 g

The four fractions were combined and thoroughly mixed to obtain a 190±0.2 g test sample. The blended sample was allowed to flow freely through a standardized funnel from a fixed height into a calibrated 100-mL cylindrical measure without vibration or compaction, as shown in Figure 20. After filling, excess material was struck off in a single smooth motion using a straight-edged spatula held horizontally. The mass of the filled measure was recorded. The procedure was repeated, and the average of the two measurements was reported. The uncompacted void content (FAA value) was calculated using Equation (13). Higher uncompacted void content values indicate more angular and rough-textured particles, whereas lower values indicate more rounded and smoother particles.

$$U = \frac{(V - F/G)}{V} \times 100 \quad (13)$$

Where:

U = uncompacted void content (percent)
V = volume of the cylindrical measure (mL)
F = net mass of fine aggregate in the measure (g)
G = bulk dry specific gravity of the fine aggregate fraction



Source: AAPTP

Figure 20. Fine Aggregate Angularity (FAA) Test

3.6 Analysis of Test Results

For Part 1 of the experimental plan, laboratory performance results for In-Spec 1 and Out-Spec 1 were compared to evaluate the effect of shifting the gradation outside the current P-401 gradation limits while maintaining similar volumetric properties and binder content. This direct comparison isolated the influence of gradation position relative to the gradation limits. If Out-Spec 1 demonstrated performance comparable to or better than the corresponding In-Spec 1, the findings would support consideration of increased flexibility in the P-401 gradation limits while maintaining compliance with existing volumetric and performance requirements.

For Part 2 of the experimental plan, volumetric properties and performance results were evaluated across the four gradation designs: In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2. The analysis focused on quantifying the effects of systematic gradation shifts on VMA and key performance indicators, including rutting resistance and cracking resistance. Comparisons were made both within individual mix designs and across the full dataset to identify consistent trends.

Based on the observed relationships, predictive models were developed to characterize the effects of gradation changes on mixture behavior. These models were not intended to predict absolute values of VMA or performance for a given mixture. Rather, they were formulated to estimate relative changes in volumetric properties and laboratory performance as a function of gradation differences and aggregate characteristics. The modeling framework included a volumetric prediction component and performance prediction components focused on rutting resistance and intermediate-temperature cracking.

Chapter 4. Development of Proposed Revisions to P-401 Gradation Limits

This chapter presents the results of Part 1 of the experimental program and evaluates potential revisions to the P-401 gradation limits for 12.5 mm-NMAS mix design. The analysis compares the laboratory performance of In-Spec 1 and Out-Spec 1, which were both designed with comparable volumetric properties and asphalt binder contents. Based on these results, proposed modifications to the gradation limits were developed.

The chapter is organized into four sections. Section 4.1 summarizes the execution of Part 1 of the experimental plan. Section 4.2 presents the summary of the mixture designs. Section 4.3 provides test results and statistical analysis. Section 4.4 discusses proposed revisions to the P-401 gradation limits.

4.1 Execution of Part 1 of Experimental Plan

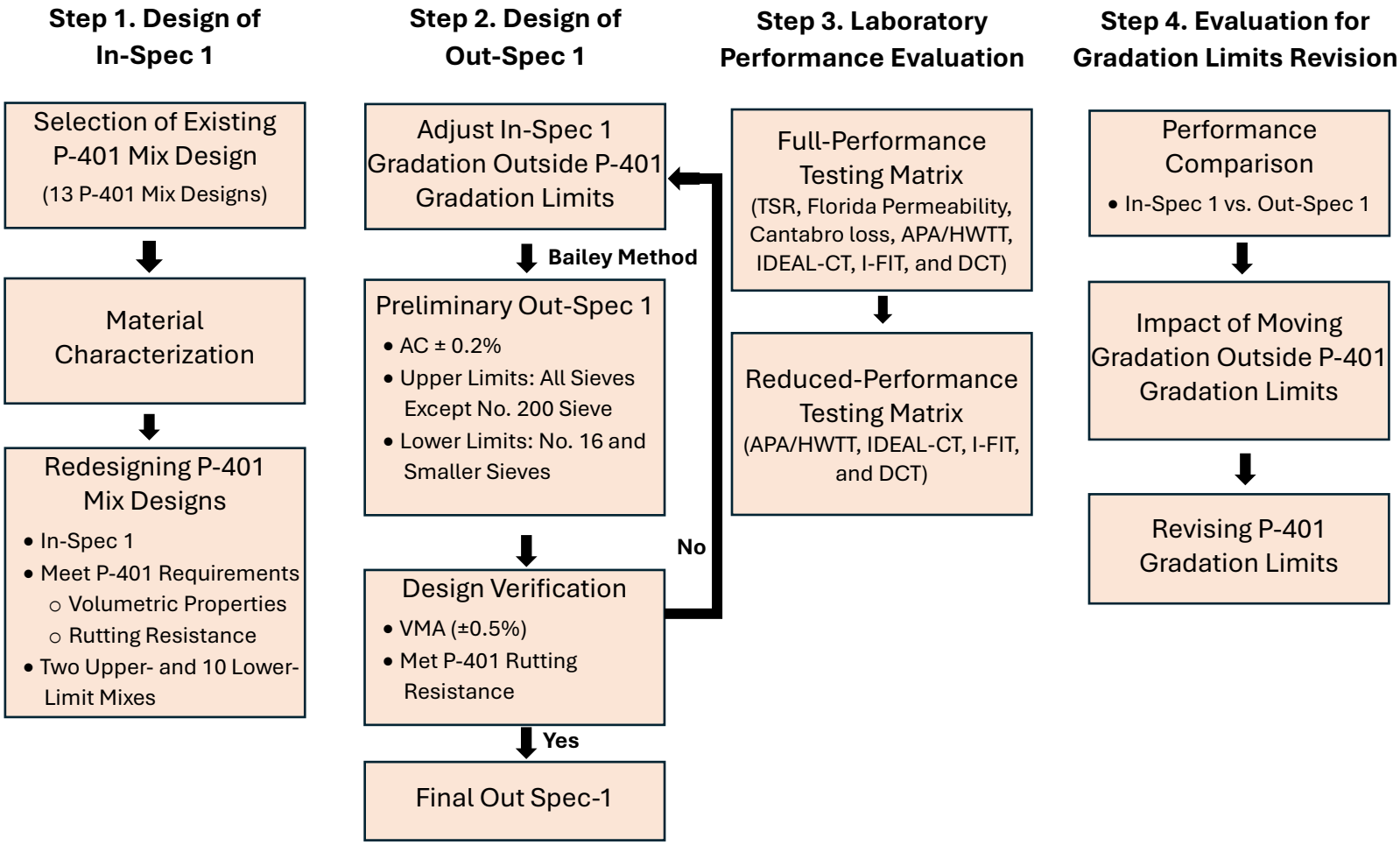
Figure 21 summarizes the execution of the Part 1 experimental program. Detailed mixture design procedures and testing methods are described in Chapter 3. This section highlights the specific design decisions and evaluation steps used to assess potential revisions to the P-401 gradation limits.

4.1.1 Step 1: Design of In-Spec 1

Raw materials from each selected P-401 mix design were sampled and characterized prior to design. Because aggregate gradation and bulk specific gravity may vary over time due to stockpile variability, the measured properties did not necessarily match the original job mix formula (JMF). Therefore, a baseline mixture, designated as In-Spec 1, was developed for each selected mix design using measured aggregate properties while fully complying with current P-401 volumetric and performance requirements.

When the original JMF gradation was positioned near the upper or lower gradation limits, the In-Spec 1 gradation was maintained near those same limits. When the JMF gradation was centered within the gradation limits, the In-Spec 1 gradation was intentionally shifted toward either the upper or lower limits. The direction of this shift was guided by engineering judgment, primarily that the resulting VMA should not be excessively different from that of the JMF mix design. This approach positioned the mixtures near gradation limits to enable controlled adjustment outside the gradation limits in subsequent steps.

To maintain confidentiality and avoid associating performance outcomes with specific airports, P-401 mix designs were anonymized and labeled as Mix 1 through Mix 13. Two mix designs (Mix 1 and Mix 2) were positioned near the upper limits, and the remaining 11 mix designs (Mix 3 through Mix 13) were positioned near the lower gradation limits. These adjustments were made based on the individual stockpile gradations for each mix design and the need to maintain similar VMA values for In-Spec 1 and Out-Spec 1.



AC = asphalt content.

Source: AAPTP

Figure 21. Experimental Design (Part 1) for Revising P-401 Gradation Limits

4.1.2 Step 2: Design of Out-Spec 1

For each In-Spec 1 gradation, a corresponding Out-Spec 1 gradation was developed by shifting the gradation outside the current P-401 gradation limits while maintaining comparable volumetric properties. Gradation adjustments were performed using the Bailey Method, as described in Section 2.6, Chapter 2.

For upper-limit mix designs, adjustments targeted the upper gradation limits at most sieve sizes (except the No. 200 sieve). For lower-limit mix designs, adjustments focused primarily on the No. 16 and smaller sieves while avoiding coarsening at the No. 8 and larger sieves. This constraint limited the potential increase in interconnected AV and helped control permeability and durability [191].

Each preliminary Out-Spec 1 was verified through laboratory measurements. The design process was repeated until the measured VMA was within ±0.5 percent of the corresponding In-Spec 1 and satisfied the P-401 rutting requirement. Asphalt binder content was maintained within ±0.2 percent of the In-Spec 1 design to isolate the influence of gradation. Once these criteria were satisfied, the mixture was accepted as the final Out-Spec 1 design.

4.1.3 Step 3: Laboratory Performance Evaluation

Both In-Spec 1 and Out-Spec 1 mixtures were evaluated using the testing program described in Chapter 3. Initially, the full testing matrix (Table 27) was applied to six airport mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11). After preliminary analysis indicated no meaningful differences in TSR, permeability, and Cantabro mass loss between In-Spec 1 and Out-Spec 1, these tests were discontinued for the remaining mix designs.

The reduced testing matrix focused on performance indicators most sensitive to gradation adjustments, including:

- Rutting resistance (APA or HWTT based on equipment availability)
- Intermediate-temperature cracking (IDEAL-CT and I-FIT)
- Low-temperature cracking (DCT)

This refinement improved testing efficiency while preserving the ability to detect gradation-related performance effects.

4.1.4 Step 4: Evaluation for Gradation Limits Revision

Laboratory results from In-Spec 1 and Out-Spec 1 were compared to assess the impact of shifting gradation outside the current P-401 gradation limits. Results are presented using plots with error bars representing ±1 standard deviation. Because mixtures were designed with comparable VMA and binder contents, observed performance differences were attributed primarily to gradation changes.

A Student's t-test was conducted at a 0.05 significance level to evaluate statistical differences between paired mixtures. Additionally, the performance test results for these paired mixtures were compared with the allowable difference between the two test results, known as the single-operator d2s limit [192], based on the corresponding standard test methods or thresholds adopted by agencies. The findings from these comparisons were used to propose potential revisions to the gradation limits for the P-401 mixture. This evaluation included twelve 12.5-mm NMAS mixtures, consisting of two upper-limit designs and 10 lower-limit designs. Based on results of the evaluation, the Out-Spec 1 gradation, defined by sieve sizes and percentage passing values outside the current limits, can serve as the basis for proposing revised upper and lower P-401 gradation limits.

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

Table 28 summarizes the aggregate gradations, FAA, asphalt binder contents, and volumetric properties of the 13 In-Spec 1 and Out-Spec 1 pairs. Sieve sizes of Out-Spec 1 that fall outside the current P-401 gradation limits are highlighted.

Across all mixture pairs, the Out-Spec 1 gradations were positioned outside the gradation limits at selected sieve sizes while maintaining similar volumetric properties. Binder contents were held within ±0.2 percent of the corresponding In-Spec 1. AV content was at the target value of 3.5 percent.

The VMA values of Out-Spec 1 were within ±0.5 percent of the corresponding In-Spec 1, which aligns with the volumetric control strategy described in Section 4.1. All mixtures met the minimum VMA requirement of 15.0 percent for 12.5-mm NMAS and 14.0 percent for 19-mm NMAS, as shown in Table 28. Additionally, the effective binder contents (V_{be}) were comparable between the paired mixtures.

FAA values did not show systematic differences between In-Spec 1 and Out-Spec 1 gradations, indicating that the gradation adjustments did not substantially alter FAA characteristics.

Table 28. Design Summary of In-Spec 1 and Out-Spec 1 (Mixes 1 through 6)

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
No. 4	73	74	67	66	60	56	63	58	69	66	53–73	59	53	45–67
No. 8	60	61	55	55	45	46	47	40	47	43	38–60	39	34	32–54
No. 16	45	48	45	46	30	28	30	25	30	27	26–48	24	21	22–44
No. 30	34	39	37	39	21	18	19	16	19	17	18–38	16	14	15–35
No. 50	25	31	27	30	13	10	11	10	11	9	11–27	10	8	9–25
No. 100	12	16	14	16	7	6	6	6	7	6	6–18	7	6	6–18
No. 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
AV (%)	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
V_{be} (%)	12.0	12.1	13.2	13.0	12.7	13.2	11.7	11.6	13.0	13.3	–	11.1	11.1	–

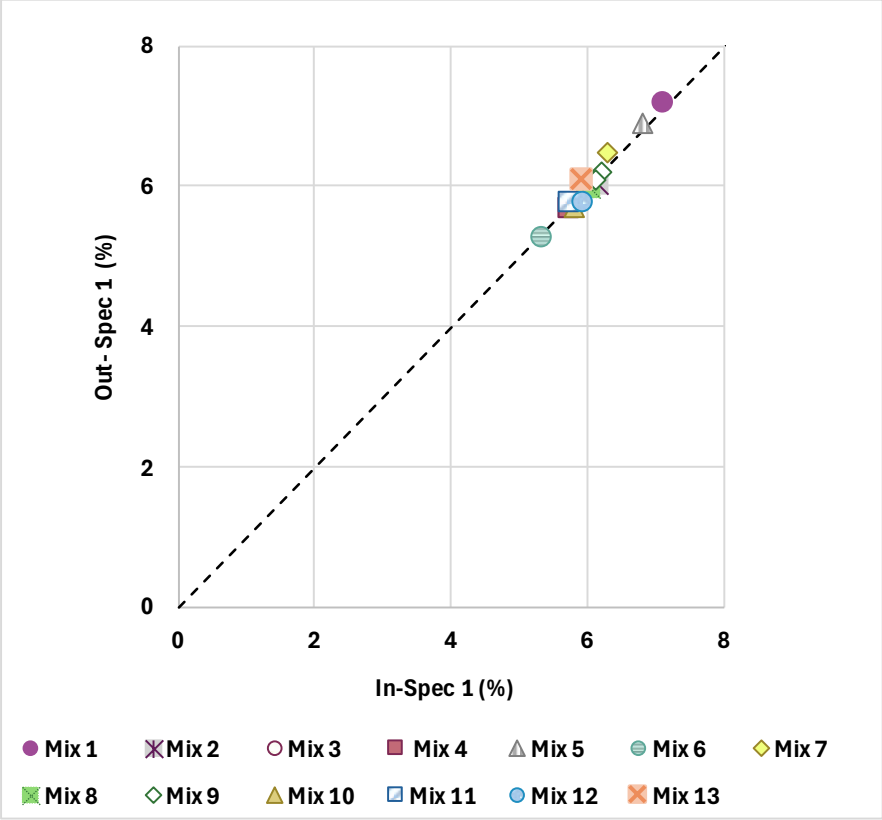
AC = asphalt content; IS 1 = In-Spec 1; OS 1 = Out-Spec 1

Table 28. Design Summary of In-Spec 1 and Out-Spec 1 (Mixes 7 through 13)

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
No. 4	58	58	53	53	65	60	69	62	61	57	55	54	66	61	53–73
No. 8	38	38	40	39	43	41	49	42	42	38	39	38	43	38	38–60
No. 16	26	25	29	23	27	25	32	26	28	25	26	25	30	25	26–48
No. 30	18	17	21	16	18	16	21	16	18	16	18	17	20	17	18–38
No. 50	11	10	14	11	12	10	13	10	11	10	13	12	13	12	11–27
No. 100	7	7	9	9	8	7	6	5	6	6	8	8	9	9	6–18
No. 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
V_{be} (%)	11.8	11.6	11.8	11.7	11.9	11.9	12.0	12.0	11.5	11.6	13.5	13.5	12.2	12.3	–

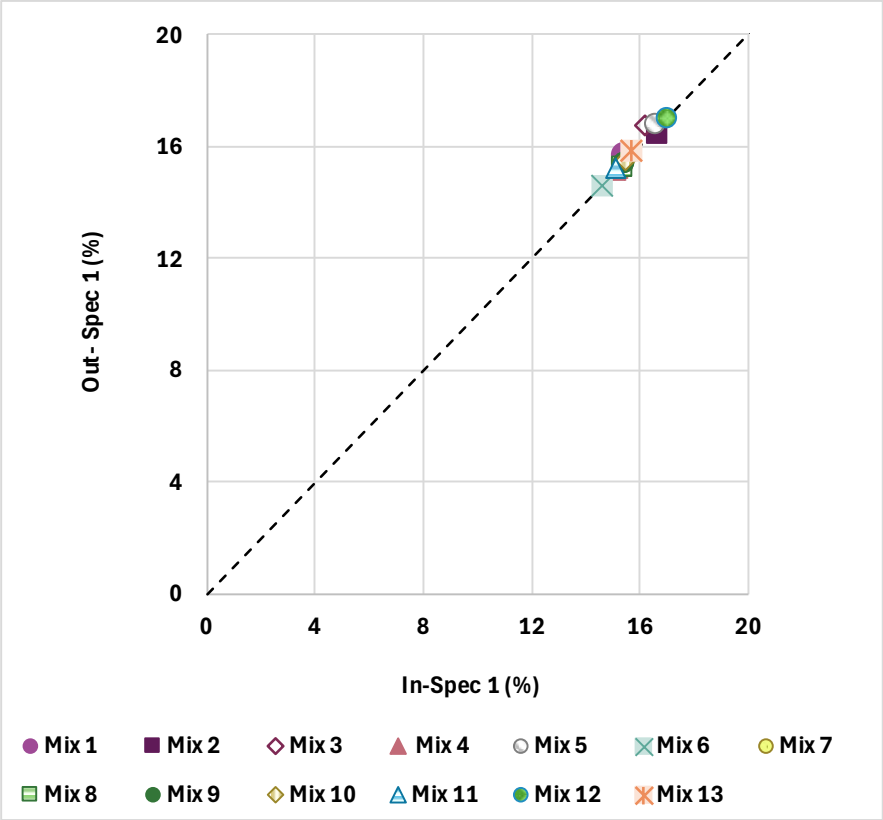
AC = asphalt content; IS 1 = In-Spec 1; OS 1 = Out-Spec 1

Figure 22 through Figure 24 present comparisons of asphalt binder content, VMA, and V_{be} for the corresponding In-Spec 1 and Out-Spec 1. Data points cluster closely along the equality line, confirming that volumetric properties were effectively controlled between paired designs. This similarity supports subsequent performance comparisons by minimizing the influence of binder content and volumetric differences.



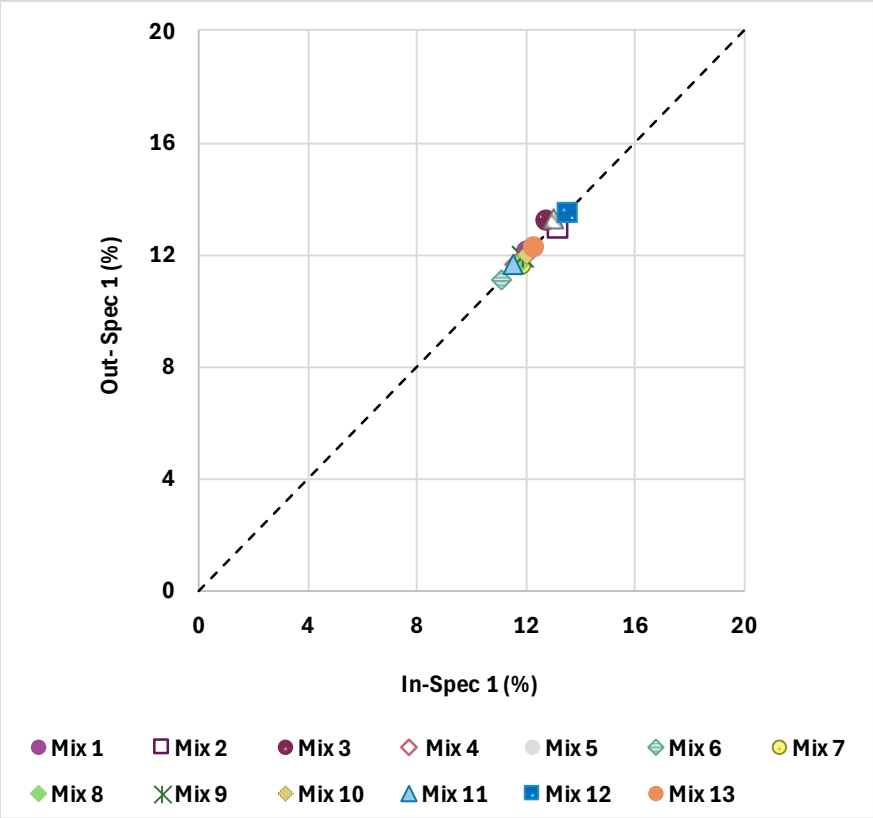
Source: AAPTP

Figure 22. Asphalt Binder Contents of In-Spec 1 and Out-Spec 1



Source: AAPTP

Figure 23. VMA of In-Spec 1 and Out-Spec 1



Source: AAPTP

Figure 24. V_{be} of In-Spec 1 and Out-Spec 1

4.3 Test Results and Data Analysis

This section presents laboratory performance results for In-Spec 1 and Out-Spec 1 and evaluates the effect of shifting gradation outside the current P-401 band. TSR, Florida permeability, and Cantabro loss were evaluated for six airport mix designs, while rutting (APA/HWTT) and cracking (IDEAL-CT, I-FIT, and DCT) were evaluated for all 13 airport mixtures.

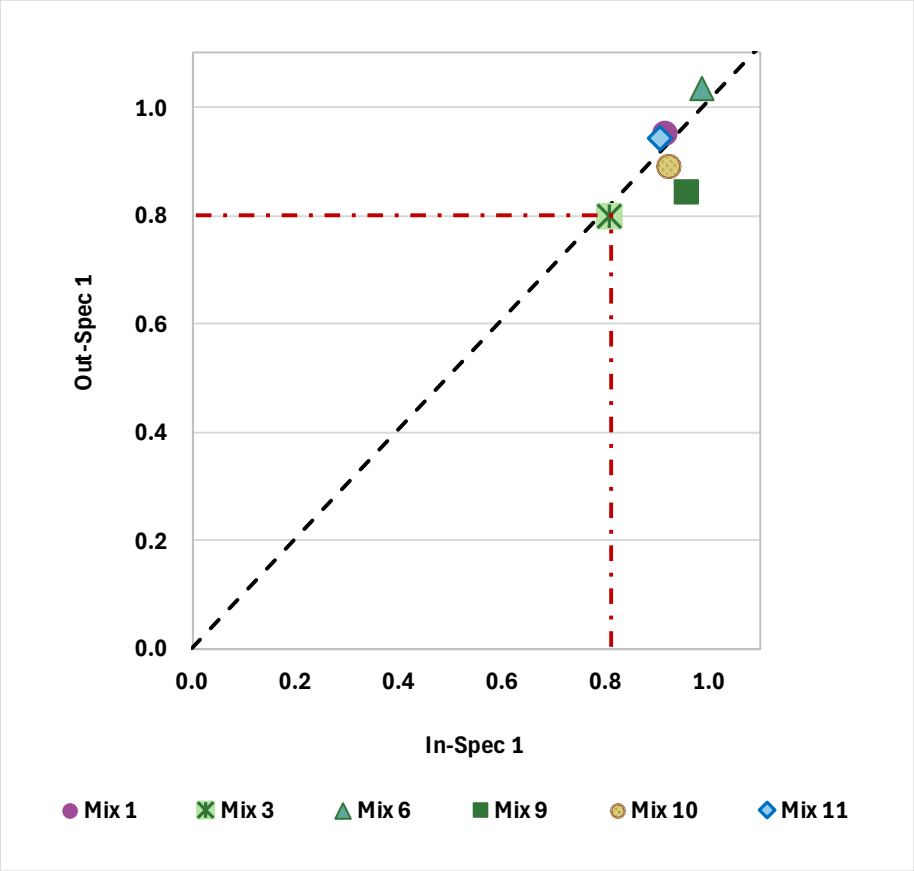
Results are presented in plots with error bars of ±1 standard deviation. A Student's t-test at a significance level of 0.05, together with comparison to the corresponding d2s limit, was conducted to assess both statistical and practical differences between paired mixtures. These comparisons provide the basis for evaluating potential revisions to the P-401 gradation limits.

4.3.1 Moisture Susceptibility Evaluation

Figure 25 presents TSR results for six mixture pairs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11), and Table 29 summarizes the TSR differences between In-Spec 1 and Out-Spec 1. All mixtures met the P-401 moisture resistance requirement, with TSR values exceeding 0.80. Since a single TSR value was determined for each mixture, differences in TSR values

between paired mixtures were assessed using the single-operator d2s limit (i.e., 0.093 [193]). For five of the six mix designs, the differences in TSR values between the paired mixtures were within the single-operator d2s limit. Only Mix 9 showed a difference exceeding the single-operator d2s limit.

Overall, shifting gradation outside the current specification limits did not produce systematic changes in moisture resistance. Because these mixtures were FG and moisture susceptibility is primarily governed by binder characteristics and aggregate–binder interaction rather than minor gradation adjustments [194, 195], this outcome was expected. Given the absence of practical differences, TSR testing was discontinued for the remaining mix designs.



Source: AAPTP

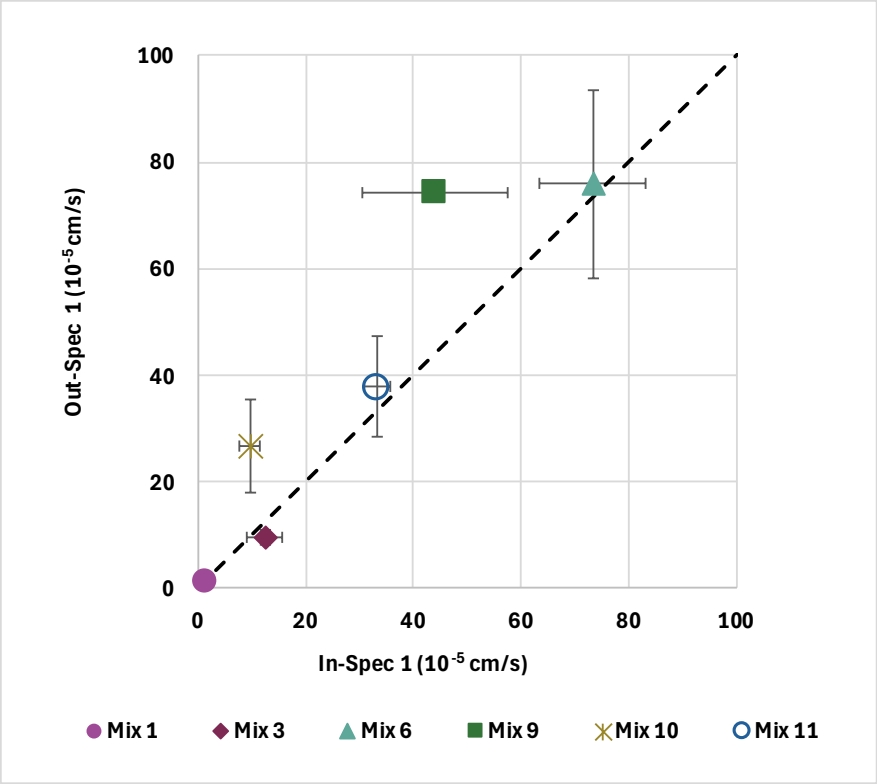
Figure 25. TSR Test Results of In-Spec 1 and Out-Spec 1

Table 29. TSR Difference Results

Mix ID	d2s = 0.093	Exceeding the Single-Operator d2s Limit?
In-Spec 1 vs. Out-Spec 1 (Upper Specification Limits)		
Mix 1	0.03	No
In-Spec 1 vs. Out-Spec 1 (Lower Specification 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 26 presents permeability results for six mix designs (Mix 1, Mix 3, Mix 6, Mix 9, Mix 10, and Mix 11), and Table 30 summarizes the corresponding p-values. Out-Spec 1 mixtures generally exhibited slightly higher permeability than their In-Spec 1 counterparts. Statistical analysis showed no significant differences ($p > 0.05$) for four of the six mix designs, while Mix 9 and Mix 10 exhibited statistically significant differences ($p < 0.05$). However, all measured permeability coefficients were below a commonly used threshold of 125×10^{-5} cm/s, indicating that the mixtures remained effectively impermeable [196]. Overall, shifting the gradation outside the current specification limits did not cause permeability issues. Given these findings, permeability testing was discontinued for the remaining mix designs.



Source: AAPTP

Figure 26. Permeability Test Results of In-Spec 1 and Out-Spec 1

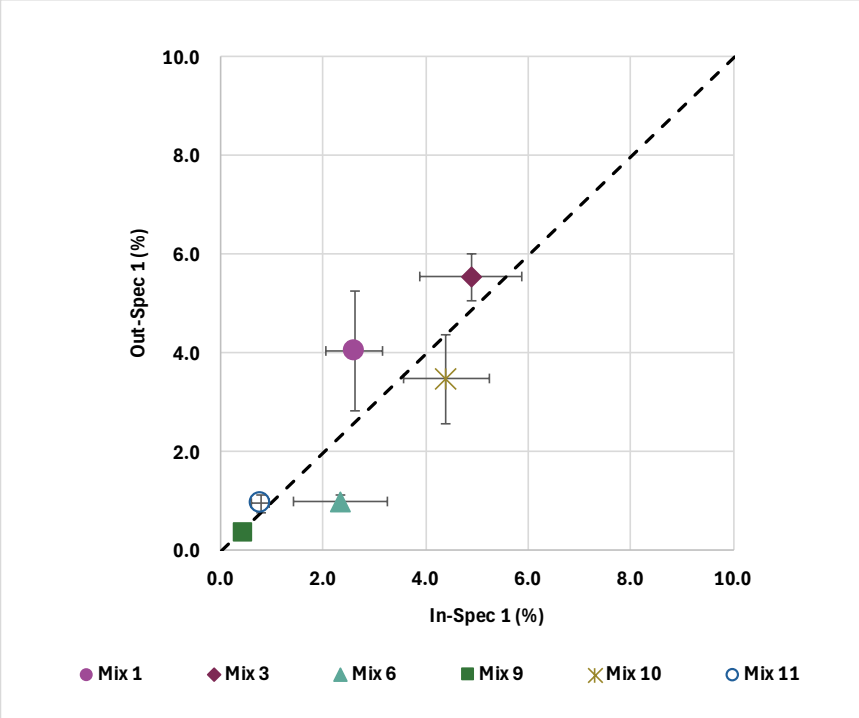
Table 30. P-Values of Permeability Test Results

Mix ID	Observed Difference	p-value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification Limits)			
Mix 1	−0.3	0.30	No
In-Spec 1 vs. Out-Spec 1 (Lower Specification 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 27 presents Cantabro mass loss results for six mix designs, and Table 31 summarizes the associated p-values. Five of the six mix designs showed no significant differences between In-Spec 1 and Out-Spec 1 ($p > 0.05$). Mix 9 was the only mix design with a statistically significant difference ($p < 0.05$); however, the magnitude of the difference was minimal (0.1 percent) and not considered practically significant. Additionally, all mixtures exhibited Cantabro mass loss values below the 6.3-percent threshold adopted by the Virginia Department of Transportation [197], indicating that

gradation shifts outside the current P-401 limits did not adversely affect mixture durability. Based on these findings, Cantabro testing was discontinued for the remaining mix designs.



Source: AAPTP

Figure 27. Cantabro Mass Loss of In-Spec 1 and Out-Spec 1

Table 31. P-Values of Cantabro Mass Loss

Mix ID	Observed Difference	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification Limits)			
Mix 1	1.4	0.16	No
In-Spec 1 vs. Out-Spec 1 (Lower Specification Limits)			
Mix 3	0.7	0.22	No
Mix 6	1.5	0.10	No
Mix 9	0.1	0.02	Yes
Mix 10	0.9	0.25	No
Mix 11	0.2	0.30	No

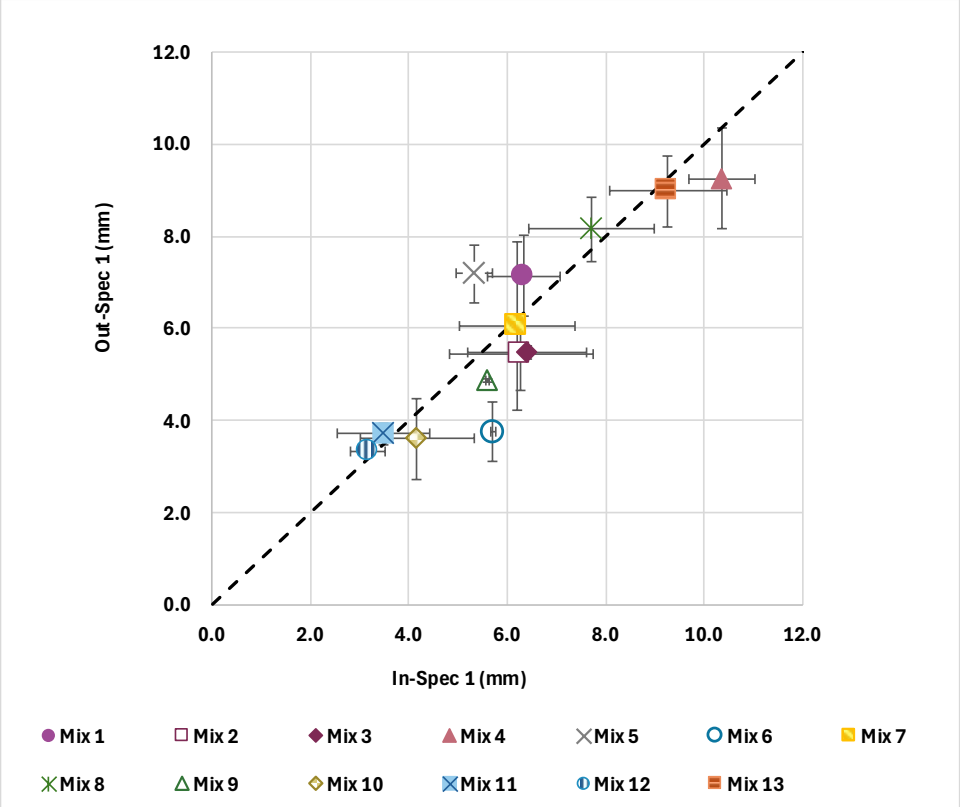
4.3.4 Rutting Resistance

Figure 28 presents rut depth results for 13 mixture pairs, and Table 32 summarizes the corresponding p-values. Nine mix designs were evaluated using the APA test and four using the HWTT. Both tests are allowed in the P-401 specifications.

All Out-Spec 1 mixtures satisfied the P-401 rutting resistance criteria, with rut depths less than 10 mm after 4,000 passes in the APA test or 20,000 passes in the HWTT. For eight of

the 13 mix designs, Out-Spec 1 exhibited lower rut depths than the corresponding In-Spec 1. However, 12 mix designs showed no significant differences between In-Spec 1 and Out-Spec 1 ($p > 0.05$). Mix 9 was the only mix design that showed a statistically significant difference; however, the absolute rut-depth difference was approximately 0.7 mm, which was not considered practically significant in the context of specification compliance.

Overall, shifting the gradation outside the current P-401 limits generally did not result in statistically or practically significant changes in rutting resistance. These results indicate that gradation adjustments, when volumetric properties are controlled, do not adversely affect rutting performance.



Source: AAPTP

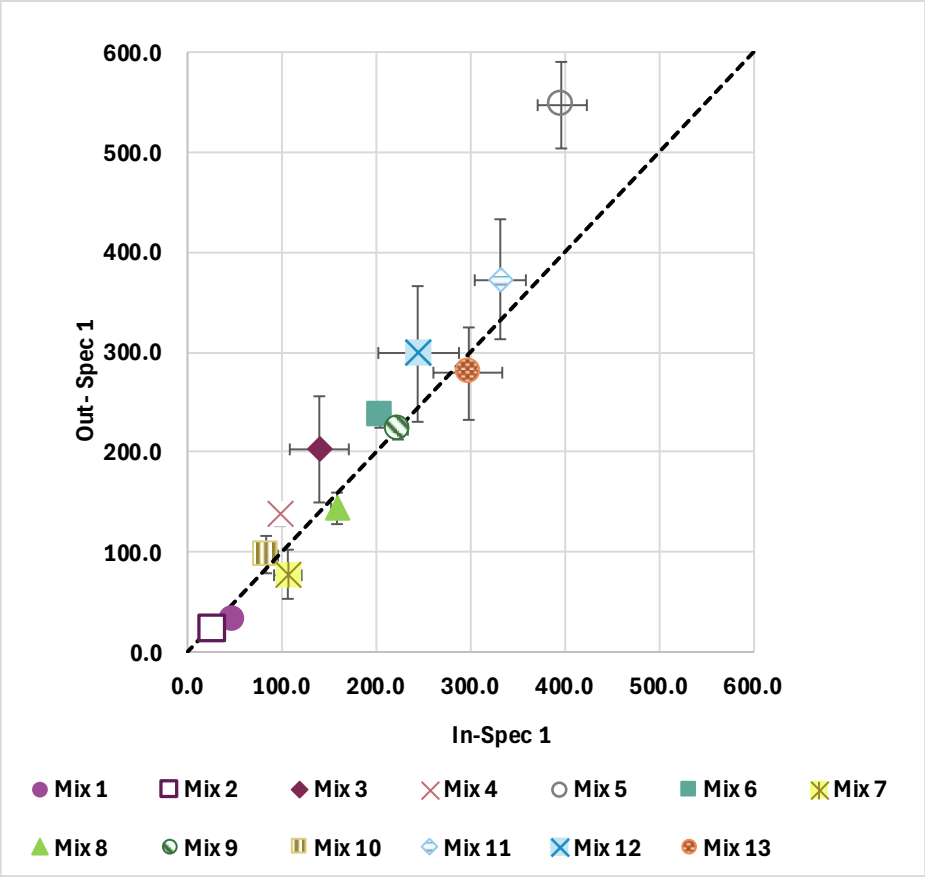
Figure 28. APA/HWTT Test Results of In-Spec 1 and Out-Spec 1

Table 32. P-Values of Rut Depths

Mix ID	Observed Difference	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification Limits)			
Mix 1 (APA 250 psi)	0.8	0.29	No
Mix 2 (APA 250 psi)	0.8	0.45	No
In-Spec 1 vs. Out-Spec 1 (Lower Specification Limits)			
Mix 3 (APA 250 psi)	0.9	0.34	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 (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 (HWTT)	0.3	0.76	No

4.3.5 Intermediate-Temperature Cracking Resistance

Figure 29 compares CT_{Index} results for 13 mix designs, and Table 33 summarizes the associated d2s and p-values. Among the 11 lower-limit mix designs, eight showed higher CT_{Index} values for Out-Spec 1 than for In-Spec 1, four of which were statistically significant ($p < 0.05$). The remaining three lower-limit mix designs (Mix 7, Mix 8, and Mix 13) showed slightly lower CT_{Index} values for Out-Spec 1; however, these differences were not statistically significant. In contrast, both upper-limit mix designs (Mix 1 and Mix 2) showed reduced CT_{Index} values when gradation was shifted outside the upper limits. The reduction was statistically significant for Mix 1 but not for Mix 2. Importantly, none of the 13 mix designs showed differences that exceeded the d2s limits. The results indicate that, while there were significant differences in CT_{Index} results between Out-Spec 1 and In-Spec 1 among mix designs, these differences fell within the test variability, suggesting they were not practically significant. Moving gradation outside the lower limits generally maintained or improved intermediate-temperature cracking resistance, whereas shifting outside the upper limits tended to reduce CT_{Index}.



Source: AAPTP

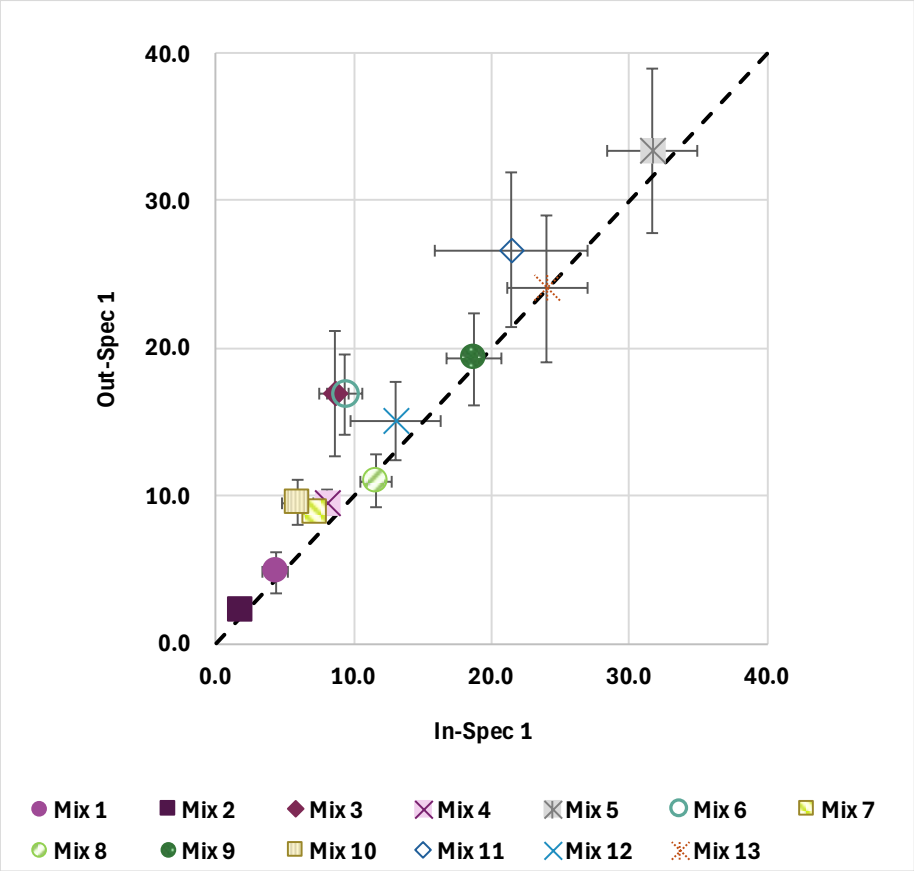
Figure 29. CT_{Index} Results for In-Spec 1 and Out-Spec 1

Table 33. D2s and P-Values of CT_{Index} for In-Spec 1 and Out-Spec 1

Mix ID	Observed Difference	Single-Operator d2s (51.1% [198])	Exceeding Single-Operator d2s Limit?	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification Limits)					
Mix 1	−16.1	19.8	No	0.00	Yes
Mix 2	−3.0	12.6	No	0.43	No
In-Spec 1 vs. Out-Spec 1 (Lower Specification Limit)					
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 30 presents FI results for the same mix designs, with corresponding d2s and p-values provided in Table 34. Among the 11 lower-limit mix designs, 10 showed higher FI values for Out-Spec 1 than for In-Spec 1, and three of these increases were statistically significant (p-values <0.05). Mix 8 showed a lower FI value for Out-Spec 1; however, the difference was not statistically significant. For the two upper-limit mix designs, FI values were slightly higher for Out-Spec 1 than for In-Spec 1; however, these differences were not statistically significant. As for the CT_{Index} results, the differences in the FI results were below the d2s limits (i.e., the test variability), suggesting they were not practically significant.

Overall, FI results indicate that shifting gradation outside the lower limits did not reduce intermediate-temperature cracking resistance and, in several cases, provided an improvement. For mix designs near the upper limits, FI did not indicate a reduction in cracking resistance, although CT_{Index} results suggest reduced cracking resistance when gradation was shifted outside the upper limits.



Source: AAPTP

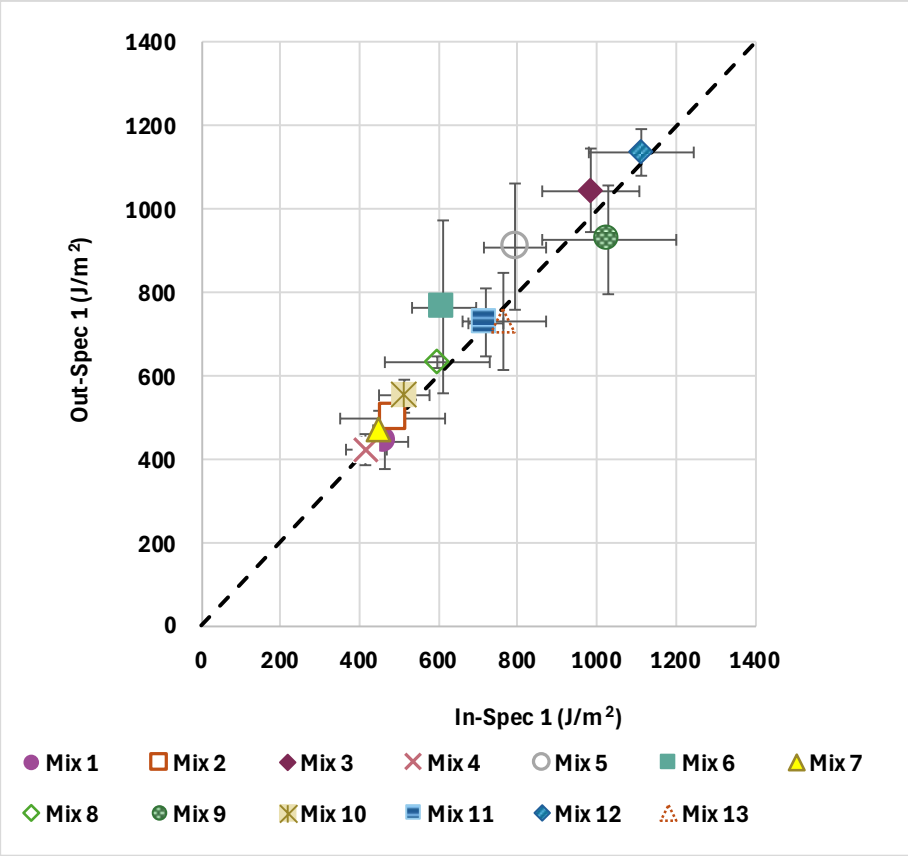
Figure 30. FI Values of In-Spec 1 and Out-Spec 1

Table 34. D2s and P-Values of FI of In-Spec 1 and Out-Spec 1

Mix ID	Observed Difference	Single-Operator d2s (75.9% [187])	Exceeding Single-Operator d2s Limit?	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification 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 Specification Limit)					
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 31 presents fracture energy results from the DCT test, which was used as the primary low-temperature cracking performance indicator, for the 13 mix designs, and Table 35 summarizes the corresponding d2s and p-values. Most mix designs exhibited similar or slightly higher fracture energy values for Out-Spec 1 compared to In-Spec 1, indicating that shifting gradation outside the current P-401 limits did not reduce low-temperature cracking resistance. However, statistical analysis showed that all mix designs indicated no statistically significant differences between paired mixtures (p-values > 0.05) and were within the range of testing variability (i.e., the d2s limit). These results suggest that gradation adjustments, when volumetric properties and binder contents are controlled, do not adversely affect low-temperature fracture performance.



Source: AAPTP

Figure 31. Fracture Energy from DCT of In-Spec 1 and Out-Spec 1

Table 35. 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 [188])	Exceeding Single-Operator d2s Limit?	P-Value	Statistical Significance?
In-Spec 1 vs. Out-Spec 1 (Upper Specification 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 Specification Limit)					
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

This section evaluated the performance of In-Spec 1 and Out-Spec 1 to assess the effects of shifting gradation outside the current P-401 limits while maintaining comparable volumetric properties, effective binder contents, and total binder contents. Based on laboratory testing and statistical analysis, the following observations were made:

- Moisture susceptibility, permeability, durability, rutting resistance, and low-temperature cracking resistance were not significantly affected by moderate gradation shifts outside the gradation limits. All mixtures met the applicable performance requirements, and observed differences between paired designs were neither statistically nor practically significant.
- Intermediate-temperature cracking resistance exhibited directional sensitivity. For mix designs positioned near the lower gradation limits, shifting gradation outside the band generally maintained or improved cracking resistance, with several statistically significant increases in CT_{Index} and FI. In contrast, for mix designs near the upper limits, CT_{Index} values tended to decrease as gradation shifted outside the upper limits, although FI results did not indicate a statistically significant change.

Overall, the results indicate that expanding flexibility within the 12.5-mm NMAS P-401 gradation limits is achievable without compromising key performance indicators, provided that volumetric properties and asphalt binder content are controlled.

4.4 Proposed Revisions to P-401 Gradation Limits

Of the 13 P-401 mix designs evaluated in this study, 12 were 12.5-mm NMAS mixtures and one was a 19.0-mm NMAS mixture. Accordingly, potential adjustments focus on the current P-401 gradation limits for 12.5-mm NMAS mixtures. The 12 mix designs consisted of two mix designs close to the upper gradation limits and 10 mix designs close to the lower gradation limits. Proposed revisions were derived from Out-Spec 1 gradations that demonstrated acceptable performance while positioned outside the existing gradation limits.

4.4.1 Upper-Limit Evaluation

Adjustments to the upper limits were evaluated using the two upper-limit mix designs (Mix 1 and Mix 2). Table 36 summarizes the number of mixes and percentage deviations from upper limits by sieve size.

Gradation shifted primarily at the No. 4, No. 8, No. 30, and No. 50 sieves. At the No. 4 and No. 8 sieves, one mixture exceeded the upper limit by 1 percent. At the No. 30 sieve, both mixtures exceeded the upper limit by 1 percent. At the No. 50 sieve, one mixture exceeded the limit by 3 percent and the other by 4 percent.

Table 36. 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 rutting and durability performance remained acceptable near the upper gradation limits, intermediate-temperature cracking resistance was reduced when gradation moved outside the upper gradation limits. Additionally, because contractors have not reported difficulty meeting the current upper limits and because cracking sensitivity was observed near the coarse boundary, retaining the existing upper limits for the P-401 gradation limits is recommended.

4.4.2 Lower-Limit Evaluation

Lower-limit adjustments were evaluated using the 10 lower-limit mix designs. Table 37 summarizes the number of mixes and percentage deviations from lower limits by sieve size.

Table 37. 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

Gradation shifts were concentrated at the No. 16, No. 30, and No. 50 sieves. At the No. 16 sieve, seven mixtures exceeded the lower limit (six by 1 percent and one by 3 percent). At the No. 30 sieve, nine mixtures exceeded the lower limit (four by 1 percent and five by 2 percent). At the No. 50 sieve, seven mixtures exceeded the lower limit (six by 1 percent and one by 2 percent). One mixture exceeded the No. 100 sieve by 1 percent.

Performance evaluation confirmed that no adverse effects on moisture resistance, permeability, durability, rutting resistance, or low-temperature cracking resistance were observed for lower-limit mix designs. Intermediate-temperature cracking resistance was generally maintained or improved.

Based on these findings, reducing the lower limits at the No. 16, No. 30, and No. 50 sieves by 2 percent is recommended. The proposed revisions are presented in Table 38. The 2-percent value corresponds to the extent to which the gradations could be shifted outside the current lower limits while keeping the mixtures FG, holding the volumetric

requirements unchanged, and remaining within the range permitted by the aggregate fractions available in the selected mix designs. It is not a maximum or optimal adjustment; larger shifts were not achievable under these constraints with the aggregates evaluated. Because the recommended reductions reflect the range of gradations evaluated in this study, it is recommended that agencies adopting the revised limits monitor mix designs and field performance during initial implementation to confirm that the laboratory-based performance trends observed in this study are also achieved under field conditions.

Table 38. Current and Revised P-401 Gradation Limits for 12.5-mm NMAS Mix Designs

Sieve Size	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
No. 4	53–73	53–73
No. 8	38–60	38–60
No. 16	26–48	24–48
No. 30	18–38	16–38
No. 50	11–27	9–27
No. 100	6–18	6–18
No. 200	3.0–6.0	3.0–6.0

4.4.3 Engineering Rationale for Proposed Adjustments to Lower Limits

Reducing the lower limits at the No. 16, No. 30, and No. 50 sieves provides both constructability and performance benefits. Under the current specification, natural sand is often used to meet the lower-limit requirements at these sieve sizes. Natural sand typically contains a higher percentage of material passing through the No. 16, No. 30, and No. 50 sieves than manufactured sand. Incorporating approximately 10-percent natural sand can increase the percent passing at these sieves by about 2 percent, facilitating compliance with lower limits. However, replacing manufactured sand with natural sand generally reduces FAA, which may lower VMA and make it more difficult to satisfy minimum VMA requirements while maintaining rutting resistance.

By reducing the lower limits by 2 percent at the No. 16, No. 30, and No. 50 sieves, mixtures can be designed with greater flexibility in fine aggregate selection, enabling increased use of manufactured sand. This approach supports maintaining or increasing VMA without compromising rutting performance.

Furthermore, cracking resistance is strongly influenced by effective binder content and binder properties. Future cracking test criteria may require higher effective binder volumes than those typically used in current P-401 mixtures. Increasing VMA is a practical mechanism for accommodating these higher binder contents. Reducing the lower

gradation limits at the critical fine aggregate sieves facilitates higher VMA and supports the development of mixtures with improved potential for cracking resistance.

These revisions provide increased gradation flexibility for 12.5-mm NMAS mixtures while maintaining demonstrated compliance with moisture, durability, rutting, and cracking performance expectations.

Chapter 5. Effect of Gradation Changes on Volumetric and Performance Properties

This chapter presents the results of Part 2 of the experimental program, which evaluates the effects of gradation changes on the volumetric properties and laboratory performance of P-401 asphalt mixtures. While Chapter 4 focused on assessing potential revisions to the P-401 gradation limits, this chapter examines how controlled gradation shifts influence VMA and key performance indicators, including rutting resistance and cracking resistance, across multiple gradation conditions. The findings provide quantitative insight into relationships among gradation, volumetrics, and performance, supporting the development of the predictive modeling framework discussed in Chapter 6.

This chapter is organized into four sections. Section 5.1 summarizes the execution of Part 2 of the experimental plan. Section 5.2 presents the mix design summary for the In-Spec 2 and Out-Spec 2 gradations, which are combined with the In-Spec 1 and Out-Spec 1 gradations for analysis in Section 5.3. Section 5.3 provides the test results and comparative analysis, including the effects of gradation changes on volumetric properties and laboratory performance. Section 5.4 summarizes the key findings.

5.1 Execution of Part 2 of Experimental Design

Figure 32 summarizes the execution of Part 2 of the experimental program. Part 2 was designed to evaluate the effects of systematic gradation shifts on volumetric properties and laboratory performance across multiple gradations.

Of the 13 mix designs included in this study, 10 were selected for gradation adjustment in Part 2. These consisted of two mix designs near the upper limits (Mix 1 and Mix 2) and eight mix designs near the lower limits (Mix 3, Mix 4, Mix 5, Mix 6, Mix 7, Mix 8, Mix 12, and Mix 13). Selection was based on material availability, project resources, and the need to maintain representation across aggregate sources and climatic regions.

For each selected mix design, two additional gradations were developed relative to the In-Spec 1 and Out-Spec 1 established in Part 1:

- In-Spec 2, with gradation close to the center of the P-401 gradation limits; and
- Out-Spec 2, with gradation positioned further outside the gradation limits than Out-Spec 1.

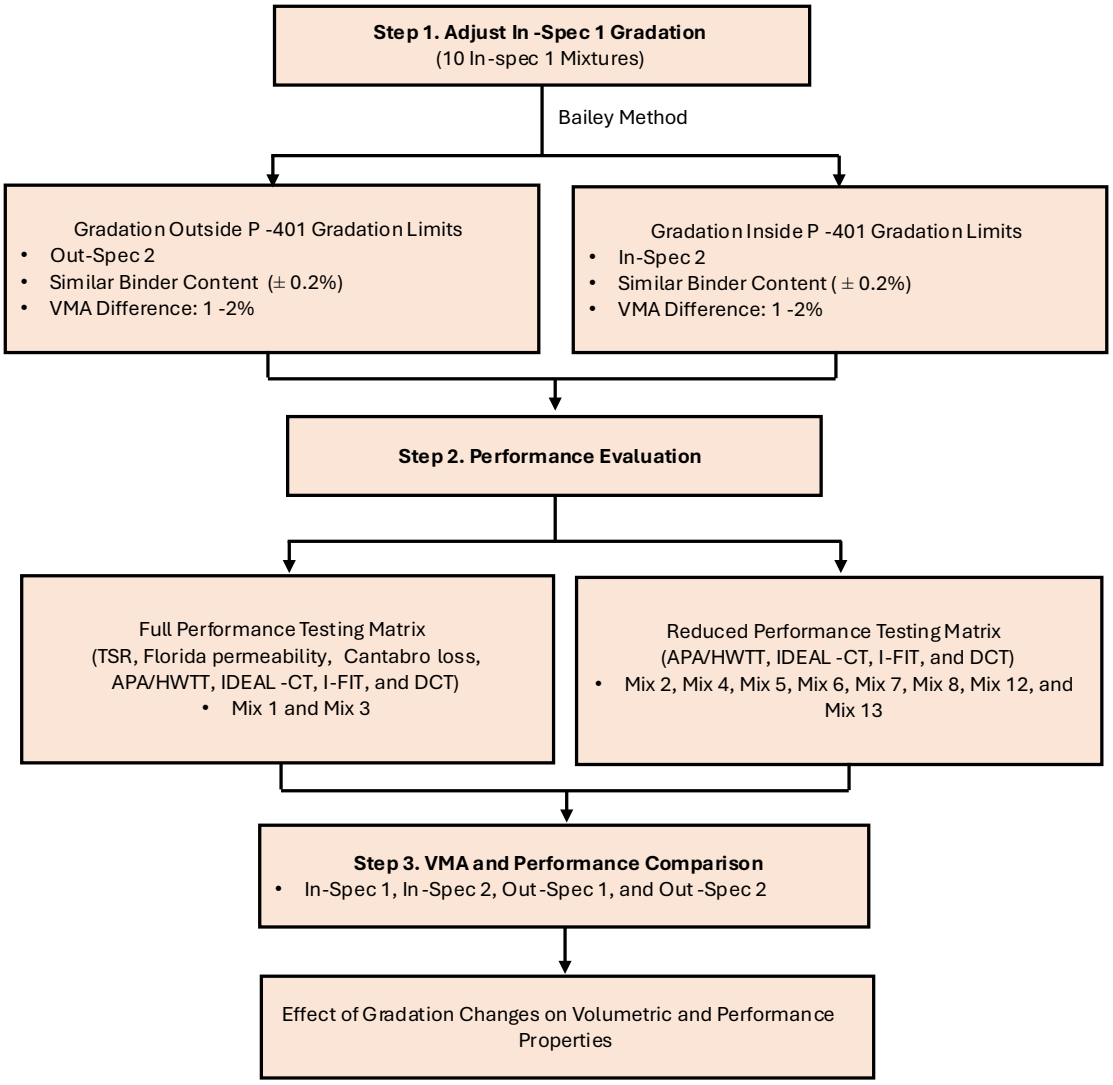
Gradation adjustments were guided by the Bailey Method to ensure controlled changes in the aggregate packing structure. For all adjusted gradations, the VMA values for In-Spec 2 and Out-Spec 2 were selected to be 1- to 2-percent different from those of the corresponding In-Spec 1 and Out-Spec 1, respectively. Their asphalt binder contents were selected to be the same as those of the corresponding In-Spec 1 and Out-Spec 1,

respectively. This approach minimized the effect of binder content and isolated the influence of gradation on volumetric and performance responses.

Comprehensive laboratory testing was initially planned for all four gradations. However, preliminary comparisons of Cantabro mass loss, permeability, and TSR for two mix designs (Mix 1 and Mix 3) indicated no significant differences and confirmed adequate durability, impermeability, and moisture resistance. Accordingly, these tests were not performed for the remaining mix designs; instead, additional aggregate sources (i.e., P-401 mix designs) were evaluated. The reduced testing matrix included rutting resistance (APA or HWTT), intermediate-temperature cracking resistance (IDEAL-CT and I-FIT), and low-temperature cracking resistance (DCT).

Volumetric properties (VMA and AV) and performance results were compared across the four gradations (In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2) for each selected mixture. Test data were presented as mean ±1 standard deviation. Statistical significance among gradations was evaluated using the Games–Howell test at the 0.05 significance level. Gradation designs sharing the same statistical grouping were considered not significantly different.

This dataset enabled a systematic evaluation of the correlation among gradation, volumetrics, and performance, serving as the basis for the predictive modeling analyses presented in Chapter 6.



Source: AAPTP

Figure 32. Experimental Design (Part 2) for Evaluating the Effect of Gradation Changes on Volumetric and Performance Properties

5.2 Mix Design Summary

Table 39 summarizes the gradation, FAA, binder content, and volumetric properties of In-Spec 2 and Out-Spec 2 that were developed for the 10 selected mix designs. The In-Spec 2 gradation was adjusted toward the center of the P-401 gradation limits, whereas the Out-Spec 2 gradation was positioned further outside the gradation limits relative to Out-Spec 1 gradation in Part 1.

As shown in Table 39, systematic gradation shifts primarily affected the intermediate sieve sizes (No. 16, No. 30, and No. 50), resulting in noticeable changes in aggregate packing structure. However, for this experiment, asphalt binder content was maintained within

±0.2 percent of the corresponding In-Spec 1, consistent with the design approach described in Section 5.1.

The volumetric results indicate that gradation shifts produced measurable changes in AV and VMA. Out-Spec 2 exhibited higher VMA than In-Spec 2, reflecting the changes in packing characteristics associated with gradation shifts. These controlled variations in volumetric structure provide a basis for evaluating their effects on laboratory performance, as discussed in Section 5.3.

Table 39. Design Summary of In-Spec 2 and Out-Spec 2 (Mixes 1 through 5)

Mix ID	1		2		3		4		5	
Sieve	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2
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
No. 4	69	80	59	71	57	48	66	58	70	62
No. 8	53	68	46	60	46	35	49	39	46	41
No. 16	39	54	38	50	34	22	32	23	29	26
No. 30	30	43	31	43	25	14	22	15	19	15
No. 50	23	33	23	32	15	8	13	9	11	7
No. 100	12	16	11	17	6	5	8	5	8	4
No. 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
AV (%)	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
V_{be} (%)	12.4	12.6	13.2	12.9	13.2	13.2	11.4	11.3	12.9	13.5

AC = asphalt content; IS 2 = In-Spec 2; OS 2 = Out-Spec 2

Table 39. Design Summary of In-Spec 2 and Out-Spec 2 (Mixes 6 through 13)

Mix ID	6		7		8		12		13	
Sieve	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2	IS 2	OS 2
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
No. 4	63	50	63	60	64	50	55	48	69	59
No. 8	43	30	43	38	49	35	39	32	48	36
No. 16	28	18	30	25	31	20	27	21	35	24
No. 30	18	11	20	16	23	14	19	14	23	17
No. 50	11	7	11	9	14	10	14	10	13	11
No. 100	8	5	7	6	9	8	9	7	8	8
No. 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
AV (%)	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
V_{be} (%)	10.9	11.0	11.8	11.7	11.7	11.9	13.6	13.5	12.2	12.3

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

5.3 Test Results and Data Analysis

This section evaluates the effects of systematic gradation shifts on volumetric properties (VMA and AV) and laboratory performance of the 10 selected P-401 mix designs. Comparisons were conducted across the four gradations: In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2.

For the upper-limit mix designs (Mix 1 and Mix 2), gradation adjustments from In-Spec 2 toward Out-Spec 2 produced a progressively finer aggregate structure. In contrast, for the lower-limit mix designs (Mix 3, Mix 4, Mix 5, Mix 6, Mix 7, Mix 8, Mix 12, and Mix 13), gradation adjustments from In-Spec 2 toward Out-Spec 2 resulted in a coarser gradation. These controlled shifts enabled evaluation of how systematic changes in aggregate packing influence volumetric structure and performance response.

Section 5.3.1 evaluates the effects of gradation change on volumetric properties. Section 5.3.2 examines the corresponding effects on rutting resistance and cracking resistance.

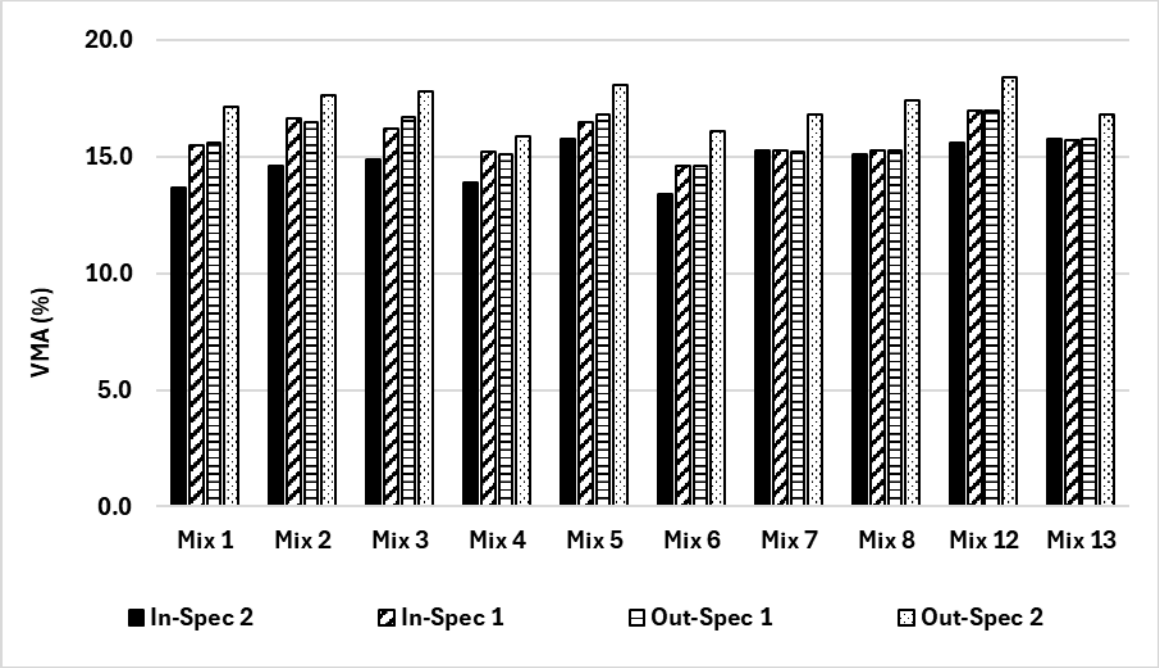
5.3.1 Effects of Gradation Changes on Volumetric Properties

Figure 33 and Figure 34 present the VMA and AV values for the four gradations (In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2) across the evaluated mix designs. A consistent trend was observed among the mix designs. Out-Spec 2 generally exhibited the highest VMA, followed by Out-Spec 1, In-Spec 1, and In-Spec 2. A similar pattern was observed for AV.

The similarity between In-Spec 1 and Out-Spec 1 confirms that the Part 1 design approach successfully controlled volumetric properties while shifting gradation outside the gradation limits. In contrast, the systematic gradation adjustments introduced in Part 2 produced measurable and progressive changes in volumetric properties.

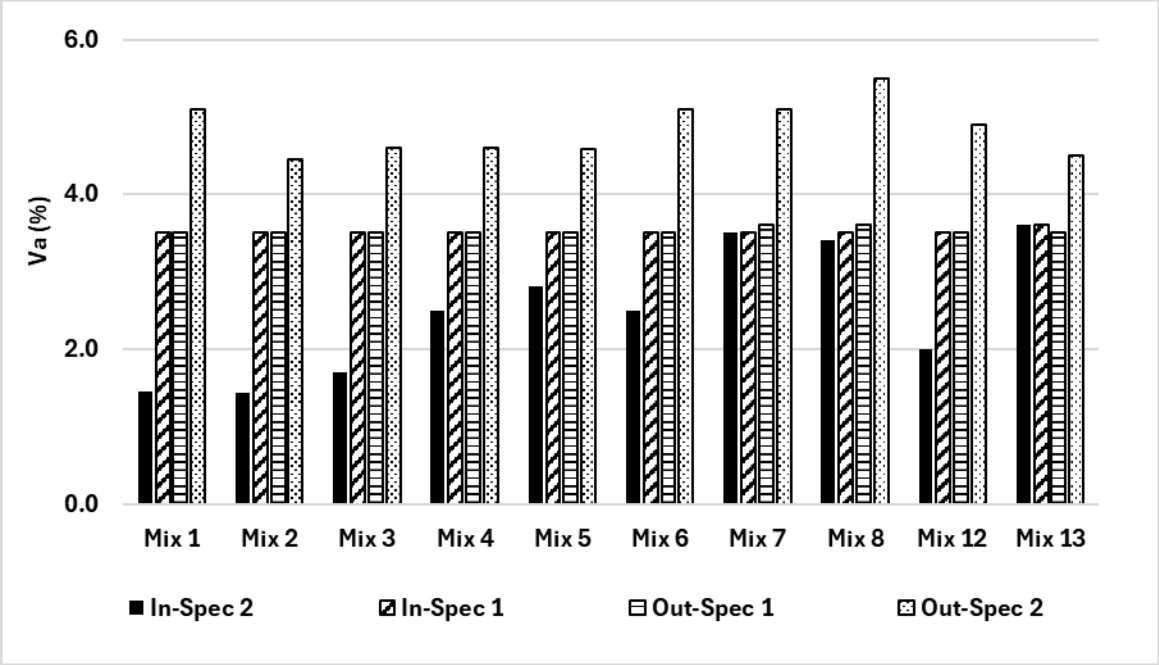
The observed trend indicates that moving the gradation curve farther from the MDL, whether toward the upper or lower limits, reduces aggregate packing efficiency and increases void space within the mixture. This reduction in packing density results in higher VMA and, under similar compaction conditions and binder contents, a corresponding increase in AV content.

These controlled variations in volumetric structure provide a quantitative basis for evaluating how changes in gradation influence laboratory performance, as discussed in Section 5.3.2.



Source: AAPTP

Figure 33. Comparison of VMA of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2



Source: AAPTP

Figure 34. Comparison of AV of In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2

5.3.2 Effects of Gradation Changes on Laboratory Performance

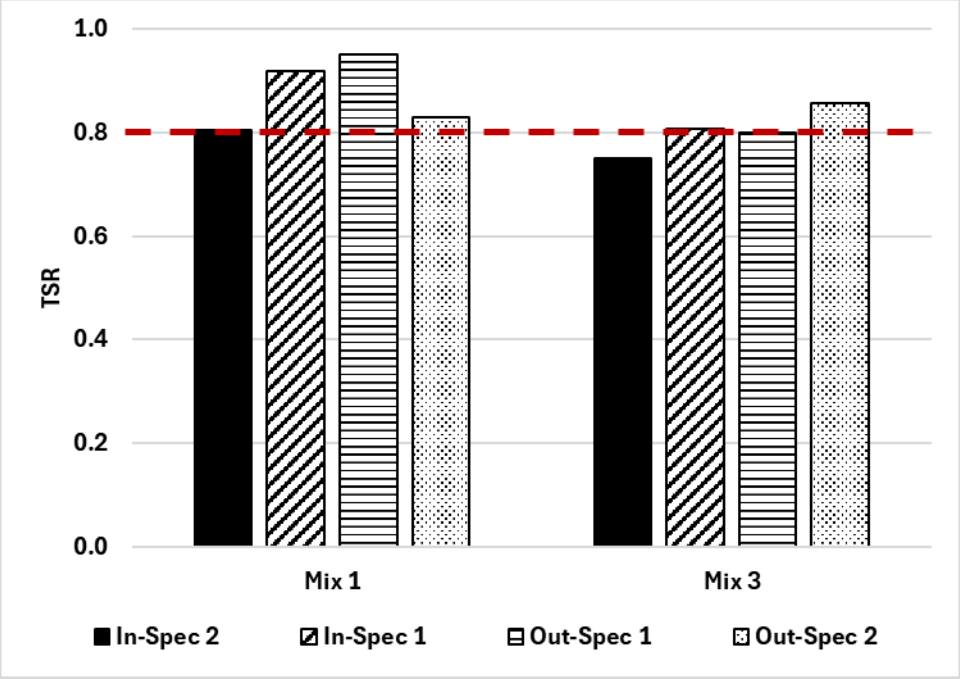
This section evaluates the effects of systematic gradation shifts on laboratory performance properties, including moisture susceptibility, permeability, durability, rutting resistance,

intermediate-temperature cracking resistance, and low-temperature cracking resistance. Performance comparisons were conducted across the four gradations to determine whether the volumetric changes observed in Section 5.3.1 translated into measurable differences in mechanical response.

5.3.2.1 Moisture Susceptibility

Figure 35 presents the TSR results for Mix 1 and Mix 3 with the four gradations. For both mix designs, TSR values for In-Spec and Out-Spec designs were generally equal to or greater than 0.80, satisfying the moisture resistance requirement specified in the P-401 specifications. The only exception was In-Spec 2 of Mix 3, which exhibited a TSR value of 0.75, slightly below the minimum threshold.

These results suggest that adjustments toward finer or coarser gradations could affect moisture susceptibility. However, the two gradations with volumetric properties meeting the P-401 specifications (i.e., In-Spec 1 and Out-Spec 1) satisfied the minimum TSR requirement of 0.8. The In-Spec 2 gradations, which were closer to the MDL and had AV contents of 1.5 and 1.7 percent and VMA values of 13.9 and 14.9 percent, yielded the lowest TSR values, underscoring the influence of volumetric properties on moisture resistance. These findings suggest that volumetric compliance plays an important role alongside binder properties and aggregate mineralogy in governing the moisture susceptibility of asphalt mixtures [199].



Source: AAPTP

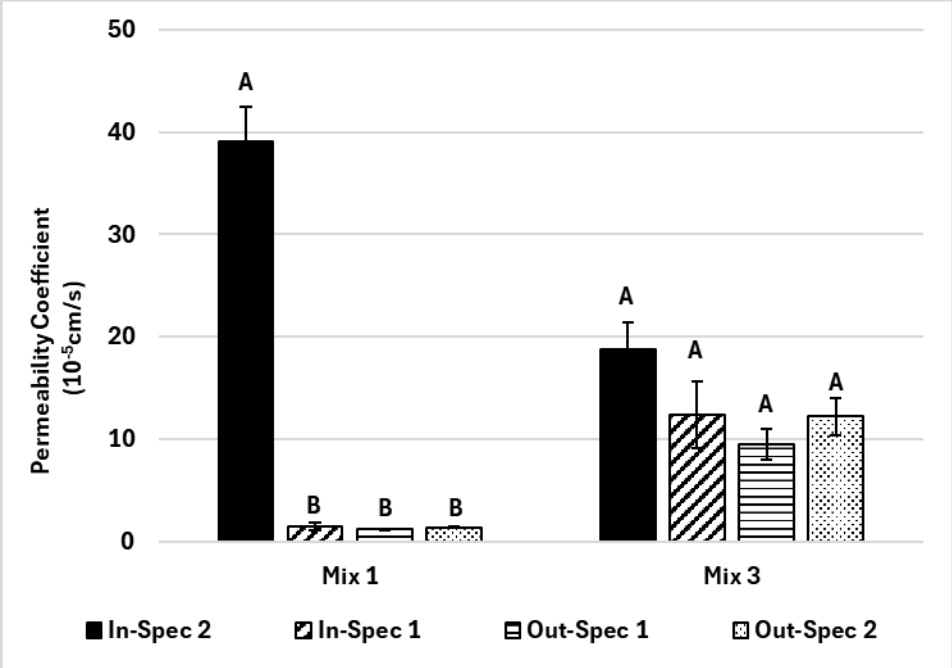
Figure 35. TSR Test Results of Mix 1 and Mix 3

5.3.2.2 Permeability

Figure 36 presents the permeability coefficients and statistical groupings for Mix 1 and Mix 3 with the four gradations. All measured permeability values were below 125×10^{-5} cm/s, confirming that the mixtures satisfied commonly accepted impermeability criteria for airfield asphalt mixtures [196].

For Mix 1, statistical grouping indicated that In-Spec 1, Out-Spec 1, and Out-Spec 2 were not significantly different from each other, while In-Spec 2 exhibited a significantly higher permeability. For Mix 3, no statistically significant differences were observed among the four gradation conditions.

The elevated permeability observed for In-Spec 2 in Mix 1 likely corresponded to its higher interconnected voids and altered packing structure, as discussed in Section 5.3.1. However, even under this condition, the mixture remained within impermeable limits. Overall, the results indicate that systematic gradation adjustments from In-Spec 2 toward Out-Spec 2 did not compromise impermeability within the evaluated range of volumetric conditions.



Source: AAPTP

Figure 36. Permeability Test Results of Mix 1 and Mix 3

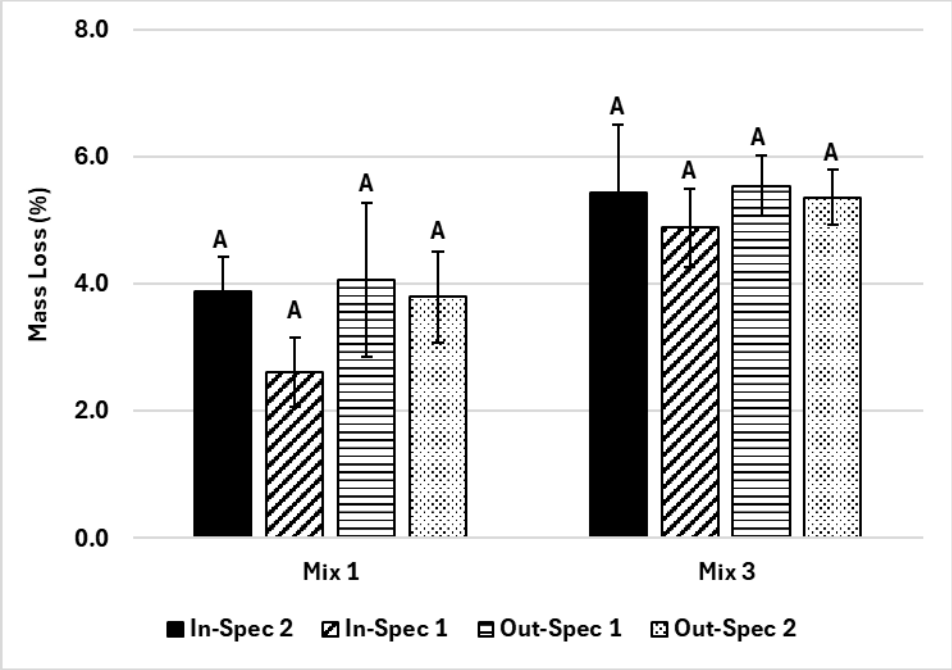
5.3.2.3 Durability

Figure 37 presents the Cantabro mass loss results and statistical groupings for Mix 1 and Mix 3 across the four gradations. For both mix designs, Cantabro mass loss values were

below the 6.3-percent threshold commonly used for durability evaluation [197], indicating acceptable resistance to raveling.

Statistical analysis showed no significant differences among the gradation conditions for either mixture. Although minor variations in mass loss were observed, they were not statistically significant.

The results indicate that systematic gradation adjustments from In-Spec 2 toward Out-Spec 2 did not adversely affect mixture durability within the evaluated volumetric range. This outcome is consistent with the understanding that Cantabro resistance is influenced primarily by binder content and aggregate–binder adhesion, which were maintained at comparable levels across the four gradations [200].



Source: AAPTP

Figure 37. Cantabro Test Results for Mix 1 and Mix 3

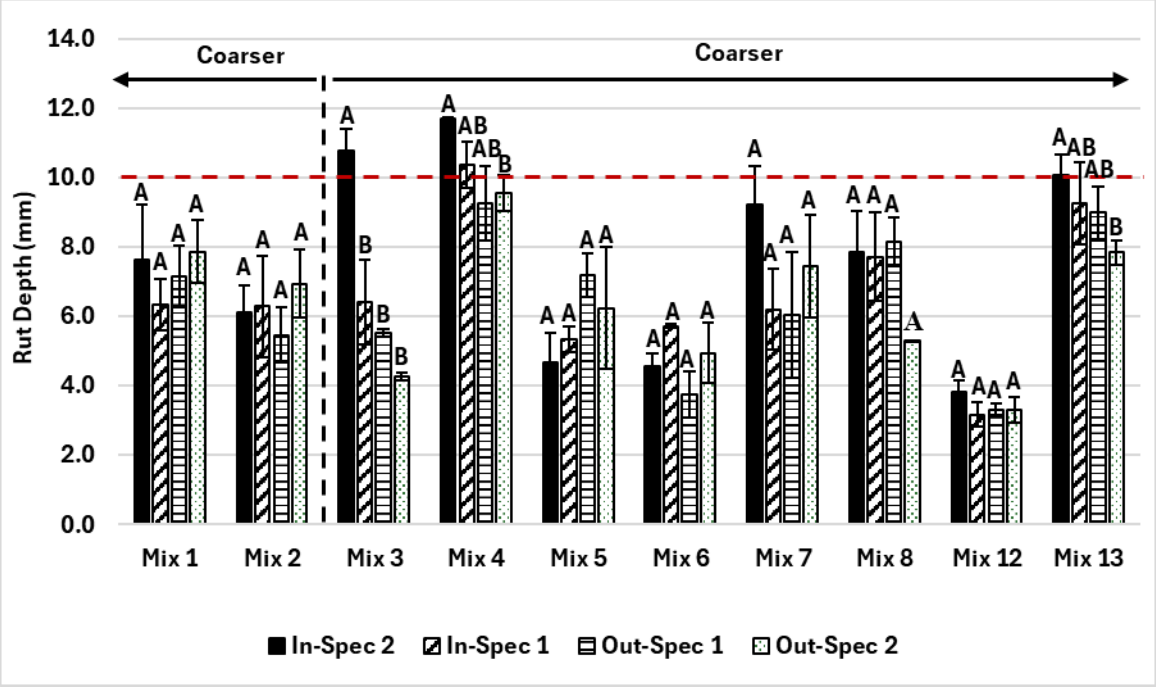
5.3.2.4 Rutting Resistance

Figure 38 presents the APA and HWTT rut depth results and statistical groupings across the four gradations. The dashed line at 10 mm represents the maximum allowable rut-depth threshold. Mixes 6, 8, and 13 were evaluated using HWTT, while the remaining mixtures were evaluated using APA.

For the upper-limit mix designs (Mix 1 and Mix 2), all gradation conditions exhibited rut depths below the 10 mm threshold. Statistical analysis indicated no significant differences among the four gradations, suggesting that systematic gradation shifts did not statistically affect rutting resistance for mixtures positioned near the upper gradation limits.

For the lower-limit mix designs (Mixes 3–8,12, and 13), rut depths were also generally below the 10-mm threshold. A consistent trend was observed in which rut depth decreased as the gradation shifted from In-Spec 2 toward Out-Spec 2, corresponding to a progressively coarser aggregate structure. This trend was statistically significant for Mixes 3, 4, and 13. For the remaining mixtures (Mixes 5, 6, 7, 8, and 12), no statistically significant differences were detected among gradations.

The observed improvement in rutting resistance with coarser gradations is consistent with enhanced aggregate interlock and stone-on-stone contact, which increase mixture shear resistance [90, 91]. Overall, shifting the gradation toward a coarser structure either maintained or improved rutting performance within the evaluated volumetric range.



Source: AAPTP

Figure 38. APA/HWTT Test Results

5.3.2.5 Intermediate-Temperature Cracking Resistance

Intermediate-temperature cracking resistance was evaluated using both the IDEAL-CT (CT_{Index}) and I-FIT (FI) tests. These complementary indicators quantify the mixture resistance to cracking at intermediate service temperatures. The results were analyzed across the four gradations to determine whether the systematic gradation shifts influenced fracture tolerance.

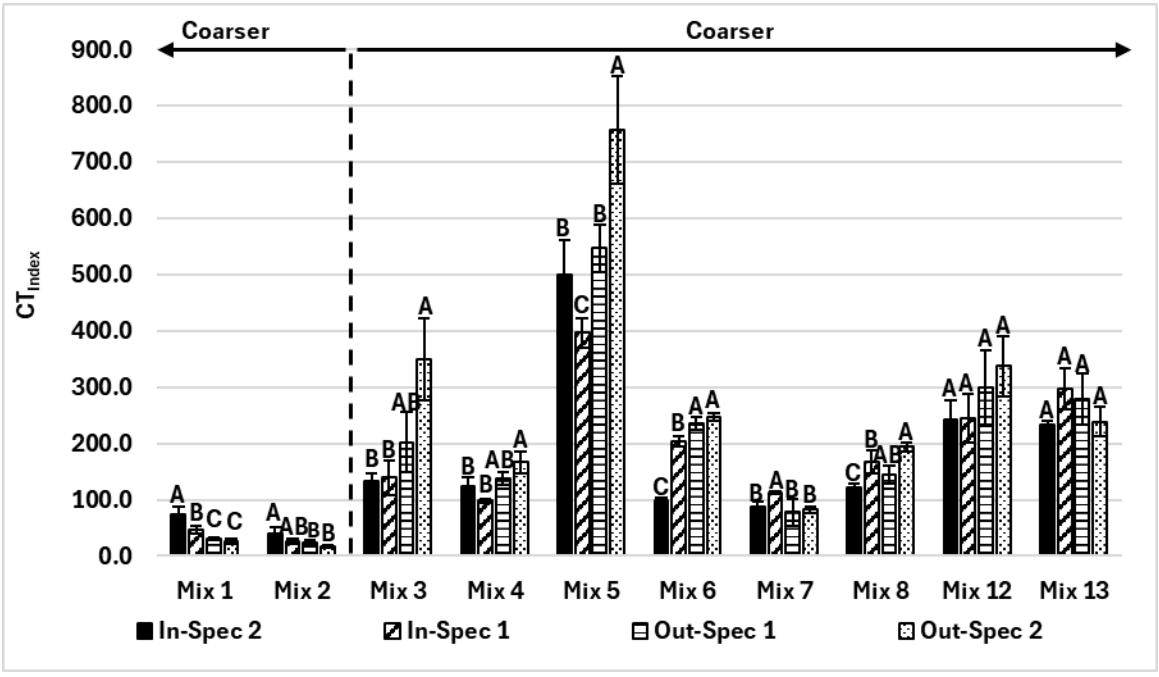
5.3.2.5.1 IDEAL-CT Results

Figure 39 presents the IDEAL-CT results (CT_{Index}) and statistical groupings for the four gradation conditions. For the upper-limit mix designs (Mix 1 and Mix 2), shifting the

gradation from Out-Spec 2 to In-Spec 1 and Out-Spec 1 and to In-Spec 2, indicating a coarser aggregate structure, produced higher CT_{Index} values. The increases were statistically significant, indicating improved intermediate-temperature cracking resistance with coarser gradation near the upper gradation limits.

For several lower-limit mix designs (Mixes 3, 4, 5, 6, and 8), CT_{Index} values generally increased as the gradation shifted from In-Spec 2 toward Out-Spec 2, indicating a coarser aggregate structure. In these cases, Out-Spec 2 often exhibited statistically higher CT_{Index} values compared to finer gradations. This trend can be attributed to how gradation affects the thickness of the binder film. All four gradations were designed with the same binder content; however, a coarser gradation, which has less fine content and a reduced total aggregate surface area, leads to thicker asphalt binder films surrounding the aggregate particles. Thicker binder films are typically linked to better resistance against the initiation and propagation of cracks [201, 202]. For Mixes 7, 12, and 13, no statistically significant differences were observed among the four gradations.

Overall, IDEAL-CT results indicate that coarser gradations tended to maintain or enhance intermediate-temperature cracking resistance within the evaluated range of volumetric conditions.



Source: AAPTP

Figure 39. IDEAL-CT Results

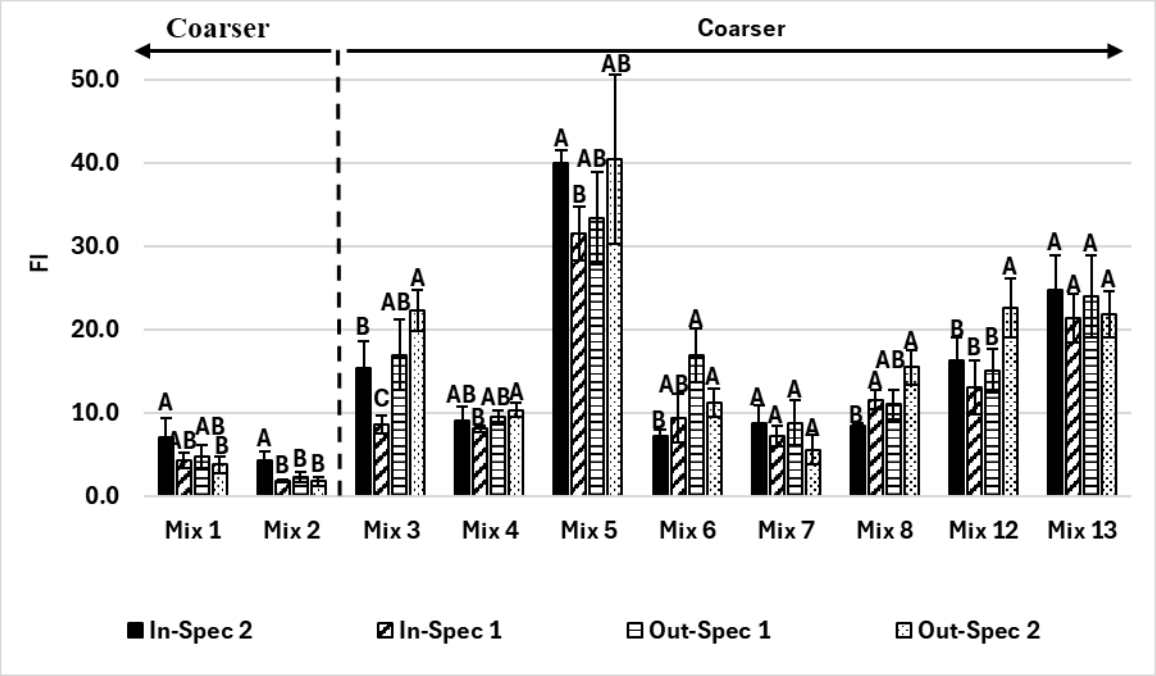
5.3.2.5.2 I-FIT Results

Figure 40 presents the I-FIT results (FI) and statistical groupings. For the upper-limit mix designs, FI values increased as the gradation shifted from Out-Spec 2 toward In-Spec 2,

consistent with the IDEAL-CT trend. Statistical grouping confirmed that these increases were significant.

For several lower-limit mix designs (Mixes 3, 6, 8, and 12), Out-Spec gradations exhibited higher FI values than their corresponding In-Spec conditions, and the improvements were statistically significant. For Mixes 4, 5, 7, and 13, no statistically significant differences were detected.

Across the IDEAL-CT and I-FIT results, the overall trend indicates that gradation shifts producing a coarser aggregate structure (Section 5.3.1) were associated with similar or improved resistance to intermediate-temperature cracking.



Source: AAPTP

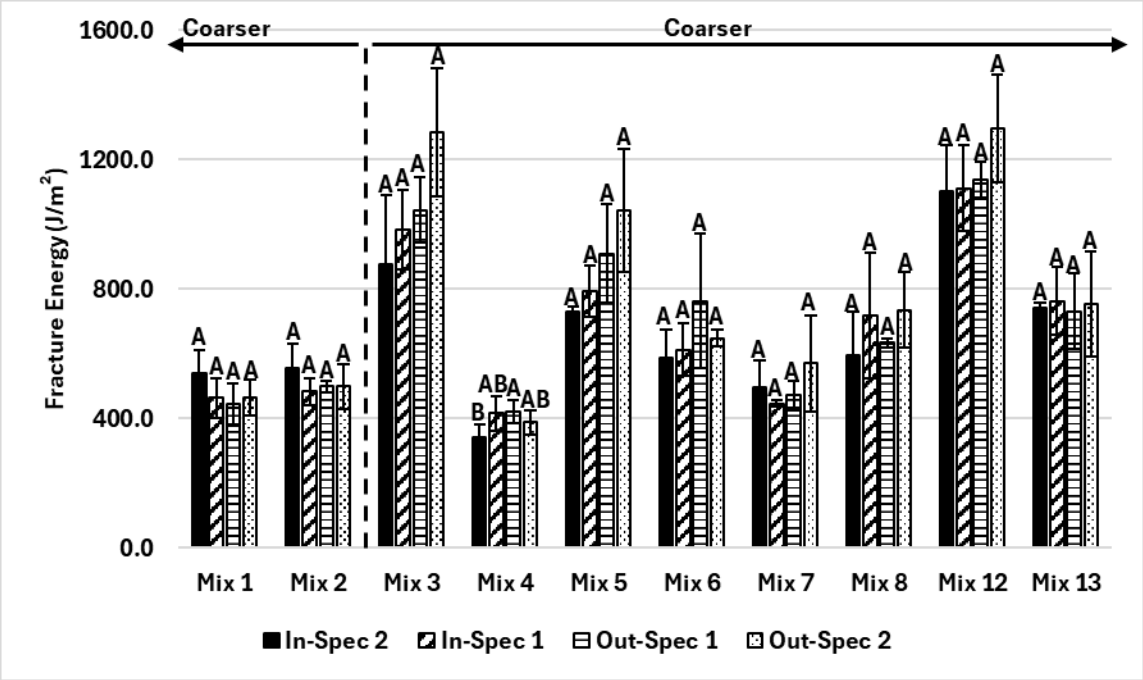
Figure 40. I-FIT Results

5.3.2.6 Low-Temperature Cracking Resistance

Low-temperature cracking resistance was evaluated using the DCT test, with fracture energy used as the primary performance indicator. Figure 41 presents the fracture energy results and statistical groupings for the four gradations. Across most mix designs, shifting the gradation toward a coarser aggregate structure was associated with higher average fracture energy values. However, statistical analysis indicated no significant differences among the four gradations for any of the evaluated mixtures.

Although increases in fracture energy were observed with coarser gradations, the lack of statistical significance suggests that systematic gradation adjustments within the evaluated range did not significantly influence low-temperature cracking resistance. This

outcome indicates that low-temperature fracture performance is primarily governed by binder properties rather than by variations in aggregate packing structure [33, 203].



Source: AAPTP

Figure 41. DCT Results

5.4 Summary

This chapter evaluated the effects of systematic gradation shifts on volumetric properties and laboratory performance of P-401 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 relative to Out-Spec 1 gradations. Asphalt binder content was maintained within ±0.2% between In-Spec 2 and Out-Spec 2. Note that, unlike those of In-Spec 1 and Out-Spec 1, the volumetric properties of 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 packing influence mixture behavior. Based on the test results, the following conclusions were drawn:

- **Volumetric Properties:** Moving the gradation farther from the MDL, whether toward the lower or upper gradation limits, generally reduced aggregate packing efficiency and increased VMA and AV content, given the same compaction effort and binder content.

- **Moisture Susceptibility and Durability:** Systematic gradation adjustments did not significantly affect TSR or Cantabro mass loss within the evaluated range. Moisture resistance and durability performance were generally maintained.
- **Permeability:** All mixtures satisfied impermeability criteria. Gradation shifts did not compromise permeability performance.
- **Rutting Resistance:** For lower-limit mix designs, shifting the gradation toward a coarser aggregate structure generally reduced rut depth and, in several cases, significantly improved rutting resistance. For upper-limit mix designs, rutting resistance was maintained across gradation conditions.
- **Intermediate-Temperature Cracking Resistance:** IDEAL-CT and I-FIT results indicated that coarser gradations generally maintained or improved intermediate-temperature cracking resistance. Several mixtures exhibited statistically significant improvements.
- **Low-Temperature Cracking Resistance:** DCT results showed no statistically significant differences among gradations, indicating that moderate gradation shifts did not significantly affect low-temperature fracture performance.

Overall, the results demonstrate that controlled gradation adjustments influence mixture volumetric structure and can enhance rutting and intermediate-temperature cracking resistance without adversely affecting durability, moisture resistance, or low-temperature cracking performance. These findings establish the basis for the predictive modeling framework presented in Chapter 6.

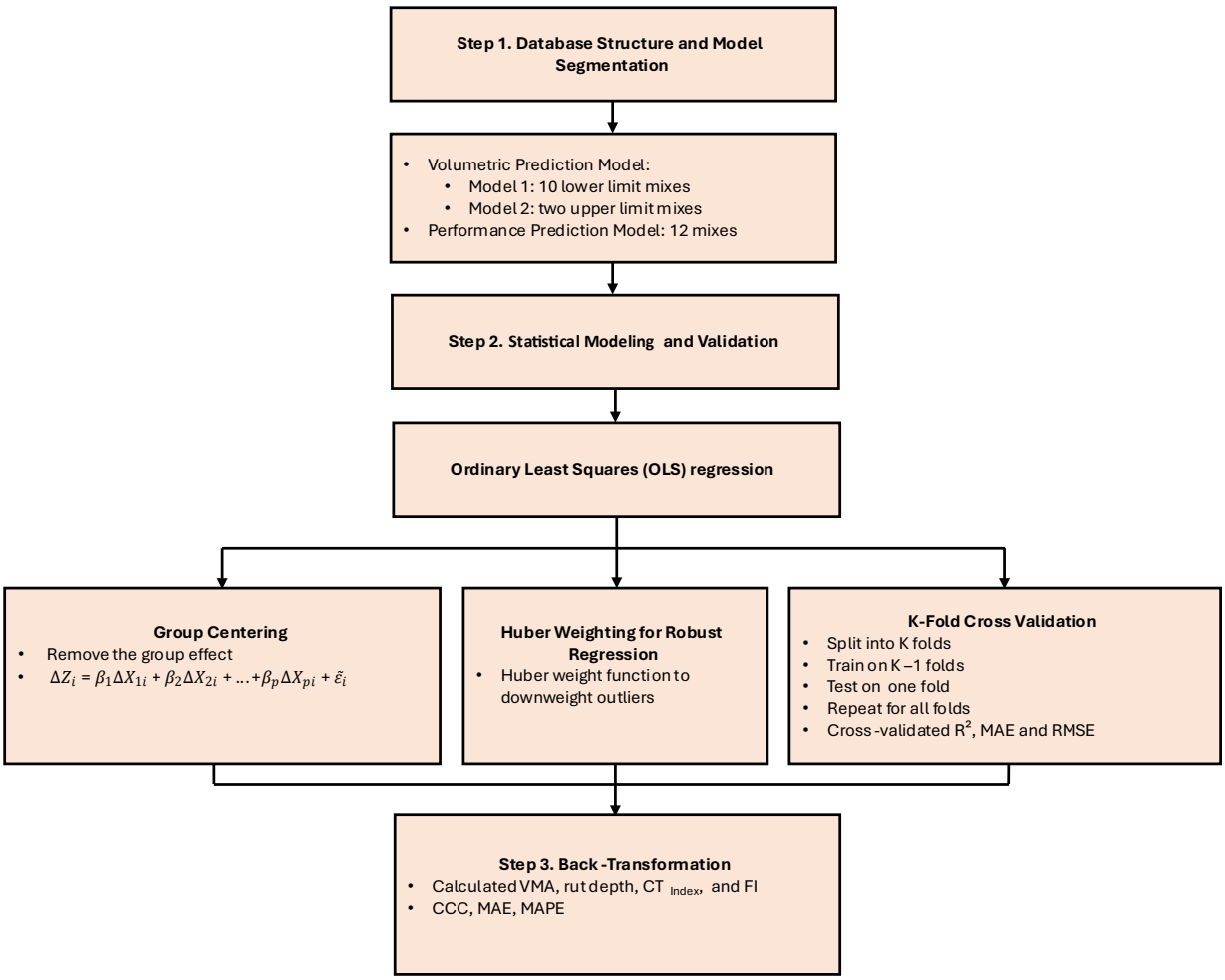
Chapter 6. Prediction Models for Volumetric and Performance Properties

The previous chapters demonstrated that changes in aggregate gradation can affect the volumetric properties of mixtures and certain performance indicators. To translate these findings into practical mix design tools, this chapter presents prediction models that quantitatively characterize relationships between changes in aggregate gradation and volumetric and performance properties. The models were developed following a procedure that included statistical analysis, statistical modeling, and validation. These models were subsequently deployed as a web-based tool, enabling users to estimate VMA and mixture performance for new gradations based on previously tested gradations and aggregate properties. The new gradations would have the same NMAS of 12.5 mm and use the same binder type, binder content, and aggregate sources.

This chapter is organized into six sections. Section 6.1 describes the modeling methodology and the dataset used for model development. Section 6.2 presents two models for predicting volumetric properties. Section 6.3 introduces a model for predicting rut depth. Section 6.4 develops 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 proposed models in mixture design and specification evaluation.

6.1 Methodology for Model Development

This section describes the statistical framework used to develop and validate prediction models for volumetric and performance properties. The modeling procedure, illustrated in Figure 42, consists of three primary components: (1) database selection, (2) statistical modeling and validation, and (3) back-transformation to determine calculated responses.



Source: AAPTP

Figure 42. Workflow for the Development of the Prediction Models

6.1.1 Step 1: Database Structure and Model Segmentation

The modeling database was developed from the laboratory results of this study. Of the 13 P-401 mix designs evaluated, 12 have a 12.5-mm NMAS, while one has a 19.0-mm NMAS. For model development, only the 12.5-mm NMAS mix designs were included to ensure consistency in aggregate packing structure, gradation limits, and volumetric behavior.

The database includes measured gradation percentages, aggregate properties (including G_{sb} and FAA), volumetric properties, and laboratory performance test results. Each P-401 mix design was assigned a unique group identifier. Each group included all mixtures developed based on a selected P-401 mix design, meaning they shared the same aggregates and binder. Within each group, several gradation trials were conducted, producing mixtures with variations suitable for developing models.

6.1.1.1 Volumetric Prediction Models

Chapter 5 demonstrated that mixtures located near the upper and lower P-401 gradation limits exhibited opposite VMA trends when gradation was adjusted. Because a single model could not represent both directional behaviors, two separate VMA prediction models were developed:

- **Upper-limit model:** Two unique sources of aggregate and/or binder; 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.

  For mixtures located near the upper gradation limits, moving the gradation toward the MDL (coarser direction) reduced VMA.

- **Lower-limit model:** Ten unique sources of aggregate and/or binder; 52 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.

  For S-shaped gradations near the upper limits at coarse sieve sizes and the lower limits at fine sieve sizes, moving the gradation coarser increased VMA.

For both models, the response variable was the change in measured VMA, and the predictor variables included:

- Percent of aggregate passing at the 1/2-inch, 3/8-inch, No. 4, No. 8, No. 16, No. 30, No. 50, No. 100, and No. 200 sieves
- Bailey Method parameters (PCS, New CA Ratio, and New FA_c Ratio) describing aggregate packing characteristics
- Aggregate properties, including G_{sb} and FAA

This segmentation ensures that each model captures a consistent directional relationship between gradation change and volumetric response, avoiding bias that would arise from combining different trends into a single regression model.

6.1.1.2 Performance Prediction Models

For rutting and intermediate-temperature cracking, a consistent directional trend was observed across all 12 P-401 mix designs: coarsening the gradation generally improved performance. Therefore, all 12 P-401 mix designs with 12.5-mm NMAS were included in the performance prediction models.

The performance database consisted of the following:

- 42 rut-depth observations (APA or HWTT)

- 39 CT_{Index} observations (IDEAL-CT)
- 39 FI observations (I-FIT)

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

Each P-401 mix design was treated as an independent statistical group to account for shared material characteristics such as aggregate source, binder type, and production conditions. This grouping structure enabled separation of gradation-induced effects from baseline differences between aggregate and/or binder sources.

6.1.2 Step 2: Statistical Modeling and Validation

6.1.2.1 General Model

Both linear and nonlinear model forms were evaluated during the preliminary analysis. Although nonlinear models were explored to capture potential curvature in the relationships between gradation parameters and mixture responses, they did not provide meaningful improvements in predictive accuracy. In addition, nonlinear models typically require substantially larger datasets to ensure stable parameter estimation and to reduce the risk of overfitting. Given the moderate sample size in this study and the primary objective of developing practical prediction tools rather than complex mechanistic models, a linear regression framework was selected. The general model form is expressed in Equation (14):

$$\boldsymbol{Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi} + \alpha_{g(i)} + \varepsilon_i} \quad (14)$$

Where:

Y_i = response variable (VMA, rut depth, CT_{Index}, or FI) for mixture *i* prepared based on P-401 mix design *g*

X_{ji} = *j*-th predictor variable for mixture *i* (aggregate gradation sieve sizes and aggregate properties) with *j* = 1, 2, … , p

β_j = regression coefficient associated with predictor X_{ji}, with *j* = 1, 2, … , p

p = number of predictor variables

$\alpha_{g(i)}$ = group-specific baseline effect for P-401 mix design *g*, which is constant for all mixtures prepared based on P-401 mix design *g*

ε_i = random error term representing laboratory and material variability

In this study, a mixture corresponds to an individual gradation prepared from a selected P-401 mix design. Mixtures that share the same aggregate and binder sources within a given P-401 mix design are considered to belong to the same group. The group-specific term $\alpha_{g(i)}$

captures baseline differences among P-401 mix designs, including variations associated with aggregate sources, binder types, and production conditions.

Within each group, the regression coefficients quantify the relationships between the predictor variables (gradation characteristics and aggregate properties) and the corresponding mixture response. This allows the model to isolate the effects of gradation changes while accounting for systematic differences between P-401 mix designs.

Overall, this model structure provides a transparent and physically interpretable representation of the relationship between gradation characteristics and mixture properties, while maintaining sufficient flexibility to support predictions across different P-401 mix designs.

6.1.2.2 Treatment of Group Effects (Fixed-Effects Transformation)

Mixtures prepared based on the same P-401 mix design share baseline characteristics, including aggregate source, binder type, and production conditions. As a result, systematic differences in baseline volumetric and performance levels were observed across P-401 mix designs. If not addressed, these between-group shifts could confound the relationship between gradation adjustments and mixture responses.

To isolate the influence of gradation changes within each P-401 mix design, a fixed-effects (within-group) transformation was applied. This approach removes group-specific baseline effects and focuses the regression on within-group variation.

Group-centered variables were defined as:

$$\boldsymbol{\Delta Y_i = Y_i - \bar{Y}_g} \tag{15}$$

$$\boldsymbol{\Delta X_{ji} = X_{ji} - \bar{X}_{jg}} \tag{16}$$

Where:

ΔY_i = group-centered response of mixture *i,* representing the deviation of the response of mixture *i* from its group mean

ΔX_{ji} = group-centered predictor X_j of mixture *i* for *j* = 1,...,p

$\bar{Y}_g$ = mean response value for mixtures prepared based on P-401 mix design *g*

$\bar{X}_{jg}$ = mean value of predictor X_j for mixtures prepared based on P-401 mix design *g*

Substituting these definitions into the regression framework eliminates the group-specific term $\alpha_{g(i)}$, yielding:

$$\boldsymbol{\Delta Y_i = \beta_1 \Delta X_{1i} + \cdots + \beta_p \Delta X_{pi} + \tilde{\varepsilon}_i} \tag{17}$$

Where:

$\tilde{\varepsilon}_i$ = the transformed error term

This transformation ensures that model coefficients reflect the effect of gradation adjustments on the volumetric and performance properties of mixtures prepared based on a P-401 mix design, independent of baseline differences across P-401 mix designs (i.e., sources of aggregate and binder). Consequently, the resulting models are better suited for predicting how changes in gradation influence volumetric and performance properties under varying material and production conditions.

6.1.2.3 Robust Estimation

Ordinary least squares (OLS) regression minimizes the sum of squared residuals and is therefore sensitive to outliers. A small number of extreme observations can disproportionately influence coefficient estimates and predictive performance, particularly in moderate-sized datasets such as the one used in this study.

To improve model stability while retaining all observations, Huber-weighted regression was implemented. The Huber weighting scheme treats observations with small residuals similarly to OLS while downweighting observations with large residuals. The robust regression procedure was implemented iteratively in four steps.

In the first step, an initial regression model was fitted, and residuals were calculated using Equation (**18**).

$$\boldsymbol{r_i = \Delta Y_i - \widehat{\Delta Y}_i} \tag{18}$$

Where:

r_i = residual of mixture *i*
$\boldsymbol{\Delta Y_i}$ = measured group-centered response of mixture *i*
$\boldsymbol{\widehat{\Delta Y}_i}$ = predicted group-centered response of mixture *i*

In the second step, residuals from the initial regression fit were scaled using the median absolute deviation, as shown in Equation (19), which was converted to an estimate of the standard deviation using Equation (20). Scaled residuals were then computed using Equation (21).

$$\boldsymbol{MAD = median(|r_i - median(r_i)|)} \tag{19}$$

$$\boldsymbol{\widehat{\sigma} = \frac{MAD}{0.6745}} \tag{20}$$

$$u_i = \frac{r_i}{\hat{\sigma}} \quad (21)$$

The third step involved calculating the Huber weight using the tuning constant *c*, which is used to identify outliers, as shown in Equation (22). The tuning constant *c* was set to 1.345, yielding 95-percent efficiency when the data were normally distributed while still protecting against outliers [204, 205].

$$\omega_i = \begin{cases} 1 \; if \; |u_i| \leq c \\ \frac{c}{|u_i|} \; if \; |u_i| > c \end{cases} \quad (22)$$

Where:

ω_i = Huber weight function of mixture *i*
u_i = scaled residual of mixture *i*
c = tuning constant threshold (1.345)

In the final step, the regression model was run using weights $(\omega_1, \omega_2, \ldots, \omega_n)$ for each observation. It produced a new set of coefficients and residuals. This process of recalculating the tuning constant, updating weights, and re-estimating the regression was continued iteratively until the model coefficients converged. This robust estimation procedure reduces the influence of extreme observations while preserving the full dataset, resulting in more stable, reliable predictive models.

6.1.2.4 Cross-Validation Strategy

To evaluate predictive reliability and generalization capability, grouped K-fold cross-validation was performed. Mixtures prepared based on the same P-401 mix design share the same material sources and production conditions, so they were treated as one fold. This grouped validation approach prevents mixtures from the same source of aggregate and binder from appearing in both training and testing datasets, thereby avoiding overly optimistic performance estimates.

The number of folds corresponded to the number of P-401 mix design groups included in each model:

- Upper-limit VMA model: 2-fold cross-validation
- Lower-limit VMA model: 10-fold cross-validation
- Performance models: 12-fold cross-validation

In each iteration, one P-401 mix design group was held out as the validation set while the remaining groups were used for model training. This process was repeated until every group served once as the validation set.

Cross-validated predictions were aggregated across all folds, and model accuracy was evaluated on the combined prediction set using Equations (23), (24), and (25).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\Delta Y_i - \widehat{\Delta Y}_i)^2}{\sum_{i=1}^{n}(\Delta Y_i - \bar{\Delta Y})^2} \quad (23)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\Delta Y_i - \widehat{\Delta Y}_i)^2} \quad (24)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | \Delta Y_i - \widehat{\Delta Y}_i | \quad (25)$$

Where:

R^2 = coefficient of determination
RMSE = root mean squared error
MAE = mean absolute error
ΔY_i and $\widehat{\Delta Y}_i$ = the measured and predicted values of mixture *i*, respectively
$\bar{\Delta Y}$ = mean of the measured data
n = number of mixtures in the database

By structuring validation at the group level, the evaluation reflects the model's ability to predict mixture behavior for new mix designs rather than for mixtures that share the same baseline conditions. This approach provides a realistic and unbiased assessment of predictive performance for practical implementation.

6.1.2.5 Multicollinearity Considerations

Aggregate gradation percentages are inherently correlated because changes in material passing one sieve necessarily influence adjacent sieves. As a result, multicollinearity among predictor variables was expected in the regression framework.

Multicollinearity can inflate the variance of individual coefficient estimates and reduce their statistical significance, particularly when coefficients are interpreted independently. However, when the primary objective is prediction rather than inference on individual regression parameters, multicollinearity does not necessarily degrade predictive accuracy [206]. In such cases, overall model fit and cross-validated performance metrics provide a more meaningful evaluation of model quality than individual p-values. Accordingly, all sieve variables were retained in the regression models to preserve the hierarchical structure of aggregate gradation and ensure that predictions reflect the full gradation curve.

6.1.3 Step 3: Back-Transformation to Original Response Scale

Because the regression models were estimated using group-centered variables, predicted responses were initially obtained in centered form (i.e., $\widehat{\Delta Y_i}$). To provide practically interpretable outputs, predictions were transformed back to the original engineering scale by restoring the group mean, as shown in Equation (26):

$$\widehat{\boldsymbol{Y}}_{\boldsymbol{i}} = \widehat{\boldsymbol{\Delta Y}}_{\boldsymbol{i}} + \bar{\boldsymbol{Y}}_{\boldsymbol{g}} \tag{26}$$

Where:

$\hat{Y}_i$= calculated response on the original scale
$\Delta\hat{Y}_i$= predicted group-centered response
$\bar{Y}_g$= mean observed response within group g

This back-transformation yields calculated values of VMA, rut depth, CT_{Index}, and FI expressed in their engineering units. Agreement between calculated and measured responses was evaluated using scatter plots, concordance correlation coefficient (CCC), as shown in Equation (30), MAE, as shown in Equation (25), and mean absolute percentage error (MAPE), as shown in Equation (32). It should be noted that while the model was developed using a fixed set of gradations, the plot contains considerably more data points, as the calculated values were computed from all possible gradation combinations and orders.

$$\boldsymbol{r_c} = \frac{\boldsymbol{2rS_XS_Y}}{\boldsymbol{S_X^2 + S_Y^2 + (\bar{X} - \bar{Y})^2}} \tag{27}$$

Where:

r_c = CCC
r = Pearson correlation between calculated and measured responses, as shown in Equation (28)
S_X and S_Y = standard deviation of measured and calculated responses, respectively
$\bar{X}$ and $\bar{Y}$ = mean of measured and calculated responses, respectively

$$\boldsymbol{r} = \frac{\boldsymbol{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}}{\boldsymbol{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}} \tag{28}$$

(29)

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \left| \frac{Y_i - X_i}{Y_i} \right|$$

Where:

Y_i and X_i = the calculated and measured variables, respectively
n = number of mixtures

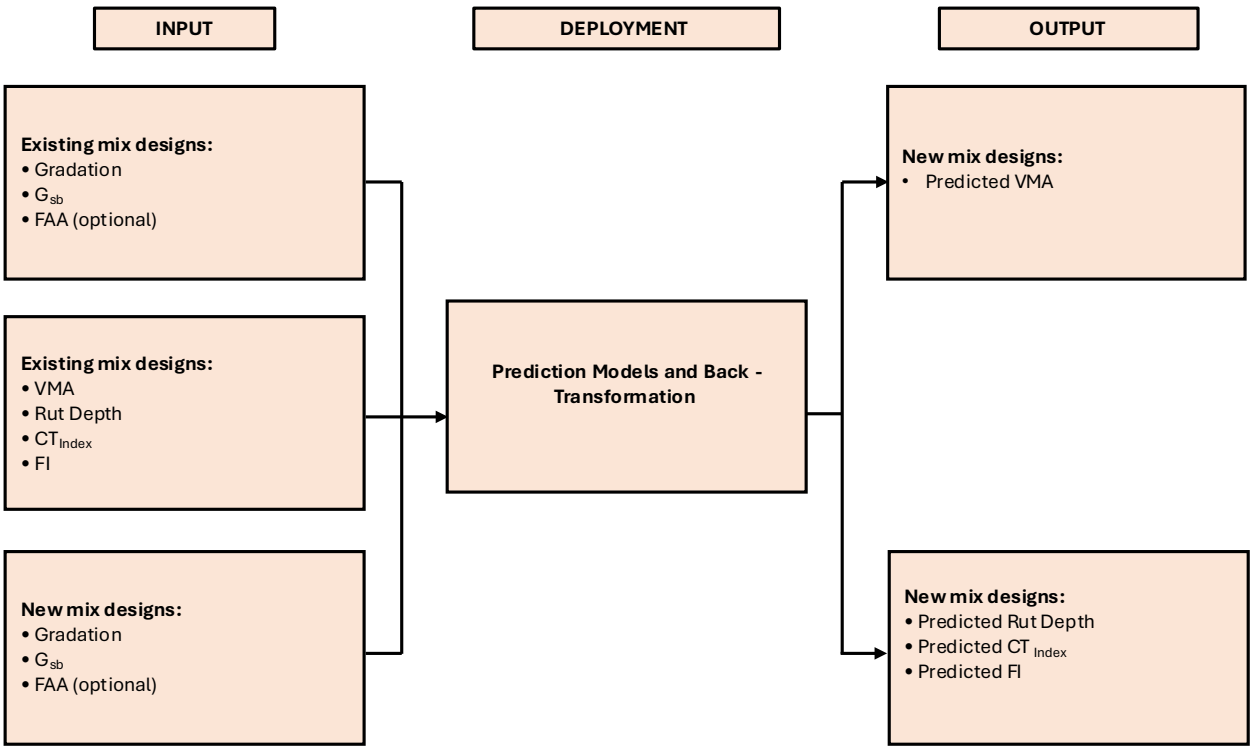
The CCC was selected because it evaluates both correlation (precision) and deviation from the 45-degree line (accuracy), thereby quantifying overall agreement between measured and calculated values. MAE provides an interpretable measure of average prediction error, while MAPE expresses error relative to the magnitude of the response. Because baseline response levels differed across P-401 mix designs, CCC was computed within each group rather than across the pooled dataset. This approach avoids artificial inflation or deflation of agreement caused by between-group mean shifts and provides a more accurate assessment of predictive performance within each design context.

6.1.4 Web-Based Deployment

The final models were implemented in a web-based tool (Figure 43). The tool uses the fixed-effects regression framework and back-transformation procedure.

Users first enter gradation, aggregate properties, and mixture characteristics for tested trial mixtures within a project to establish baseline group relationships. Next, users input the gradation and aggregate properties of a new trial mixture. The tool applies the appropriate volumetric or performance model and generates calculated VMA, rut depth, CT_{Index}, and FI.

This deployment framework enables practitioners to evaluate the expected impact of gradation adjustments before laboratory verification, supporting data-driven refinement of P-401 mix designs. The models are applied under the assumption that the mixtures being compared are produced with the same aggregate source and binder content.



Source: AAPTP

Figure 43. Workflow of the Prediction Tool

6.2 Prediction Model for Volumetric Properties

Changes in aggregate gradation can influence VMA, depending on whether the gradation is near the upper or lower P-401 limits. Separate regression models were developed to describe the relationship between gradation parameters and VMA based on their position relative to the MDL, as detailed in the following sections.

6.2.1 First Prediction Model for Volumetric Properties

The first volumetric prediction model was developed for FG mixtures with gradation curves located above the MDL. As discussed in Chapter 5, these mixtures exhibited a consistent trend: moving the gradation coarser toward the MDL reduced VMA.

For this model, Bailey Method parameters for FG mixtures (PCS, New CA Ratio, and New FA_c Ratio) were selected as predictors rather than individual sieve percentages. This selection was made for two primary reasons:

- **Mix Type:** All mixtures included in this model were FG. Bailey Method parameters were consistently defined and physically meaningful across all gradation designs.
- **Sample Size Consideration:** The model was developed using gradations derived from two upper-limit P-401 mix designs, comprising 18 observations. Including individual sieve percentages would have introduced numerous predictors, which, given the available sample size, could have increased the risk of overfitting and

reduced the model's stability. Using Bailey Method parameters reduced dimensionality while maintaining the critical packing structure of the gradation.

This approach improves model interpretability and statistical robustness under limited sample conditions.

6.2.1.1 Model Formulation

The resulting regression model, estimated using the group-centered robust regression framework described in Section 6.1, is shown in Equation (30):

$$\Delta \boldsymbol{VMA} = \mathbf{0.2158}\Delta \boldsymbol{PCS} + \mathbf{2.32}\Delta \boldsymbol{New\ CA\ Ratio} - \mathbf{1.55}\Delta \boldsymbol{New\ FA_c\ Ratio} \quad (30)$$

Where:

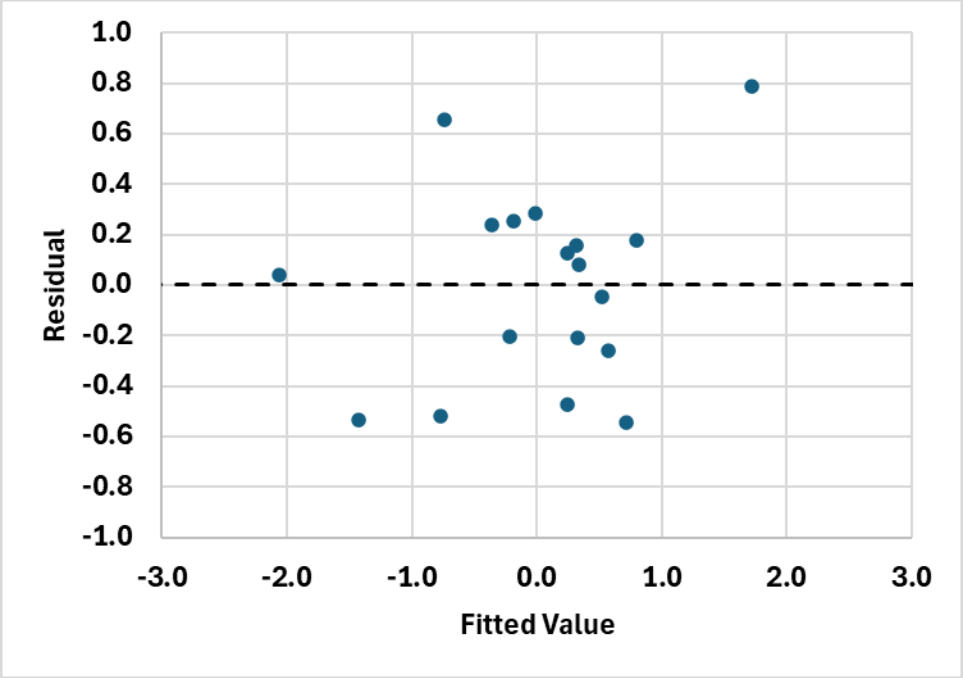
ΔVMA = group-centered value (i.e., change) of VMA

$\Delta PCS, \Delta New\ CA\ Ratio, \Delta New\ FA_c\ Ratio$ = group-centered value (i.e., changes) of Bailey Method parameters for FG mixtures

The positive coefficient for ΔPCS and $\Delta New\ CA\ Ratio$ indicates that increases in CA structure parameters are associated with increases in VMA, while the negative coefficient for ΔFAc_{New} suggests that increased fine aggregate packing reduces VMA. These relationships are consistent with aggregate packing principles and the trends observed experimentally.

6.2.1.2 Model Evaluation

Figure 44 presents the distribution of residuals in relation to the fitted values of ΔVMA from the first volumetric prediction model. The random scatter of the residuals around the fitted values indicates that a linear model was appropriate and effectively captured the relationship between changes in gradation and ΔVMA. Additionally, the vertical spread of the residuals remained relatively consistent across the central cluster of fitted values, suggesting that the important assumption of homoscedasticity was generally satisfied, with no significant evidence of heteroscedasticity.



Source: AAPTP

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

6.2.1.3 Cross-Validation

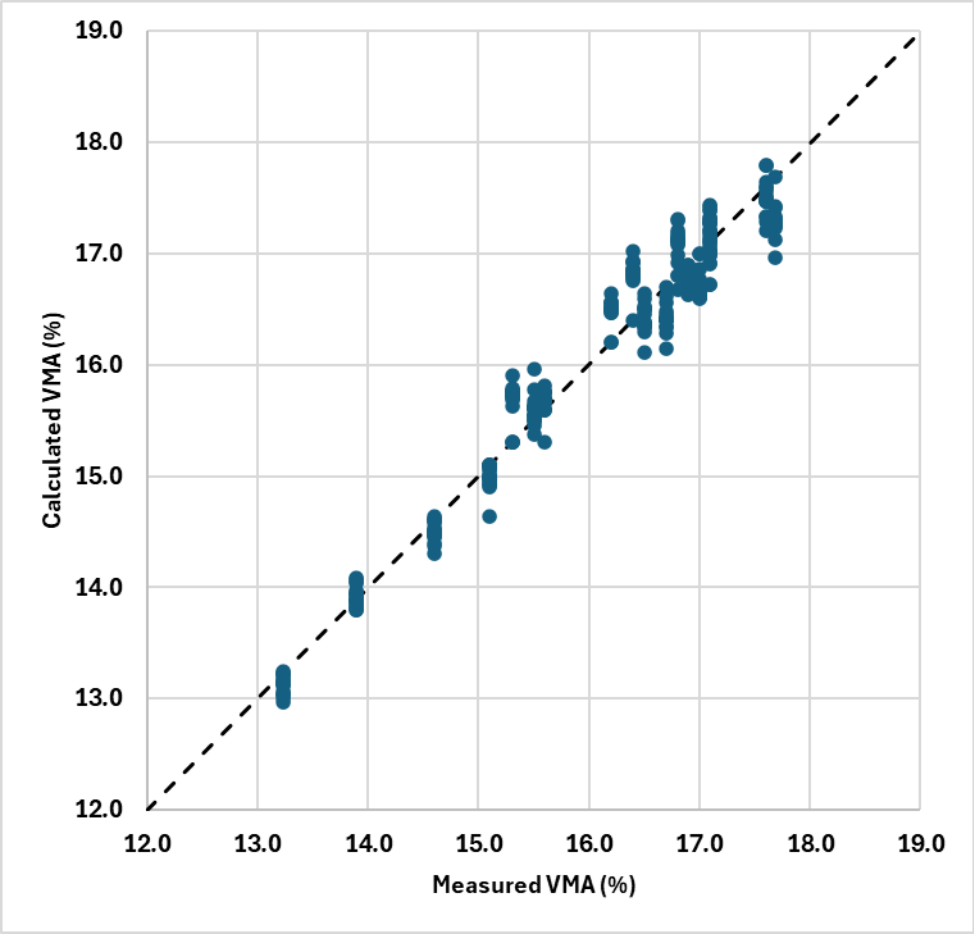
Model performance was evaluated using grouped two-fold cross-validation, with each P-401 mix design serving once as the validation set. Performance metrics are summarized in Table 40. The cross-validated R^2 value of 0.86 indicates that the model accounts for 86 percent of the variation within the group. The MAE of 0.3 percent and the RMSE of 0.4 percent indicate that typical prediction errors are relatively small compared to typical VMA magnitudes, which range from approximately 14 to 18 percent. This suggests that the model has strong predictive precision on the centered scale.

Table 40. Statistical Metrics of Predicted and Measured ΔVMA for First Volumetric Prediction Model

MAE	RMSE	K-fold R^2
0.31	0.40	0.86

6.2.1.4 Back-Transformation and Comparison to Measured VMA

Figure 45 presents measured versus predicted VMA values after back-transformation to the original scale. Most observations cluster closely around the line of equality, indicating strong agreement across the two P-401 mix designs. The average CCC was 0.96, demonstrating very strong agreement between measured and calculated VMA values. The MAE on the original scale was 0.2 percent VMA, indicating that calculated values deviated from measured values by only 0.2 percentage points on average, which is within the within-laboratory repeatability of VMA measurement of approximately 0.7 percent [207].



Source: AAPTP

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

6.2.1.5 Model Limitations

Although the model shows strong statistical performance, it is important to note that it was created using 18 observations from two P-401 mix designs. While cross-validation indicates consistent results between the two groups, the limited variety of source materials limits its broader applicability. Therefore, the first volumetric prediction model should be used with caution.

6.2.2 Second Prediction Model for Volumetric Properties

The second volumetric prediction model was developed for mixtures with S-shaped CG gradations and with fine-fraction sieve sizes below the MDL. These mixtures showed a consistent trend: coarser gradation generally increased VMA.

This model was developed from 52 observations across 10 P-401 mix designs. Since this is a larger dataset, individual sieve percentages were used directly as predictor variables to capture detailed gradation effects.

6.2.2.1 Model Formulation

The OLS regression equation of the second VMA prediction model is shown in Equation (31).

$$\begin{aligned}\Delta\mathbf{VMA} = {} & \mathbf{0.0282} + \mathbf{0.314}\Delta\mathbf{No.\,1/2} - \mathbf{0.0271}\Delta\mathbf{No.\,3/8} - \mathbf{0.0036}\Delta\mathbf{No.\,4} \\ & - \mathbf{0.044}\Delta\mathbf{No.\,8} + \mathbf{0.099}\Delta\mathbf{No.\,16} - \mathbf{0.324}\Delta\mathbf{No.\,30} + \mathbf{0.969}\Delta\mathbf{No.\,50} \\ & - \mathbf{0.430}\Delta\mathbf{No.\,100} - \mathbf{1.074}\Delta\mathbf{No.\,200} + \mathbf{87.7}\Delta\mathbf{G_{sb}} + \mathbf{0.135}\Delta\mathbf{FAA}\end{aligned} \quad (31)$$

Where:

$\Delta\mathrm{VMA}$ = group-centered value of VMA

$\Delta\mathrm{No.\,1/2}$, $\Delta\mathrm{No.\,3/8}$ … $\Delta\mathrm{No.\,200}$ = group-centered values of percentage of aggregate passing at No. 1/2, No. 3/8 … No. 200 sieves

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

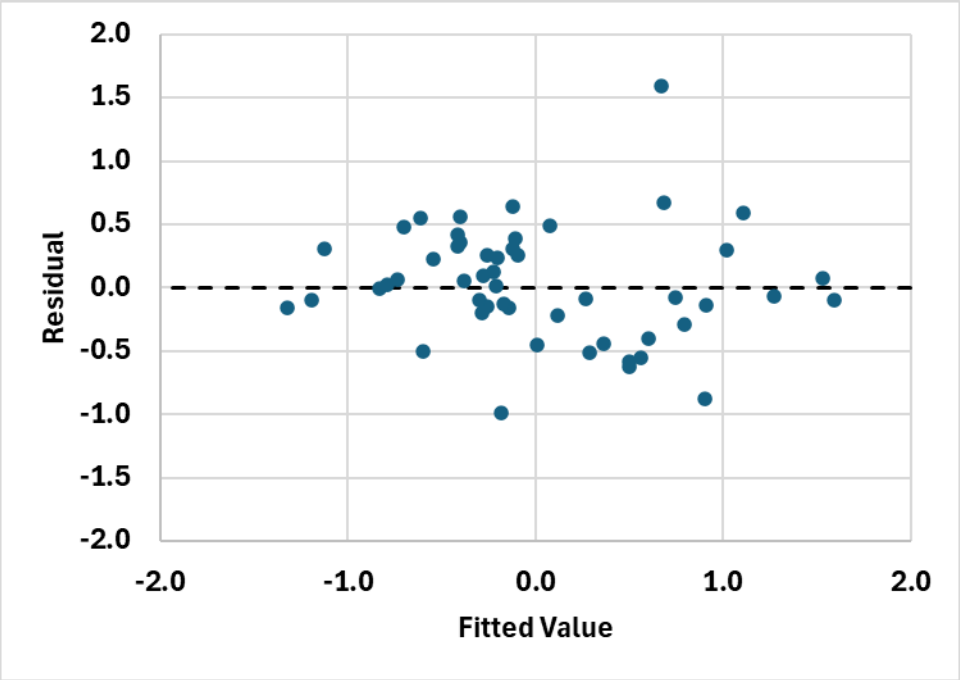
$\Delta\mathrm{FAA}$ = group-centered value of fine aggregate angularity of aggregate blend

Because sieve percentages are inherently correlated, individual coefficients should not be interpreted independently. Instead, the model should be interpreted as a combined representation of the full gradation profile.

The large coefficient associated with ΔG_{sb} reflects VMA's physical sensitivity to aggregate specific gravity, which directly influences volumetric calculations.

6.2.2.2 Model Evaluation

Figure 46 illustrates the distribution of residuals in relation to the fitted values of ΔVMA in the second volumetric prediction model. The random scatter of the residuals around the fitted values indicates that a linear model was appropriate and effectively captured the relationship between changes in gradation and ΔVMA. Additionally, the vertical spread of the residuals was relatively uniform throughout the central cluster of fitted values. This suggests that the key assumption of homoscedasticity was generally satisfied, with no significant evidence of heteroscedasticity.



Source: AAPTP

Figure 46. Residuals and Fitted Values of ΔVMA in Second Volumetric Prediction Model

6.2.2.3 Cross-Validation

Model performance was evaluated using grouped 10-fold cross-validation, with each P-401 mix design serving once as the validation set. Performance metrics are summarized in Table 41.

The cross-validated R^2 of 0.69 indicates that the model explains approximately 69 percent of within-group variation in ΔVMA. While lower than the first model, this level of explanatory power is reasonable given the increased variability across 10 distinct P-401 mix designs and the greater structural diversity of the lower-limit gradations.

The MAE (0.3 percent) and RMSE (0.4 percent) demonstrate that the typical prediction error remains small relative to typical VMA values (approximately 14–18 percent), indicating acceptable predictive precision.

Table 41. Statistical Metrics of Predicted and Measured ΔVMA for Second Volumetric Prediction Model

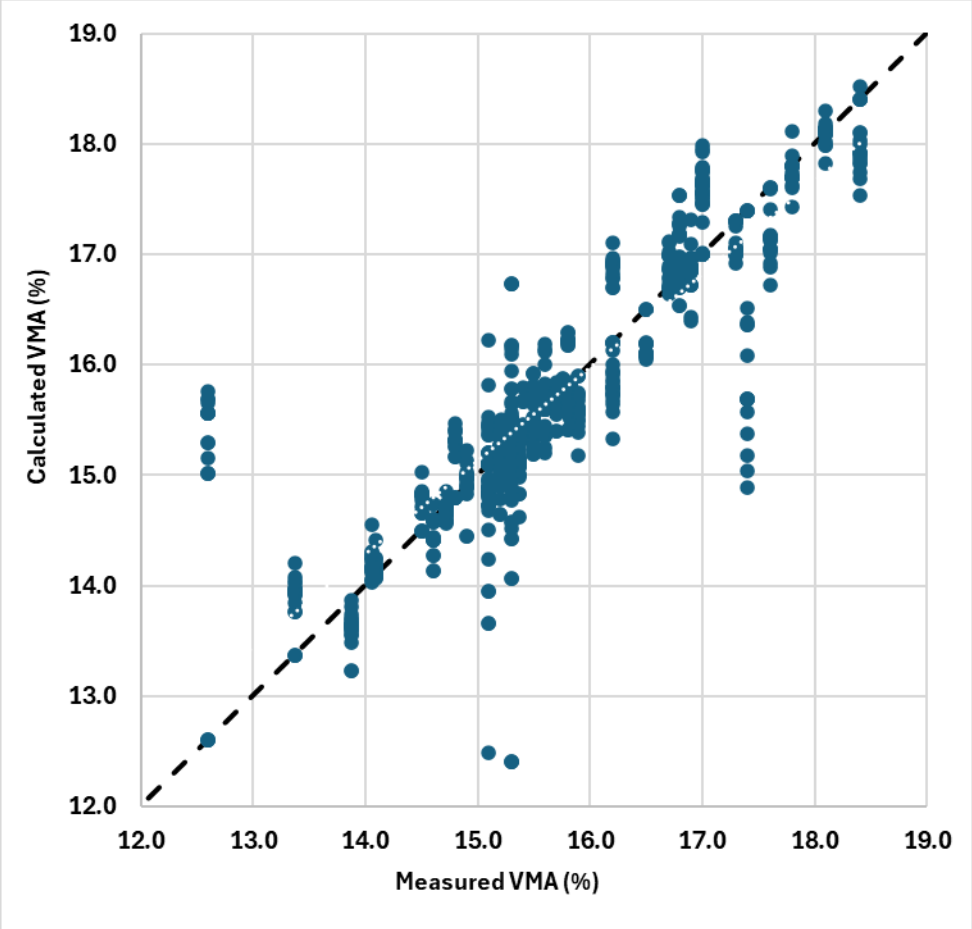
MAE	RMSE	K-fold R^2
0.30	0.40	0.69

6.2.2.4 Back-Transformation and Comparison to Measured VMA

Figure 47 presents measured versus calculated VMA values after back-transformation to the original scale. Most observations cluster around the line of equality, demonstrating that the model captures both the direction and magnitude of VMA changes across diverse P-401 mix designs.

The average CCC was 0.81, indicating strong agreement between measured and calculated VMA values. The MAE on the original scale was 0.4-percent VMA, indicating that the calculated values deviated from the measured values by 0.4 percentage points on average. This falls within the within-laboratory repeatability of VMA measurement of approximately 0.7 percent [207], indicating that model's predictive accuracy is considered acceptable.

Several observations appear visually distant from the line of equality. This outcome is expected because the robust Huber procedure down-weights influential observations without removing them. Consequently, these points remain visible in the plot but do not dominate the regression fit.



Source: AAPTP

Figure 47. Measured and Calculated VMA from Second Volumetric Prediction Model

6.2.2.5 Second Volumetric Prediction Model Without Fine Aggregate Angularity

Because the changes in FAA are relatively small among gradations within each group of P-401 mix designs, as discussed in Section 4.2, $\Delta\mathrm{FAA}$ does not appear to significantly contribute to explaining the variation in ΔVMA. Therefore, a reduced-form model excluding FAA was developed as shown in Equation (32).

(32)

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

As shown in Table 42, removing FAA did not materially change MAE, RMSE, or cross-validated R^2 (0.69). Therefore, the reduced model provides comparable predictive performance with fewer parameters and may be preferred when FAA data are unavailable.

Table 42. Statistical Metrics of Predicted and Measured ΔVMA for 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 rut depth prediction model was developed using the same framework described in Section 6.1. The objective was to quantify the relationship between gradation adjustments and changes in rutting resistance while using the same asphalt binder when adjusting gradations for each airport mix design.

Because rutting behavior exhibited a consistent directional trend across the 12 P-401 mix designs, in which coarsening the gradation generally reduced rut depth, all 12 mix designs were included in the model development.

6.3.1 Model Formulation

The OLS regression equation of the rutting prediction model is shown in Equation (33).

(33)

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

Where:

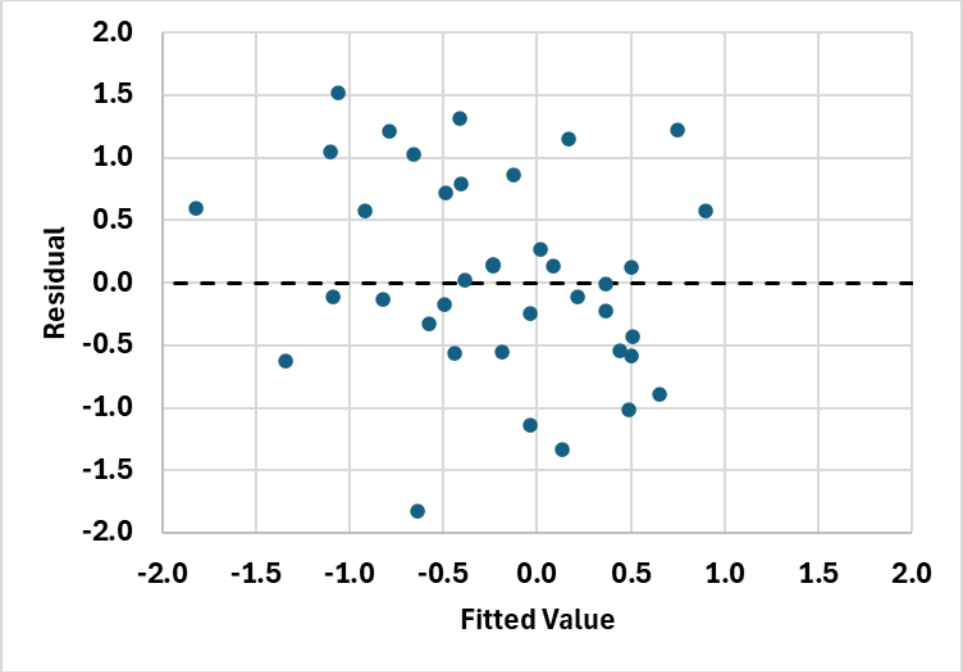
$\Delta \text{Rut depth}$ = group-centered value of rut depth
$\Delta No.\,1/2$, $\Delta No.\,3/8$ … $\Delta No.\,200$ = group-centered value of percentage of aggregate passing at No. 1/2, No. 3/8 … No. 200 sieves
ΔG_{sb} = group-centered value of specific gravity of aggregate blend
ΔFAA = group-centered value of FAA of the blend

Due to the correlation between sieve variables, the individual magnitudes of coefficients should not be interpreted in isolation. Instead, the model captures the combined effect of the gradation profile on rutting behavior.

The negative coefficient associated with the variable reflects the impact of aggregate density on the stiffness of the mixture and its resistance to rutting. Collectively, the signs of the sieve coefficients indicate that a shift toward coarser gradations is generally linked to a reduction in rut depth, which is consistent with the experimental observations presented in Chapter 5.

6.3.2 Model Evaluation

Figure 48 shows the distribution of residuals in relation to the fitted values of ΔRut. The random scatter of residuals indicates that a linear model was appropriate and effectively captured the relationship between changes in gradation and ΔRut. Additionally, the vertical spread of residuals remained relatively uniform across the central cluster of fitted values, suggesting that the crucial assumption of homoscedasticity was generally satisfied without any signs of significant heteroscedasticity.



Source: AAPTP

Figure 48. Residuals and Fitted Values of ΔRut Depth

6.3.3 Cross-Validation

Grouped 12-fold cross-validation was used, with each P-401 mix design serving once as the validation set. Performance metrics are summarized in Table 43.

Table 43. Statistical Metrics of Predicted and Measured ΔRut Depth

MAE	RMSE	K-fold R^2
0.67	0.81	0.39

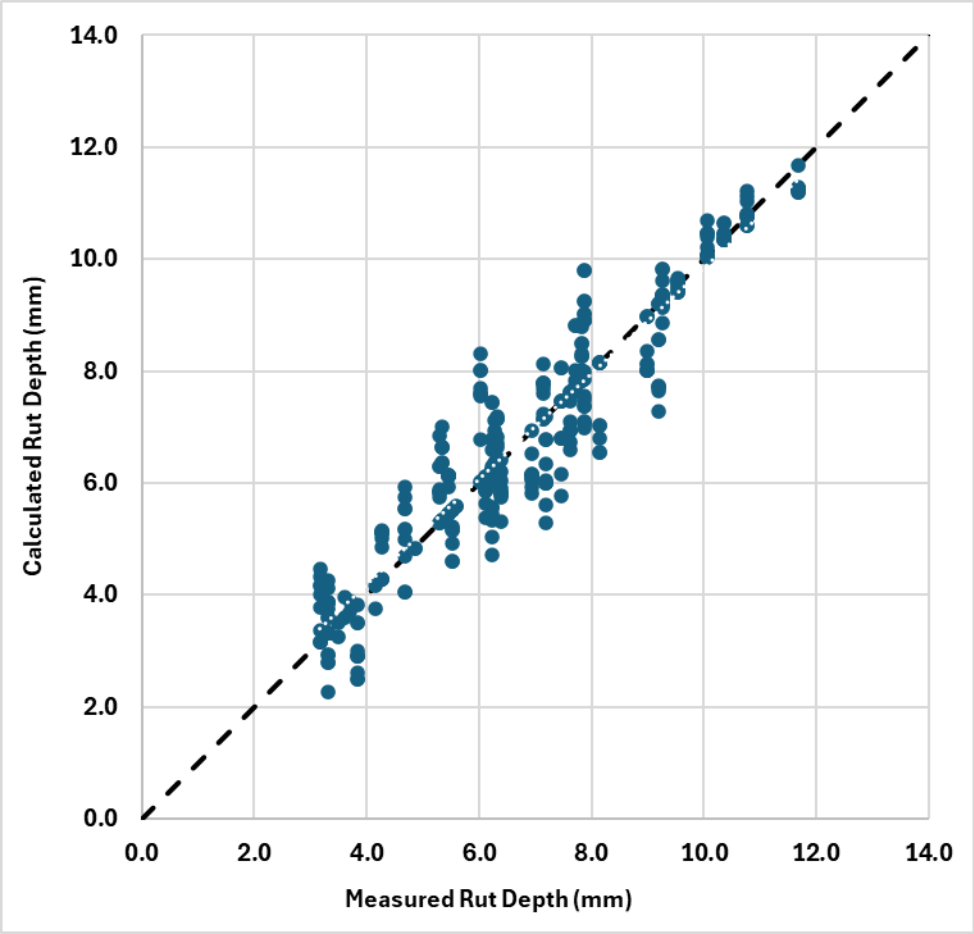
The cross-validated R^2 of 0.39 indicates that approximately 39 percent of the within-group variability in rut depth was explained by gradation and aggregate properties. While lower than the volumetric models, this level of explanatory power is reasonable given the higher experimental variability and mechanical complexity associated with rutting performance.

The MAE (0.67 mm) and RMSE (0.81 mm) indicate that the typical prediction error is less than 1 mm. Considering that rut depth values typically range from approximately 4 to 10 mm, this magnitude of error is acceptable for screening-level prediction during mix design adjustment.

6.3.4 Back-Transformation and Comparison to Measured Rut Depth

Figure 49 presents measured versus predicted rut depth after back-transformation to the original scale. The data cluster around the line of equality, though with greater dispersion compared to the volumetric models.

The average CCC was 0.53, indicating moderate agreement between measured and calculated rut depths. The MAE of 0.68 mm confirms that calculations deviate from measured values by less than 1 mm on average. The lower CCC relative to VMA models reflects the greater inherent variability of rutting performance rather than systematic model bias.



Source: AAPTP

Figure 49. Measured and Calculated Rut Depth

6.3.5 Rutting Prediction Model Without Fine Aggregate Angularity

Another version of the rut depth prediction model was developed without FAA variables, as shown in Equation (34). Table 44 provides the statistical metrics of the model without the FAA variable. Removing the FAA variable slightly improved cross-validated R^2 from 0.39 to 0.45 and reduced prediction error. This result suggests that the FAA does not provide additional explanatory power for rutting beyond sieve distribution and aggregate density. The reduced model is therefore preferred for consistency and improved predictive stability.

(34)

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

Table 44. 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 Models

Intermediate-temperature cracking resistance was evaluated using both CT_{Index} and FI. Because cracking performance reflects complex interactions among aggregate structure, binder properties, and volumetric characteristics, greater variability was expected compared to volumetric responses. The objective of this section was to develop predictive relationships that quantify how gradation adjustments affect cracking resistance in P-401 mix designs, while using the same asphalt binder when adjusting gradations for each airport mix design. The same fixed-effects, robust regression, and grouped cross-validation framework described in Section 6.1 was applied.

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

This subsection develops a prediction model for intermediate-temperature cracking resistance using CT_{Index} as the response variable. The model quantifies the relationship between gradation adjustments and CT_{Index} within P-401 mix designs, providing a practical tool for evaluating cracking sensitivity during mix design refinement.

6.4.1.1 Model Formulation

The OLS regression equation of the prediction model based on CT_{Index} is shown in Equation (35).

$$\Delta CT_{Index} = 0.04 + 4.05\Delta No.\frac{1}{2} - 4.66\Delta No.\frac{3}{8} + 27.68\Delta No.4 - 43.12\Delta No.8 + 16.1\Delta No.16 + 16.5\Delta No.30 - 32.2\Delta No.50 + 29.1\Delta No.100 - 17.0\Delta No.200 + 3050\Delta G_{sb} - 15.22\Delta FAA \quad (35)$$

Where:

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

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

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

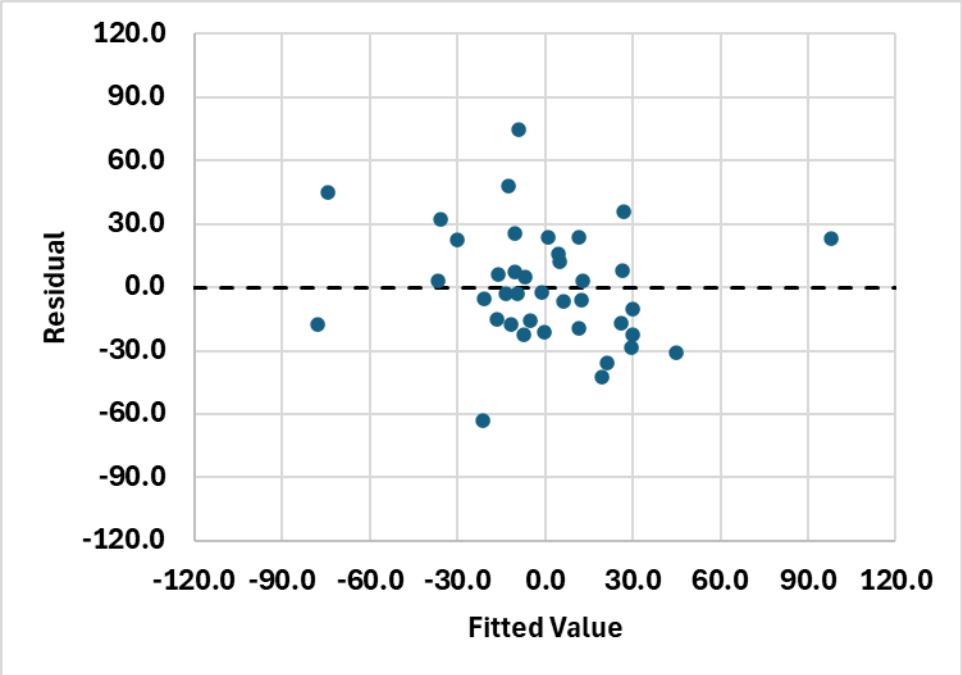
ΔFAA = group-centered FAA of blend

Because sieve percentages are inherently correlated, individual coefficients should not be interpreted independently. Instead, the model captures the combined structural effect of gradation shifts on cracking resistance.

The large coefficient associated with ΔG_{sb} reflects the sensitivity of cracking performance to aggregate packing and volumetric structure. The mixture-scale interpretation remains consistent with the experimental observations in Chapter 5: gradation adjustments influence cracking resistance through structural rearrangement of the aggregate skeleton.

6.4.1.2 Model Evaluation

Figure 50 presents the distribution of residuals against the fitted values of ΔCT_{Index}. As shown, the random scatter of residuals relative to the fitted values indicates that a linear model was appropriate and effectively captured the relationship between changes in gradation and ΔCT_{Index}. Additionally, the vertical spread of the residuals remained relatively uniform throughout the central cluster of fitted values, suggesting that the critical assumption of homoscedasticity was generally met, with no evidence of severe heteroscedasticity.



Source: AAPTP

Figure 50. Residuals and Fitted Values of ΔCT_{Index}

6.4.1.3 Cross-Validation

Grouped 12-fold cross-validation was performed using the P-401 mix design as the fold structure. Performance metrics are summarized in Table 45. The cross-validated R^2 of 0.62 indicates that approximately 62 percent of the within-group variability in CT_{Index} was

explained by gradation and aggregate properties. Given the mechanical complexity and test variability associated with cracking resistance, this level of explanatory power is considered strong for a performance prediction model.

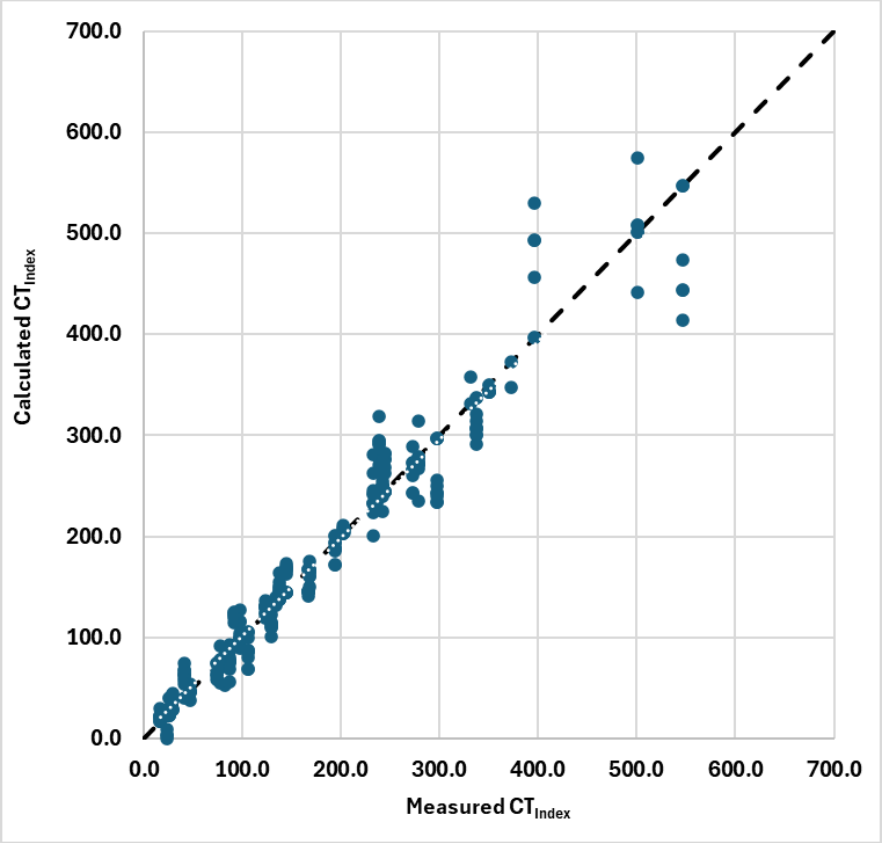
The MAE of 16 and RMSE of 20 indicate the typical magnitude of prediction error. For context, CT_{Index} values in this dataset span a wide numerical range; therefore, absolute error must be interpreted relative to test variability.

Table 45. Statistical Metrics of Predicted and Measured ΔCT_{Index}

MAE	RMSE	K-fold R^2
16.0	20.0	0.62

6.4.1.4 Back-Transformation and Comparison to Measured CT_{Index}

Figure 51 presents measured versus calculated CT_{Index} after back-transformation to the original scale. The average CCC was 0.60, indicating moderate agreement between measured and calculated values. The MAPE was 17.7 percent. Importantly, the observed MAPE is similar to the typical within-laboratory coefficient of variation for CT_{Index} testing, which is approximately 20 percent [208]. This comparison indicates that the model's predictive error falls within the inherent variability of the test method. In practical terms, the model introduces no more variability than is expected from repeated laboratory measurements. Accordingly, the model can be considered reliable for screening and comparative evaluation of gradation adjustments within mix design.



Source: AAPTP

Figure 51. Measured and Calculated CT_{Index}

6.4.1.5 CT_{Index} Prediction Model Without Fine Aggregate Angularity

Another version of the CT_{Index} prediction model was developed without the FAA variable, as shown in Equation (36). Table 46 provides the statistical metrics of the model without the FAA variable. Removing the FAA reduced the cross-validated R^2 from 0.62 to 0.50 and increased the prediction error. This indicates that the FAA provides additional explanatory value for CT_{Index} prediction beyond sieve distribution and aggregate density. Therefore, the full model, including FAA, may be preferred for implementation.

(36)

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

Table 46. 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

This subsection presents the prediction model for intermediate-temperature cracking resistance using FI as the response variable. Similar to CT_{Index}, FI reflects fracture behavior under monotonic loading; however, FI exhibited smaller within-group variation after group-centering, which influenced model diagnostics on the centered scale.

6.4.2.1 Model Formulation

The OLS regression equation for the FI prediction model is shown in Equation (37).

$$\begin{aligned}\Delta\mathbf{FI} = {} & \mathbf{0.016} + \mathbf{0.625}\Delta\mathbf{No.}\frac{\mathbf{1}}{\mathbf{2}} - \mathbf{0.431}\Delta\mathbf{No.}\frac{\mathbf{3}}{\mathbf{8}} + \mathbf{0.044}\Delta\mathbf{No.4} - \mathbf{0.985}\Delta\mathbf{No.8} \\ & - \mathbf{0.18}\Delta\mathbf{No.16} + \mathbf{3.22}\Delta\mathbf{No.30} - \mathbf{4.61}\Delta\mathbf{No.50} + \mathbf{4.34}\Delta\mathbf{No.100} \\ & - \mathbf{3.01}\Delta\mathbf{No.200} + \mathbf{44.6}\Delta\mathbf{G_{sb}} + \mathbf{0.03}\Delta\mathbf{FAA}\end{aligned} \quad (37)$$

Where:

ΔFI = group-centered FI

$\Delta No.\ 1/2$, $\Delta No.\ 3/8$ … $\Delta No.\ 200$ = group-centered percentage of aggregate passing at No. 1/2, No. 3/8 … No. 200 sieves

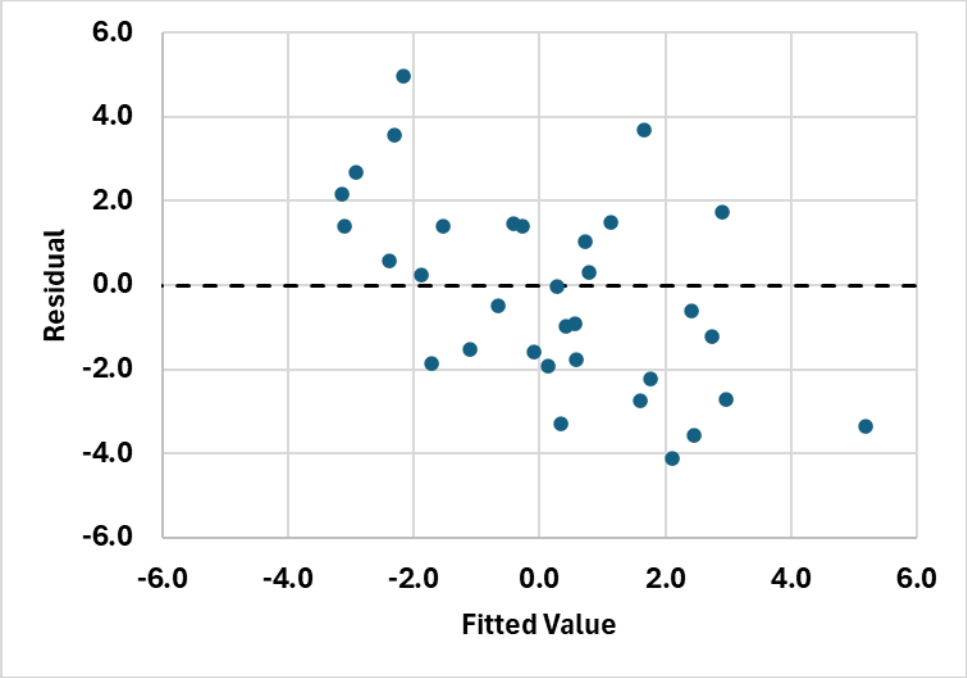
ΔG_{sb} = group-centered specific gravity of aggregate blend

ΔFAA = group-centered FAA of aggregate blend

As with other performance models, individual coefficients should not be interpreted independently due to multicollinearity among sieve variables. The model captures the combined structural effect of gradation adjustments on FI.

6.4.2.2 Model Evaluation

Figure 52 presents residuals versus fitted values of ΔFI. The residuals display no strong nonlinear trend, supporting the linear model assumption. Residual dispersion appears relatively uniform across the fitted range, with variability consistent with the known scatter of FI testing. 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.



Source: AAPTP

Figure 52. Residuals and Fitted Values of ΔFI

6.4.2.3 Cross-Validation

Cross-validated performance metrics for the ΔFI model are shown in Table 47. The cross-validated R^2 of approximately zero reflects the limited within-group variability in FI after removal of group means. In other words, once airport-specific baseline differences were removed, the remaining ΔFI values were small and dominated by experimental noise. Under these conditions, even modest prediction errors result in very low R^2 values.

This does not imply that the model lacks practical predictive value. Rather, it indicates that gradation adjustments produced relatively small within-group FI changes compared to inherent test variability.

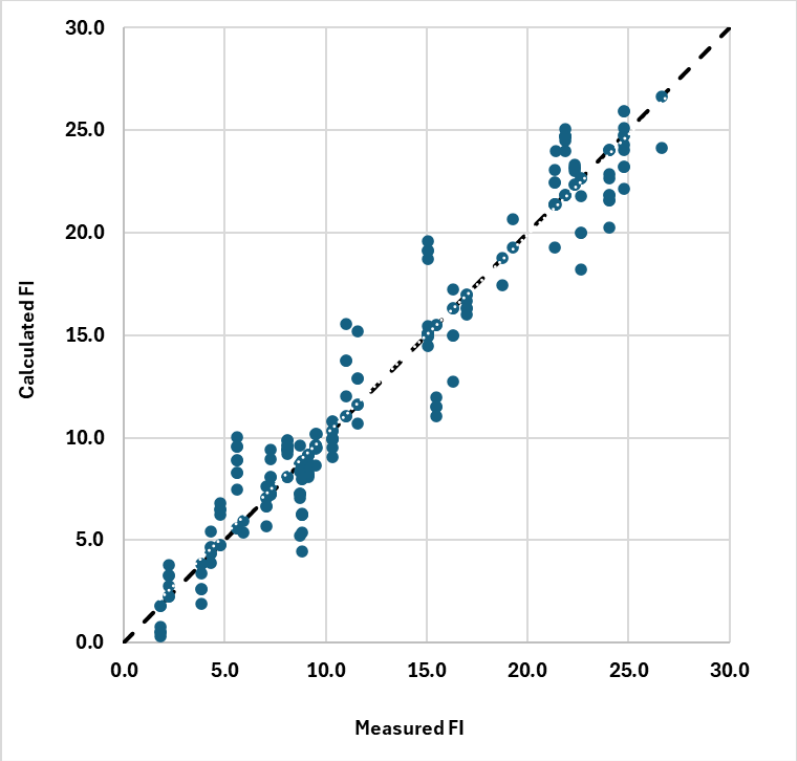
Table 47. Statistical Metrics of Predicted and Measured ΔFI

MAE	RMSE	K-fold R^2
1.9	2.4	0.0

6.4.2.4 Back-Transformation and Comparison to Measured FI

Figure 53 compares measured and calculated FI values after back-transformation to the original scale. The CCC of 0.45 indicates poor to moderate agreement between measured and predicted values, suggesting the presence of systematic bias in absolute FI predictions. Nevertheless, the results show a moderate positive linear relationship, indicating that the model captures the overall trend in FI values reasonably well even though the predicted values do not closely match the measured values in absolute terms.

The MAPE of 17.0 percent is lower than the typical within-laboratory coefficient of variation for FI testing, which is approximately 25 percent [125]. This comparison suggests that the prediction error of the model falls within the inherent experimental variability of the FI test method. The model is therefore more suitable for screening and comparative evaluation rather than precise estimation of small within-group FI shifts.



Source: AAPTP

Figure 53. Measured and Calculated FI

6.4.2.5 FI Prediction Model Without Fine Aggregate Angularity

Another version of the FI prediction model was developed without FAA variables, as shown in Equation (38). Table 48 provides the statistical metrics of the model without the FAA variable. Removing FAA slightly reduced MAE and RMSE, with no impact on R^2. Given the minimal contribution of FAA to predictive accuracy and the desire for model parsimony, the reduced model may be preferred for implementation.

$$\begin{aligned}\Delta FI = {} & 0.017 + 0.637\Delta No.1/2 - 0.434\Delta No.3/8 + 0.049\Delta No.4 - 0.986\Delta No.8 \\ & - 0.20\Delta No.16 + 3.25\Delta No.30 - 4.64\Delta No.50 + 4.38\Delta No.100 \\ & - 3.04\Delta No.200 + 45.5\Delta Gsb\end{aligned} \quad (38)$$

Table 48. Statistical Metrics of Predicted and Measured ΔFI from Prediction Model Without FAA

MAE	RMSE	K-fold R^2
1.6	2.0	~0.00

6.5 Model Performance Summary

Table 49 summarizes the cross-validated performance metrics on the group-centered scale, including K-fold R^2, RMSE, and MAE, for all prediction models with and without the FAA term. The two VMA models demonstrated the strongest predictive performance, followed by the CT_{Index} model, the rut depth model, and the FI model.

The influence of FAA varied across models. Including FAA improved the predictive performance of the CT_{Index} model, whereas removing FAA slightly improved the rut depth model. For the second VMA model and the FI model, FAA had little effect on predictive accuracy.

Table 49. Cross-Validated Performance Metrics of Prediction Models (Group-Centered Scale)

Model	No. of Mix Designs	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%	–	–	–
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 50 summarizes prediction accuracy on the original response scale after back-transformation, using CCC together with MAE or MAPE. Consistent with the group-centered results, the VMA models showed the strongest agreement between measured and calculated values. The CT_{Index} model demonstrated moderate agreement, followed by the rut depth model, while the FI model showed the lowest agreement.

Table 50. Back-Transformed Prediction Accuracy of Models with FAA

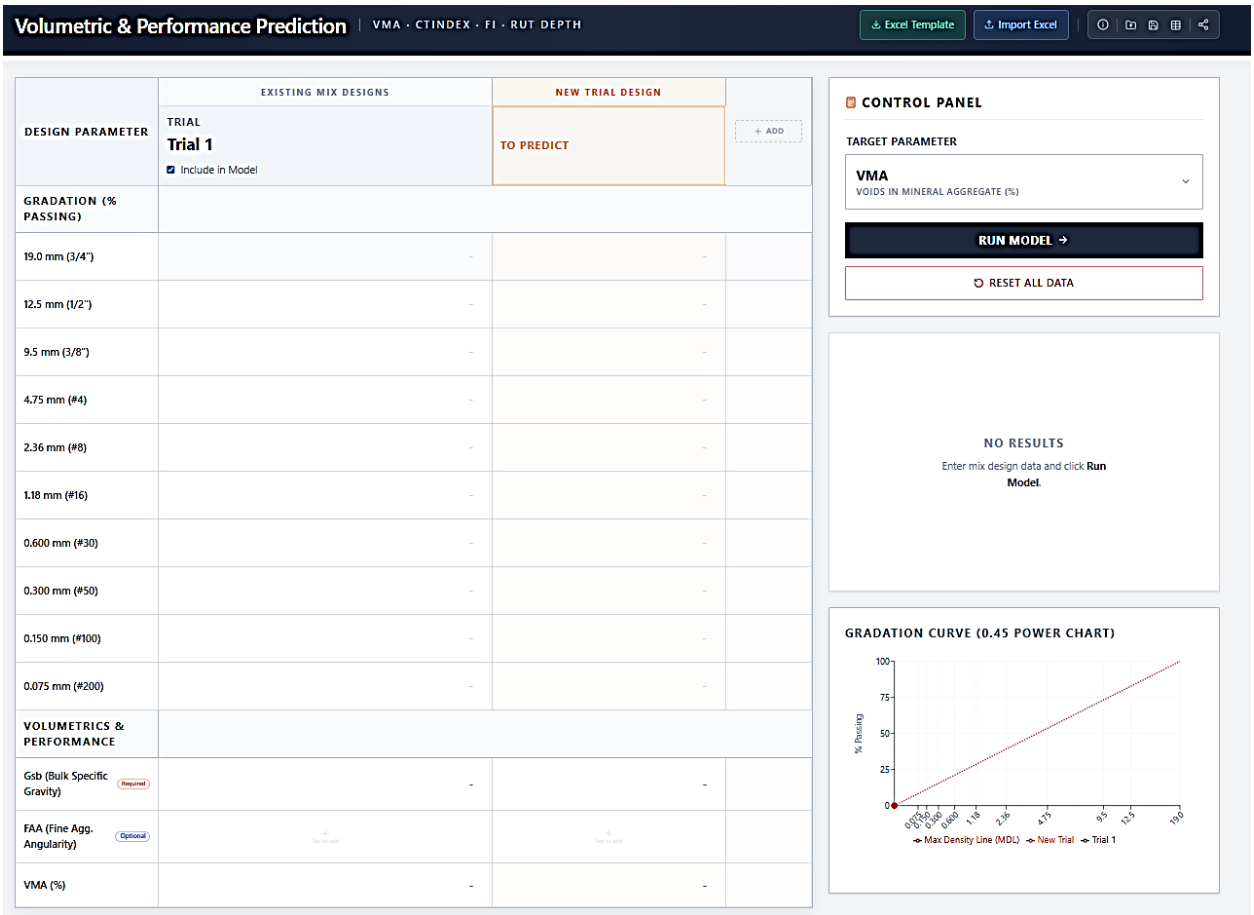
Model	Number of Mix Designs	CCC	MAE/MAPE
VMA First Model	2	0.96	MAE = 0.20%
VMA Second Model	10	0.81	MAE = 0.40%
Rut Depth	12	0.53	MAE = 0.68 mm
CT_{Index}	12	0.60	MAPE = 17.7%
FI	12	0.45	MAPE = 17.0%

Although the performance models showed lower agreement than the VMA models, their prediction errors remained within or below the inherent variability of the corresponding laboratory test methods. Consequently, the models are appropriate for screening and comparative evaluation of gradation adjustments during mix design. Overall, the proposed framework provides a practical tool for estimating the effects of gradation changes on the volumetric and performance properties of P-401 mixtures.

6.6 Web-Based Decision Tool

Figure 54 shows the interface of the Volumetric and Performance Prediction tool, a web-based application developed from the prediction models and back-transformation procedure described in this chapter.

The tool allows users to estimate volumetric properties and mixture performance for a new trial design using gradation, aggregate properties, and laboratory results from previously tested trial designs. The tool is intended for situations in which the existing and new trial mixtures share the same aggregate source, binder type, and binder content. Predictions may not be reliable if the new trial design uses different materials or binder contents.



Source: AAPTP

Figure 54. Screenshot of the Web-Based Tool for Estimating VMA, Rutting, and Cracking Properties

The following steps describe how to use the tool.

Step 1: Select the Target Parameter

Open the tool and locate the Control Panel on the right side of the interface. In the navigation menu, select the response variable to be estimated. Available options include VMA, rut depth, CT_{Index}, and FI. Each target parameter requires a different set of input data.

Step 2: Enter Input Data for Previously Tested and New Trial Designs

In the Design Parameters, Volumetrics, and Performance table, enter the information for the previously tested mixtures under the Existing Mix Design column. Required inputs include aggregate gradation, aggregate bulk specific gravity, and FAA, when available.

Depending on the parameter selected in Step 1, additional data may be required. For example, measured VMA values from previously tested trial designs must be provided when predicting VMA for a new trial design, while laboratory performance results (rut depth, CT_{Index}, and FI) from previously tested trial designs are required for performance prediction.

The tool allows users to enter results from multiple previously tested trial designs to improve predictions for a new trial design by clicking the Add Trial option. Individual trials can be included or excluded from the prediction model by selecting or clearing the corresponding checkbox.

Next, enter the gradation and aggregate properties for the new trial design under the New Trial Design column. These inputs include the aggregate gradation, G_{sb}, and optional FAA values for the new trial design.

Step 3: Run the Prediction Model

After all required data have been entered, click the Run Model button to execute the prediction. The tool processes the inputs using the developed prediction models and displays the predicted results for the selected response variable. If FAA values are not provided, the tool automatically applies the model version that excludes these values.

For VMA prediction, the tool automatically determines whether the first or second VMA prediction model should be applied based on the gradation position relative to the MDL.

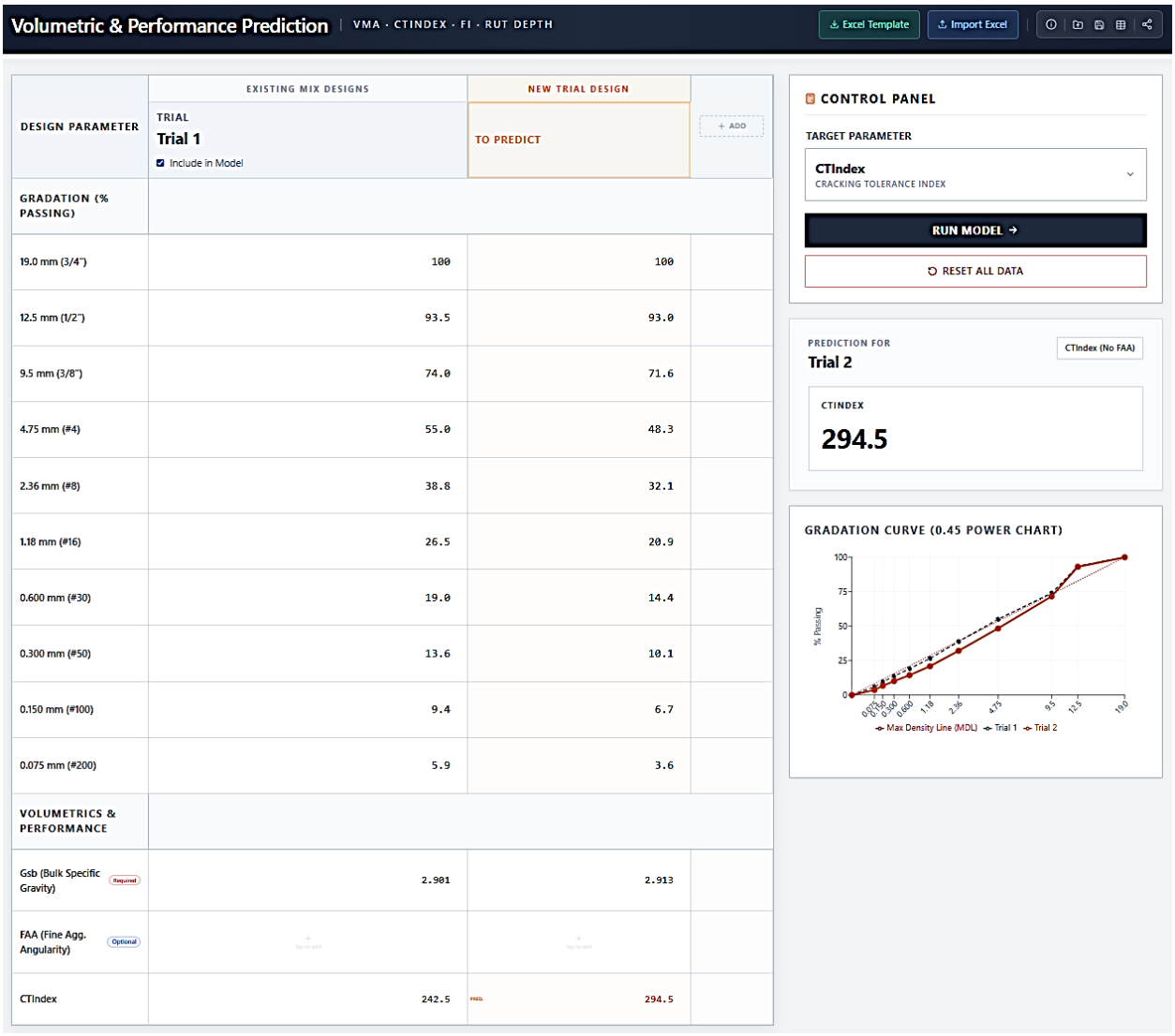
Step 4: Save, Load, or Share the Project

The tool also provides several options 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.

These functions allow users to document trial designs, compare alternative gradation scenarios, and share prediction results during mix design development.

Figure **55** presents an example of using the tool 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 FAA was not available, the tool applied the model without FAA. Based on the inputs, the predicted CT_{Index} for the new trial is 294.5.



Source: AAPTP

Figure 55. Example of Prediction Tool for CT_{Index}

Chapter 7. Conclusions and Recommendations

This study evaluated whether the current Federal Aviation Administration P-401 aggregate gradation limits for airfield asphalt mixtures could be adjusted while maintaining required volumetric properties and performance. The research addressed two primary objectives:

1. Evaluate the effect of gradations outside the current P-401 gradation limits while maintaining comparable volumetric properties and asphalt binder content (including effective binder content) and determine whether revisions to the specification could be supported.
2. Evaluate the effect of systematic gradation changes on the volumetric properties and laboratory performance of airfield asphalt mixtures and develop predictive models that quantify these relationships.

The experimental program included laboratory evaluation of P-401 mixtures with systematic gradation adjustments and the development of statistical prediction models linking gradation characteristics to the volumetric and performance properties of the mixtures. The key conclusions and recommendations derived from this work are presented below.

7.1 Proposed Revisions to Current P-401 Gradation Limits

The first goal of this study was to determine whether asphalt mixtures with gradations outside the current P-401 gradation limits could perform comparably to those within the gradation limits. To isolate the effect of gradation, asphalt binder content and key volumetric properties were maintained at approximately comparable levels across the evaluated mixtures.

This assessment included 13 P-401 mix designs from across the country. Mixtures with gradations outside the current limits (Out-Spec 1) were compared with corresponding mixtures within the limits (In-Spec 1). The key findings are summarized below.

- Moisture resistance showed no statistically significant differences between gradations inside and outside the current limits.
- Permeability remained within acceptable limits for all mixtures, indicating that the evaluated gradation shifts did not compromise impermeability requirements for FG airfield asphalt pavements.
- Durability, evaluated using the Cantabro mass loss test, was comparable between In-Spec 1 and Out-Spec 1 mixtures, indicating similar resistance to raveling.
- Rutting resistance measured by HWTT and APA tests was not significantly affected by shifting the gradation outside the current limits.
- Intermediate-temperature cracking resistance, as measured by CT_{Index}, was generally maintained or improved when gradation was shifted below the lower

gradation limits but tended to decrease when gradation was shifted above the upper gradation limits.
- Intermediate-temperature cracking resistance, as measured using FI, was generally maintained or improved when gradations were adjusted outside the limits.
- Low-temperature cracking resistance, evaluated using the DCT test, showed no statistically significant differences between In-Spec 1 and Out-Spec 1 mixtures.

Overall, these results indicate that mixtures with gradations outside the current P-401 limits can achieve performance comparable to that of mixtures designed within the limits, provided volumetric properties and asphalt binder content are maintained. Based on these findings, it is recommended that the lower limits for Gradation 2 with 12.5-mm NMAS be reduced by 2 percent at the No. 16, No. 30, and No. 50 sieves. No changes are recommended to the current P-401 volumetric requirements. The existing upper limits are also retained at all sieves, because intermediate-temperature cracking resistance decreased when gradations were shifted above the upper limits and contractors have not reported difficulty meeting the current upper limits (Section 4.4.1).

The recommended 2-percent reduction reflects the range of gradations evaluated in this study. Gradations of Out-Spec 1 mixtures were shifted below the current lower limits as far as the available aggregate fractions allowed, while maintaining an FG mixture structure and similar binder content and volumetric properties as compared to the corresponding In-Spec 1 mixtures. Under these constraints, the resulting deviations generally clustered within 1 to 2 percent below the lower limits at the No. 16, No. 30, and No. 50 sieves (Section 4.4.2). Consequently, the recommendation should not be interpreted as an optimal or maximum allowable adjustment. Rather, it represents a conservative revision supported by the mixtures evaluated and tested in this study.

The recommended magnitude is also consistent with the mix design adjustment that the revision is intended to address. Under the current limits, compliance at the No. 16, No. 30, and No. 50 sieves is often achieved by incorporating natural sand, and approximately 10 percent natural sand increases the percent passing at these sieves by about 2 percent (Section 4.4.3). A 2-percent reduction in the lower limits, therefore, offsets approximately the amount of natural sand a designer would otherwise need to add, allowing greater use of manufactured sand without reducing FAA or the VMA it supports.

The recommendation aligns with prior airfield research related to mixture permeability. The proposed revision applies only to the No. 16, No. 30, and No. 50 sieves. Retaining the current requirements for other sieves, especially the PCS (i.e., No. 8) for 12.5-mm NMAS mixtures, preserves the FG characteristics of P-401 mixtures and avoids the CG gradations associated with higher permeability and greater difficulty achieving the required in-place density [8, 144, 148, 155]. At the same time, the proposed adjustment provides additional flexibility for aggregate blending and mixture volumetric control.

The recommendation should be considered within the scope of the study. Because binder content and volumetric properties were controlled to isolate the effects of gradation, the study did not evaluate every combination of gradation, aggregate source, and binder content that could potentially be developed under the revised limits. Likewise, larger reductions at the No. 16, No. 30, and No. 50 sieves were not evaluated, and this study does not establish an upper bound on the magnitude of revision that could be adopted without affecting performance.

Nevertheless, the direction of the recommended change is supported by the observed performance trends. Performance was generally maintained or improved when gradations were shifted below the lower limits, while reductions in intermediate-temperature cracking resistance were observed only when gradations were shifted above the upper limits. The proposed revision, therefore, extends the allowable gradation range in the direction in which comparable or improved performance was demonstrated.

Accordingly, agencies adopting the revised limits should monitor mix designs and field performance during initial implementation to confirm that the laboratory performance trends identified in this study are achieved under production and field conditions.

7.2 Effects of Gradation Changes on Mixture Volumetric and Performance Properties

The second objective of this study was to evaluate the effects of systematic gradation adjustments on the volumetric characteristics of the mixture and its laboratory performance. This analysis compared four gradations: In-Spec 1, In-Spec 2, Out-Spec 1, and Out-Spec 2, with In-Spec 2 and Out-Spec 2 being closer to the center or further outside of the gradation limits. The key findings are summarized below.

- Shifting the gradation toward either the lower or upper gradation limits generally increased VMA and AV content, reflecting changes in aggregate packing characteristics.
- Moisture resistance, durability, and permeability remained within acceptable limits across all mixtures evaluated.
- Mixtures with coarser gradations generally maintained or improved rutting resistance compared to mixtures with finer gradations.
- Intermediate-temperature cracking resistance, evaluated using IDEAL-CT and I-FIT, generally improved as gradation became coarser.
- Low-temperature cracking resistance, evaluated using the DCT test, did not show a statistically significant improvement associated with coarser gradations.

These results indicate that minor adjustments to aggregate gradation influence the volumetric structure of the mixture and certain performance properties while maintaining

acceptable durability and rutting resistance. The findings further support the proposed revisions to P-401 gradation limits.

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

Prediction models were developed to quantify the relationships between gradation adjustments and changes in mixture responses, utilizing the laboratory dataset generated in this study. Four models were developed to estimate the effects of gradation changes and aggregate properties on the following:

- VMA (with/without fine aggregate angularity)
- Rutting resistance (HWTT/APA rut depth)
- Intermediate-temperature cracking resistance measured by CT_{Index}
- Intermediate-temperature cracking resistance measured by FI

The modeling results demonstrate that changes in mixture volumetric properties and performance can be predicted with reasonable accuracy using aggregate gradation parameters and aggregate properties when binder type, binder content, and aggregate source are held constant.

To facilitate practical implementation, the models were incorporated into a web-based tool that allows users to estimate volumetric and performance responses for new trial designs based on previously tested trial designs. The tool provides mixture designers with a practical method for evaluating gradation adjustments and identifying gradation scenarios that satisfy both volumetric and performance requirements.

7.4 Recommendations for Future Research

Although this study provides a technical basis for proposing revisions to the P-401 gradation limits, several areas warrant further investigation.

First, the proposed gradation adjustments were evaluated mainly for Gradation 2 mixtures with 12.5-mm NMAS. Future research should expand the database to include a wider range of P-401 asphalt mixtures, including 19-mm and 9.5-mm NMAS mixtures, to determine whether similar revisions are appropriate for other gradations.

Second, the prediction models developed in this study were based on OLS regression, which performed well for the available dataset. As additional datasets become available, future research should evaluate advanced modeling approaches, including ML methods, which may better capture nonlinear relationships among mixture variables.

Third, the current modeling framework predicts changes in mixture responses relative to changes in gradation. Future work should extend this approach by developing models that

predict changes in mixture volumetric properties and performance resulting from changes in binder content and properties.

Finally, the prediction models developed in this study assume that aggregate sources remain constant during gradation adjustments. However, mix designers may occasionally decide to change one or more aggregate sources or components during these adjustments. Future research should assess the applicability of these models to mixtures for which aggregate sources or components are changed during the gradation adjustment process. Future research should also consider incorporating additional aggregate property variables, alongside gradation, into the predictive models.

References

1. Federal Aviation Administration. *Standards for Specifying Construction of Airports*. Advisory Circular 150/5370-10H. U.S. Department of Transportation, Washington, DC, 2018.
2. American Association of State Highway and Transportation Officials. *Standard Specification for Superpave Volumetric Mix Design*. AASHTO M 323. AASHTO, Washington, DC, 2022.
3. Hubbard, P., and F. C. Field. *Stability and Related Tests for Asphalt Paving Mixtures*. Research Series No. 1. Asphalt Institute, Lexington, KY, 1935.
4. Welborn, J. Y. *State-of-the-Art in Asphalt Pavement Specifications*. Report No. FHWA-RD-84-075. Federal Highway Administration, Washington, DC, 1984.
5. Bureau of Public Roads. *Aggregate Gradation for Highways: Simplification, Standardization, and Uniform Application and a New Graphical Evaluation Chart*. U.S. Department of Commerce, Bureau of Transportation Statistics, Washington, DC, 1962.
6. American Association of State Highway and Transportation Officials. *Standard Specification for Sizes of Aggregate for Road and Bridge Construction*. AASHTO M 43-05. AASHTO, Washington, DC, 2022.
7. McFadden, G., and W. C. Ricketts. Design and Control of Asphalt Paving Mixtures for Military Installations. *Proceedings of the Association of Asphalt Paving Technologists,* Vol. 17, 1948, pp. 93–113.
8. Cooley, L. A., Jr., R. C. Ahlrich, and R. S. James. *Design of Hot Mix Asphalt for Airfield Pavements Using the Superpave Gyratory Compactor.* Presented at the 2010 FAA Worldwide Airport Technology Transfer Conference, Atlantic City, NJ, 2010.
9. Goode, J. F., and L. A. Lufsey. A New Graphical Chart for Evaluating Aggregate Gradation. In *Aggregate Gradation for Highways: Simplification, Standardization, and Uniform Application and a New Graphical Evaluation Chart*. Bureau of Public Roads, U.S. Department of Commerce, Washington, DC, 1962, pp. 13–26.
10. Nijboer, L. W. *Plasticity as a Factor in the Design of Dense Bituminous Road Carpets*. Elsevier Publishing Company, New York, NY, 1948.
11. McLeod, N. W. Design of Dense Graded Asphalt Concrete Pavements. *Proceedings of the Canadian Technical Asphalt Association,* 1987.
12. Kandhal, P. S., and L. A. Cooley, Jr. *The Restricted Zone in the Superpave Aggregate Gradation Specification*. NCHRP Report 464. Transportation Research Board, National Research Council, Washington, DC, 2001.
13. Newman, J. K., and R. B. Freeman. *Evaluation of Strategic Highway Research Program (SHRP) Products for Application to Airport Pavements*. Report No. DOT/FAA/AR-99/13. Federal Aviation Administration, Washington, DC, 2000.
14. Coree, B. J., and W. P. Hislop. A Laboratory Investigation into the Effects of Aggregate-Related Factors on Critical VMA in Asphalt Paving Mixtures. *Journal of the Association of Asphalt Paving Technologists*, Vol. 70, 2001, pp. 70–131.

15. Hislop, W. P., and B. J. Coree. VMA as a Design Parameter in Hot-Mix Asphalt. In *Proceedings of the 2000 Mid-Continent Transportation Symposium,* Iowa State University, Ames, IA, 2000.
16. Aschenbrener, T., and C. MacKean. *Factors That Affect the Voids in the Mineral Aggregate in Hot-Mix Asphalt.* Final Report No. CDOT-DTD-R-92-13. Colorado Department of Transportation, Denver, CO, 1992.
17. Park, D.-W., A. Chowdhury, and J. W. Button. *Effects of Aggregate Gradation and Angularity on VMA and Rutting Resistance.* ICAR Report 201-3F. International Center for Aggregates Research, Austin, TX, 2001.
18. Purcell, E. M., and S. A. Cross. *Effects of Aggregate Angularity on VMA and Rutting of KDOT Superpave Level 1 Mixes*. Final Report, Contract C-1047. University of Kansas and Kansas Department of Transportation, Lawrence and Topeka, KS, 2001.
19. Aschenbrener, T. *How to Increase Voids in Mineral Aggregate: Guidelines to Increase VMA of Superpave Mixtures.* Presented at the Federal Highway Administration Asphalt Mixture Expert Technical Group Meeting, 1996.
20. 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*, No. 1787, 2002, pp. 90–98.
21. Bessa, I. S., V. T. F. Castelo Branco, J. B. Soares, and J. A. Nogueira Neto. Aggregate Shape Properties and Their Influence on the Behavior of Hot-Mix Asphalt. *Journal of Materials in Civil Engineering*, Vol. 27, No. 7, 2015, Article 04014212.
22. Asphalt Institute. *MS-2 Asphalt Mix Design Methods.* Manual Series No. 2, 7th ed. Asphalt Institute, Lexington, KY, 2014.
23. McGennis, R. B., R. M. Anderson, T. W. Kennedy, and M. Solaimanian. *Background of Superpave Asphalt Mixture Design and Analysis*. Publication No. FHWA-SA-95-003. Federal Highway Administration, Washington, DC, 1995.
24. Leon, L. P., and R. Charles. Aggregate Angularity on the Permanent Deformation Zones of Hot-Mix Asphalt. *Global Journal of Researches in Engineering: E Civil and Structural Engineering*, Vol. 15, No. 3, 2015, pp. 25–29.
25. Souza, L. T., Y.-R. Kim, F. V. Souza, and L. S. Castro. Experimental Testing and Finite-Element Modeling to Evaluate the Effects of Aggregate Angularity on Bituminous Mixture Performance. *Journal of Materials in Civil Engineering*, Vol. 24, No. 3, 2012, pp. 249–258.
26. 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*, Vol. 132, No. 7, 2006, pp. 580–589.
27. Renteria, D., E. M. Mahmoud, R. Yanez, and V. Burbach. Discrete Element Analysis of SCB Variability—Asphalt Mixtures. In *Geotechnical Frontiers 2017: Transportation Facilities, Structures, and Site Investigation,* Geotechnical Special Publication 277, American Society of Civil Engineers, Reston, Va., 2017, pp. 302–313.
28. Saadeh, S., and M. El Asmar. *Sensitivity Analysis of the IDEAL CT Test Using the Distinct Element Method.* Report No. 23-22. Mineta Transportation Institute, San José State University, San José, Calif., 2023.

29. Park, D.-W., and H.-S. Lee. Test Methods for Fine Aggregate Angularity Considering Resistance of Rutting. *KSCE Journal of Civil Engineering*, Vol. 6, No. 4, 2002, pp. 421–427.
30. Topal, A., and B Sengoz. Evaluation of Compacted Aggregate Resistant Test Compared with the Fine Aggregate Angularity Standards. *Construction and Building Materials*, Vol. 22, No. 5, 2008, pp. 993–998.
31. Ramli, I., H. Yaacob, N. A. Hassan, C. R. Ismail, and M. R. Hainin. *Fine Aggregate Angularity Effects on Rutting Resistance of Asphalt Mixture. Jurnal Teknologi (Sciences & Engineering)*, Vol. 65, No. 3, 2013, pp. 105–109.
32. Iuele, T., F. G. Praticò, and R. Vaiana. Fine Aggregate Properties vs Asphalt Mechanical Behavior: An Experimental Investigation. In *Pavement and Asset Management: Proceedings of the World Conference on Pavement and Asset Management,* CRC Press, London, 2019, pp. 309–315.
33. Johnson, E., X. Li, A. Zofka, M. O. Marasteanu, and T. R. Clyne. Investigation of Superpave Fine Aggregate Angularity Criterion for Asphalt Concrete. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1998, 2007, pp. 75–81.
34. Chowdhury, A., J. W. Button, V. Kohale, and D. Jahn. *Evaluation of Superpave Fine Aggregate Angularity Specification*. ICAR Report 201-1. International Center for Aggregates Research, Texas Transportation Institute, Texas A&M University System, College Station, Tex., 2001.
35. Fernandes, J. L., Jr., R. Roque, M. Tia, and L. Casanova. Uncompacted Void Content of Fine Aggregate as Quality Indicator of Materials Used in Superpave Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1723, 2000, pp. 37–44.
36. Bahia, H. U., and A. D. Stakston. *The Effect of Fine Aggregate Angularity, Asphalt Content and Performance Graded Asphalts on Hot Mix Asphalt Performance*. Report No. WHRP 03-04. Wisconsin Highway Research Program, University of Wisconsin–Madison, Madison, Wis., 2003.
37. Lee, C.-J., T.D. White, and T.R. West. *Effect of Fine Aggregate Angularity on Asphalt Mixture Performance.* Report No. FHWA/IN/JTRP-98/20. Joint Transportation Research Program, Purdue University, West Lafayette, Ind., 2000.
38. Kandhal, P. S., and F. Parker, Jr. *Aggregate Tests Related to Asphalt Concrete Performance in Pavements*. NCHRP Report 405. Transportation Research Board, National Research Council, Washington, DC, 1998.
39. 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 Record: Journal of the Transportation Research Board*, No. 1761, 2001, pp. 26–31.
40. Buchanan, M. S. *Evaluation of the Effect of Flat and Elongated Particles on the Performance of Hot Mix Asphalt Mixtures*. NCAT Report 00-03. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2000.
41. 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 and Building Materials*, Vol. 258, 2020, Article 119674.

42. Chen, J.-S., M. Chang, and K.-Y. Lin. Influence of Coarse Aggregate Shape on the Strength of Asphalt Concrete Mixtures. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, 2005, pp. 1062–1075.
43. Mukhopadhyay, A. K., B. Pitre, A. Russell, E. Arambula, C. Estakhri, and Y. Deng. *Treatments for Clays in Aggregates Used to Produce Cement Concrete, Bituminous Materials, and Chip Seals: Technical Report*. Texas Department of Transportation, Research and Technology Implementation Office, College Station, Tex., 2013.
44. Dong, Q., J. Yuan, X. Chen, and X. Ma. Reduction of Moisture Susceptibility of Cold Asphalt Mixture with Portland Cement and Bentonite Nanoclay Additives. *Journal of Cleaner Production*, Vol. 176, 2018, pp. 320–328.
45. Yin, F., E. Arambula, R. Lytton, A. E. Martin, and L. G. Cucalon. Novel Method for Moisture Susceptibility and Rutting Evaluation Using Hamburg Wheel Tracking Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2446, 2014, pp. 1–7.
46. Afsha, S., S. R. Vaid, M. A. Montoya, I. Abdallah, and S. Nazarian. Impact of *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2678, No. 10, 2024, pp. 356–366.
47. Dahir, S. H. A Review of Aggregate Selection Criteria for Improved Wear Resistance and Skid Resistance of Bituminous Surfaces. *Journal of Testing and Evaluation*, Vol. 7, No. 5, 1979, pp. 245–253.
48. 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*, No. 1638, 1998, pp. 85–93.
49. Amirkhanian, S. N., D. Kaczmarek, and J. L. 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*, No. 1301, 1991, pp. 77–86.
50. Masad, E., D. Olcott, T. White, and L. Tashman. Correlation of Fine Aggregate Imaging Shape Indices with Asphalt Mixture Performance. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1757, 2001, pp. 148–156.
51. Hassan, H. M. Z., K. Wu, W. Huang, S. Chen, Q. Zhang, J. Xie, and X. Cai. Study on the Influence of Aggregate Strength and Shape on the Performance of Asphalt Mixture. 2021. *Construction and Building Materials*, Vol. 294, 2021, Article 123599.
52. Coree, B. J., and Hislop, W. P. *The Difficult Nature of Minimum VMA: A Historical Perspective.* Center for Transportation Research and Education, Iowa State University, Ames, Iowa, 1998.
53. Hudson, S. B., and R. L. Davis. Relationship of Aggregate Voidage to Gradation. *Proceedings of the Association of Asphalt Paving Technologists*, Vol. 34, 1965, pp. 574–593.
54. Kandhal, P. S., and R. B. Mallick. *Design of New-Generation Open-Graded Friction Courses (Revised)*. NCAT Report 99-3. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 1999.

55. National Asphalt Pavement Association. *Designing and Constructing SMA Mixtures—State-of-the-Practice.* Quality Improvement Series 122. National Asphalt Pavement Association, Lanham, Md., 2002.
56. Pavement Interactive. Pavement Types/Mix Types. *Pavement Interactive.* https://pavementinteractive.org/reference-desk/pavement-types-and-history/pavement-types/. Accessed July 27, 2025.
57. Koch, J. *Effects of Aggregate Blend Gradations on Hot Mix Asphalt Volumetrics*. Department of Civil Engineering and Construction, Bradley University, Peoria, Ill., and Illinois Asphalt Pavement Association, 2021.
58. Williams, S. G., and J. K. Foreman. *Effects of Various HMA Material Properties on Pavement Performance*. Arkansas State Highway and Transportation Department, Little Rock, Ark., 2006.
59. Chadbourn, B. A., E. L. Skok, Jr., D. E. Newcomb, B. L. Crow, and S. Spindler. *The Effect of Voids in Mineral Aggregate (VMA) on Hot-Mix Asphalt Pavements.* Minnesota Department of Transportation, St. Paul, Minn., 1999.
60. Bessette, L. *Evaluation of the Effects of Aggregate Gradation and Compaction Effort on the Voids in Mineral Aggregate in Asphalt Concrete*. Master's thesis. West Virginia University, Morgantown, 2013.
61. Huber, G. A., and T. S. 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*. ASTM STP 1147, R. C. Meininger, ed., ASTM International, Philadelphia, Pa., 1992, pp. 225–251.
62. Kandhal, P. S., K. Y. Foo, and J. A. D'Angelo. Field Management of Hot Mix Asphalt Volumetric Properties. In *Quality Management of Hot-Mix Asphalt*, ASTM STP 1299, D. S. Decker, ed., ASTM International, West Conshohocken, Pa., 1996, 28–45.
63. Xie, H., L. A. Cooley, and M. Huner. *4.75 mm NMAS Stone Matrix Asphalt (SMA) Mixtures.* NCAT Report 03-05. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2003.
64. Leiva, F., and R. West. *Using the Primary Control Sieve Index to Define Gradation Type and as a Factor Related to Asphalt Mixture Properties*. NCAT Report 21-01. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2021.
65. 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, West Virginia University, Morgantown, 2006.
66. Khalifa, M. O. L. *The Behavior of Asphaltic Concrete Constructed with Large-Sized Aggregate*. Dissertation. University of Illinois at Urbana-Champaign, Urbana, 1969.
67. Liu, H., P. Hao, and J. Xu. Effects of Nominal Maximum Aggregate Size on the Performance of Stone Matrix Asphalt. *Applied Sciences*, Vol. 7, No. 2, 2017, Article 126.
68. Federal Highway Administration. *Superpave Mix Design and Gyratory Compaction Levels.* FHWA-HIF-11-031. Federal Highway Administration, Washington, DC, 2010.
69. 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*, Vol. 23, No. 5, 2011, pp. 664–672.

70. Pouranian, M. *Aggregate Packing Characteristics of Asphalt Mixtures*. Ph.D. dissertation. Purdue University, West Lafayette, Ind., 2019.
71. Zhang, H., Z. Ma, P. Ji, Y. Bi, W. Cao, and S. Liu. Establishment and Application of Theoretical-Empirical Prediction Model of VMA in Hot Mix Asphalt Mixture. *Case Studies in Construction Materials*, Vol. 18, 2023, Article e01986.
72. Roberts, F. L., P. S. Kandhal, E. R. Brown, D.-Y. Lee, and T. W. Kennedy. *Hot Mix Asphalt Materials, Mixture Design and Construction*. 2nd ed. National Asphalt Pavement Association Research and Education Foundation, Lanham, Md., 1996.
73. Graziani, A., G. Ferrotti, E. Pasquini, and F. Canestrari. An Application to the European Practice of the Bailey Method for HMA Aggregate Grading Design. *Procedia—Social and Behavioral Sciences*, Vol. 53, 2012, pp. 990–999.
74. Golalipour, A., E. Jamshidi, Y. Niazi, Z. Afsharikia, and M. Khadem. Effect of Aggregate Gradation on Rutting of Asphalt Pavements. *Procedia—Social and Behavioral Sciences*, Vol. 53, 2012, pp.440–449.
75. Scherocman, J. A. Stone Mastic Asphalt Reduces Rutting. *Better Roads,* Vol. 61, No. 11, 1991, pp. 26–27.
76. Scherocman, J. A. *Construction of SMA Test S in the US*. *Proceedings of the Association of Asphalt Paving Technologists*, Vol. 61, 1992, p.642.
77. Asi, I. M. Laboratory Comparison Study for the Use of Stone Matrix Asphalt in Hot Weather Climates. *Construction and Building Materials,* Vol. 20, No. 10, 2006, pp. 982–989.
78. Zhao, Z., J. Jiang, F. Ni, and L. Zhou. Factors Affecting the Rutting Resistance of Asphalt Pavement Based on the Field Cores Using a Multi-Sequenced Repeated Loading Test. *Construction and Building Materials,* Vol. 253, 2020, Article 118902.
79. 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*, Vol. 20, No. 1, 2016, pp. 385–392.
80. Musselman, J. A., B. Choubane, G. C. Page, and P. B. Upshaw. Superpave Field Implementation: Florida's Early Experience. *Transportation Research Record*: *Journal of the Transportation Research Board,* No. 1609, 1998, pp. 51–60.
81. Kim, Y.-R., N. Kim, and N. P. Khosla. Effects of Aggregate Structure on Hot-Mix Asphalt Rutting Performance in Low Traffic Volume Local Pavements. *Construction and Building Materials*, Vol. 23, No. 6, 2009, pp. 2177–2182.
82. Cross, S. A., A. Adu-Osei, M. R. Hainin, and R. K. Fredrichs. *Effects of Gradation on Performance of Asphalt Mixtures*. Presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, 1999.
83. Garba, R. *Permanent Deformation Properties of Asphalt Concrete Mixtures*. Norwegian University of Science and Technology, Trondheim, Norway, 2002.
84. El-Basyouny, M. M., and M. S. Mamlouk. *Effect of Aggregate Gradation on Rutting Potential of Superpave Mixes*. Presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC, 1999. Paper No. 99-0602.
85. Kandhal, P. S., and R. B. Mallick. Effect of Mix Gradation on Rutting Potential of Dense-Graded Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1767, 2001, pp. 146–151.

86. Chowdhury, A., J. W. Button, and J. D. Grau. *Effects of Superpave Restricted Zone on Permanent Deformation*. ICAR Report 201-2. International Center for Aggregates Research, Texas Transportation Institute, Texas A&M University System, College Station, Tex., 2001.
87. Hand, A. J., J. L. Stiady, T. D. White, A. S. Noureldin, and K. Galal. Gradation Effects on Hot-Mix Asphalt Performance. *Transportation Research Record: Journal of the Transportation Research Board,* No. 1767, 2001, pp. 152–157.
88. West, R., D. Timm, J. R. Willis, R. B. Powell, N. Tran, D. Watson, M. Sakhaeifar, and J. Nelson. *Phase IV NCAT Pavement Test Track Findings*. NCAT Report 12-10, National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2012.
89. Hasan, M. M. *Evaluation of Fine and Coarse Graded Asphalt Mixes*. Master's thesis. University of New Mexico, Albuquerque, 2021.
90. Hasan, M. M., M. Ahmad, M. A. Hasan, H. M. Faisal, and R. A. Tarefder. Laboratory Performance Evaluation of Fine and Coarse-Graded Asphalt Concrete Mix. *Journal of Materials in Civil Engineering,* Vol. 31, No. 11, 2019, Article 04019259.
91. Pan, Y., H. Guo, W. Guan, and Y. Zhao. A Laboratory Evaluation of Factors Affecting Rutting Resistance of Asphalt Mixtures Using Wheel Tracking Test. *Case Studies in Construction Materials*, Vol. 18, 2023, Article e02148.
92. 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),* IEEE, 2011, pp. 461–465.
93. Gao, Y., K. Hou, Y. Jia, Z. Wei, S. Wang, Z. Li, F. Ding, and X. Gong. Variability Evaluation of Gradation for Asphalt Mixture in Asphalt Pavement Construction. *Automation in Construction*, Vol. 128, 2021, Article 103742.
94. Lv, Q., W. Huang, M. Zheng, H. Sadek, Y. Zhang, and C. Yan. Influence of Gradation on Asphalt Mix Rutting Resistance Measured by Hamburg Wheel Tracking Test. *Construction and Building Materials*, Vol. 238, 2020, Article 117674.
95. Chen, J.-S., and M.-C. Liao. Evaluation of Internal Resistance in Hot-Mix Asphalt (HMA) Concrete. *Construction and Building Materials,* Vol. 16, No. 6, 2002, pp. 313–319.
96. Zhang, Y., X. Luo, I. Onifade, X. Huang, R. L. Lytton, and B. Birgisson. Mechanical Evaluation of Aggregate Gradation to Characterize Load Carrying Capacity and Rutting Resistance of Asphalt Mixtures. *Construction and Building Materials*, Vol. 205, 2019, pp. 499–510.
97. Teklu, W. *Effect of Gradation of Aggregates on the Rutting Performance of Hot Mix Asphalt*. Thesis. Addis Ababa University, Addis Ababa, Ethiopia, 2015.
98. Thompson, G. *Investigation of the Bailey Method for the Design and Analysis of Dense-Graded HMAC Using Oregon Aggregates: Final Report*. SPR 304-311. Oregon Department of Transportation Research Unit, Salem, 2007.
99. Manjunath, K. R., and N. B. Poornachandra. Design of Hot Mix Asphalt Using Bailey Method of Gradation. *International Journal of Research in Engineering and Technology*, Vol. 3, No. 6, 2014, pp. 386–393.

100. Lu, W.-J., A.-W Chen, P.-W. Hao, and J.-L. Dai. Effect of CA Ratio on Asphalt Mixture Property Based on Bailey Method. *Journal of Chang'an University (Natural Science Edition)*, Vol. 25, No. 4, 2005, pp. 5-8.
101. Gierhart, D. Analysis of Oklahoma Mix Designs for the National Center for Asphalt Technology Test Track Using the Bailey Method. In *Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing*, 2007, p. 33.
102. Aurilio, V., W. J. Pine, and P. Lum. Bailey Method – Achieving Volumetrics and HMA Compactability. In *Proceedings of the Annual Conference of the Canadian Technical Asphalt Association*, 2005.
103. 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, Ind., 2006.
104. Ma, H., C. Zhou, D. Feng, and L. Sun. Influence of Fine Aggregate Content on Low-Temperature Cracking of Asphalt Pavements. *Journal of Testing and Evaluation*, Vol. 45, No. 3, 2017, pp. 835–842.
105. Nian, T., J. Ge, P. Li, M. Wang, and Y. Mao. Improved Discrete Element Numerical Simulation and Experiment on Low-Temperature Anti-Cracking Performance of Asphalt Mixture Based on PFC2D. *Construction and Building Materials,* Vol. 283, 2021, Article 122792.
106. 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,* Vol. 9, No. 4, 2019, Article 628.
107. Tan, Y.-Q., H.-N. Xu, Z.-J. Dong, and W.-Q. Gong. Effect of Aggregate Grading on Low Temperature Cracking Resistance in Asphalt Mixtures Base on Mathematical Statistic. In *Proceedings of the Eighth International RILEM Conference on Cracking in Pavements,* 2008.
108. Chen, J., Z. Li, Z. Zhao, X. Huang, J. Chen, and Z. Liu. Investigation on Temperature Shrinkage Characteristics of the Combined Structure in Asphalt Pavement. *Frontiers in Materials,* Vol. 9, 2022, Article 1055641.
109. 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,* Vol. 39, 2014, pp. 7577–7590.
110. Li, G.-Y., M. Zhong-Chuan, and G. Ze-Yu. Experimental Study on Aggregate Gradation Influence on the Performance-Related Properties of Structural-Layer Asphalt Mixtures. *Highway,* No. 5, 2019, pp. 219–225.
111. Lv, S., Z. Wang, X. Zhu, J. Yuan, and X. Peng. Research on Strength and Fatigue Properties of Asphalt Mixture with Different Gradation Curves. *Construction and Building Materials*, Vol. 364, 2023, Article 129872.
112. Saleh, M. F. Effect of Aggregate Gradation and Percentage of Air Voids on the Fatigue Behaviour of Hot Mix Asphalts. *Road & Transport Research*, Vol. 21, No. 3, 2012, pp. 53–63.
113. Nejad, F. M., E. Aflaki, and M. A. Mohammadi. Fatigue Behavior of SMA and HMA Mixtures. *Construction and Building Materials,* Vol. 24, No. 7, 2010, pp. 1158–1165.

114. 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,* American Society of Civil Engineers, Reston, Va., 2019.
115. Hainin, M. R., A. A. Abdulridha, H. Yaacob, Z. Haron, and N. I. M. Yusoff. Fatigue Life of Malaysian Hot Mix Asphalt Mixtures. *Malaysian Journal of Civil Engineering*, Vol. 25, No. 1, 2013, pp. 94–105.
116. Lou, K., X. Wu, P. Xiao, A. Kang, Z. Wu, and Y. Xia. Comprehensive Study About Effect of Basalt Fiber, Gradation, Nominal Maximum Aggregate Size and Asphalt on the Anti-Cracking Ability of Asphalt Mixtures. *Applied Sciences,* Vol. 11, No. 5, Article 2289.
117. Shen, S., and S. Carpenter. Development of an Asphalt Fatigue Model Based on Energy Principles. *Asphalt Paving Technology: Association of Asphalt Paving Technologists—Proceedings of the Technical Sessions*, Vol. 76, 2007, p. 525.
118. Abo-Qudais, S., and I. Shatnawi. Prediction of Bituminous Mixture Fatigue Life Based on Accumulated Strain. *Construction and Building Materials,* Vol. 21, No. 6, 2007 pp. 1370–1376.
119. 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,* Vol. 21, No. 10, 2020, pp. 1275–1280.
120. Etheridge, R. A., Y. D. Wang, S. S. Kim, and Y. R. Kim. Evaluation of Fatigue Cracking Resistance of Asphalt Mixtures Using Apparent Damage Capacity. *Journal of Materials in Civil Engineering,* Vol. 31, 2019, Article 04019257.
121. Christensen, D. W., Jr. and R. F. Bonaquist. *Volumetric Requirements for Superpave Mix Design*. NCHRP Report 567, Transportation Research Board, Washington, DC, 2006.
122. Ali, Y., M. Irfan, S. Ahmed, S. Khanzada, and T. Mahmood. Sensitivity Analysis of Dynamic Response and Fatigue Behaviour of Various Asphalt Concrete Mixtures. *Fatigue & Fracture of Engineering Materials & Structures*, Vol. 38, No. 10, 2015, pp. 1181–1193.
123. Bari, J. *Development of a New Revised Version of the Witczak E* Predictive Models for Hot Mix Asphalt Mixtures*. Dissertation, Arizona State University, Tempe, Ariz.,
124. 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*, Vol. 30, No. 11, 2018, Article 04018298.
125. Abbas, A. R., R. B. Mallick, J. D. H. Guo, and T. Quasem. *Crack Resistance and Durability of Ohio DOT Asphalt Mixtures Using I-FIT & IDEAL-CT: Phase 2*. Report No. FHWA/OH-2022-01. Ohio Department of Transportation, Office of Statewide Planning and Research, 2021.
126. Haryanto, I., and O. Takahashi. Crack Initiation Assessment of Wearing Course Asphalt Mixtures Using Aggregate Gradation Characteristic. *Journal of Engineering and Technological Sciences*, Vol. 39, No. 1, 2007, pp. 28–42.

127. Sousa, J. B., J. C. Pais, M. Prates, P. A. Barros, and A.-M. Leclerc. Effect of Aggregate Gradation on Fatigue Life of Asphalt Concrete Mixes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1630, 1998, pp. 62–68.
128. Garcia, V. M., L. Barros, J. Garibay, I. Abdallah, and S. Nazarian. Effect of Aggregate Gradation on Performance of Asphalt Concrete Mixtures. *Journal of Materials in Civil Engineering*, Vol. 32, No. 5, 2020, Article 04020102.
129. Lu, Q., P. Fu, and J. Harvey. *Laboratory Evaluation of the Noise and Durability Properties of Asphalt Surface Mixes.* University of California Pavement Research Center, Davis and Berkeley, Calif., 2009.
130. Shiva Kumar, G., and S. J. Suresha. *State of the Art Review on Mix Design and Mechanical Properties of Warm Mix Asphalt. Road Materials and Pavement Design*, 2019, Vol. 20, No. 7, 2019, pp. 1501–1524.
131. Watson, D., N. A. Qureshi, and Z. Xie. *Improvements of Open-Graded Friction Course in Alabama*. Project No. ALDOT 930-790. Alabama Department of Transportation, Montgomery, 2020.
132. Xie, Z., N. Tran, D. E. Watson, and L. D. Blackburn. Five-Year Performance of Improved Open-Graded Friction Course on the NCAT Pavement Test Track. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, 2019, pp. 544–551.
133. Tran, T. *Open-Graded Friction Courses Suitable for Suburban Environments*. Master's thesis. Department of Civil and Environmental Engineering, Auburn University, Auburn, Ala., 2023.
134. Bennert, T., and L. A. Cooley, Jr.. *Evaluate the Contribution of the Mixture Components on the Longevity and Performance of FC-5.* FDOT Final Report No. BDS15 977-01. Florida Department of Transportation, Tallahassee, 2014.
135. Ji, T., P. Hao, H. She, K. Yang, H. Li, D. Wang, R. Kang, and J. Liu. Durability Behavior of Asphalt Mixtures in Regard to Material Properties and Gradation Type. *Frontiers in Materials*, Vol. 10, 2023, Article 1151479.
136. 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,* Vol. 13, No. 5, 2012, pp. 451–458.
137. 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,* Mehran University of Engineering and Technology, Pakistan, 2017.
138. 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, Engineering and Technology,* Vol. 6, 2012, pp. 12–28.
139. 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: Journal of the Transportation Research Board*, Vol. 2673, No. 1, 2019, pp. 240–245.

140. 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,* Vol. 125, 2016, pp. 426–433.
141. Cooley, L. A., Jr., E. R. Brown, and S. Maghsoodloo. Developing Critical Field Permeability and Pavement Density Values for Coarse-Graded Superpave Pavements. *Transportation Research Record: Journal of the Transportation Research Board, No. 1761*, 2001, pp. 41–49.
142. Gogula, A. K., M. Hossain, and S. A. Romanoschi. *A Study of Factors Affecting the Permeability of Superpave Mixes in Kansas*. Kansas Department of Transportation, Topeka, 2004.
143. Brown, E. R., D. S. Decker, R. B. Mallick, and J. Bukowski. Superpave Construction Issues and Early Performance Evaluation. *Journal of the Association of Asphalt Paving Technologists,* Vol. 68, 1999.
144. Brown, E. R., M. R. Hainin, L. A. Cooley, Jr., and G. Hurley. *Relationships of HMA In-Place Air Voids, Lift Thickness, and Permeability.* NCHRP Report 531. Transportation Research Board, Washington, DC, 2004, pp. 9–27.
145. Cooley, L. A., Jr., B. D. Prowell, and E. R. Brown. *Issues Pertaining to the Permeability Characteristics of Coarse-Graded Superpave Mixes.* NCAT Report 02-06. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2002.
146. Mallick, R. B., L. A. Cooley, Jr., M. R. Teto, R. L. Bradbury, and D. Peabody. *An Evaluation of Factors Affecting Permeability of Superpave Designed Pavements*. NCAT Report 03-02. National Center for Asphalt Technology, Auburn University, Auburn, Ala., 2003.
147. Chen, J., J. Yin, H. Wang, Y. Ding, and L. Huang. 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,* Vol. 27, No. 11, 2015, Article 04015023.
148. Hardiman, C. The Improvement of Water Drainage Function and Abrasion Loss of Conventional Porous Asphalt. *Proceedings of the Eastern Asia Society for Transportation Studies,* Vol. 5, 2005, pp. 671–678.
149. Momm, L., E. Meurer Filho, and L. L. Bariani Bernucci. Study of the Aggregate for the Pervious Asphalt Concrete. In *Performance Testing and Evaluation of Bituminous Materials: Proceedings of the 6th International RILEM Symposium, Zurich, Switzerland, April 14–16, 2003,* 2003, pp. 237–243.
150. 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: Journal of the Transportation Research Board*, 2023, Article 03611981231172749.
151. Westerman, J. R. AHTD's Experience with Superpave Pavement Permeability. Presented at the *Arkansas Superpave Symposium,* 1998.
152. 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,* Vol. 67, 1998.

153. Khosla, N. P., and S. Sadasivam. *Determination of Optimum Gradation for Resistance to Permeability, Rutting and Fatigue Cracking*. Report No. FHWA/NC/2005-11. Center for Transportation Engineering Studies, North Carolina State University, Raleigh, 2005.
154. Hainin, M. R., L. A. Cooley, Jr., and B. D. Prowell. An Investigation of Factors Influencing Permeability of Superpave Mixes. *International Journal of Pavements,* Vol. 2, No. 2, 2003 pp. 41–52.
155. Bhasin, A., J. W. Button, A. Chowdhury, and E. Masad. *Analysis of South Texas Aggregates for Use in Hot Mix Asphalt*. Report No. FHWA/TX-05/0-4203-4. Texas Transportation Institute, Texas A&M University System, College Station, 2005.
156. Christensen, D. W., A. Hanz, R. Velasquez, A. Arshadi, and H. Bahia. *Refinement of Current WisDOT HMA Mixture Application Guidelines Related to NMAS and Aggregate Characteristics*. Report No. 0092-12-01. Wisconsin Highway Research Program, University of Wisconsin-Madison, 2013.
157. Hafeez, I., M. A. Kamal, M. W. Mirza, and A. A. Khanzada. Investigating the Effects of Maximum Size of Aggregate on Rutting Potential of Stone Mastic Asphalt. *Pakistan Journal of Engineering and Applied Sciences*, Vol. 10, 2012, pp. 89–96.
158. Christensen, D. W. and R. F. Bonaquist. *Evaluation of Indirect Tensile Test (IDT) Procedures for Low-Temperature Performance of Hot Mix Asphalt*. NCHRP Report 530. Transportation Research Board, Washington, DC, 2004.
159. William, W. R., G. A. Huber, W. J. Pine, S. H. Carpenter, and R. Bailey. Bailey Method for Gradation Selection in Hot-Mix Asphalt Mixture Design. *Transportation Research Circular E-C044*, Transportation Research Board, Washington, DC, 2002.
160. Vavrik, W. R., W. J. Pine, G. A. Huber, S. H. Carpenter, and R. Bailey. The Bailey Method of Gradation Evaluation: the Influence of Aggregate Gradation and Packing Characteristics on Voids in the Mineral Aggregate (with Discussion). *Journal of the Association of Asphalt Paving Technologists*, Vol. 70, 2001, pp. 132–175.
161. Zhang, H. *Asphalt Mixture Volume Index and Anti-Rutting Performance Unified Prediction Model and Its Application*. Master's thesis. School of Civil Engineering, Shandong University, 2009.
162. Liu, H. Y. *Research on Asphalt Mixture Volume Parameter Characteristics and Gradation Design Method*. Master's thesis. Highway School, Chang'an University, Xi'an, China, 2007.
163. 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, Changsha, China, 2006.
164. Inoue, T., Y. Gunji, and H. Akagi. Rational Design Method of Hot Mix Asphalt Based on Calculated VMA. In *Proceedings of the Third Eurasphalt and Eurobitume Congress, Vienna, Austria, May 2004*, 2004.
165. Yu, A. B., and N. Standish. A Study of the Packing of Particles with a Mixture Size Distribution. *Powder Technology*, Vol. 76, No. 2, 1993, pp. 113–124.
166. 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 and Building Materials*, Vol. 188, 2018, pp. 292–304.

167. Li, X., and Z. He. Prediction Model for Connected Voids Ratio of the Porous Asphalt Mixture. *Advances in Civil Engineering*, Vol. 2020, 2020, Article 3760435.
168. Othman, K. Prediction of the Hot Asphalt Mix Properties Using Deep Neural Networks. *Beni-Suef University Journal of Basic and Applied Sciences*, Vol. 11, 2022, Article 40.
169. Shaikh, S., A. Gupta, V. Raj, and A. Singh. Data-Driven Predictive Models for Evaluating Optimum Binder Content and Volumetric Properties of Bituminous Mixtures Using Design Variables. *Measurement*, Vol. 253, 2025, Article 117411.
170. Ezzat, E. N., and A. H. Abed. Prediction of Modified Asphalt Concrete Rutting Depth Using Statistical Model. *Materials Today: Proceedings*, Vol. 42, 2021, pp. 1784–1790.
171. Majidifard, H., B. Jahangiri, P. Rath, L. U. Contreras, W. G. Buttlar, and A. H. Alavi. Developing a Prediction Model for Rutting Depth of Asphalt Mixtures Using Gene Expression Programming. *Construction and Building Materials*, Vol. 267, 2021, Article 120543.
172. Ma, W. A Method for Preliminary Prediction of Asphalt Mixture Permanent Deformation Test Results. In *Airfield and Highway Pavements 2025,* 2025, pp. 238–245.
173. 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: Journal of the Transportation Research Board*, Vol. 2679, No. 10, 2025, pp. 870–887.
174. Erten, K. M., and R. Gürfidan. Regression-Based Performance Prediction in Asphalt Mixture Design and Input Analysis with SHAP. *Applied Sciences*, Vol. 15, No. 19, 2025, Article 10779.
175. Tong, B., W. Huang, J. Habbouche, I. Boz, Q. Guo, S. D. Diefenderfer, and G. W. Flintsch. Machine Learning–Based Prediction and Optimization of Balanced Mixture Design Performance Indices. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2679, No. 7, 2025, pp. 156–173.
176. Cooper, S. B., Jr., L. N. Mohammad, and M. A. Elseifi. Development of a Predictive Model Based on an Artificial Neural Network for the Semicircular Bend Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2576, 2016, pp. 83–90.
177. Nguyen, L. N., T.-H. Le, L. Q. Nguyen, and V. Q. Tran. Machine Learning Approaches for Predicting Cracking Tolerance Index (CTIndex) of Asphalt Concrete Containing Reclaimed Asphalt Pavement. *PLOS ONE*, Vol. 18, No. 10, 2023, Article e0287255.
178. Rad, S. M., M. Sadeghi, J. Bausano, and M. Elkashef. Evaluation of Asphalt Mixture Cracking Resistance and Development of a Machine Learning-Based Application for Cracking Tolerance Index Prediction. *Construction and Building Materials*, Vol. 490, 2025, Article 142519.
179. Mirzaiyanrajeh, D., E. V. Dave, and J. E. Sias. Developing a Prediction Model for Low-Temperature Fracture Energy of Asphalt Mixtures Using Machine Learning Approach. *International Journal of Pavement Engineering*, Vol. 24, No. 2, 2023, Article 2024185.
180. Hajj, E. Y., and N. G. Elias. *Balanced Mix Design: Rutting Performance Tests*. National Asphalt Pavement Association, Greenbelt, Md., 2025.

181. ASTM International. *ASTM D4867: Standard Test Method for Effect of Moisture on Asphalt Mixtures*. ASTM International, West Conshohocken, Pa., 2017, 2024.
182. American Association of State Highway and Transportation Officials. *AASHTO T 340: Standard Method of Test for Determining Rutting Susceptibility of Asphalt Mixtures Using the Asphalt Pavement Analyzer (APA)*. Washington, DC, 2023.
183. American Association of State Highway and Transportation Officials. *AASHTO T 324: Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures*. Washington, DC, 2019.
184. American Association of State Highway and Transportation Officials. *AASHTO T 401: Standard Method of Test for Cantabro Abrasion Loss of Asphalt Mixture Specimens*. Washington, DC, 2022.
185. Florida Department of Transportation. *FM 5-565: Measurement of Water Permeability of Compacted Asphalt Paving Mixtures*. State Materials Office, Gainesville, Fla., 2015.
186. ASTM International. *ASTM D8225-19: Standard Test Method for Determination of Cracking Tolerance Index of Asphalt Mixture Using the Indirect Tensile Cracking Test at Intermediate Temperature*. West Conshohocken, Pa., 2019.
187. American Association of State Highway and Transportation Officials. *AASHTO T 393: Standard Method of Test for Determining the Fracture Potential of Asphalt Mixtures Using the Illinois Flexibility Index Test (I-FIT)*. Washington, DC, 2022.
188. ASTM International. *ASTM D7313: Standard Test Method for Determining Fracture Energy of Asphalt Mixtures Using the Disk-Shaped Compact Tension Geometry*. West Conshohocken, Pa., 2020.
189. American Association of State Highway and Transportation Officials. *AASHTO T 304: Uncompacted Void Content of Fine Aggregate*. Washington, DC, 2022.
190. American Association of State Highway and Transportation Officials. *AASHTO T 27: Standard Method of Test for Sieve Analysis of Fine and Coarse Aggregates*. Washington, DC, 2024.
191. Cooley, L. A., Jr., R. C. Ahlrich, R. S. James, B. D. Prowell, E. R. Brown, and A. Kvasnak. Implementation of Superpave Mix Design for Airfield Pavements. In *Proceedings of the 2007 FAA Worldwide Airport Technology Transfer Conference*, Atlantic City, N.J., 2007.
192. ASTM International. *ASTM C670: Standard Practice for Preparing Precision and Bias Statements for Test Methods for Construction Materials*. West Conshohocken, Pa., 2025.
193. Azari, H., *Precision Estimates of AASHTO T 283: Resistance of Compacted Hot Mix Asphalt to Moisture-Induced Damage*. NCHRP Web-Only Document 166, Transportation Research Board, Washington, DC, 2010.
194. Kim, S., and B. J. Coree. *Evaluation of Hot Mix Asphalt Moisture Sensitivity Using the Nottingham Asphalt Test Equipment*. Center for Transportation Research and Education, Iowa State University, Ames, 2005.
195. 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: Journal of the Transportation Research Board, No. 2209, 2011, pp. 70–81.

196. West, R. C., M. A. Heitzman, D. M. Rausch, and G. Julian. *Laboratory Refinement and Field Validation of 4.75 mm Superpave Designed Asphalt Mixtures.* NCAT Report 11-01. National Center for Asphalt Technology, Auburn, Ala., 2011.
197. Habbouche, J., I. Boz, S. D. Diefenderfer, B. S. Underwood, and O. Othman. *Mechanistic-Based Evaluation of Performance Thresholds for Balanced Mix Design Asphalt Surface Mixtures.* Report No. FHWA/VTRC 26-R08. Virginia Transportation Research Council, Charlottesville, 2025.
198. 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, Charlottesville, 2022.
199. Hicks, R. G. *Moisture Damage in Asphalt Concrete*. NCHRP Synthesis of Highway Practice 175. Transportation Research Board, Washington, DC, 1991.
200. Alvarez, A. E., A. Epps Martin, and C. Estakhri. A Review of Mix Design and Evaluation Research for Permeable Friction Course Mixtures. *Construction and Building Materials*, Vol. 25, No. 3, 2011, pp. 1159–1166.
201. Oliver, J. W. H. The Effect of Binder Film Thickness on Asphalt Cracking and Ravelling. *Road & Transport Research*, Vol. 20, No. 3, 2011, pp. 3–13.
202. Elseifi, M. A., I. L. Al-Qadi, S.-H. Yang, and S. H. Carpenter. Validity of Asphalt Binder Film Thickness Concept in Hot-Mix Asphalt. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2057, 2008, pp. 37–45.
203. Jung, D., and T. S. Vinson. *Low-Temperature Cracking: Test Selection*. SHRP-A-400. Strategic Highway Research Program, National Research Council, Washington, DC, 1994.
204. Heritier, S., E. Cantoni, S. Copt, and M.-P. Victoria-Feser. *Robust Methods in Biostatistics*. John Wiley & Sons, Chichester, United Kingdom, 2009.
205. Huber, P. J. Robust Estimation of a Location Parameter. In *Breakthroughs in Statistics: Methodology and Distribution* (S. Kotz and N. L. Johnson, eds.), Springer, New York, 1992, pp. 492–518.
206. Kutner, M. H., C. J. Nachtsheim, J. Neter, and W. Li. *Applied Linear Statistical Models,* 5th ed., McGraw-Hill/Irwin, New York, 2005.
207. Hand, A. J. T., and A. Epps. Effects of Test Variability on Mixture Volumetrics and Mix Design Verification. *Asphalt Paving Technology*, Vol. 69, 2000, pp. 635–674.
208. Zhou, F. *Development of an IDEAL Cracking Test for Asphalt Mix Design, Quality Control and Quality Assurance.* NCHRP-IDEA Program Project Final Report 195. Transportation Research Board, 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 this 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 (15) and (16).

Table 51 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 of aggregate passing at No. 1/2, No. 3/8 … No. 200 sieves ($\Delta\text{No. } 1/2$, $\Delta\text{No. } 3/8$ … $\Delta\text{No. } 200$), group-centered value of the specific gravity of the aggregate blend ΔG_{sb}, and group-centered value of fine aggregate angularity of the blend (ΔFAA).
- Response variable: the group-centered value of VMA (ΔVMA).

Table 51. 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

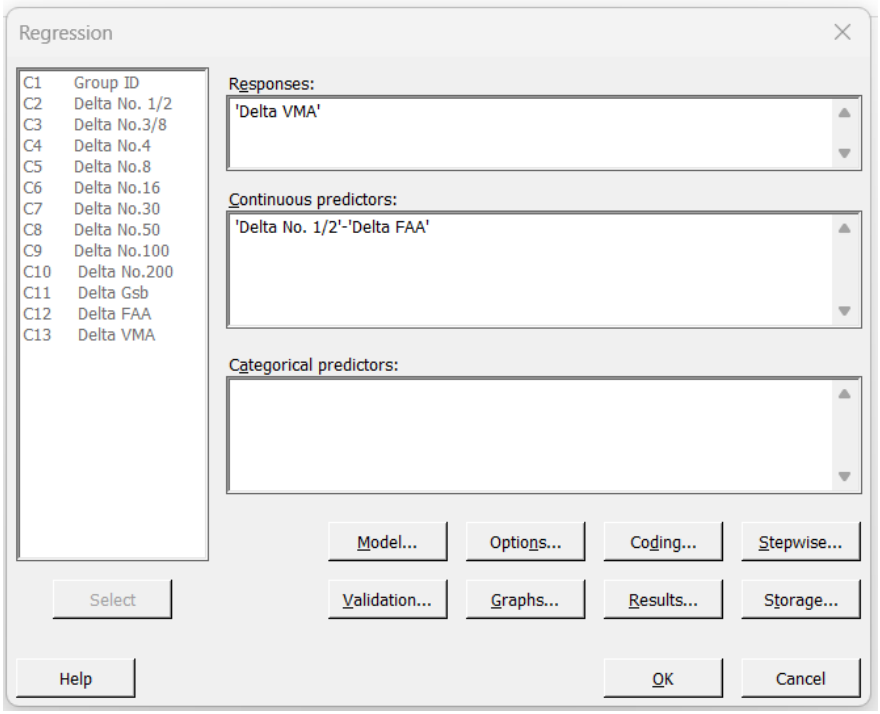
This first regression uses equal weights of 1 (standard OLS) to obtain an initial set of residuals.

2.1. Open Minitab and import the Excel file from Step 1 using File → Open → Excel.

2.2. Navigate to Stat → Regression → Regression → Fit Regression Model.

2.3. In the regression dialog box, as shown in Figure 56:

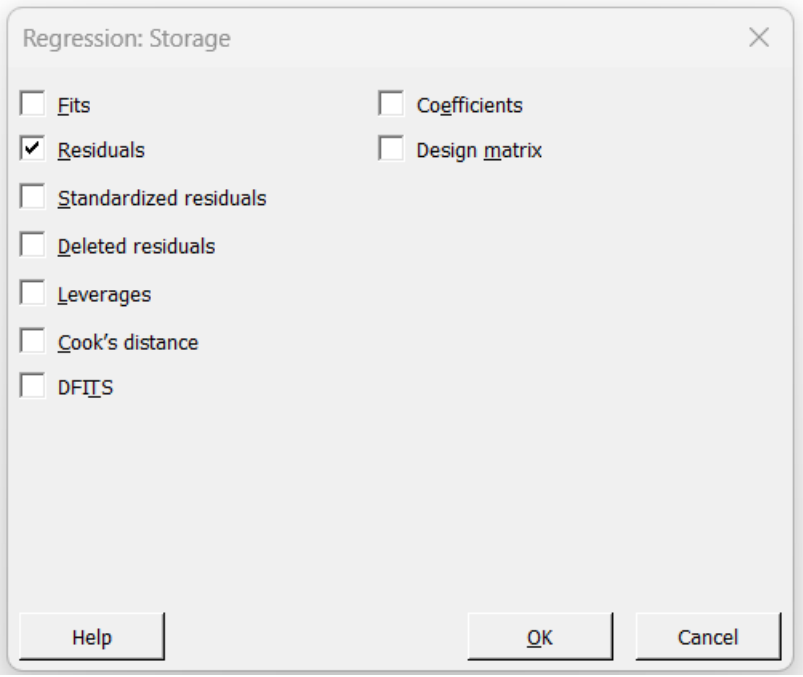
- Responses: Set to ΔVMA
- Continuous predictors: Set to $\Delta\text{No. } 1/2$, $\Delta\text{No. } 3/8$ … $\Delta\text{No. } 200$, ΔG_{sb}, and ΔFAA (if the model includes FAA).



Source: AAPTP

Figure 56. Minitab Regression Dialog Box

2.4. Click the Storage button in the regression dialog box (Figure 56) to open the Regression: Storage dialog box (Figure 57), check the Residuals box, and click OK. Running the regression will create a new column labeled RESI in the Minitab worksheet.



Source: AAPTP

Figure 57. Minitab Regression: Storage Options

2.5. Click OK in the main regression window (Figure 56) to run the regression. Minitab will produce the initial regression output and store residuals in column RESI (C14), as shown in Table 52. The residuals represent the difference between the measured and fitted group-centered response values.

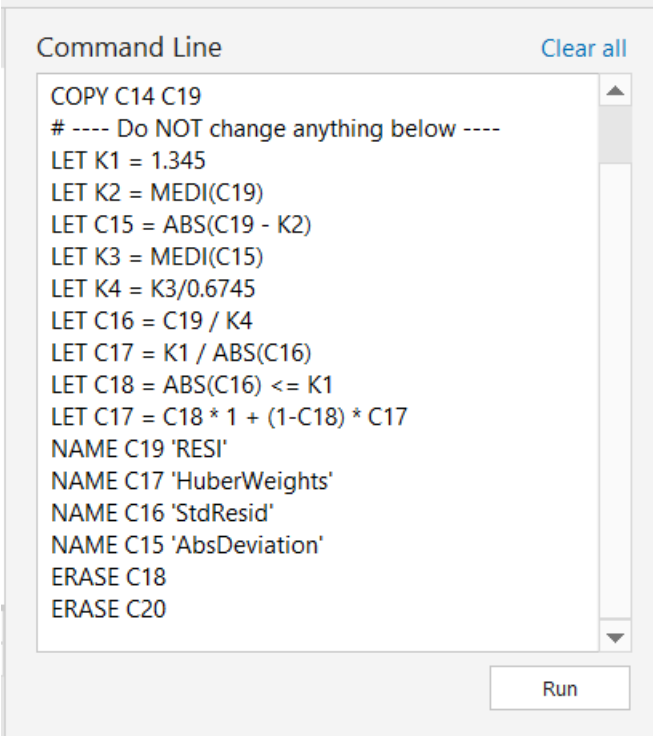
Table 52. Minitab Initial Regression Output with Residual in Column C14

	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.

3.1. Navigate to View → Command Line/History. A Command Line window will appear. Add a code that was used to calculate Huber weights, built based on Equations (19)–(22), to the Command Line window, as shown in Figure 58. Note that only the first line needs to be changed in each iteration.



Source: AAPTP

Figure 58. Command Line with Huber Weight Calculation Code

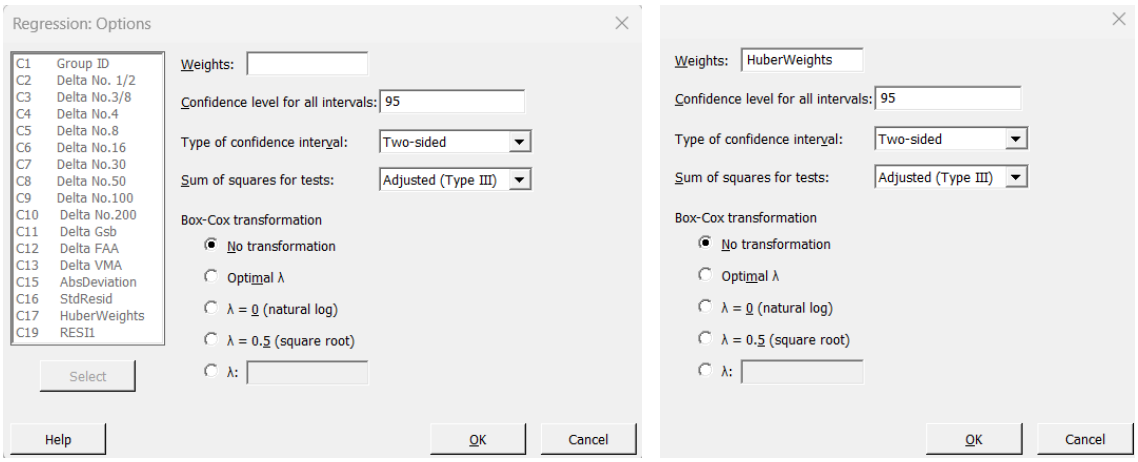
3.2. Click Run in the Command Line window (Figure 58). The Huber weights will be calculated and stored in column C17 (HuberWeights).

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.

4.1. Repeat Step 2.2 to open the regression dialog box (Figure 56) and keep all predictors and response settings identical to Step 2.

4.2. Click Options in the main regression window (Figure 56) to open the Regression: Options dialog box (Figure 59). Locate the Weights field, select C17 HuberWeights, and click Select. Then click OK to close the dialog box.



Source: AAPTP

Figure 59. Minitab Regression: Options—Selecting Huber Weights

4.3. Click OK in the main regression window (Figure 56) to run the weighted regression. A residual that now accounts for Huber weights will be stored in column C20.

Step 5. Recalculate Huber Weights and Iterate

5.1. 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.

5.2. Repeat steps 4 and 5.1 until the regression coefficients stabilize across iterations.

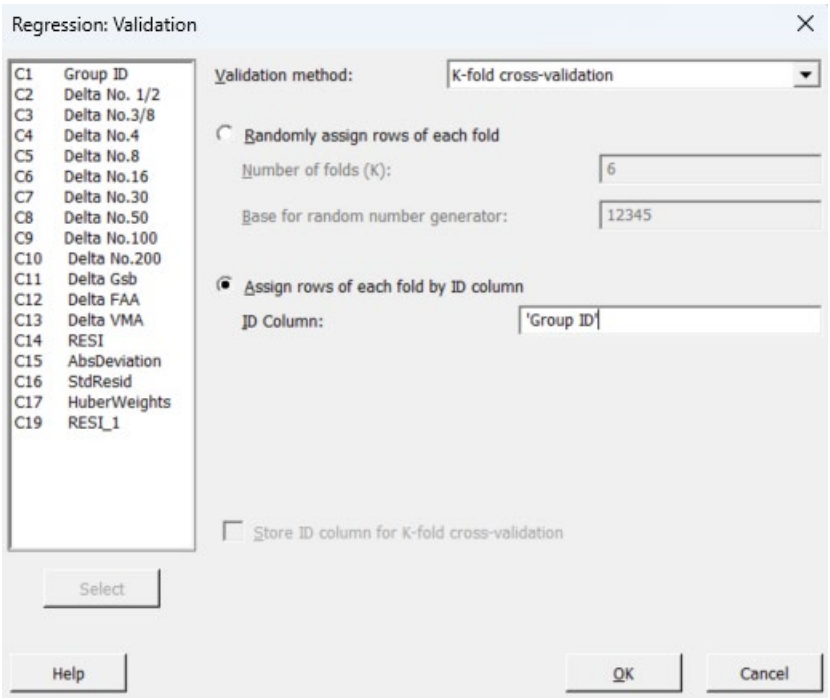
Step 6. K-fold Cross-Validation

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

6.1. Navigate to Stat → Regression → Regression → Fit Regression Model to open the main regression window (Figure 56).

6.2. Click Validation to open the Regression: Validation dialog box. Under Validation method, select K-fold cross-validation. Select the "Assign rows 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 60.

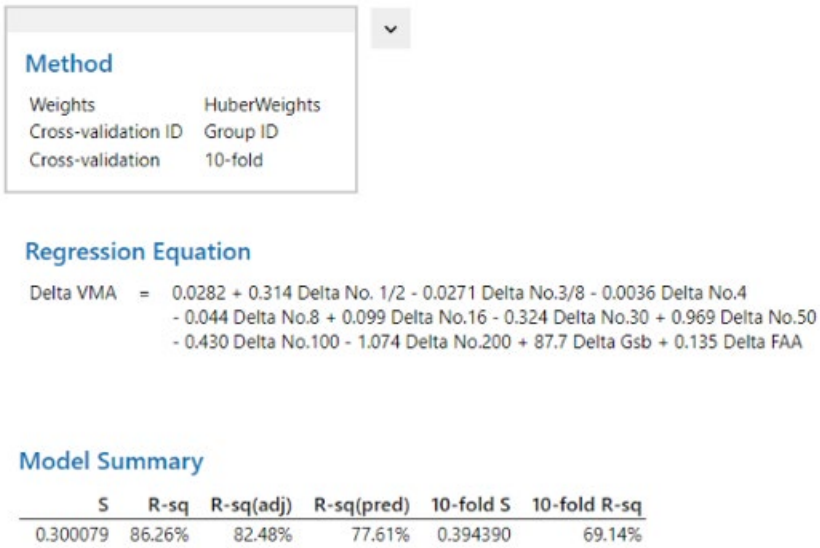
6.3. Click OK in the main regression window (Figure 56) to run the cross-validation.



Source: AAPTP

Figure 60. K-fold Cross-Validation Using Group ID

The results, including the regression equation and model performance metrics, are shown in Figure 61.



Source: AAPTP

Figure 61. Minitab Results of Second VMA Prediction Model