Balanced Asphalt Mix Design for North Carolina

NCDOT Project 2023-02
FHWA/NC/2023-02
August 2025

NC STATE UNIVERSITY

B. Shane Underwood, Ph.D., et al.
Civil, Construction, and Environmental Engineering,
North Carolina State University

STATE OF NORTH CAROLINA DEPARTMENT OF TRANSPORTATION

RESEARCH & DEVELOPMENT

Balanced Asphalt Mix Design for North Carolina

FINAL REPORT

Submitted to:
North Carolina Department of Transportation
Research and Development Unit
(Research Project No. RP2023-02)

Submitted by

B. Shane Underwood, Ph.D., Cassie Castorena, Ph.D., Y. Richard Kim, Ph.D.Pablo Vestena, and Jaime Preciado
Civil, Construction, and Environmental Engineering,
North Carolina State University
915 Partners Way
Raleigh NC 27606
919-515-8632
shane.underwood@ncsu.edu

North Carolina State University

October 2025

Technical Documentation Page

1. Report No. *FHWA/NC/2023-02*	2. Government Accession No. No	3. Recipient's Catalog No. …
4. Title and Subtitle Balanced Asphalt Mix Design for North Carolina		5. Report Date August, 2025
		6. Performing Organization Code
7. Author(s) B. Shane Underwood, Ph.D., Cassie Castorena, Ph.D., Y. Richard Kim, Ph.D., Pablo Vestena, and Jaime Preciado		8. Performing Organization Report No. None
9. Performing Organization Name and Address Civil, Construction, and Environmental Engineering, North Carolina State University 915 Partners Way Raleigh NC 27606		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address North Carolina Department of Transportation Research and Development Unit 1549 Mail Service Center Raleigh, North Carolina 27699-1549		13. Type of Report and Period Covered August 1, 2022 – July 31, 2025
		14. Sponsoring Agency Code RP2023-02

Supplementary Notes:

16. Abstract

Traditional asphalt mixture design in North Carolina, which is based primarily on volumetric properties and rutting performance, may not sufficiently ensure long-term durability across diverse pavement conditions. While the Asphalt Pavement Analyzer (APA) has been used to mitigate surface rutting, and the Tensile Strength Ratio (TSR) to manage moisture susceptibility, mix designs lack the capability to address cracking, leaving a critical gap in achieving balanced performance. This research addresses that gap by evaluating and proposing a framework for integrating both cracking and rutting performance tests into the mix design and quality assurance processes—a methodology known as Balanced Mix Design (BMD). A comprehensive experimental program was conducted on 14 surface mixtures obtained across North Carolina. Laboratory testing included Indirect Tension Cracking Test (IDT-CT) for cracking resistance, APA for rutting, Indirect Tensile Strength at High Temperature (IDT-HT), Cyclic Fatigue (CF), Dynamic Modulus (DM), and Stress Sweep Rutting (SSR). Performance predictions were made using FlexPAVE™ to link laboratory results with anticipated field performance. The study found substantial variability in both rutting and cracking resistance among mixtures, with performance strongly tied to plant source. Cracking resistance, in particular, showed complicated relationships across tests, with no single test perfectly predicting the simulated field performance. Nonetheless, the IDT-CT test demonstrated promise as a practical and efficient measure of cracking resistance, with a preliminary threshold value of 14 suggested for North Carolina surface mixtures. The IDT-HT test emerged as a cost-effective and efficient rutting evaluation tool. Preliminary minimum strength thresholds of approximately 150 kPa, 180 kPa, and 215 kPa were identified for RS9.5B, RS9.5C, and RS9.5D mixtures, respectively. Importantly, performance test results aligned well with simulated rutting but were less predictive of cracking outcomes, reinforcing the need for multiple tests to ensure a comprehensive understanding of mixture behavior. Survey responses from both contractors and NCDOT personnel confirmed that current APA-based approvals rarely lead to test failures or quality concerns, suggesting that the opportunity may exist to adjust mixtures to improve IDT-CT results while producing mixtures that meet existing APA thresholds. While concerns regarding training, equipment, and lab capacity were noted, personnel expressed a willingness to adopt a phased implementation strategy. This study concludes with recommendations to incorporate IDT-CT and IDT-HT tests into North Carolina's asphalt mix design and QA framework. Doing so may provide a more balanced assessment of mixture performance, reduce variability in field outcomes, and extend pavement life. Future work should include refinement of performance thresholds and a pilot implementation to validate specification updates in real-world conditions.

17. Key Words Asphalt Concrete, Balanced Mix Design, IDEAL-CT, pavement performance, FlexPAVE	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 204	22. Price … …

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

Disclaimer

The contents of this report reflect the views of the author who is responsible for the facts and the accuracy of the data presented herein. The contents of the report do not reflect the official views or policies of the North Carolina Department of Transportation. This report does not constitute a standard, specification, or regulation."

Acknowledgments

The research team would like to express their gratitude and appreciation to the North Carolina Department of Transportation (NCDOT) for the provided funding needed to conclude this research study.

TABLE OF CONTENTS

Table of Contents ... i
List of Figures ... iii
List of Tables ... vii
Executive Summary ... 1
1. Introduction ... 2
  1.1. Overview ... 2
  1.2. Status of the Literature ... 4
    1.2.1. Summary of BMD Test Methods ... 4
    1.2.2. State Efforts with BMD ... 7
    1.2.3. Integrating BMD into Quality Assurance ... 7
    1.2.4. Test Costs ... 9
    1.2.5. Knowledge Gaps ... 10
  1.3. Operational Implications of BMD in North Carolina ... 10
    1.3.1. Survey with Contractors ... 11
    1.3.2. Survey with North Carolina DOT ... 12
  1.4. Report Organization ... 13
2. Materials and Methods ... 15
  2.1. Study Overview ... 15
  2.2. Material Selection and Sampling ... 15
  2.3. Test Methods ... 18
    2.3.1. Specimen Fabrication ... 18
    2.3.2. Indirect Tension Cracking Test (IDT-CT) ... 19
    2.3.3. Asphalt Pavement Analyzer (APA) ... 20
    2.3.4. Indirect Tension-High Temperature Test (IDT-HT) ... 20
    2.3.5. AMPT Testing ... 20
3. Experimental Results ... 22
  3.1. Overview ... 22
  3.2. Benchmarking ... 22
    3.2.1. BMD Testing ... 22
    3.2.2. AMPT Testing ... 25
    3.2.3. Summary of Benchmarking ... 29
  3.3. IDT-HT Feasibility Study ... 30
  3.4. Comparison Laboratory Aging Protocol ... 31
  3.5. Short-Term Aging Protocol Variation on RS9.5D mixes ... 32
  3.6. Sensitivity of Cracking Index Results to Asphalt Supplier ... 35
  3.7. Performance Evaluation ... 40
    3.7.1. Rutting Predictions ... 41
    3.7.2. Fatigue Damage Predictions ... 42
    3.7.3. Summary of Performance Evaluation ... 44
4. Implementation Implications ... 47
  4.1. Overview ... 47
  4.2. Mix Design Specifications ... 47
  4.3. Impacts on Quality Assurance ... 50
  4.4. Efficiency Opportunities ... 51

5. Conclusions and Recommendations 53
5.1. Conclusions 53
5.2. Recommendations 54
6. Implementation and Technology Transfer Plan 56
7. References 57
Appendix A. Detailed Literature Review 60
Appendix B. Detailed Analysis of Contractor Survey 129
Appendix C. Detailed Summary of BMD experiments 133
Appendix D. Detailed AMPT Experimental Results 150
Appendix E. Detailed Performance Evaluation Summary 169
Appendix F. Draft Specification on Mix Design 187

LIST OF FIGURES

Figure 1. Predicted performance of three different asphalt mixtures of the same class: (a) RS9.5C mixtures with a thick ABC pavement structure and (b) RS9.5B mixtures with an intermediate deep strength structure. 3
Figure 2. Flowchart of the research. 15
Figure 3. Geographical distribution of sampled component materials. 16
Figure 4. IDT-CT results for benchmarking study. 22
Figure 5. APA results for benchmarking study. 25
Figure 6. DM results for benchmarking study. 26
Figure 7. CF results for benchmarking study. 28
Figure 8. SSR results for benchmarking study. 29
Figure 9. IDT-HT results for selected mixtures. 30
Figure 10. IDT-HT results correlations with other rutting tests: (left) APA; (right) SSR. 31
Figure 11. Comparison of CT_{index} values from PMLC and lab-mixed, lab-compacted specimens for FB and HC mixtures. 32
Figure 12. Comparison of APA values from PMLC and lab-mixed, lab-compacted specimens for FB and HC mixtures. 32
Figure 13. IDT-CT and S_{app} results for RS9.5D STOA study. 33
Figure 14. DM test results for RS9.5D STOA study. 33
Figure 15. IDT-CT results for asphalt supplier experiment. 37
Figure 16. DM test results for asphalt supplier experiment. 38
Figure 17. CF results for asphalt suppler study. 39
Figure 18. Rutting in the surface layer from FlexPAVE™ simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 41
Figure 19. Top percent damage from FlexPAVE™ simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 42
Figure 20. Top percent damage from FlexPAVE™ additional simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 43
Figure 21. Comparison of pavement performance simulations showing grouping for ABCtk pavements. 44
Figure A.1. Overview of I-FIT test; (a) instrumentation and (b) specimen geometry. (Rivera-Pérez *et al.* 2021). 62
Figure A.2. IDT-CT overview; (a) load-displacement curve and (b) test setup and conditions (Zhou *et al*. 2017b). 67
Figure A.3. BBF overview; (a) test configuration and (b) specimen cross section (Kim *et al.* 2011). 73
Figure A.4. OT overview; (a) test configuration and (b) specimen dimensions (Walubita *et al.* 2013). 74
Figure A.5. HWTT overview; (a) typical results and (b) gyratory mounting (AASHTO 2023a). 82
Figure A.6. IDEAL-RT Test apparatus and stresses (Yin *et al*. 2020b, Zhou 2021). 86
Figure A.7. Alternative HWTT data analysis based on CRD_{20k} and *SN* (West *et al.* 2021b). 109
Figure C.1. Plant F (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 133
Figure C.2. Plant Y (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 134

Figure C.3. Plant W (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 135
Figure C.4. Plant S (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 136
Figure C.5. Plant U (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 137
Figure C.6. Plant H (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 138
Figure C.7. Plant Z (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 139
Figure C.8. Plant V (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 140
Figure C.9. Plant X (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 141
Figure C.10. Plant U (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 142
Figure C.11. Plant R (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 143
Figure C.12. Plant V (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 144
Figure C.13. Plant T (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 145
Figure C.14. Plant S (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve. 146
Figure D.1. Plant F (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve... 150
Figure D.2. Plant F (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 151
Figure D.3. Plant W (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 152
Figure D.4. Plant S (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve... 153
Figure D.5. Plant U (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve... 154
Figure D.6. Plant H (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve... 155
Figure D.7. Plant Z (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right)

Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 156
Figure D.8. Plant V (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 157
Figure D.9. Plant X (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 158
Figure D.10. Plant U (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 159
Figure D.11. Plant R (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 160
Figure D.12. Plant V (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 161
Figure D.13. Plant T (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 162
Figure D.14. Plant S (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve. 163
Figure E.1. Damage contours FB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 169
Figure E.2. Damage contours YB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 170
Figure E.3. Damage contours WB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 171
Figure E.4. Damage contours SB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 172
Figure E.5. Damage contours UB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 173
Figure E.6. Damage contours HC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 174
Figure E.7. Damage contours ZC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 175
Figure E.8. Damage contours VC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 176
Figure E.9. Damage contours XC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 177
Figure E.10. Damage contours UC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 178

Figure E.11. Damage contours RC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 179
Figure E.12. Damage contours VD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 180
Figure E.13. Damage contours VD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 181
Figure E.14. Damage contours SD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick. 182
Figure E.15. Rutting prediction for UC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 183
Figure E.16. Rutting prediction for VC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 183
Figure E.17. Rutting prediction for VD structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 184
Figure E.18. Rutting prediction for WB structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 184
Figure E.19. Rutting prediction for YB structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 185
Figure E.20. Rutting prediction for ZC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick. 185

LIST OF TABLES

Table 1. Comparison of Different Cracking Test Methods (Adapted from Zhou *et al*. 2017a). 5
Table 2. Comparison of Different Rutting Test Methods (Adapted from (Zhou *et al.* 2020). 6
Table 3. State DOTs BMD Cracking Tests Summary. 8
Table 4. State DOTs BMD Rutting Tests Summary.. 8
Table 5. Summary of Testing Equipment Prices. 10
Table 6. Asphalt Content Decision. 17
Table 7. Mixtures Information. 17
Table 8. Summary of Performance Tests.. 18
Table 9. Individual Mixture Performance.. 29
Table 10. Binder Characterization for Binder Supplier Sensitivity Study................................. 35
Table 11. Experimental Program of Binder Supplier Sensitivity Study. 36
Table 12. Summary of Statistical Analysis of Binder Supplier Sensitivity Study....................... 37
Table 13. Pavement Structures used in Pavement Simulations. 40
Table 14. Traffic Characteristics used in Pavement Simulations. 40
Table 15. Material Characteristics used in Pavement Simulations. 40
Table 16. Summary of Rutting Ranking. 44
Table 17. Summary of Fatigue Damage and Durability Ranking. 45
Table 18. Summary of Responsibilities from Alternative Mix Design Scenarios....................... 48
Table 19. Summary of Time, Cost and Control from Alternative Mix Design Scenarios per Mix Design. 50
Table A.1. Summary of Performance Indicators from I-FIT. 63
Table A.2. Summary of Findings using the I-FIT Method. 64
Table A.3. Summary of Findings using the SCB-LTRC Method... 66
Table A.4. Summary of Findings using the IDT-CT Method... 68
Table A.5. Summary of Findings using the Cantabro Method. 70
Table A.6. Summary of Findings using the Uniaxial CF Method. 72
Table A.7. Summary of Findings using the BBF Method. 74
Table A.8. Summary of Findings using the Overlay Test. 76
Table A.9. Comparison of Different Cracking Test Methods (Adapted from Zhou *et al*. 2017a). 78
Table A.10. Summary of Findings using the APA Method... 81
Table A.11. Summary of Findings using HWTT Method. 83
Table A.12. Summary of Findings using the IDT-HT Method. 85
Table A.13. Summary of Findings using the Confined FN Method... 88
Table A.14. Summary of Findings using the SSR Method... 90
Table A.15. Comparison of Different Rutting Test Methods (Adapted from Zhou *et al.* 2020).. 91
Table A.16. Summary of Additional Steps for Specimen Preparation Prices. 94
Table A.17. Summary of Testing Equipment Prices. 96
Table A.18. Eight Potential Tasks for BMD Implementation (FHWA 2022). 98
Table A.19. BMD Implementation for Different States (NAPA 2020)..................................... 100
Table A.20. ALDOT IDT-CT Criteria (ALDOT 2020). 101
Table A.21. Caltrans HWTT Thresholds (Caltrans 2022). 101
Table A.22. GDOT HWTT Criteria (GDOT 2021). 102
Table A.23. GDOT Proposed IDT-CT Criteria (Sala *et al.* 2022).. 102

Table A.24. Illinois Modified AASHTO T 324 Requirements (IDOT 2022). 102
Table A.25. Illinois Modified AASHTO T 393 Requirements (IDOT 2022). 103
Table A.26. Nebraska DOT Recommended G-stability Criteria (Nsengiyumva *et al.* 2020). ... 104
Table A.27. New Jersey DOT Performance Tests Criteria (Bennert 2022). 105
Table A.28. NJDOT Proposed IDT-CT and IDT-HT Criteria (Bennert 2022). 105
Table A.29. Performance Test Criteria for the Performance Tests used in the NYSDOT BMD Study (Bennert *et al.* 2022). 106
Table A.30. ODOT Proposed CT_{index} Limits (Abbas *et al.* 2021). 107
Table A.31. ODOT HWTT Criteria (ODOT 2019). 107
Table A.32. WisDOT Recommended Performance Tests Thresholds (West *et al.* 2021b). 109
Table A.33. State DOTs BMD Rutting Tests Summary. 110
Table A.34. State DOTs BMD Cracking Tests Summary. 111
Table B.1. Contractor Questions Related to Demographic and General Information. 129
Table B.2. Contractor Questions Related to Operational Issues. 130
Table B.3. Contractor Questions Related to Training. 131
Table B.4. Contractor Questions Related to General Practices. 132
Table C.1. Analysis of Variance for IDT-CT on Individual Mixtures. 147
Table C.2. Analysis of Variance for IDT-CT on Mix Type. 147
Table C.3. Analysis of Variance for IDT-CT on Mix Type (Excluding WB and RS9.5D). 147
Table C.4. Mix Type and Source Region Mixed Model for IDT-CT. 147
Table C.5. Analysis of Variance for APA on Individual Mixtures. 148
Table C.6. Analysis of Variance for APA on Mix Type. 148
Table C.7. Analysis of Variance for APA on Mix Type (Excluding YB and RS9.5D). 148
Table C.8. Mix Type and Source Region Mixed Model for APA. 148
Table C.9. APA Rut Depths Difference to Reported in the Job Mix Formula. 149
Table D.1. Analysis of Variance for DM (4°C, 1Hz) on Individual Mixtures. 164
Table D.2. Analysis of Variance for DM (20°C, 1Hz) on Individual Mixtures. 164
Table D.3. Analysis of Variance for DM (40°C, 1Hz) on Individual Mixtures. 165
Table D.4. Analysis of Variance for DM (4°C, 1Hz) on Mix Type. 165
Table D.5. Analysis of Variance for DM (20°C, 1Hz) on Mix Type. 165
Table D.6. Analysis of Variance for DM (40°C, 1Hz) on Mix Type. 165
Table D.7. Mix Type and Source Region Mixed Model for DM (4°C, 1Hz). 166
Table D.8. Mix Type and Source Region Mixed Model for DM (20°C, 1Hz). 166
Table D.9. Mix Type and Source Region Mixed Model for DM (40°C, 1Hz). 166
Table D.10. Analysis of Variance for S_{app} on Individual Mixtures. 167
Table D.11. Analysis of Variance for D^R on Individual Mixtures. 167
Table D.12. Analysis of Variance for S_{app} on Mix Type. 168
Table D.13. Analysis of Variance for D^R on Mix Type. 168
Table D.14. Mix Type and Source region Mixed Model for S_{app}. 168
Table D.15. Mix Type and Source Region Mixed Model for D^R. 168
Table E.1. Asphalt Mixtures Fatigue and Shift Factor Inputs for FlexPAVE^{TM}. 186
Table E.2. Asphalt Mixtures 2S2P1D Inputs for FlexPAVE^{TM} 186
Table E.3. Asphalt Mixtures Rutting Inputs for FlexPAVE^{TM}. 186

EXECUTIVE SUMMARY

Research conducted over the past two and a half decades has found that reliance on volumetric properties alone is insufficient for asphalt mixture design, as it can result in varied and unpredictable performance. This may be associated with the increased percentage of recycled content and/or the use of binder modifiers. North Carolina previously addressed this shortcoming with respect to surface rutting by integrating asphalt pavement analyzer (APA), testing to mix design approvals. However, focusing on a single performance test can compromise the mixture's resistance to other distresses, such as fatigue cracking. This gap has been identified in several national research projects over the past decade, as well as locally during project HWY2002-07.

To ensure asphalt mixtures exhibit sufficient and balanced performance, it is necessary to integrate tests addressing both rutting and fatigue cracking into mix design practices. The chosen tests should also be amenable to quality assurance and quality control activities. The lack of a consolidated performance test for durability assessment represents a critical gap in achieving an appropriate balance in North Carolina. This research addresses this need by identifying an appropriate cracking test and establishing a framework for its implementation.

For the study described in this report, a comprehensive laboratory evaluation was conducted on fourteen mixtures sourced from across the state, utilizing Indirect Tensile Asphalt Cracking Test (IDT-CT), Cyclic Fatigue, APA, Dynamic Modulus, and Stress Sweep Rutting tests. The study also included pavement performance simulations using FlexPAVE™ to bridge the gap between laboratory results and expected field performance. The key findings from this study are as follows;

- Performance Variation: The laboratory testing identified significant performance differences among current NCDOT mixtures. Testing consistently identified a 'worst' mixture with respect to cracking and a 'worst' mixture with respect to rutting from among the RS9.5B mixes. The RS9.5D mixtures showed exceptional rutting resistance.
- Lab vs. Expected Pavement Performance: While simulated rutting performance correlated well with lab tests, the correlation for cracking was more complex. No single laboratory cracking test perfectly predicted the simulated pavement performance, highlighting the need for a balanced approach rather than reliance on a single index.
- Influence of Material Sources: The plant source was found to be the dominant factor controlling mixture performance. The specific binder source, based on a limited dataset, did not have a statistically significant impact on cracking.
- Current Practices: Surveys with contractors confirmed that mix designs rarely fail the existing APA rutting criteria, and quality-related disputes are infrequent.

Based on the findings of this study, the IDT-CT shows promise as a primary performance test for evaluating cracking resistance. Based on the statistical analysis conducted herein, a preliminary threshold value of 14 was identified. To complement this finding, the IDT-HT test was shown to have promise as for gaining efficiency in asphalt mixture rutting evaluation. A small pilot study identified a potential minimum strength criterion of approximately 150 kPa, 180 kPa, and 215 kPa for RS9.5B, RS9.5C, and RS9.5D mixtures, respectively. Further evaluation would be beneficial to explore how these tests and criteria could be integrated into a formal Balanced Mix Design specification through a phased implementation plan, with the goal of enhancing long-term pavement durability.

1. Introduction

1.1. Overview

Asphalt mixture is an engineered material that is subjected to different traffic and climatic conditions throughout its service life. Different exposure conditions, in combination with changes in the material's properties with time, make asphalt mixtures susceptible to distresses such as fatigue cracking, thermal cracking, rutting, raveling, and others, affecting the functional and structural performance of the asphalt pavement. Different methods have been developed in the past to design asphalt mixtures to yield satisfactory pavement performance (NAPA 1982, Asphalt Institute 1997).

The current Superpave mix design method was developed during the Strategic Highway Research Program (SHRP) as a response to the awareness that pavement materials were not providing satisfactory performance (McDaniel *et al.* 2012). Considering the shortcomings of the empirically based Marshall and Hveem test methods, Contract A-003A *Performance-Related Testing and Measuring of Asphalt-Aggregate Interaction and Mixtures* of SHPR was aimed at developing reliable and reproducible test methods that could be used to characterize asphalt mixtures in terms of fundamental engineering properties. The SHRP efforts resulted in a hierarchical mix design system intended to encompass material characterization, volumetric mix design, and performance testing analysis depending on the design traffic level (Cominsky *et al.* 1994). Level 1 mix design consisted of a volumetric design with strict attention to the selection of the asphalt binder and aggregates. Levels 2 and 3 also considered volumetrics in addition to performance tests and predictions that would allow for the mix to be optimized with respect to one more distress. However, Levels 2 and 3 were, by and large, never implemented presumably due to complicated test protocols, equipment costs, and the general belief that volumetrics and binder performance could ensure adequate pavement performance (McDaniel *et al.* 2012, Diefenderfer and Bowers 2019). In the early 2000s, gaps in the use of Level 1 mix design started to become evident through observations of inadequate pavement performance, the incorporation of recycled asphalt materials into asphalt mixtures, and other factors. These gaps renewed the interest in integrating performance tests in mix design protocols (also known as performance mix design), and multiple research efforts have been conducted to allow for a successful implementation. In NCHRP Project 09-19, Witczak *et al.* (2002) developed simplified performance test (SPT) methods to address permanent deformation, fatigue cracking, and low-temperature cracking. However, the test protocols and attention to specific distress have varied across state agencies.

Since the initial deployment of Superpave mixture design, the current procedures for North Carolina asphalt mixture design have evolved. Currently, the NCDOT requires contractors to conform to volumetric requirements on the air void content, voids filled with asphalt, voids in mineral aggregate, and other parameters at a fixed, traffic- and layer-specific compaction effort. Recently, RP2019-20 has tested the dynamic modulus, cracking, and rutting performance of asphalt mixtures from across North Carolina produced under the most recent mixture design guidelines. These mixtures spanned multiple contractors and regions and included surface mixtures (at both the low and moderate traffic categories), intermediate mixtures, and base mixtures. In all instances, highly variable mixture properties were identified. These differences were substantial enough that pavement performance predicted by AASHTO Pavement ME Design (for a single weather station) varied by 80% or more with respect to predicted fatigue cracking and up to 40% with respect to predicted asphalt concrete layer rutting when the range of material properties for a single mixture category was considered. Two examples of the differences are shown in Figure 1 using a thick full depth asphalt (FDA) structure (17 total inches of asphalt concrete mixture) and an intermediate thickness deep strength pavement (10 total inches of asphalt concrete mixture and 8 inches of aggregate base course). In both cases only the surface course varied. Such variation in performance is, in itself, not a major problem as long as all mixtures meet a minimum level of acceptable performance. Unfortunately, under the current standards and practices, testing is not performed to ensure that this is the case with respect to cracking.

Differences like these are not currently captured in North Carolina mixture design standards and practices, nor are they captured in quality assurance and quality control procedures because direct assessment of mixture performance is not incorporated into either activity. However, as the data in Figure 1 demonstrates, the lack of such capabilities can have substantial implications in design, performance, and long-term management of roadways.

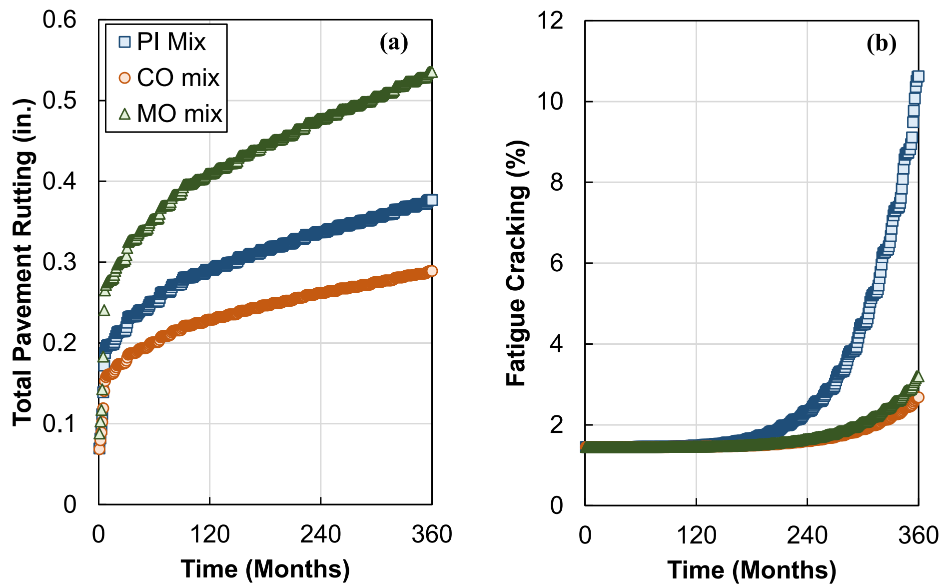


Figure 1. Predicted performance of three different asphalt mixtures of the same class: (a) RS9.5C mixtures with a thick ABC pavement structure and (b) RS9.5B mixtures with an intermediate deep strength structure.

The research described in this report was performed after realizing that current NCDOT practices, which rely on volumetrics, an assessment of permanent deformation characteristics, and an assessment of moisture damage potential, may have some shortcomings. The potential problem with one primary performance test focused on rutting is that the mixture's performance with respect to other distresses may be compromised. This issue has been identified in several national research projects over the past decade, as well as more locally during HWY2002-07. When these differences occur, it is necessary to integrate both rutting and cracking tests in order to properly 'balance' mixture performance.

The most certain way to ensure that asphalt mixtures will show sufficient and balanced performance with respect to both fatigue cracking and rutting is to integrate both performance tests into mix design practices and do so with tests that are also amenable to quality assurance and quality control activities. The lack of a performance test for cracking (durability) assessment is a critical gap in achieving an appropriate balance. Other needs that suggest the importance of performing this research include;

- Identifying the most appropriate test (or tests) and/or methodology for this purpose, given the equipment, limitations, and existing practices of the state and its contractors, as well as potential for production assessment.
- A test limit/threshold should exist to permit quantitative assessment of the quality of the materials and to judge whether a mixture is or is not acceptable.
- The lack of a formal process for integrating the test methods into mix design and quality assurance/control practices and recognizes the breadth of mixtures and categories currently used and how this integration will affect costs and productivity.

1.2. Status of the Literature

There has been substantial work conducted to develop and evaluate BMD. The efforts are documented in both literature and practice documents and focus on the following topics:

- Test methods (broadly divided into methods to address durability and methods to address stability),
- States efforts to implement BMD,
- Frameworks for implementing test methods into mix design and quality assurance, and
- Evaluations of test implementation costs.

A complete review of the work conducted in each of these areas is given in Appendix A of this report. A summary of the relevant findings with respect to each of these four areas follows.

1.2.1. Summary of BMD Test Methods

The literature review identified many test methods that have been developed and used to evaluate the durability of asphalt mixtures. Table 1, adapted from Zhou *et al.* (2017a), summarizes the most relevant and commonly used methods, including their advantages and limitations. A second version of this table with more information also appears in Table A.9 in Appendix A. Note that the costs presented in the table are an estimate for the equipment used for the given test alone.

Most studies suggest that the simpler test methods (I-FIT, IDT-CT, and SCB-LTRC) have the easiest pathway for BMD adoption. These three test methods yield index values that can be used to rank different materials and also assess the acceptability of a mix design. However, the results of the test are not amenable to conducting pavement performance predictions. The literature review also found that multiple studies have concluded that I-FIT and IDT-CT results tend to yield higher index values with higher air void content, which is not the expected trend. While intuitively incorrect, studies go on to make the point that this result may not be an issue for mix design if the testing is conducted at a fixed air void content. On the other hand, this characteristic does make the tests non-ideal for QA processes that seek to evaluate in-place density (Jeong *et al.* 2022). Trends for the SCB-LTRC test are also inconsistent.

Table 1. Comparison of Different Cracking Test Methods (Adapted from Zhou *et al*. 2017a).

Test name	Test standard	Parameter	Correlation to field performance	Typical Test Variability	Key Advantages	Key Drawbacks
IDT-CT	ASTM D8225	CT_{index}	Fair-to-Good (Zhou *et al*. 2017b; West *et al*. 2021a).	Medium-to-High	• No coring, cutting, and gluing is required. • The test duration is less than 5 minutes.	• The trends with air voids are not consistent with expectations. • The test is not amenable to conduct pavement performance predictions.
SCB-LTRC	ASTM D8044	J_c	Fair-to-Poor (Kim *et al*. 2012; West *et al*. 2021a)	Medium	• The analysis theory identifies a fundamental fracture parameter.	• At least twelve specimens are required for a full characterization of the mixture.
I-FIT	AASHTO TP 124	*FI, CRI* and *BCI*	Fair-to-Good (Ozer *et al*. 2018; West *et al*. 2021a)	Medium-to-High	• The test duration is amenable for routine characterization	• The test requires cut and notched specimens. • A minimum of six replicates are needed given the test variability.
Cantabro	AASHTO TP 108	*ML*	Unclear due to variability	Medium	• No cutting require; mix design specimens can be used as test specimens; relatively short test duration;	• The test does not explicitly evaluate any fundamental fracture mechanics/continuum damage properties.
Uniaxial Cyclic Fatigue	AASHTO T 411 AASHTO T 400	S_{app} , D^R	Promising-to-Good (West *et al*. 2021a)	Low	• Allows the prediction of fatigue performance at various conditions. • Results can be combined with pavement analysis model. • Only three specimens are required.	• Fabrication of the test specimens requires coring, cutting and gluing. • Requires specialized software for analysis. • Uncertainty in the test duration can be problematic for routine characterization.
Overlay Tester	Tex-248-F	N_f,	Promising-to-Good (West *et al*. 2021a).	Medium-to-High	• Interpretation of the results is relatively simple.	• Requires five sample due to inherent test variability. • Cutting and gluing is necessary.
Bending Beam Fatigue	AASHTO TP 321	N_f	Good as per SHRP-404.	High	• The test mimics bottom-up fatigue cracking and can be used for pavement performance predictions.	• Fabrication of specimens and test setup can be cumbersome. • The test specimens require significant amounts of material. • Typical analysis requires multiple tests at different strain levels and temperatures.

Repeated load tests attempt to simulate the process of fatigue damage accumulation. The main consequence of performing tests in this manner is that the resulting analysis generally provides input that can be used to perform pavement performance prediction, in addition to providing an index to rank or assess the acceptance of asphalt mixtures. The literature suggests that repeated load tests, bending beam fatigue (BBF), cyclic fatigue (CF), and the overlay test (OT) are able to capture the effect of mix design properties on the fatigue performance of asphalt mixtures. However, some of the main challenges associated with the implementation of repeated load tests for BMD purposes are the uncertainty in the time duration of the test. Another challenge is sample preparation, which can be lengthy due to coring, cutting, and gluing. Further, the training required to conduct the tests (setting up the equipment, mounting the specimen, etc.) and the data analysis are relatively more specialized than monotonic tests. Finally, the implementation cost and laboratory space for the equipment for BBF, CF, or OT is notably higher compared to the monotonic tests, which may represent a roadblock for state DOTs and contractors if multiple test machines are needed to supply the demand for testing.

Beyond durability testing, there are also several ways to assess an asphalt mixture's stability. To properly relate to what is expected from the material, a reliable field performance correlation should exist, but this method should also be simple enough to conduct that it can be implemented. Table 2 shows an overview regarding the six tests most commonly identified, with respect to costs of equipment, sample preparation, and testing time. A more detailed version of this table is shown in Table A.15 in Appendix A. Broadly, these six tests can be divided into three main categories: 1) small-scale simulation tests, 2) diametrical monotonic testing, and 3) repeated load testing.

Table 2. Comparison of Different Rutting Test Methods (Adapted from (Zhou *et al.* 2020).

Test name	Test Standard	Parameter | Criteria	Specimen Preparation
APA	AASHTO T 340	Rut depth | Pass/fail	Four samples, no cutting, nor gluing
HWTT	AASHTO T 324	Rut depth, *SIP* | Pass/fail	Four samples, one cut per specimen
HT- IDT	AASHTO T 283	Peak Strength | Pass/fail	Three samples, no coring/cutting
IDEAL-RT	ASTM D8360	RT_{Index} | Pass/fail	Three samples, no coring/cutting
Confined FN	AASHTO T 378	*FN* | Pass/fail or predictive model	Three samples, one coring per sample, two cuts per sample, membrane
SSR	AASHTO TP 134	RSI, p_1, p_2, d_1, d_2 | Pass/fail or FlexPAVE™	Three samples, one coring per sample, two cuts per sample, membrane

The small-scale simulation tests, Asphalt Pavement Analyzer (APA) and Hamburg Wheel Tracking Test (HWTT), evaluate rutting performance by applying repeated wheel load passes on the surface of an asphalt sample that is conditioned in air or conditioned under water. The tests produce results that are fairly simple to understand and have known repeatability expectations. For BMD purposes, these tests are most often performed to either a critical rutting level or until a preset number of cycles. These tests generally have (or are believed to have) a strong correlation to field performance, and so these types of tests are often used as a benchmark when generating new rutting test methodologies.

Indirect diametrical monotonic loading tests, the indirect tension-high temperature test (IDT-HT) and IDEAL Rutting Test (IDEAL-RT), are relatively simple to perform and have become increasingly popular due to the low equipment costs, easy sample preparation, and overall testing efficiency. The tests can take less than a day, or even six hours, from sample collecting, preparation, and testing, which is a more acceptable interval for quality assurance and control protocols. The analysis is based on the maximum shear induced in the sample, or on the energy curve for displacement versus shear strength. From the two described tests, cohesion and friction angle are the main characteristics assessed. Also, promising correlations to wheel-loading tests have been shown.

Mechanistic characterization tests, Confined Flow Number (CFN) and Stress Sweep Rutting (SSR), can be analyzed to yield material parameters amenable to pavement performance predictions. The repeated axial tests focus heavily on the stresses directly below the tire but can mimic both densification and shear phenomenon. Usually, there are performance prediction tools associated with the tests, which improve the benefit of the methodology to rank materials from their index, and to predict an expected rutting in the field. The biggest shortcoming of these types of tests is the cumbersome sample preparation protocols and the total testing time, which can take more than a day. The equipment is also more expensive, and the training requirements are greater.

1.2.2. State Efforts with BMD

Many states have completed and/or have active investigations into the development or deployment of BMD. The most common tests being adopted for rutting and cracking susceptibility are the HWTT (followed by APA) and the IDT-CT (followed by SCB-type tests), respectively. Out of the 18 state DOTs that were identified as developing or implementing BMD, 14 are using or considering the use of the HWTT, and 11 are using or considering the use of the IDT-CT in their BMD specifications. The temperatures at which these two most common tests are conducted are fairly consistent, typically 50°C for HWTT and 25°C for the IDT-CT. The most common air-void level at which all rutting and cracking performance tests are conducted is 7.0%. A summary of the durability and stability test details adopted by different states is summarized in Table 3 and Table 4, respectively (note that more comprehensive versions of these tables appear in Appendix A in Table A.34 and Table A.33 respectively). In these tables, the tests used are listed along with the temperature at which they are conducted, air-void level, and set or recommended criteria. Testing conditions and criteria used or recommended by the states differ for each test and can be affected by many variables, such as the binder PG, aging level, traffic, modifier usage, mix type, and number of design gyrations. To account for that and have an overall summary, the tables often contain a range of temperatures, air-void levels, or criteria.

1.2.3. Integrating BMD into Quality Assurance

Although necessary, an adequate mix design alone is not a guarantee of satisfactory pavement performance, and efforts must be made to integrate performance tests with specifications that assure that reasonable process and procedures are being followed in the production of the final engineered product. Depending on the quality characteristics used for acceptance, these specifications are often called quality assurance (QA) specifications or performance-related specifications (PRS).

Table 3. State DOTs BMD Cracking Tests Summary.

State	Test	Criteria
Alabama	IDT-CT	Design $CT_{index} \geq$ 55-110, Production $CT_{index} \geq$ 50-100
California	BBF	50% Loss of flexural stiffness at 10 Hz
Georgia	IDT-CT	$CT_{index} \geq$ 15-150
Illinois	I-FIT	$FI \geq$ 5.0-16
Louisiana	SCB-LTRC	$J_c >$ 0.5-0.6 kJ/m^2
Maine	IDT-CT	$CT_{index} \geq$ 150
Nebraska	SCB	$FI \geq$ 6.0
New Jersey	OT BBF	Cycles to 93% Load Reduction > 100-700 >100,000 cycles at 1500 microstrain, 10 Hz
New York	OT IDT-CT I-FIT	Cycles to 93% Load Reduction > 250 $CT_{index} >$ 135 $FI \geq$ 8.0
Ohio	IDT-CT BBF	$CT_{index} \geq$ 60-80 >100,000 cycles at 1500 microstrain, 10 Hz
Oklahoma	IDT-CT I-FIT	$CT_{index} \geq$ 80 $FI \geq$ 8.0
Oregon	I-FIT	$FI \geq$ 6-8.0
Texas	OT IDT-CT	CFE $\geq$ 1 in-lb/in^2 and CPR $\leq$ 0.45 $CT_{index} \geq$ 105
Utah	IDT-CT	Not Assigned
Virginia	IDT-CT	$CT_{index} \geq$ 70
Wisconsin	IDT-CT	$CT_{index} \geq$ 40-80

Table 4. State DOTs BMD Rutting Tests Summary.

State	Test	Criteria
Alabama	HWTT IDT-HT	Passes to 10-mm rut depth ≥ 10,000-20,000 IDT-HT strength ≥ 20 psi
California	HWTT	Passes to 12.5-mm rut depth ≥ 10,000-25,000
Georgia	HWTT	Passes to 12.5-mm rut depth and to *SIP* > 15,000-20,000
Illinois	HWTT	Passes to 12.5-mm rut depth ≥ 5,000-20,000
Louisiana	HWTT	Rut depth at 20,000 cycles < 6-10 mm
Maine	HWTT	Rut depth at 20,000 cycles < 12.5 mm
Nebraska	G-Stability	G-Stability ≥ 5.55-64.17 kN
New Jersey	APA	Rut depth at 8,000 cycles < 3-7 mm
New York	APA HWTT IDT-HT	Rut depth at 8,000 cycles < 4 mm Rut depth at 20,000 cycles < 12.5 mm IDT-HT strength > 30 psi
Ohio	APA	Rut depth at 8,000 cycles < 3-5 mm
Oklahoma	HWTT	Passes to 12.5-mm rut depth ≥ 10,000-20,000
Oregon	HWTT	Rut depth at 20,000 cycles ≤ 2.5-3.0 mm
Texas	HWTT IDT-HT	Passes to 12.5-mm rut depth ≥ 10,000-20,000 IDT shear strength ≥ 1.02 MPa
Utah	HWTT	Passes to 10-mm rut depth ≥ 10,000-20,000
Virginia	APA	Rut depth at 8,000 cycles ≤ 8 mm
Wisconsin	HWTT	$CRD_{20k} \leq$ 6-8 mm[b], SN ≥ 2,000

The integration of performance tests into QA practices can represent a challenge for state DOTs due to the adjustments and resources to conduct performance tests on a routine basis. Further, guidance is needed for determining minimum sampling frequency (and risks associated), and for sampling, fabrication and testing procedures that could be unknown to DOTs personnel. These challenges motivated the development of TFRS 01: *Quality Assurance Aspects of Performance Related Specifications* and NCHRP 10-107: *Guide for Implementing Performance Specifications*. The general objective of TFRS 01 is to integrate PASSFlex^{TM} (system of tools based on fundamental engineering properties for PRS) within a QA system to ensure long-term asphalt performance and reduce life cycle costs (NCAT 2021). Specifically, TFRS 01 addresses:

(a) the use of the cyclic fatigue S_{app} and rutting *RSI* test parameters, index thresholds, and acceptance limits in support of performance engineered mix design approaches and to facilitate further implementation of the tests and performance predictions
(b) material selection and mixture design changes that can impact the test results (cyclic fatigue, SSR, and their index parameters) and trends associated with owner agency specified performance thresholds, and
(c) the major elements of a QA system per TRB E-Circular 235 and associated buyer/seller and payment risks (TRB 2018).

In addition to the TFRS project, NCHRP 10-107 researchers are developing a guide that will assist state DOTs with performance specification implementation, including integrating tests into QA systems. The guide will "specifically address, but not be limited to, implementation activities such as using pilot projects and shadow specifications, establishing appropriate control and specification limits, gaining buy-in from agency and industry personnel, and managing risk." (NCHRP 2022).

1.2.4. Test Costs

One important factor for conducting BMD tests is the total cost of performing the respective test. One key component for the total cost is the cost of the testing equipment, which is summarized in Table 5. The numbers shown here represent a sample of national territory sellers from the start of 2023. None of the displayed values covers taxes, crating, shipping, discounts that may apply, or installation. With respect to the testing equipment, there are three main categories: stand-alone versions, electromechanical loading-frame compatible tests, and AMPT compatible fixtures. This consideration should be carefully taken into account, since some tests can be performed in more than one machine setup. For example, kits exist for some tests like the SCB, I-FIT, and IDT-CT/IDEAL-RT that can leverage modern era Marshall load frames. Also, at least one AMPT manufacturer includes a range of kits for multiple tests (i.e., CF, SSR/FN, IDT-based tests, OT, etc.). The costs for these different options are all summarized in Table 5.

In addition to the direct costs of the equipment, other considerations include yearly equipment calibrations, the need for additional support equipment (saws, coring machines, compactor, etc.), and personnel training should also be considered. With respect to personnel costs, it is noted that as test complexity increases, the cost of training would increase, but that resources exist for all of the tests discussed here and that examples of successful training regimes also exist for each of the tests. A more detailed summary of the equipment and cost is provided in Appendix A.

Table 5. Summary of Testing Equipment Prices.

Test name	Machine Cost	EM Load-Frame Compatible	Additional Accessories	AMPT Compatible	Additional Accessories
I-FIT	$9,000[a]	yes	$700	yes	$1,000
SCB	$9,000[a]	yes	$700	yes	$1,000
IDT-CT	$9,000[a]	yes	$550	yes	$600
Cantabro	$8,000	no	-	no	-
CF	>$72,000[b]	no	-	yes	$10,615
BBF	>$32,200	no	-	no	-
OT	$55,000	no	-	yes	$4,000
APA	>$66,000	no	-	no	-
HWTT	>$50,000	no	-	no	-
IDT-HT	$9,000[a]	yes	$550	yes	$600
IDEAL-RT	$9,000[a]	yes	$900	yes	$900
Confined FN	>$72,000[b]	no	-	yes	$305
SSR	>$72,000[b]	no	-	yes	$305

[a] Cost for general purpose EM load-frame capable of BMD testing
[b] Cost for standard AMPT equipment without accessories

1.2.5. Knowledge Gaps

The literature review shows that considerable effort has been made to evaluate various rutting and fatigue cracking test protocols for integration into BMD. However, key aspects related to the selection, compatibility, and implementation of these tests remain unresolved. First, there is no consensus on the most appropriate rutting or cracking test to adopt—agencies must balance simplicity, cost, fundamental insight, and state-specific conditions when making this decision. Second, most studies have evaluated test protocols in isolation, focusing on either cracking or rutting. This approach overlooks the need to assess these protocols simultaneously to identify compatible test methods and opportunities for optimization, particularly when similar specimen geometries or test equipment can be used for both distress types. Third, while BMD test methods can indicate performance, many proposed protocols lack a robust mechanistic framework that links test outcomes to fundamental engineering properties—an essential requirement for integrating mixture and structural pavement design. Fourth, ambiguity remains regarding which of the AASHTO PP 105 (2024a) approaches (A–D) is most effective in improving asphalt mixture performance. While Approaches A and B retain traditional volumetric design with added performance checks, Approaches C and D provide contractors greater flexibility but require confidence in test-based performance prediction. Fifth, best practices for incorporating performance testing into quality assurance (QA) are still evolving. Ongoing efforts, including NCHRP 10-107 and TFRS01, emphasize that test simplicity and quick turnaround are essential for QA viability. Lastly, there are no universally accepted performance thresholds for BMD tests, and results must be calibrated to local materials and environmental conditions to ensure validity and effectiveness.

1.3. Operational Implications of BMD in North Carolina

As part of this research effort, contractors within North Carolina were surveyed to understand their practices with respect to quality control and the potential impact of implementing BMD testing into mix design and quality control. At the same time, the research team also conducted an interview with North Carolina DOT personnel to understand time constraints and other issues that may be relevant to BMD implementation.

1.3.1. Survey with Contractors

The contractor survey included questions related to contractor demographics, general knowledge, and opinions about BMD, mix design, and quality control operations, and finally, information regarding training and personnel. A total of six contractors submitted their responses, with their identifications not disclosed to the research team. A full discretization of the answers can be seen in Appendix A.

The contractors surveyed varied from small companies, with three plants doing 300,000 tons of asphalt mix per year, to large corporations, with fifteen plants doing 1.5 million tons per year. Some reported working only in North Carolina, while others reported also having contracts for Virginia or South Carolina. In general terms, seminars and informal talks are the main source of knowledge about BMD among contractors, while some were aware from actively working with it for Virginia contracts, two of the six contractors indicated that they were not very familiar with BMD. There is an understanding that BMD might help improve the quality of mixes, but contractors seek more freedom to design based on performance, with smaller constraints from actual volumetric design. Some of the concerns listed by contractors, if implementation occurs, focused on lab space, new equipment, training personnel, and rushed work with having both performance and volumetric samples to produce and test.

The survey also provided information regarding how contractors develop a new design or what triggers them to update the existing ones. Usually, when production trend lines are moving away from the JMF, they might consider developing or updating a new design. Other contributors to these are changes in aggregate and sand sources due to price or availability. Increase in RAP content, management of existing stockpiles, having a backup design, or the ability to run multiple designs were also cited as desirable strategies. The general idea is to keep the design, which is working for each division, and maximize the use of recycled material. Generally, each plant would have backup JMFs per mix type that would be switched in case of RAS use, or need to reduce RAP content.

In a usual year, there are about four to ten new designs per company, and it takes about a week to develop each, if no big issues happen. The target points for asphalt content is usually set based on experience, on gravities, or on past mixes, with little change. Four of the six contractors said that only one employee is responsible for doing the designs for the entire company and works in a centralized lab. The other two answered that more than one senior tech would have the capacity to work on designs inside the company, and they also reported not having only one central lab facility, but not all plants can do design testing. It takes from one to three years to reach the level of being fully trained on mix design.

For Quality Control operations, contractors reported taking about three hours to finish a full series. The number of techs each company has varies from a three to nineteen employees for production. To finalize training for Quality Control, it can vary from three to five weeks to almost a year. Some reported having a specific training program, but most of them are using hands-on practice with senior techs as the main source of information. The usual process proceeds from moisture concepts, sample fabrication, testing, and becoming familiar with NCDOT specifications. There is generally a yearly refresher training, with NCDOT helping with an organized class program. Some companies have a higher rotation of employees, while some had only one new tech in the last couple of years.

In the actual NCDOT procedure, contractors are asked to submit specimens for APA test and TSR results for mix design approval. If there are substantial changes in a specific JMF, contractors are asked to resubmit them, which is also rare to occur according to contractors. According to the survey, they have not seen a mix design failing APA limits, so only the contractor that works with Virginia contracts pre-tests the mix design for APA before submitting to North Carolina DOT. TSR testing seems to be a bigger concern to contractors when updating a JMF, usually due to a change in anti-strip suppliers.

1.3.2. Survey with North Carolina DOT

Similarly to the process done with contractors, a set of questions was developed to guide a meeting with the NCDOT, in order to familiarize the research team with their process of conducting new mix designs, updating existing JMFs, quality assurance practices, regional and central labs workflow, and personnel.

On a typical year, there are about 50 new mix designs for mix type RS9.5C, from contractors across the state. The DOT would prefer to add a revision number and keep the original year of the JMF if no big changes occur. Ideally, there would be up to three or four revisions before issuing a new mix design number. The actual practice is to update to a current year number on the JMF. The updates happen throughout the year, with a higher intensity of new mixes from fall to spring timeframe. A comment was made to some contractors having their production trends and quarries so consistent that they hardly need to change a JMF.

Several tests are required as part of a new mix design. While most of them are done by the contractor, like verifying aggregate gravities, asphalt content of recycle material, total asphalt content design and mix gravities, the DOT performs the APA rutting testing. For TSR testing, contractors are required to submit all files generated by the test equipment. To complete testing and approving a new mix design can take five to ten working days in late spring, summer, and fall, or three to five days in winter. APA testing is what drives the approval time for mix designs. The turnaround time is about three days for intermediate and base mixes, or five to ten days for surface mixes. Recently, the DOT acquired another APA testing device, which significantly improved their capacity for testing.

Updating a JMF often requires submitting test specimens when significant changes occur, such as a change in quarries, deleting an aggregate source from the blend, or changing the anti-stripping supplier. Altering the sand source or recycle content typically results in the issuance of a new number of JMF. The standard procedure is to re-verify APA rutting and non-production TSR, and production TSR after production begins, but this may vary case by case. Minor updates, like adjusting gravities, stockpile blend percentages, or small increments on asphalt content, usually only require paperwork, resulting in a new revision number. According to the agency, about 5% of revisions require resubmission of APA samples. The typical turnaround time to update an existing JMF is about two to three days.

Regarding APA testing, contractors produce the samples and ship it to the agency, while the testing is done only at the central Materials and Tests Unit facility in Division 5. The DOT corroborated the contractors' input that the mix designs rarely fail the test. Regional labs have the capacity to operate TSR testing for JMF update and verification.

Quality control is routinely performed by the contractor, while quality assurance is the agency's process of assuring quality of the mix. According to the NCDOT Asphalt Quality Management

System, there are seven approaches for conducting QA: 1) testing verification samples of 20% of QC sample frequency; 2) testing split samples obtained by contractor of 5% of QC frequency; 3) periodically observing sampling and testing procedures done by contractor; 4) directing the contractor to obtain additional samples at any time or location during production; 6) by conducting audits; or 7) by any combination of the above. According to the agency, there is no preferred approach from the seven listed above, nor one that is most used.

The testing and verification of quality assurance is done independently of the contractors' tests, by either the central Materials and Tests Unit, or one of the 10 regional labs across the state. The central laboratory takes over testing for Division 5, while regional labs support one or two divisions. A full QA series takes about three working hours, with multiple mixes being tested at once. The usual turnaround time is one to three working days, depending on the demand.

The current sampling rate of quality assurance can pose a challenge to personnel. The central lab has four employees, fully dedicated to asphalt testing with some specialized by test, while regional labs usually have two techs that can perform any type of testing, even outside of the asphalt range. There is little turnover of personnel, with mostly agency-contracted employees, but some promotions can happen. Generally, there are one or two new hires per year, taking up to six months to train them in the regional labs or four weeks in the central lab. Retraining can occur only if issues are observed. As part of the validation of the work, AASHTO accreditation is performed yearly.

When values do not meet the specification for Quality Assurance, the mix would go to a process called dispute, through which a separate sample would be tested by a third party, usually another QA lab in close proximity. According to the agency, disputes happen with a frequency of less than 3%. Some of the reasons for mixes going to dispute might be VMA, #8 sieve, #200 sieve, air voids, or asphalt content being out of assurance limits. Also, according to the agency, contractors are usually proactive regarding production trend lines, so they would perform adjustments on their own before getting to a dispute resolution. The possible outcomes are a reduction in pay, or on rare occasions, removal and replacement.

1.4. Report Organization

This report is organized into seven chapters and six appendices.

- Chapter 1 introduces Balanced Mix Design (BMD) concepts, relevant performance tests, and efforts by other states. It also summarizes the findings from surveys conducted with contractors and NCDOT regarding BMD.
- Chapter 2 details the materials and methods used in the experimental program, showing the processes used to fabricate specimens, choose mixtures, and sample the mixtures. It also lists the characteristics of the study materials. The specific procedures for the test methods utilized in this report are also included in this chapter.
- Chapter 3 presents the experimental plan and the analysis of the results. Its subsections cover benchmarking for BMD and AMPT tests, a feasibility study of the IDT-HT test, a comparison of laboratory aging protocol impacts on asphalt mixture behaviors, the protocol for short-term aging of RS9.5D mixtures, cracking sensitivity to asphalt binder sources, and the pavement performance evaluation using FlexPAVE™.

- Chapter 4 discusses the potential outcomes and implications of implementing BMD testing into mix design and quality assurance. Possible scenarios are discussed, with potential pros and cons of each being described.
- Chapter 5 lists the final conclusions and recommendations of the study. Chapter 6 outlines a plan for implementation and technology transfer. Chapter 7 contains the references cited in this report.

The appendices provide supplementary information: Appendix A contains a detailed literature review; Appendix B details the survey findings; Appendices C and D show individual mixture testing results and statistical analysis; Appendix E details the contour results from the FlexPAVE™ simulations, damage and rutting progressions; and Appendix F presents a draft specification for NCDOT regarding balanced mix design.

2. Materials and Methods

2.1. Study Overview

This study was conducted to address the need for a performance-based test to evaluate the cracking resistance of asphalt mixtures in North Carolina, a critical step toward implementing a comprehensive Balanced Mix Design (BMD) framework. The primary objectives were to identify a suitable test protocol, benchmark the performance of current mixtures, and establish preliminary performance thresholds to ensure long-term pavement durability. Figure 2 presents a flowchart of the report's structure. The gray boxes outline the selection of materials and methods; the orange boxes detail the main benchmarking experimental program; the blue boxes describe side studies; and the green boxes highlight the primary outcomes of this report.

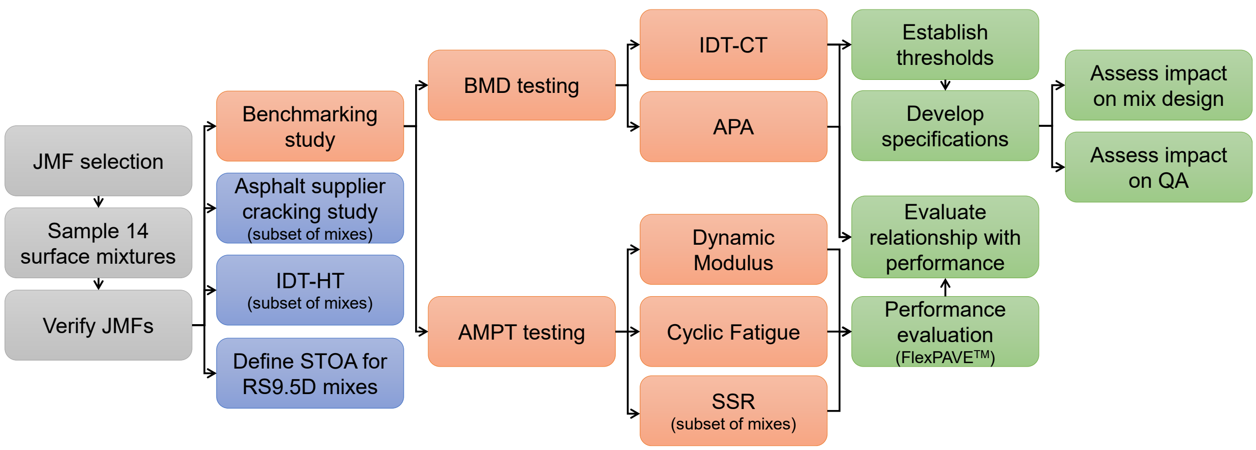


Figure 2. Flowchart of the research.

2.2. Material Selection and Sampling

This research study involved testing 14 different surface mixtures (RS9.5B, RS9.5C, and RS9.5D traffic level denominations) distributed across the state. To select these mixtures, the research team first worked with NCDOT personnel to identify all active JMFs developed between 2019 and 2022 (515 total JMFs). The limit of considering only JMFs issued since 2019 was selected due to an update on JMFs in 2018, which changed some mix types definitions. The list was then categorized by mix type (RS9.5B, RS9.5C, and RS9.5D), county, division, region (Mountains, Piedmont, and Coastal Plains), and RAP content. The objective was to understand which counties had the most active JMFs per mix type, and their representative RAP contents.

Additional filters were then added to eliminate mixtures having VMA or VFA values close to the specification limits, recycled binder replacement below 30%, and to eliminate JMFs using RAS materials. Finally, a list of five RS9.5B, six RS9.5C, and three RS9.5D mixes was identified based on spatial distribution and variation in material source to sample as shown in Figure 3. Asphalt binder was sampled from the appropriate terminal source, and for the same asphalt binder grade currently in production, according to plant operators. For the purposes of this report, plants were anonymized as shown in Table 6 below.

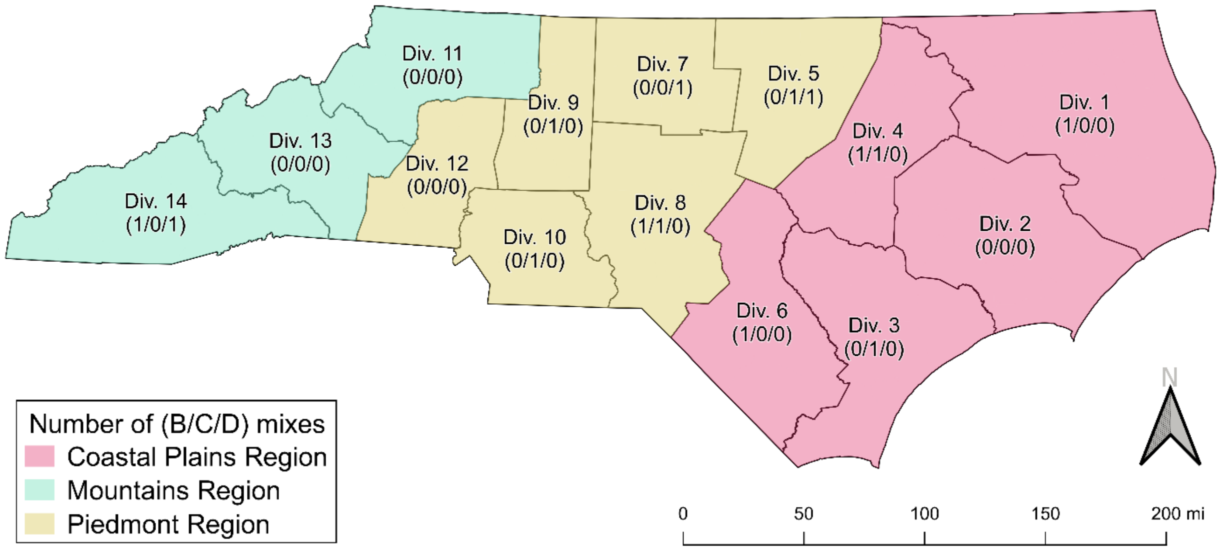


Figure 3. Geographical distribution of sampled component materials.

Plants F (RS9.5B) and H (RS9.5C) were verified as part of the RP2021-06, and no change in asphalt content was needed to achieve the required volumetrics. For the other plants, a number of mixture verification steps were performed. First, the RAP asphalt content was verified using the ignition oven (after applying the calibration coefficient suggested in the NCDOT QMS manual). Then the gradations of the individual stockpiles (including RAP) were measured and used to compute the blended gradation. Next, two gyratory mix-design samples were fabricated and compacted at each of the JMF asphalt content and at asphalt contents plus and minus 0.5% from the design content. The bulk specific gravity of each sample was measured in accordance with AASHTO T 166 (2024d). At least two G_{mm} samples were mixed at the binder content specified in the JMF, in order to calculate the volumetric properties. These G_{mm} values were used to compute the design specimen air void contents. Finally, aggregates specific gravities were updated to reflect the mixture's measured maximum specific gravity.

The mix was considered verified if the gradation blend was close to the one specified, all limits of VMA, VFA, and dust proportion were met, and the air void content was within ±1.0% of the air void content requirement. If the design asphalt content did not meet the volumetric requirements at the design gyration number (50 for RS9.5B, 65 for RS9.5C, 100 for RS9.5D), an adjustment to the asphalt content in accordance with the measured results to yield between 3 and 5% air and pass the other volumetric requirements was made. Ultimately, the mix designs for Plants F (RS9.5B), W (RS9.5B), U (RS9.5B and RS9.5C mixes), H (RS9.5C), R (RS9.5C), and V (RS9.5D only) were verified without change. However, adjustments to the asphalt content for Plant Z (RS9.5C), Plant Y (RS9.5B), Plant V (RS9.5C), Plant T (RS9.5D), Plant S (RS9.5B and RS9.5D), and Plant X (RS9.5C) were necessary to achieve the requirements described in the NCDOT QMS manual. These revised asphalt contents are summarized in Table 6. An adjustment to Plant V (RS9.5D) stockpile proportions was made to meet the JMF combined gradation, and as a result, there was no change in asphalt content.

Table 6. Asphalt Content Decision.

Plant	Mix Type	Location [a]	Original AC (%)	New AC (%)	Gradation Change
F	RS9.5B	Div. 4 | CO	6.30	same	same
Y	RS9.5B	Div. 6 | CO	6.80	6.87	same
W	RS9.5B	Div. 1 | CO	6.30	same	same
S	RS9.5B	Div. 14 | MO	6.50	6.77	same
U	RS9.5B	Div. 8 | PI	6.20	same	same
H	RS9.5C	Div. 10 | PI	6.00	same	same
Z	RS9.5C	Div. 3 | CO	6.00	6.20	same
V	RS9.5C	Div. 5 | PI	5.60	5.30	same
X	RS9.5C	Div. 4 | CO	5.60	5.90	same
U	RS9.5C	Div. 8 | PI	5.60	same	same
R	RS9.5C	Div. 9 | PI	6.10	same	same
V	RS9.5D	Div. 5 | PI	5.80	same	proportions
T	RS9.5D	Div. 7. | PI	5.60	5.24	same
S	RS9.5D	Div. 14 | MO	5.80	6.00	same

[a] Regions: Coastal Plains (CO), Piedmont (PI), and Mountains (MO)

The naming convention for each mix consists of a letter indicating the plant source, a letter indicating NCDOT traffic designation (B, C, or D), and the short-term aging used in the production of the samples. For example, "ZC (2hCT)" refers to the plant Z high traffic (RS9.5C) mixture that was conditioned for two hours at the compaction temperature as the short-term aging procedure. Information regarding volumetric characteristics of the mixtures can be found in Table 7, in terms of: virgin binder PG grade, effective asphalt content (P_{be}), percentage of coarse aggregate (CA), fine aggregate (FA), natural sand (NS), recycled material (RAP), recycled binder replacement (RBR), gyrations at design condition (N_{design}), voids in mineral aggregate (VMA), voids filled with asphalt (VFA), maximum specific gravity (G_{mm}), and dust proportion (DP).

Table 7. Mixtures Information.

Mixture	Binder Grade (Virgin)	P_{be} [a] (%)	CA/FA/NS (%)	RAP (%)	RBR [a]	N_{design}	VMA [b]	VFA [b]	G_{mm} [a]	DP [a]
FB	58-22	6.16	15/11/34	40	0.32	50	18.0	77.8	2.427	0.83
YB	64-22	6.79	22/23/25	30	0.22	50	18.9	78.8	2.404	1.02
WB	64-22	6.18	20/24/26	30	0.23	50	17.4	77.0	2.422	1.25
SB	64-22	6.72	15/46/9	30	0.23	50	18.5	77.8	2.425	0.99
UB	64-22	5.71	23/27/20	30	0.26	50	17.2	76.7	2.498	0.98
HC	64-22	5.41	22/33/10	35	0.28	65	16.7	76.0	2.524	0.84
ZC	64-22	6.09	22/18/30	30	0.25	65	17.3	76.9	2.437	1.02
VC	64-22	5.20	24/46/-	30	0.26	65	16.5	76.6	2.440	1.28
XC	64-22	5.81	25/45/-	30	0.26	65	16.6	75.7	2.447	0.94
UC	64-22	5.28	32/22/16	30	0.29	65	16.3	75.5	2.536	1.04
RC	64-22	6.07	30/30/10	30	0.29	65	17.2	76.7	2.455	0.94
VD	76-22	5.56	22/59/-	19	0.15	100	16.5	75.8	2.436	1.17
TD	76-22	5.09	30/50/-	20	0.17	100	17.1	76.6	2.618	1.25
SD	76-22	5.67	31/41/13	15	0.13	100	16.5	75.8	2.450	0.84

[a] Calculated by the NCSU laboratory, [b] Specified in the job mix formula

2.3. Test Methods

Based on the findings from the literature review and current NCDOT practice, the Indirect Tension Asphalt Cracking Test (IDT-CT) and the Asphalt Pavement Analyzer (APA) were selected as the main durability and stability tests for this study, respectively. This study also evaluated the possibility of benchmarking Indirect Diametral Tension at High Temperature (IDT-HT) test results on selected mixtures. The IDT-HT test was chosen for this purpose because it has been adopted by other agencies (e.g., VDOT) and has been found to be the most sensitive to compositional changes in the asphalt mixture compared to the APA, IDEAL-RT, and Marshall Stability test (Boz *et al*., 2025). In addition to the aforementioned BMD tests, AMPT Dynamic Modulus (DM), Cyclic Fatigue (CF), and Stress Sweep Rutting (SSR) tests were conducted to compare more fundamental mechanical results to the index test results. Table 8 presents a summary of the testing done at each mixture of this study.

Table 8. Summary of Performance Tests.

Plant	Mix Type	BMD Testing	AMPT Testing
F	RS9.5B	IDT-CT, APA	DM, CF
Y	RS9.5B	IDT-CT, APA, IDT-HT	DM, CF, SSR
W	RS9.5B	IDT-CT, APA	DM, CF, SSR
S	RS9.5B	IDT-CT, APA	DM, CF
U	RS9.5B	IDT-CT, APA	DM, CF
H	RS9.5C	IDT-CT, APA	DM, CF
Z	RS9.5C	IDT-CT, APA, IDT-HT	DM, CF, SSR
V	RS9.5C	IDT-CT, APA, IDT-HT	DM, CF, SSR
X	RS9.5C	IDT-CT, APA	DM, CF
U	RS9.5C	IDT-CT, APA, IDT-HT	DM, CF, SSR
R	RS9.5C	IDT-CT, APA	DM, CF
V	RS9.5D	IDT-CT, APA, IDT-HT	DM, CF, SSR
T	RS9.5D	IDT-CT, APA	DM, CF
S	RS9.5D	IDT-CT, APA	DM, CF

2.3.1. Specimen Fabrication

For the mixtures tested in this study, the fabrication process followed the normal standard method used in the NCSU asphalt mixture laboratory. First, binder was collected in 5-gallon buckets. The sealed buckets were heated in an oven at 70°C for 8 hours so that they would be thoroughly heated and initially liquid. The lids were then opened, and the oven temperature was increased to approximately 125°C for an additional two hours. Next, the binder was poured into 1-gallon cans. An anti-stripping agent was added in percentages specified by the job mix formula and thoroughly combined using a powered stirrer. Finally, the treated binder was apportioned into smaller cans, each containing enough material for a single specimen plus a small safety margin.

Prior to sieving, virgin aggregates were placed in large pans and dried in an oven at 110°C until a constant mass was achieved. The dried aggregates were then sieved and separated into individual size fractions on pre-determined sieves. The pre-determined sieves, referred to as cut sieves, were established to balance batching efficiency and quality control. For the Reclaimed Asphalt Pavement (RAP), the material was spread in a separate pan to a thickness not exceeding 2 inches. It was then dried at a lower temperature of 60°C and stirred frequently until reaching a constant mass. Mechanical splitters were used to obtain RAP samples for batching.

Each specimen was mixed and compacted individually to ensure consistency. The individual virgin aggregate fractions were batched to their target weights and heated overnight in an oven set 10°C above the required mixing temperature. The pre-weighed RAP was placed in a separate pan at room temperature. Two hours before mixing, the binder cans were placed in an oven set 3°C above the mixing temperature. All mixing equipment (metal pans, buckets) and compaction equipment were preheated to their respective mixing and compaction temperatures in separate ovens.

Forty minutes prior to mixing, the RAP was added to the pan containing the hot virgin aggregates, stirred thoroughly, and returned to the oven. To begin the mixing process, the hot aggregate-RAP blend was transferred to a mixing bucket. The bucket was placed on a scale, and the precise amount of preheated binder was poured over the aggregates. The bucket was then secured in a mechanical mixer and mixed for approximately 1-2 minutes, until all aggregate particles were uniformly coated. Immediately after mixing, the loose-mix asphalt was spread into a pan, batched to the final specimen mass, and placed in an oven for short-term aging as specified by the aging protocol.

Once the aging period was complete, the material was moved into a transfer device preheated to the compaction temperature and then carefully poured into a preheated mold. To ensure a uniform distribution of material, a spatula was inserted approximately eight times around the circumference of the mold, with its flat face held against the mold wall. It was then inserted into the center of the mix 10-12 times in a cross-hatch (#) pattern. For larger specimens (180-mm height), the material was added in two layers, and the spatula process was performed after each lift was placed. Finally, the prepared mold was placed in a gyratory compactor and compacted to the target specimen height. After compaction, the finished specimen was extruded from the mold and allowed to cool to room temperature before handling.

2.3.2. Indirect Tension Cracking Test (IDT-CT)

The IDT-CT test procedure followed in this research was consistent with the ASTM D8225 (2019) standard. Specimens were conditioned in an air chamber at 25°C prior to testing. Five specimens were tested for each mix, with the highest and lowest results trimmed from the analysis. CT_{index} was then calculated for each specimen separately using Equation (1) and then averaged. The testing was repeated if the observed range in test results was not below the allowable limits calculated from the Virginia Transportation Research Council Round Robin study (Habbouche *et al*. 2021a, Boz *et al*. 2022). Test specimens were 150 mm in diameter and 62 mm in height and were fabricated with 7 ± 0.5 percent air voids. Before compacting, specimens were subjected to short-term oven aging by placing the mixture in an oven for 4 hours at the mixture compaction temperature (143°C for RS9.5B and RS9.5C mixes). An exception was made for RS9.5D mixes, where the mix was first aged for four hours at 143°C, but was then heated for an additional 50 minutes to reach the compaction temperature of 163°C. The use of 4 hours at compaction temperature for the short-term aging protocol was consistent with practices in Virginia and was decided on in consultation with the project steering committee. It was also closest to the AASHTO R 30-19 protocol for mix design that was in place at the outset of the project (AASHTO 2021). The extended time, from 2 hours in AASHTO R 30-19 to 4 hours, had been found in Virginia to aid in managing workflow where limited oven space existed, and specimen fabrication for volumetric analysis, APA testing, Cantabro testing, and IDT-CT testing was performed.

$$CT_{index} = \frac{t}{62} \times \frac{l_{75}}{D} \times \frac{G_f}{|m_{75}|} \times 10^6 \tag{1}$$

Where t is the specimen thickness (mm), D is the specimen diameter (mm), l_{75} is the displacement at 75% of the peak load after the peak (mm), G_f is the failure energy (J/m^2), and m_{75} is the slope at 75% peak load after the peak (N/m).

2.3.3. Asphalt Pavement Analyzer (APA)

The APA test procedure followed the AASHTO T 340 (2023e) standard and NCDOT requirements. Two sets of specimens were tested at 64°C, and the results were reported as the average rut depth after 8,000-wheel repetitions. Test specimens were 150 mm in diameter and 75 mm in height and fabricated with 4 ± 0.5 percent air voids. The short-term aging protocol followed the NCDOT practice of two hours at compaction temperature before compacting for all mix types.

2.3.4. Indirect Tension-High Temperature Test (IDT-HT)

The IDT-HT test procedure follows an adaptation of AASHTO T 283 (2022a). Specimens were conditioned in a chamber at 51°C to match the SSR-test high temperature. Four specimens were tested for selected mixes, and the results are described in terms of the peak stress strength reached during the test. The test geometry was the same as the IDT-CT geometry (150 mm diameter and 62 mm height at 7 ± 0.5 percent air voids). Finally, the short-term aging protocol was the same as the other rutting testing, so prior to compaction, mixtures stayed at two hours at compaction temperature.

2.3.5. AMPT Testing

Large gyratory samples of 180 mm height and 150 mm diameter were compacted for fabricating AMPT tests. For DM and CF tests, four 110-mm height and 38-mm diameter samples (5 ± 0.5% air voids) were cored and cut from the larger gyratory sample. Like for IDT-CT testing, RS9.5B and RS9.5C mixtures were short-term aged in an oven for four hours at 143°C prior to compaction, while RS9.5D mixtures were short-term aged at 143°C for four hours and then heated to 163°C for 50 minutes prior to compaction. For the SSR test, each large gyratory sample resulted in a single cored and cut specimen of 150 mm height and 100 mm diameter (7 ± 0.5% air voids). For the SSR testing, the short-term aging protocol followed the previous rutting performance test samples (two hours at 143°C prior to compacting RS9.5B and RS9.5C, and two hours at 163°C for RS9.5D).

Dynamic Modulus (DM) Test

The DM test procedure followed the AASHTO TP 132 (2023g) standard for each combination of frequency (0.1, 1, 10 Hz) and temperature (4, 20, 40°C). The acceptability of the test results was confirmed by checking the data quality indicators and d2S limits identified by Underwood *et al.* (2023). FlexMAT™ version 2.1.4.6 was used to analyze the data and fit to 2S2P1D model.

Cyclic Fatigue (CF) Test

CF test was performed in accordance with AASHTO T 411 (2024c) standard. The testing temperature was chosen based on the binder grade (18°C for RS9.5B and RS9.5C, and 21°C for RS9.5D). Both DM-prior tested and non-tested specimens were used to characterize each material, as long as three replicates per mixture met the required repeatability limits. Afterwards, S_{app} index values were calculated using FlexMAT™ version 2.1.4.6 by setting the location as Raleigh, NC.

Stress-Sweep Rutting (SSR) Test

The SSR test procedure followed the AASHTO TP 134 (2023h) standard. Two replicates per temperature per mix were tested. The difference between the two replicates at the end of the tests was less than 25% for all the results reported. Based on the analysis conducted in in Underwood *et al*. (2021) the high- and low-test temperatures were selected as 51°C and 29°C, respectively. FlexMAT^{TM} version 2.1.4.1 was used to calculate RSI index, using Raleigh, NC, as the location for climatic conditions.

3. EXPERIMENTAL RESULTS

3.1. Overview

The experimental program, as described in the previous chapter, was carried out using the mixtures described in Section 2.2. The results of these experiments are presented in the following sections and in the following order: benchmarking evaluation of BMD and AMPT tests, IDT-HT feasibility study, short-term aging study with RS9.5D mixtures, binder source cracking sensitivity, and FlexPAVE™ performance evaluation.

3.2. Benchmarking

This section presents the results of the comprehensive laboratory testing program conducted on the fourteen laboratory mixed compacted asphalt mixtures evaluated in this study. The experimental plan was designed to assess mixture performance with respect to cracking resistance, rutting, and modulus. The following subsections detail the results from the Indirect Tension Cracking Test (IDT-CT), Asphalt Pavement Analyzer (APA), Dynamic Modulus (DM), Cyclic Fatigue (CF), and Stress Sweep Rutting (SSR) tests.

3.2.1. BMD Testing

Indirect Tensile Asphalt Cracking Test

The results from the IDT-CT test for all fourteen studied mixtures are summarized in Figure 4. The bars show the CT_{index} from three replicates (trimmed from five results), with the error bars indicating the range between maximum and minimum values. As seen in this figure, both the RS9.5B and RS9.5C groups show a wide range of results, from a high of 25.8 to a low of 9.3. The WB mixture showed the lowest value, while the RS9.5D group presented consistent results but at a lower scale than the other mixtures.

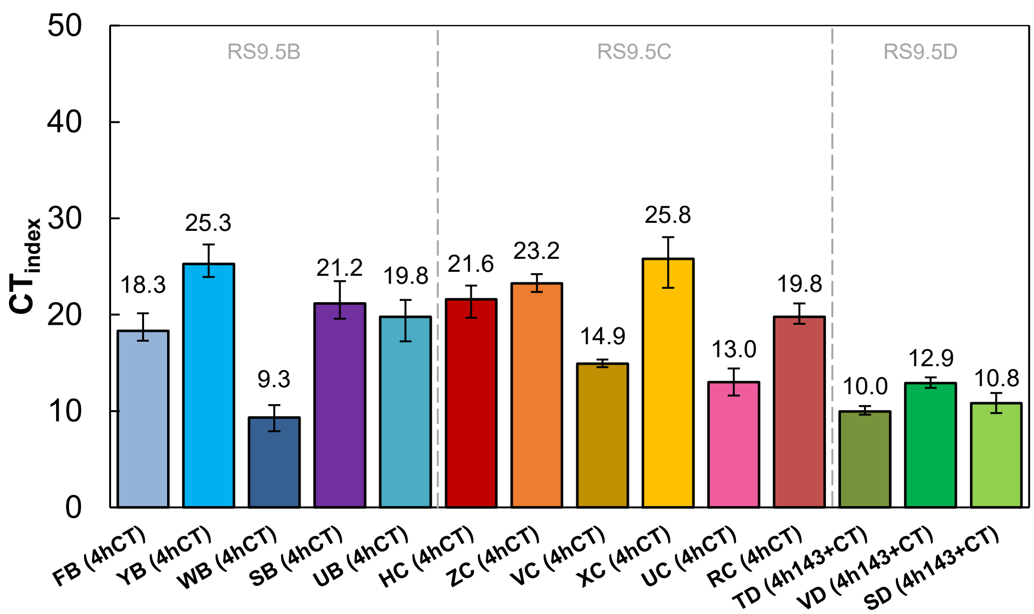


Figure 4. IDT-CT results for benchmarking study.

A non-parametric box plot was used to identify potential outliers within the dataset. Based on the median and interquartile range (IQR), the outlier thresholds were defined as any value falling outside the calculated boundaries. The upper boundary is defined as the third quartile plus 1.5×IQR, and the lower boundary is defined as the first quartile minus 1.5×IQR. For the analysis focusing on the RS9.5B and RS9.5C mixtures, no data points were flagged as outliers using this method.

Based on total and effective asphalt content (Table 6 and Table 7), it was expected that RS9.5B mixtures would achieve slightly better performance than RS9.5C mixtures, which was not observed. In addition to that, the lower performance of the polymer-modified RS9.5D mixtures is likely not due to their binder content either, since these values were similar in range to the RS9.5C mixtures.

The following paragraphs present a statistical discussion of the IDT-CT results; detailed statistical outputs can be found in Appendix C. As part of the standard procedure, a preliminary test for equality of variances was conducted prior to each one-way ANOVA. If the equal variance assumption was not met, a Welch's ANOVA was performed for the statistical comparison. The one-way ANOVA test (α=0.05) confirmed significant statistical differences among the individual mixtures (p-value < 0.001), and a subsequent Tukey HSD test identified seven statistical groups. The top-performing group included mixtures XC, YB, ZC, and HC, while the lowest-performing group consisted of UC, WB, and all RS9.5D mixtures.

To further investigate the IDT-CT results, mix type was selected as a treatment factor in a separate ANOVA, which verified significant differences among the three types (p-value < 0.001). The post-hoc Tukey HSD test identified only two statistical groups: one containing RS9.5B and RS9.5C mixtures (average CT_{index} of 18.5), and another containing RS9.5D mixtures (average CT_{index} of 11.2). Note that these results are different from FHWA/NC 2023-03 project, which did identify statistical differences between RS9.5B and RS9.5C mixtures. The differences could be due to the fact that the FHWA/NC 2023-03 project leveraged plant mix specimens, which had an overall higher value than the lab mixed samples tested here. Also, the FHWA/NC 2023-03 dataset has a mix of high and low RBR mixes, including some where the RBR was high enough to warrant a binder grade change in the JMF. Finally, the dataset gathered in this project and the one gathered in FHWA/NC 2023-03 are relatively small (5 RS9.5B mixes and 6 RS9.5C mixes in the current project and 4 RS9.5B mixes and 3 RS9.5C mixes in FHWA/NC 2023-03), and these differences could simply be related to random sampling issues. This result validated the visual observation that the RS9.5B and RS9.5C groups were statistically similar. This outcome was double-checked by excluding the WB mixture and RS9.5D mixtures and re-running the test, which again found no statistical differences between RS9.5B and RS9.5C mixtures (p-value = 0.382). Additionally, a model including both mix type and source region as factors found no statistical differences among source regions for the fourteen mixtures in this study (p-value = 0.196).

A comparison of these CT_{index} results and findings from other states confirms that IDT-CT performance is regionally dependent (see Table 3 from Section 1.2.2). There are a few possible explanations for why the values in the North Carolina dataset are lower than other states. First, the aging protocol adopted is somewhat more severe than most other states. Aging at four hours at the compaction temperature results in properties that appear to be more extreme than plant production. FHWA/NC 2023-03, which has been performed in parallel to the current effort, but evaluated plant mix materials has found a consistently higher CT_{index} value from plant-mix materials. In fact, mixtures FB and HC were included in both studies, and in this case, the CT_{index} values from plant mixed materials were approximately 60.1 and 40.8, respectively. These values are higher than those obtained from the lab-mixed, lab-compacted values reported in Figure 4 by a factor of 3.2 and 1.9, respectively. The lab aging protocol is not the only reason for the low values. Despite this study adopting Virginia's short-term aging protocols, the results were unexpectedly different. Virginia has also adopted the four hour on temperature protocol and has established a threshold value of 70 (Bowers and Diefenderfer 2018). In the Virginia work, this value corresponded to approximately the 15th percentile of the benchmark mixtures they tested. Using a similar 15th percentile would yield a preliminary threshold of approximately 14 for North Carolina.

Another potential explanation for the low values is that the selected RS9.5B and RS9.5C mixtures had high RBR values, ranging from 0.23 to 0.32. These high RBRs are common in North Carolina, but would represent relatively high values in most other states. For context, 73 percent of states specify a maximum allowable RAP content in surface mixtures between 0.20 and 0.30, placing the materials used in this study at the upper end of national practice (Castorena and Costa, 2024). At these RBR values, the NCDOT allows contractors to use virgin PG 64-22 asphalt, but many other states require contractors to reduce the binder grade. Furthermore, the FHWA/NC 2023-03 study found that, on average, plant-mixed, lab-compacted mixtures with a PG 58-28 and RBR above 0.30 had higher CT_{index} values than mixes using PG 64-22 at RBR of 0.30 or lower. This difference was found to be statistically significant. FHWA/NC 2024-14, currently underway as of this writing, has reported data on lab-mixed, lab-compacted specimens that shows when a PG 58-28 binder is used in a given mixture that the CT_{index} values are higher than when a PG 64-22 binder is used in the same mixture.

Asphalt Pavement Analyzer

The APA test was also conducted on the fourteen mixtures of this study. Figure 5 shows the average result of the left and right wheel path rut depth from the APA test obtained at cycle number 8000. The dashed lines indicate the NCDOT thresholds for the mix types. The error bars indicate maximum and minimum values. A lower rut depth means a more rutting-resistant mixture. All mixtures successfully passed their respective thresholds, and most would also succeed even under the strictest limit of 4.5 mm (for RS9.5D). Different from IDT-CT results, there is a higher consistency among the results within each mix type group.

The same non-parametric box plot procedure was used to identify potential outliers within the APA rut depth dataset. For the analysis focusing on the RS9.5B and RS9.5C mixtures, YB mixture was verified to be an outlier to the dataset. Visually, RS9.5C mixtures perform slightly better than RS9.5B mixtures; however, if YB is treated as an outlier, both groups perform similarly, with rut depths between 2.3 mm and 4.1 mm. The RS9.5D mixtures seem to behave differently, showing superior performance.

A clear trend exists between total and effective asphalt content (Table 6 and Table 7) and rut depth. The YB mixture, with the highest asphalt content, performed the worst. For the RS9.5C mixtures, there is a nearly linear correlation where higher asphalt contents correspond to greater rut depths. This trend is not that clear for RS9.5D mixtures, possibly because the polymer-modified binder used in them enhances rut resistance, potentially masking the effects of varying binder content.

The following paragraphs present a statistical discussion of the APA rut depth results; detailed statistical outputs can be found in Appendix C. As part of the standard procedure, a preliminary test for equality of variances was conducted prior to each one-way ANOVA. If the equal variance assumption was not met, a Welch’s ANOVA was performed for the statistical comparison.

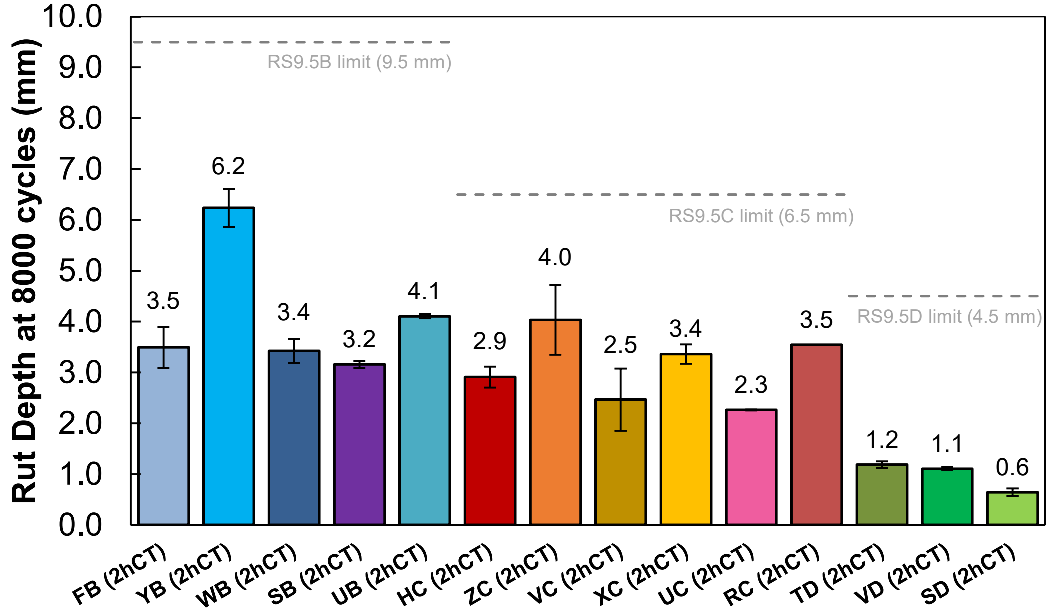


Figure 5. APA results for benchmarking study.

A one-way ANOVA test (α=0.05) confirmed statistically significant differences among the individual mixtures (*p*-value < 0.001). A subsequent Tukey HSD test identified seven distinct statistical groups. The highest rutting group contained only YB mixture, while the top-performing group (with the lowest rut depth) consisted of UC and all three RS9.5D mixtures. As expected, an ANOVA using mix type as a treatment factor confirmed statistical differences among them (*p*-value < 0.001). The post-hoc Tukey HSD test revealed that all three mix types were statistically different from each other, with average values of 4.1 mm, 3.1 mm, and 1.0 mm to RS9.5B, RS9.5C, and RS9.5D mixtures, respectively. To validate this, a follow-up ANOVA excluding outliers (the YB and RS9.5D mixtures) was performed, which then revealed no statistical differences between RS9.5B and RS9.5C groups (*p*-value = 0.153). Additionally, a final model including mix type and source region found no statistical differences among source regions for the fourteen mixtures in this study (*p*-value = 0.163).

3.2.2. AMPT Testing

<u>Dynamic Modulus</u>

The DM test was conducted on the fourteen mixtures at three temperatures (4, 20, 40°C) and three loading frequencies (0.1, 1, 10 Hz). Figure 6 presents the resulting dynamic modulus ($|E^*|$) master curves in both log-log and semi-log plots. In the figure, the lines represent the 2S2P1D model fit, while the markers indicate the experimental data points.

A visual analysis of the dynamic modulus master curves reveals distinct characteristics for each mix type. Within the RS9.5B group, which shows a wide range of results, the WB mixture exhibited a higher modulus than the other four mixtures (FB, YB, SB, and UB). For the RS9.5C group, the ZC, HC, VC, and RC mixtures form a main cluster, while UC was distinct with a higher modulus. In the RS9.5D group, SD and VD mixtures behaved nearly identically, while TD mixture had a higher modulus. A comparison between the groups suggests that while the RS9.5B and RS9.5C types were similar, the RS9.5D mixtures appeared to be stiffer overall.

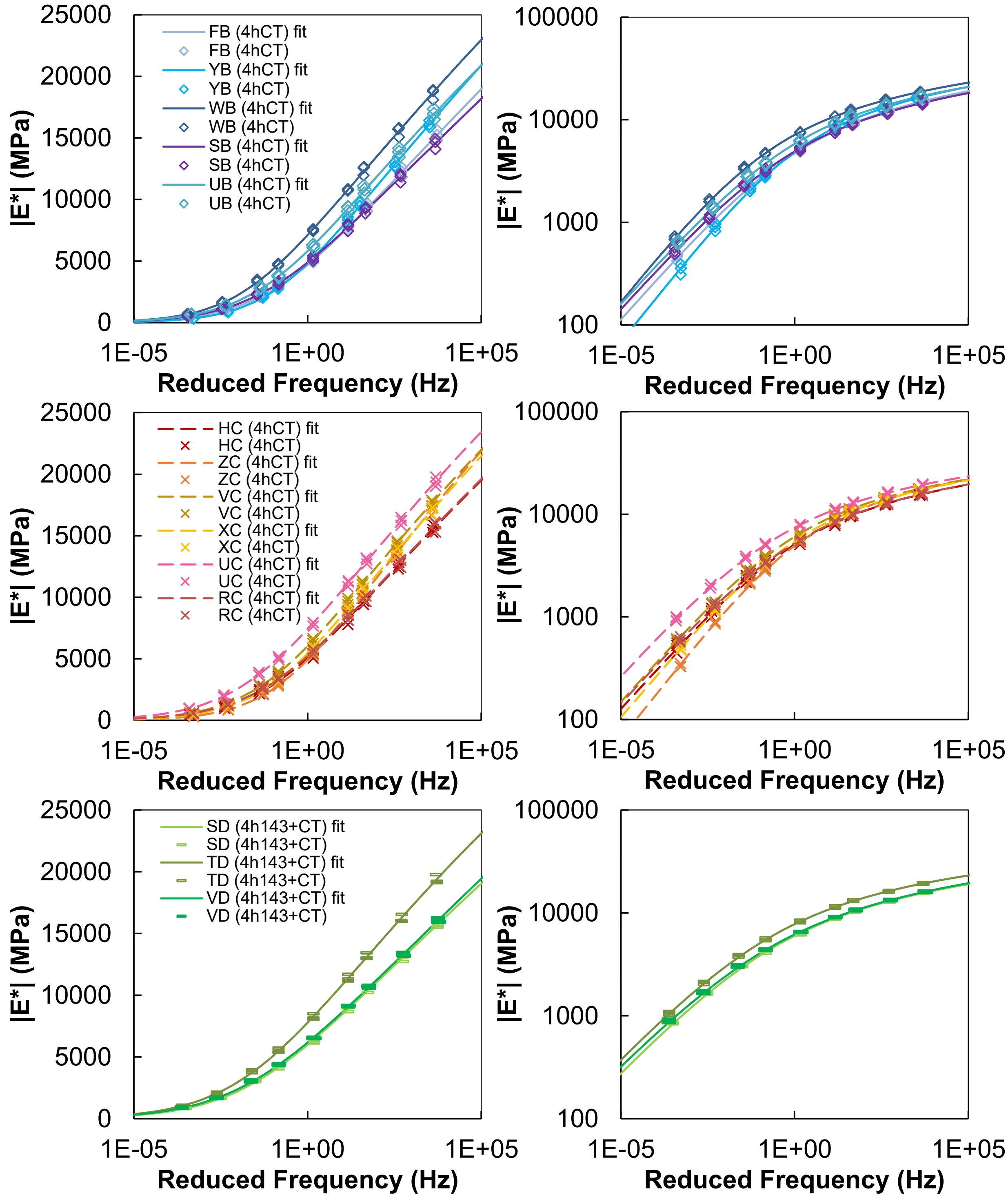


Figure 6. DM results for benchmarking study.

The following paragraphs present a statistical discussion of the DM results; detailed statistical outputs can be found in Appendix D. As part of the standard procedure, a preliminary test for equality of variances was conducted prior to each one-way ANOVA. If the equal variance assumption was not met, a Welch's ANOVA was performed for the statistical comparison.

A one-way ANOVA ($\alpha = 0.05$) was conducted on the dynamic modulus data at 1 Hz for each test temperature (4°C, 20°C, and 40°C) to identify differences among the mixtures. The analysis found statistically significant differences at all three temperatures (p-value < 0.001 for all). A subsequent Tukey HSD post-hoc test revealed several key groupings. The TD and UC consistently exhibited a statistically higher modulus than the other mixtures at all temperatures. The WB mixture was statistically similar to this top group at 4°C but performed in a lower statistical group at 20°C and 40°C. At the lower end of the performance spectrum, the group with the lowest modulus contained the HC, FB, and SB mixtures at both 4°C and 20°C. At 40°C, the lowest modulus group consisted of the FB, YB, and ZC mixtures.

An ANOVA was conducted with mix type as the treatment factor to identify differences in dynamic modulus at each temperature. The analysis found no statistical difference among the mix types at 4°C (p-value = 0.216). However, at 20°C and 40°C, significant differences were observed (p-value = 0.014 and p-value < 0.001, respectively). A subsequent Tukey HSD test for these temperatures revealed that the RS9.5D mixtures exhibited a statistically higher modulus than the RS9.5B and RS9.5C mixtures. This evaluation confirmed that the RS9.5B and RS9.5C mixtures have a statistically similar modulus across all tested temperatures.

A final model including both mix type and region from within North Carolina, where the material was sourced as factors, found that the regional source had a statistically significant effect on the dynamic modulus at all three temperatures: 4°C (p-value = 0.018), 20°C (p-value = 0.032), and 40°C (p-value = 0.007). The post-hoc Tukey HSD analysis provided more detail on these regional differences. At both 4°C and 20°C, materials from the piedmont region were statistically different from those in the mountains region. At 40°C, materials from the piedmont region were statistically different from those in the coastal plains region. This regional inconsistency is in agreement with findings from RP2019-20 (Underwood *et al*., 2021) for North Carolina surface mixtures, in which piedmont region materials were statistically different from the others at 4°C and 20°C, while coastal plains materials were statistically different from the others at 40°C.

Cyclic Fatigue

Cyclic Fatigue test was performed on the fourteen mixtures, with the resulting S_{app} and D^R indexes shown in Figure 7. The numbers express the representative values, while the error bars represent maximum and minimum values. A higher index indicates a better fatigue performance. Similar to the IDT-CT results, the RS9.5B group presented varied performance, ranging from 18.3 to 25.0 for S_{app}, and 0.431 to 0.578 for D^R index. In contrast, the RS9.5D group showed consistently high values for both S_{app} and D^R. The top performing group of mixtures for this test was FB, HC and RC, which differed from CT_{index} results, where YB, XC, and ZC mixtures achieved higher performance. On the other hand, the WB mixture was consistently the worse-performing mixture across all cracking indexes (CT_{index}, S_{app}, and D^R).

The following paragraphs present a statistical discussion of the CF results; detailed statistical outputs can be found in Appendix D. As part of the standard procedure, a preliminary test for equality of variances was conducted prior to each one-way ANOVA. If the equal variance assumption was not met, a Welch's ANOVA was performed for the statistical comparison.

A one-way ANOVA test (α=0.05) confirmed statistically significant differences among the individual mixtures for both S_{app} and D^R indexes (p-value < 0.001 for both). A subsequent Tukey HSD test identified five distinct statistical groups in each case, with significant overlap between the intermediate levels. For example, HC, VD, and UB mixtures were contained within all five

Tukey groups for the S_{app} index, indicating their performance was not statistically distinguishable from a wide range of other mixtures. When the analysis was grouped by mix type, the ANOVA found no significant differences either S_{app} (p-value = 0.092) or D^R (p-value = 0.184).

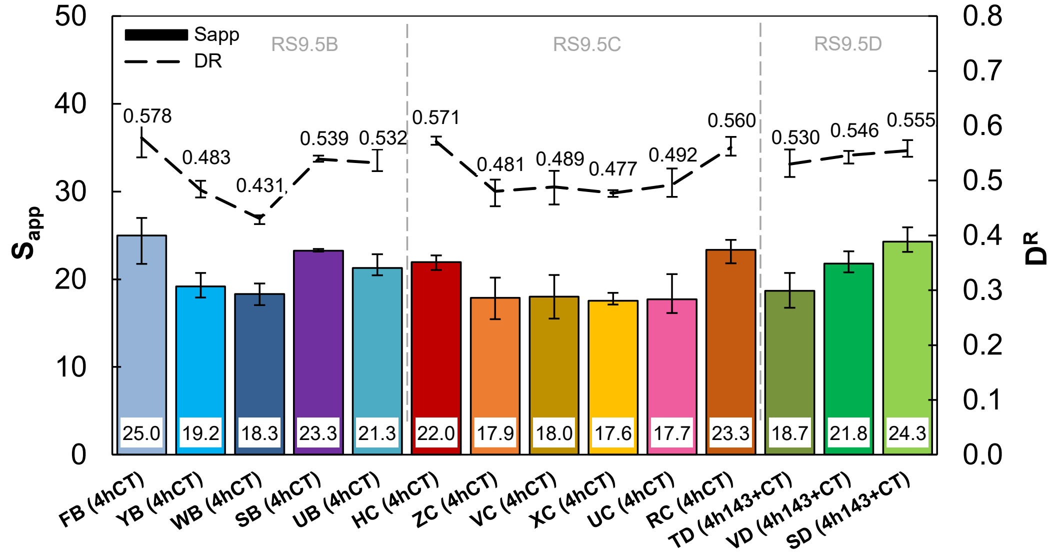


Figure 7. CF results for benchmarking study.

A final model, including both mix type and source region, found that source region had a statistically significant effect on both S_{app} (p-value = 0.048) and D^R (p-value = 0.018). The post-hoc Tukey HSD analysis revealed that for S_{app} index, the mountains region differed from those in the coastal plains region. For D^R index, materials from the piedmont region differed from those in coastal plains region. RP2019-20 (Underwood *et al*., 2021) findings also suggest that there are regional differences when evaluating CF test results for surface mixtures in North Carolina, in which all regions were different from each other.

<u>Stress Sweep Rutting</u>

As shown in Table 8, the SSR test was conducted on a selected group of six mixtures to provide a more mechanistic and fundamental evaluation of rutting resistance. The results, presented as the RSI index in Figure 8, indicate that a higher index corresponds to a greater susceptible mixture to rutting. Consistent with the APA results, the YB mixture exhibited a distinctly poor performance.

The performance ranking of the mixtures based on the RSI was identical to the ranking from the APA test, showing agreement between the two tests. When YB mixture was treated as an outlier, no clear difference was observed between RS9.5B and RS9.5C mix types; however, the VD mixture (the only RS9.5D mix tested) clearly performed better than the others. Since the test yielded only one index value per mixture, a statistical comparison among the mixtures was not performed.

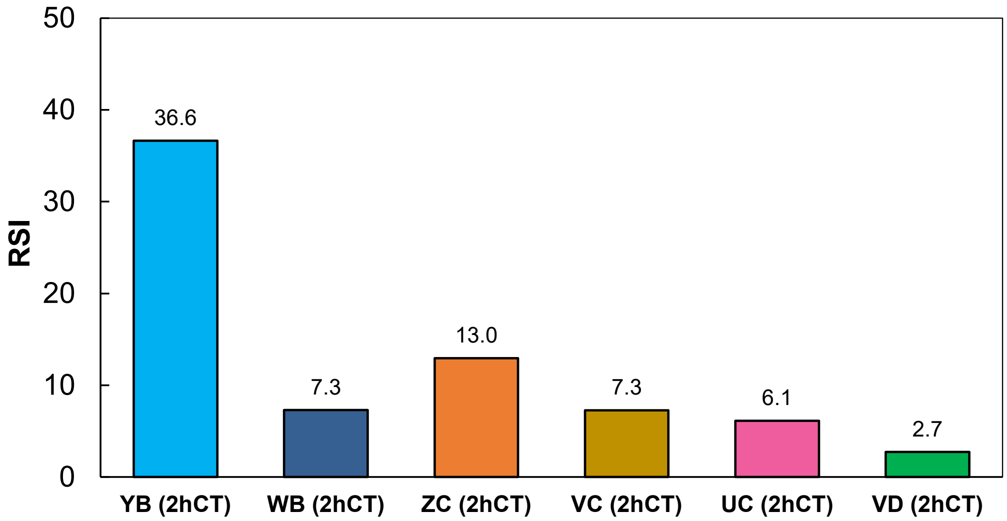


Figure 8. SSR results for benchmarking study.

3.2.3. Summary of Benchmarking

In total, 14 mixtures from different locations in North Carolina were tested using the IDT-CT, APA, DM, and CF tests. A subset of six of these mixtures was also tested using the SSR test method. The performance conclusions for each test, based on statistical Tukey HSD groupings (Appendix C and D), are summarized in Table 9. The table lists the mixtures contained in the top and worst-performing Tukey groups for each respective test result. Regarding cracking performance, the WB mixture was consistently identified as a poor performer across multiple tests, while the HC mixture was consistently among the top performers. For rutting resistance, the YB mixture consistently performed poorly, whereas the RS9.5D mixtures were consistently identified as having excellent performance.

Table 9. Individual Mixture Performance.

Index	High Performing Criteria	Mixtures	Criteria	Mixtures
CT_{index}	highest index	XC, YB, ZC, HC	lowest index	WB, TD, SD, VD, UC
S_{app} and D^R	highest index	FB, SD, RC, HC, SB	lowest index	WB, XC, ZC, VC
APA rut depth	lowest rut depth	SD, VD, TD	highest rut depth	YB
RSI	lowest index	VD	highest index	YB
DM at 4°C	highest modulus	TD, UC, WB	lowest modulus	SB, FB, HC
DM at 20°C	highest modulus	TD, UC	lowest modulus	FB, YB, SB, ZC, HC
DM at 40°C	highest modulus	TD, UC	lowest modulus	ZC, YB, FB

Regarding the mix types, a primary finding was that after removing statistical outliers, no test method identified a significant difference between the RS9.5B and RS9.5C mixtures. The distinction for RS9.5D mixtures was less consistent; while most tests clearly differentiated them, the CF test and DM at 4°C did not find a statistical difference between the RS9.5D group and the other two.

The influence of the source region was also not consistent across all performance tests. Neither the IDT-CT nor the APA test found any significant differences among the regions. However, the DM and CF tests did identify some specific differences. DM results differentiated the piedmont from the mountains (at 4°C and 20°C) and from the coastal plains (at 40°C), while the D^R index also distinguished the piedmont from the coastal plains. The S_{app} index was the only metric to identify a difference between the mountains and coastal plains regions.

3.3. IDT-HT Feasibility Study

The IDT-HT test is a simple and rapid alternative for evaluating rutting resistance. Its advantages include inexpensive equipment and a streamlined procedure using gyratory-compacted specimens with no cutting or gluing required. A diametric load is applied at a constant rate, with the resulting peak tensile strength used as the performance metric. Due to its simplicity, the IDT-HT is well-suited for quality assurance testing and is often used as a pass/fail criterion that correlates well with other rutting performance tests. Additionally, the IDT-HT test has been found to be more sensitive to changes in mixture composition than other stability tests such as the APA, IDEAL-RT, and Marshall Stability test (Boz *et al.*, 2025).

Figure 9 shows the IDT-HT strength results for the five mixtures selected for this evaluation, where a higher strength value indicates better resistance to rutting. The error bars indicate the range between maximum and minimum values. Consistent with the other rutting tests, the YB mixture performed worse than the other mixtures, although the difference was less pronounced here. The overall ranking of the mixtures was in agreement with the APA and SSR test results, demonstrating the high consistency of this simpler test method.

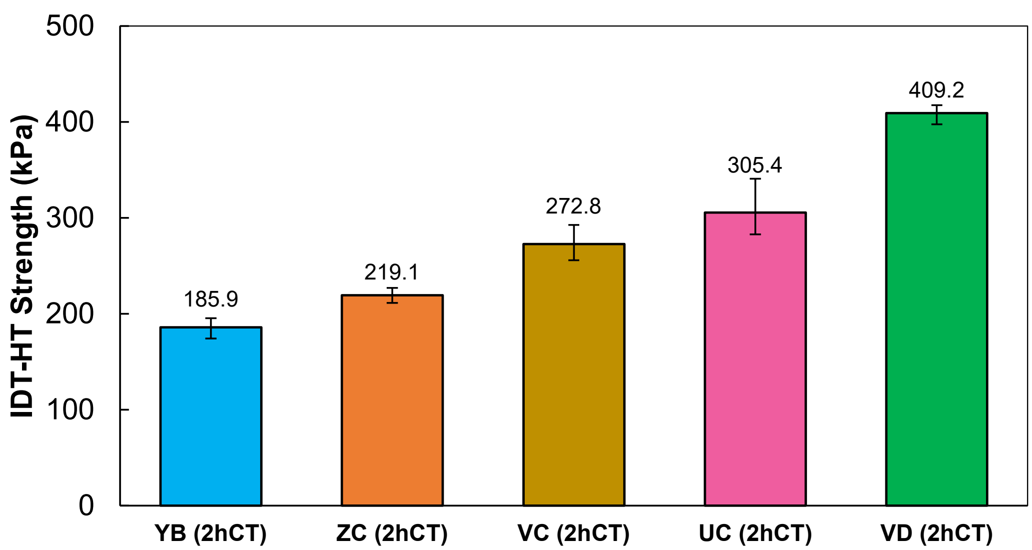


Figure 9. IDT-HT results for selected mixtures.

A one-way ANOVA test (α=0.05) confirmed statistically significant differences among the individual mixtures (*p*-value < 0.001). Interestingly, a subsequent post-hoc Tukey HSD test showed that each mixture of the five mixtures was statistically distinct from each other. This level of differentiation was superior to the APA test, which, due to its higher variability, could not distinguish between certain similarly performing mixtures. This consistency was also proven by the high R^2 values when plotting the test results against each other, as shown in Figure 10.

Finally, based on the strong correlation between the APA and IDT-HT tests, the NCDOT rutting thresholds were translated into preliminary IDT-HT strength criteria. The APA limits of 9.5 mm, 6.5 mm, and 4.5 mm for RS9.5B, RS9.5C, and RS9.5D mixtures corresponded to minimum IDT-

HT strengths of approximately 150 kPa, 180 kPa, and 215 kPa, respectively. It is noted that further work to refine and evaluate these preliminary limits should be done. These current values do not consider issues like variability in the two tests, nor the ability to discriminate mixtures. In addition, the power-law relationship implies that the IDT strengths do not change much with APA rut depth increases above approximately 6 mm, which needs to be considered in light of the inherent test variability of the IDT-HT.

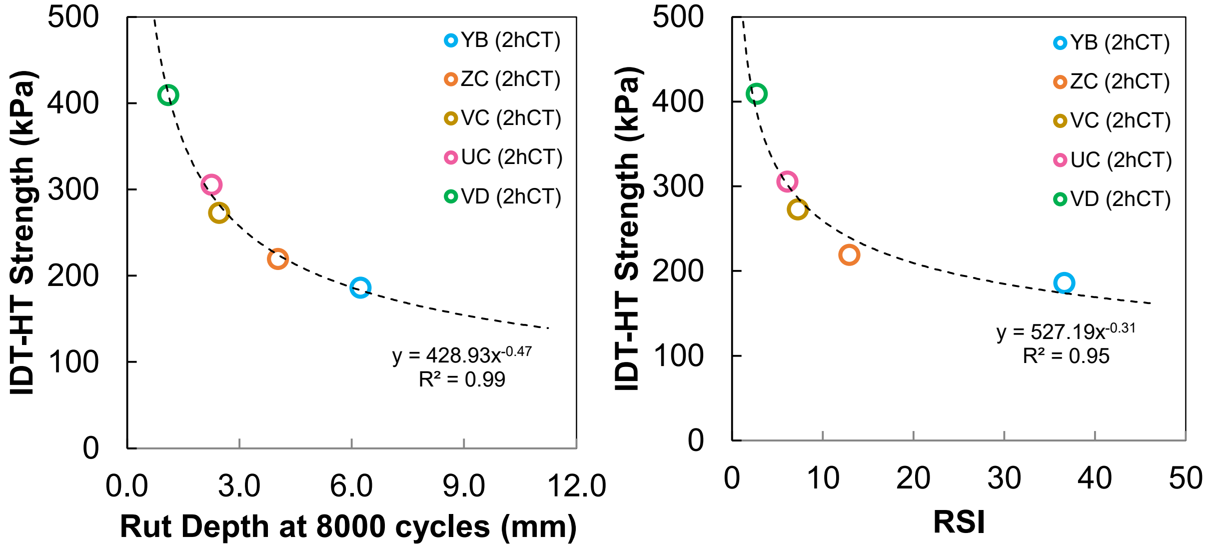


Figure 10. IDT-HT results correlations with other rutting tests: (left) APA; (right) SSR.

3.4. Comparison Laboratory Aging Protocol

Through coordination efforts between this project and FHWA/NC 2023-03, mixtures FB and HC were evaluated using both lab-mixed, lab-compacted specimens and plant-mixed, lab-compacted specimens (PMLC). In both cases, the aggregate and RAP materials were obtained on the same day that the loose-mixture was sampled, and the asphalt binder was sampled as close as possible to the binder acquisition date for the day of production mixture. Note that for the PMLC case, samples of loose mixture were transported to NCSU, where they were then reheated and compacted to the appropriate air void content.

Figure 11 shows the results of CT_{index} values for these two mixtures. As shown, the lab-mixed specimens aged for 4 hours at compaction temperature showed notably lower CT_{index} values than the PMLC specimens (approximately 52% lower in the case of HC and 70% lower in the case of FB). These results suggest that 4 hours of aging at the compaction temperature is a much harsher condition than what occurs in the plant. Figure 12 shows the results from the APA testing. In this case, the aging condition of the lab mixed samples was 2 hours at compaction temperature, and the differences are much smaller. Analysis in this case shows that while the FB mixture and HC mixtures are different, there is no statistically significant difference between the PMLC specimens and the lab-mixed specimens aged for 2 hours.

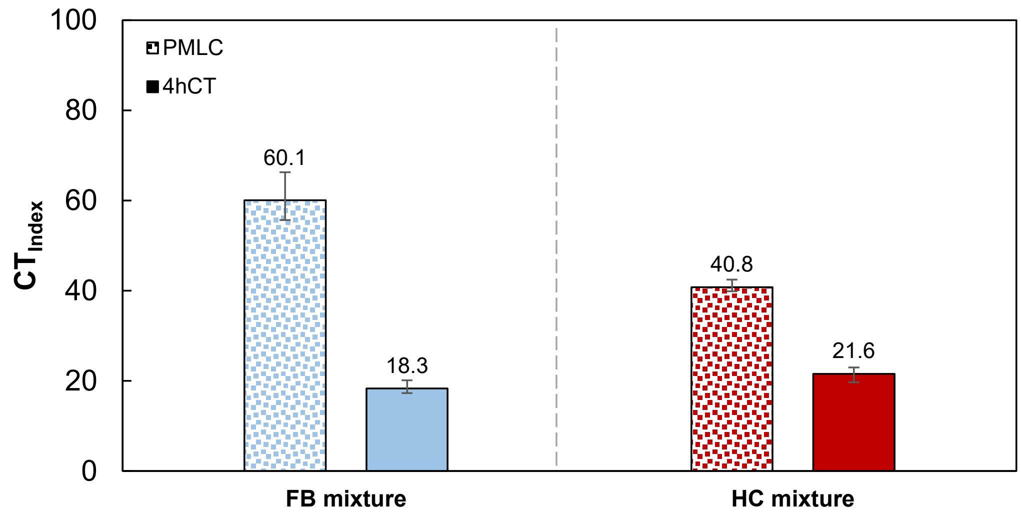


Figure 11. Comparison of CT_{index} values from PMLC and lab-mixed, lab-compacted specimens for FB and HC mixtures.

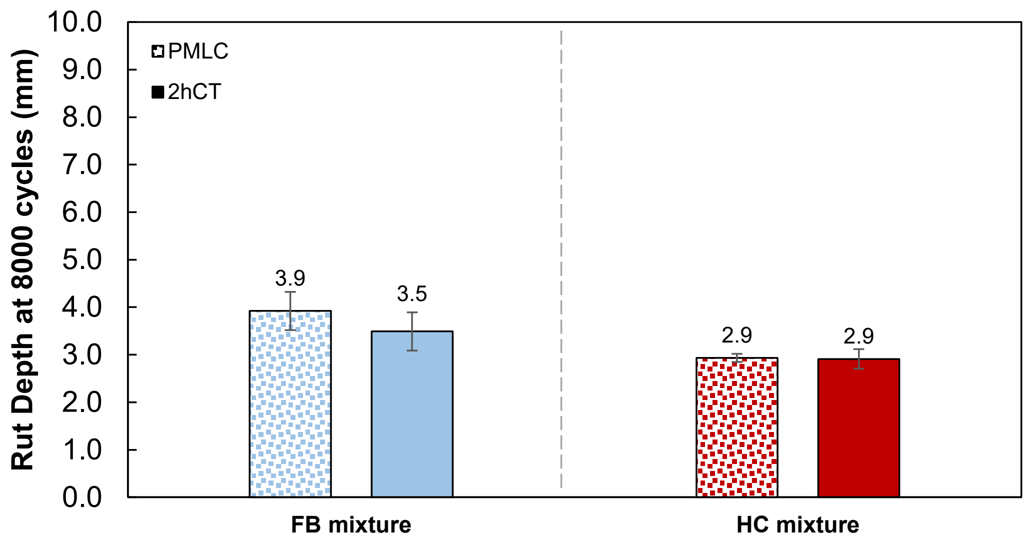


Figure 12. Comparison of APA values from PMLC and lab-mixed, lab-compacted specimens for FB and HC mixtures.

3.5. Short-Term Aging Protocol Variation on RS9.5D mixes

NCDOT job mix formulas specify different compaction temperatures, and consequently, different short-term aging conditions for non-modified and polymer-modified surface mixtures. Typically, the compaction temperature for non-polymer modified mixtures is 143°C and for polymer-modified mixtures is 163°C. Thus, protocols based on aging at compaction temperature are likely to have a greater effect on the properties of polymer-modified mixtures than they will on non-polymer-modified mixtures. In order to more equitably compare both mix types, a separate study was conducted to identify the effects of aging polymer-modified surface mixtures at a different temperature than the non-polymer-modified mixtures and to propose a short-term aging procedure for polymer-modified mixtures.

This investigation focused on two RS9.5D mixtures, sourced from Plant T and Plant V. For the Plant T mixture, short-term aging was conducted using three distinct protocols: 1) aging for four hours at 143°C and then compacting at 143°C (4h143); 2) aging for four hours at 143°C followed by an additional 50 minutes at 163°C (4h143+CT) before compaction; and 3) aging for four hours at 163°C (4h163) and then compacting. The Plant V mixture was only evaluated using only the 4h163 and 4h143+CT conditions. After compaction, the mixtures were subjected to IDT-CT, DM, and CF tests, with the detailed procedures for these tests described in Section 2.3. CT_{index} and S_{app} results are shown in Figure 13, with the DM results shown in Figure 14.

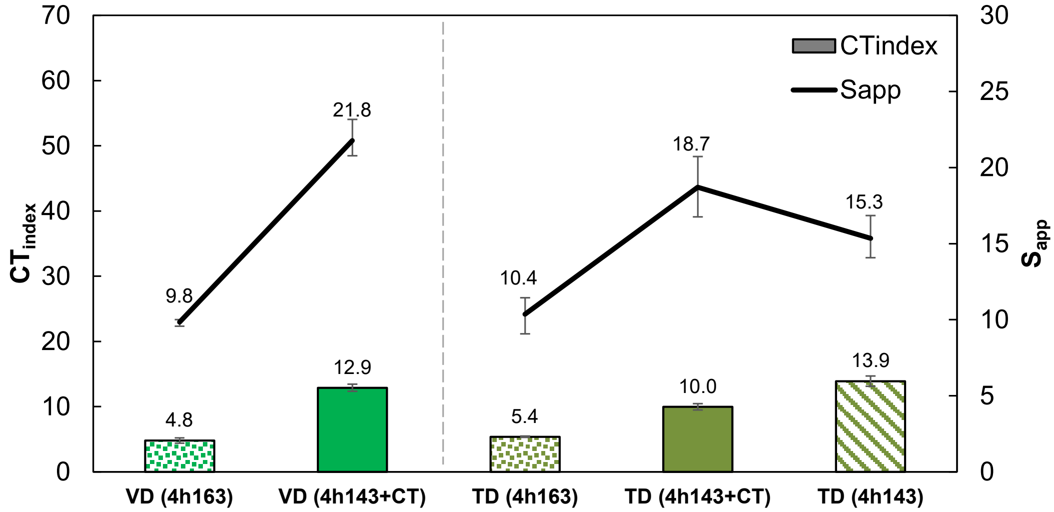


Figure 13. IDT-CT and S_{app} results for RS9.5D STOA study.

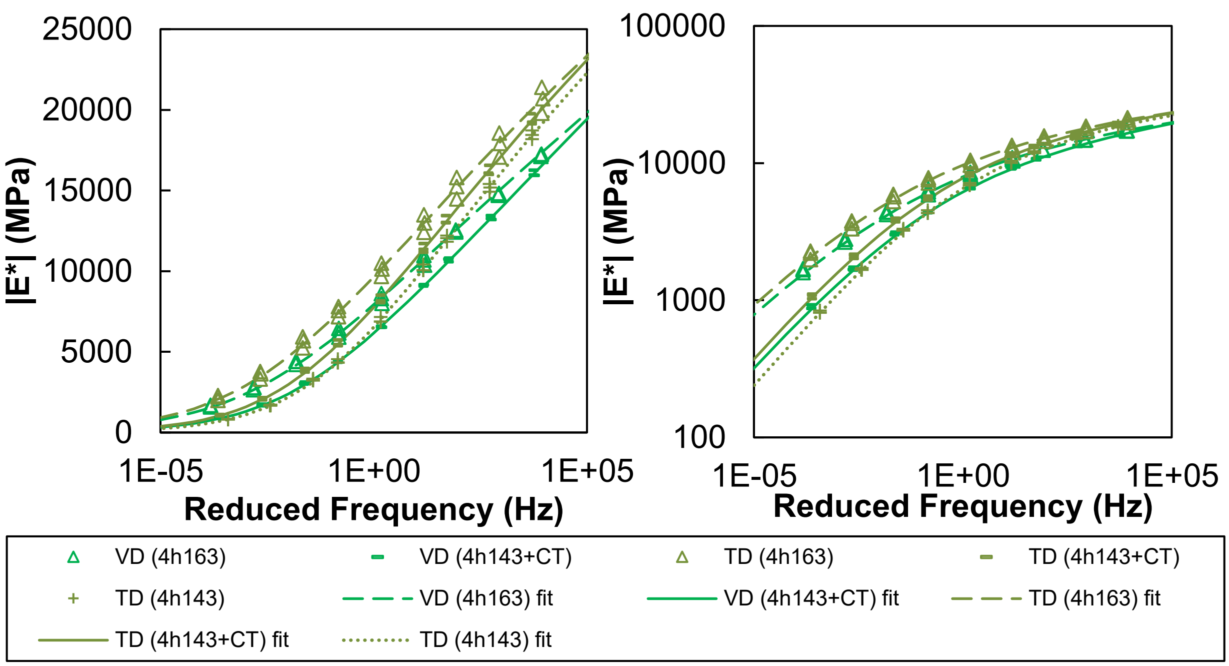


Figure 14. DM test results for RS9.5D STOA study.

Regarding IDT-CT test results, the CT_{index} decreases with the more severe aging conditions for both plant sources. A t-test ($\alpha = 0.05$) was performed for each mixture separately and found that the 4h143+CT condition results in a statistically higher CT_{index} than 4h163 for both VD (85% higher) and TD (169% higher) plants. The Plant T mixture at the 4h143 also had a statistically higher result than TD (4h143+CT). Interestingly, the pairwise t-test also suggested that there is no statistical difference between VD (4h163) and TD (4h163) CT_{index} results; however, for the 4h143+CT condition, the Plant V mix yields a 29% higher CT_{index} compared to Plant T, with the difference being statistically significant.

For S_{app}, the trend is similar to that for CT_{index}. Results were proportional with the aging time, with 4h143+CT results being statistically higher than 4h163 for both VD (122% higher) and TD (47% higher). Statistically, TD (4h143) and TD (4h143+CT) have the same S_{app} results, according to a pairwise t-test. As seen before with IDT-CT results, there is no statistical difference between TD (4h163) and VD (4h163). On the other hand, for 4h143+CT condition, Plant V mix yielded a 42% higher S_{app} index result than Plant T mix, statistically significantly different.

For the DM test results, the modulus at the 1 Hz frequency and at each tested temperature (4°C, 20°C, and 40°C) was compared using a pairwise t-test. The trend is also clear, aging at 4h163 conditions results in a statistically higher modulus than the other conditions for both VD and TD and for all temperatures. The TD (4h143+CT) mixture also shows a pairwise statistically higher modulus value than the TD (4h143) condition at all temperatures. When comparing both plants, TD (4h163) had statistically higher modulus than VD (4h163), and TD (4h143+CT) had statistically higher modulus than VD (4h143+CT) at all temperatures.

For a broader evaluation of the effects of short-term aging conditioning on all the generated data, a two-way Analysis of Variance (ANOVA) test ($\alpha = 0.05$) was conducted to evaluate the effect of both independent variables: plant source and aging condition. The statistical test was done for CT_{index}, S_{app}, and DM at 1 Hz for 4°C, 20°C, and 40°C independently. For the aging condition treatment, all five analyses showed that there were statistical differences among the aging conditions (p-value < 0.001 for all). A Tukey HSD post-hoc test was conducted on each independent variable separately, showing that each aging condition was in a different group from each other. For CT_{index}, the order from highest to lowest result was 4h143, 4h143+CT, and 4h163. For S_{app}, the order from higher to lower results was 4h143+CT, 4h143, and 4h163. For all three temperatures, the order from higher to lower modulus was 4h163, 4h143+CT, and 4h143.

Regarding plant source treatment, for both cracking indexes, there was no statistical differentiation for plant source treatment (p-value of 0.078 and 0.187 for CT_{index} and S_{app} respectively), meaning that both plants exhibited index results that are statistically the same. On the other hand, for DM results, all three temperatures found statistical significance (p-value < 0.001) for plant source. A Tukey HSD post-hoc test was conducted, which identified TD as resulting in a higher modulus than VD for all three temperatures.

In summary, the choice of short-term aging conditions significantly impacted the results. While mixes from Plant T and Plant V showed similarity in cracking performance (CT_{index} and S_{app}), the modulus results showed that they were not the same. Applying the AASHTO R 30-19 standard aging condition of four hours at compaction temperature (163ºC) proved to be too severe, as it aged the materials to a point that the cracking performance was statistically indistinguishable. On the other hand, the less harsh condition of four hours at 143°C (plus 50 minutes to reach compaction temperature) revealed statistical differences between Plant T and V mixes' cracking performances. The current AASHTO R 30-22 practice of aging at 135°C for 2 hours was not evaluated based on input from the project steering committee. Based on these findings, the research team decided to apply 4h143+CT as a condition to subsequent testing of RS9.5D mixes for this project.

3.6. Sensitivity of Cracking Index Results to Asphalt Supplier

According to NCDOT procedures, contractors are allowed to switch binder suppliers without notifying the NCDOT as long as the anti-stripping supplier and PG grade do not also change. However, it is not clear if this practice has impact on the cracking potential, given the variability of asphalt binder supplies in North Carolina. Therefore, following a recommendation from NCDOT, the research team conducted a separate experiment to evaluate the effect of asphalt supplier changes on surface mixture properties. The experiment consisted of using the five terminal sources (bA, bB, bC, bD, and bE) of PG 64-22 sampled during the project and from three different companies. The earliest binder sampling was done in September of 2022, while the last one was done in July of 2023. Each source was sampled only once.

To establish a baseline for subsequent analysis, this study characterized the rheological properties of the PG 64-22 binders using a Dynamic Shear Rheometer (DSR) with a parallel plate geometry. The evaluation included testing the binders in both their original (unaged) state and after short-term aging via the Rolling Thin-Film Oven (RTFO) method (AASHTO T 240, 2023c). Low temperature testing characteristics were not verified. The testing protocol consisted of Performance Grade (PG) verification and Multiple Stress Creep Recovery (MSCR) tests, conducted according to AASHTO T 315-24 (2024b) and AASHTO T 350-19 (2023f), respectively. Table 10 presents a detailed summary of the characterization, presenting the average results of two replicates. The repeatability was verified to meet all specified limits. As shown all binders are graded with a high temperature grade of PG 64 according to the AASHTO M 320 system. In the AASHTO M 332 system bA and bE are considered PG 64H while all others are PG 64S.

Table 10. Binder Characterization for Binder Supplier Sensitivity Study.

Test Condition	Binder Source				
	bA	bB	bC	bD	bE
Original binder, $\lvert G^*\rvert/\sin\delta$ (kPa) at 64°C	2.020	1.763	1.953	1.768	1.764
Original binder, $\lvert G^*\rvert/\sin\delta$ (kPa) at 70°C	0.966	0.864	0.948	0.871	0.850
Original binder, Failing Temperature	69.7°C	68.8°C	69.6°C	68.8°C	68.7°C
RTFO binder, $\lvert G^*\rvert/\sin\delta$ (kPa) at 64°C	5.084	4.200	4.802	3.939	6.232
RTFO binder, $\lvert G^*\rvert/\sin\delta$ (kPa) at 70°C	2.380	1.988	2.270	1.869	2.908
RTFO binder, $\lvert G^*\rvert/\sin\delta$ (kPa) at 76°C	1.161	-	1.117	-	1.413
RTFO binder, Failing Temperature	70.7°C	69.2°C	70.3°C	68.7°C	72.3°C
RTFO binder, $J_{nr3.2}$ (kPa^{-1}) at 64°C	1.988	2.193	2.055	2.517	1.520
Performance Grade (AASHTO M 320)	PG 64	PG 64	PG 64	PG 64	PG 64
Performance Grade (AASHTO M 332)	PG 64H	PG 64S	PG 64S	PG 64S	PG 64H

The job mix formula of four different RS9.5B mixtures was tested using three different binders (no changes in gradation or binder content), following a statistically designed experimental plan, summarized in Table 11. The research team generated this design by first using JMP's Custom Design to create an A-optimal design to estimate the main effects of the two categorical factors: binder source and plant source. The team structured the experiment in four blocks of size three, ignoring the interaction effect for this analysis. The design ensured that the original JMF condition for each plant source was included in the experiment; for example, the Plant Y mixture with binder bA had already been tested as part of the RP2023-03 project.

Table 11. Experimental Program of Binder Supplier Sensitivity Study.

Plant	Mix Type	Binder Source				
		bA	bB	bC	bD	bE
Y	RS9.5B	JMF	X			X
W	RS9.5B	JMF		X	X	
S	RS9.5B		X	JMF		X
U	RS9.5B		JMF		X	X

The CT_{index} average results from the variants shown in Table 11 are shown in Figure 15, with error bars representing maximum and minimum values. As shown, the results vary substantially across the different combinations. Mixtures from Plants Y, S, and U consistently produced higher CT_{index} results, ranging from about 20 to 25 units. In contrast, the Plant W group was a clear underperformer, with all three mixes registering CT_{index} values below 10 regardless of the binder source. Visually, no binder source was clearly outstanding; however, source bA seems to result in slightly higher CT_{index} values, while source bB seems to result in slightly lower CT_{index} values.

A two-way Analysis of Variance (ANOVA) test ($\alpha = 0.05$) was conducted to evaluate the effect of both independent variables: plant source and binder source. The results are shown in Table 12. The test concluded that there are differences among plant source (p-value < 0.001) and also binder source (p-value of 0.008). Since differences were seen, a post-hoc Tukey honestly significance difference (HSD) test was performed to identify the statistical differentiation among treatment factors. For the first treatment, there are two statistically different groups: one containing Plants U, Y, and S, and another one containing Plant W. For the binder source treatment, the Tukey HSD test revealed that binder bA produced a statistically higher CT_{index} than binder bB. No other binder pairings showed a significant difference.

Figure 16 shows the dynamic modulus |E*| master curves for all combinations of plant and binder sources. From both semi-log and log-log plots, there is a visual consistency within each plant group, indicating little influence from binder source, especially apparent for Plants S and U groups. The overall mixture source (i.e.: production plant) seems to be the dominant factor controlling the material's stiffness properties compared to the binder source.

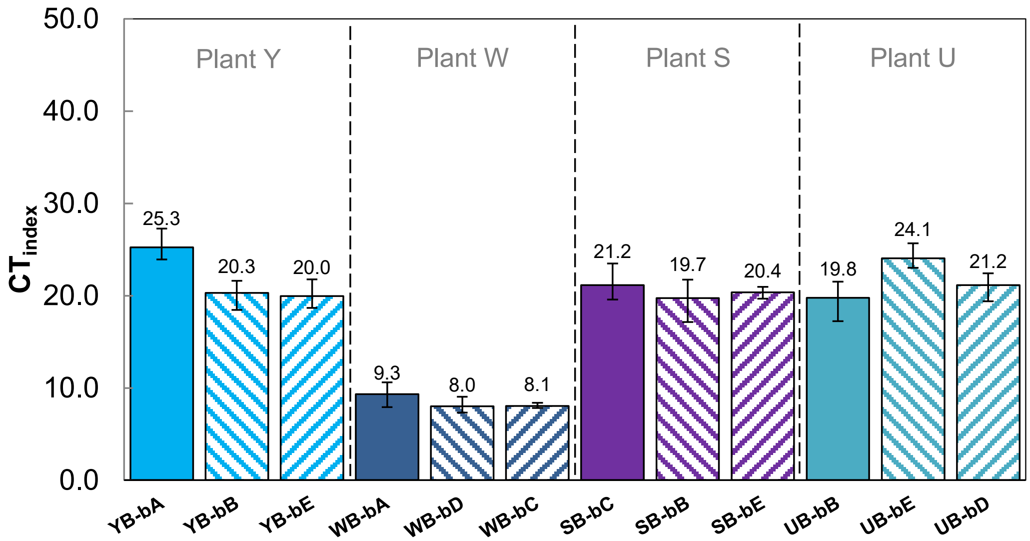


Figure 15. IDT-CT results for asphalt supplier experiment.

Table 12. Summary of Statistical Analysis of Binder Supplier Sensitivity Study.

Attribute	**Difference in Means, two-way ANOVA (α=0.05) Prob>F**					
	CT_{index}	**DM (4°C)**	**DM (20°C)**	**DM (40°C)**	S_{app}	D^R
Plant Source	<0.0001	<0.0001	<0.0001	<0.0001	0.0484	0.0002
Binder Source	0.0077	0.0010	0.0182	0.0007	0.8107	0.1553
Tukey HSD Letter Grouping						
Binder Source	CT_{index}	**DM (4°C)**	**DM (20°C)**	**DM (40°C)**	S_{app}	D^R
Binder bA	A	A	A	B C		
Binder bC	A B	A B	A B	A B C		
Binder bD	A B	B	A B	A	-[1]	-[1]
Binder bE	A B	B	A B	A B C		
Binder bB	B	B	B	B		
Tukey HSD Letter Grouping						
Plant Source	CT_{index}	**DM (4°C)**	**DM (20°C)**	**DM (40°C)**	S_{app}	D^R
Plant Y	A	C	D	D	A B	A
Plant S	A	C	C	C	A	A
Plant U	A	B	B	B	A B	A
Plant W	B	A	A	A	B	B

[1] Statistical Analysis not shown.

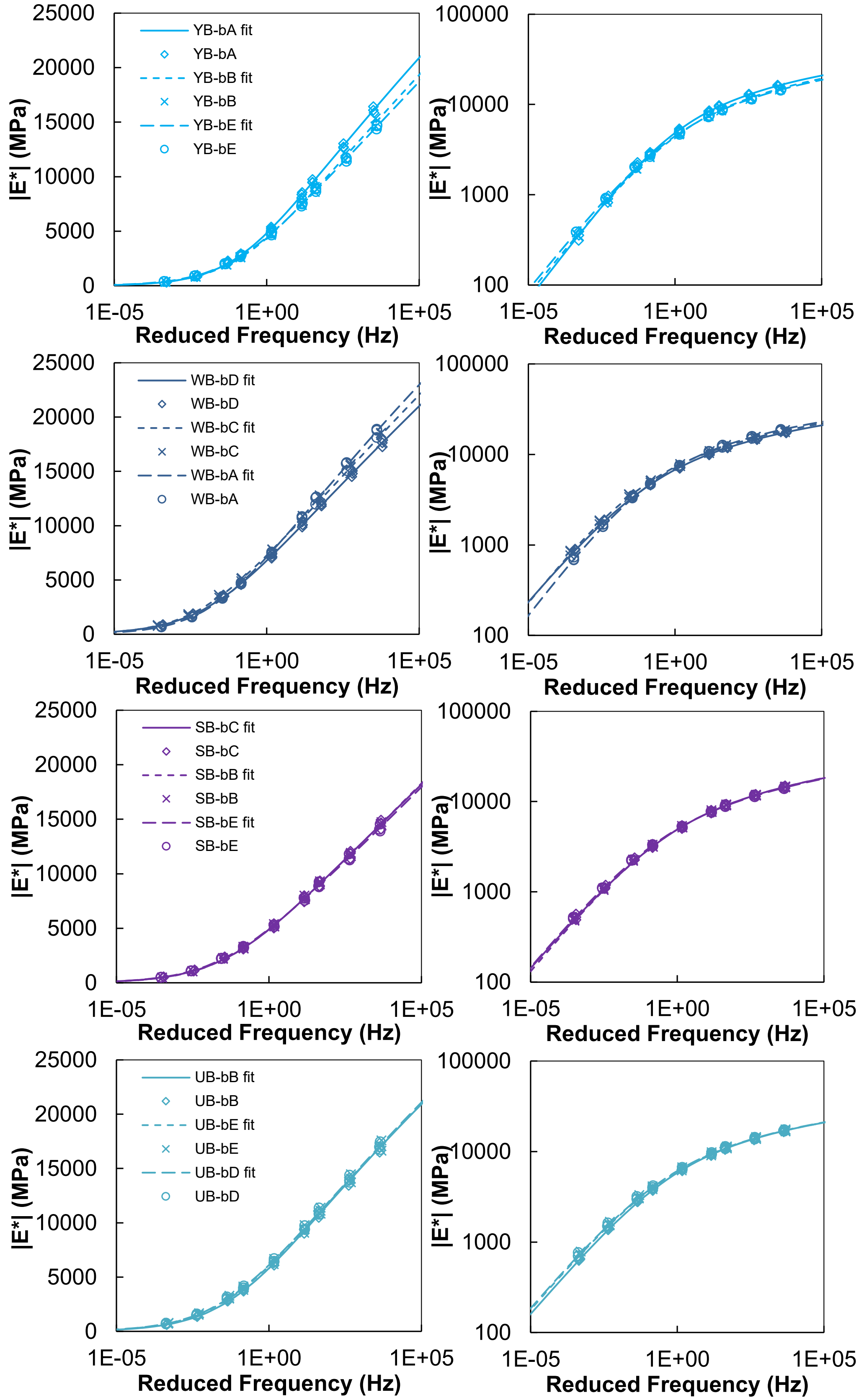


Figure 16. DM test results for asphalt supplier experiment.

Using experimental data at 1 Hz and from the 4°C, 20°C, and 40°C temperatures, a two-way ANOVA test (α = 0.05) was conducted to identify differences among both plant and binder sources. The analysis showed that the plant source had a statistically significant effect on all temperatures (p-value < 0.001 for all). The binder source also showed significant, though less pronounced, effects at 4°C (p-value = 0.001), 20°C (p-value = 0.018), and 40°C (p-value < 0.001).

A subsequent Tukey HSD post-hoc test clarified statistically different groups among the treatments. For the plant source, a consistent trend showed up at all temperatures. Plant W had the highest modulus, followed by Plant U, with Plant S and Y having the lowest. For the binder source, the effect was not consistent across all temperatures. At 4°C and 20°C, binder bA was the stiffest, while at 40°C, binder bD had a higher modulus. At 4°C, bA had a higher modulus than bB, bD and bE, with no other pairings showing a significant difference. At 20°C, bA had statistically higher modulus than bB, with no other binder pairings showing a significant difference. At 40°C, bD had a higher modulus than bB and bA, with no other binder pairings showing a significant difference. This lack of a consistent trend explains why the binder effect was not visually apparent on the overall master curves and confirms plant source as a predominant factor controlling material stiffness.

Figure 17 shows the resulting D^R and S_{app} representative index results from the CF test, while the error bars represent maximum and minimum values. Like IDT-CT results, both D^R and the S_{app} index indicate that the Plant W group underperformed compared to the other groups, with average values about 0.440 for D^R and 18.0 S_{app}. For D^R, the Plant Y group showed the highest variability, but also present the highest single value, 0.565 with the bE binder. Plant S and Y groups stand out with the highest values for S_{app}, while the Plant U group showed a consistent performance with results around 20.

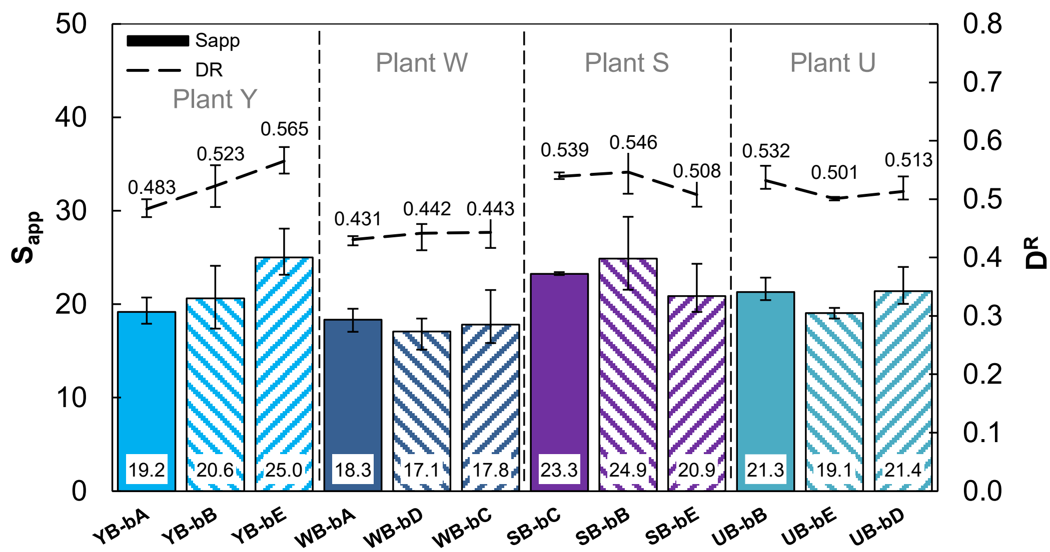


Figure 17. CF results for asphalt suppler study.

A two-way ANOVA test (α = 0.05) was conducted to evaluate the effect of both plant and binder source on D^R and S_{app}. The analysis confirmed that the plant source had a statistically significant effect on both D^R (p-value < 0.001) and S_{app} (p-value = 0.048). In contrast, the binder source had no significant effect on either index value. A subsequent Tukey HSD post-hoc detailed plant source differences. For D^R, Plant W was statistically lower than all other plants. For S_{app} index, Plant W produced a statistically lower value than Plant S, however, no other binder pairings showed a significant difference.

In summary, the plant source was a more dominant factor in mixture performance than the binder source. Plant W consistently showed worse potential cracking performance and a higher modulus at all temperatures. Among other mixtures (Plants Y, S, and U), there is no clear indication that one is better than the other. The binder source, which was the focus of this study, had a limited impact. Its effect was either not statistically significant (D^R and S_{app} index values), showed inconsistent trends across different temperatures (dynamic modulus), or only revealed significant differences between binders bA and bB. Therefore, for the binders sampled in this project, there is little evidence to suggest that swapping between binder sources significantly affects the performance testing results. This conclusion should be considered with caution due to the narrow time frame of binder sampling.

3.7. Performance Evaluation

While evaluating mixture properties on their own is useful, its performance can be either magnified or minimized when applied in a pavement structure. Therefore, to estimate the service life of the selected materials, pavement performance simulations were conducted using the FlexPAVE™ 2.2 performance model. Two characteristic structures were chosen for the analysis: a full-depth asphalt (FDA) pavement and a pavement with an aggregate base course (ABC), each with a thick and thin variation. The specific layer thicknesses, traffic inputs, and material characteristics used in the simulations are detailed in Table 13, Table 14, and Table 15, respectively. In total, twenty-four simulations were performed using the climatic file for Raleigh, NC.

Table 13. Pavement Structures used in Pavement Simulations.

Structure type	Abbreviation	Thickness (in.)					
		Surface AC	Intermediate AC	Base AC	ABC	Subgrade	Total
FDA thin	FDAtn	3	-	4	-	Infinite	7
FDA thick	FDAtk	3	4	10	-	Infinite	17
ABC thin	ABCtn	3	-	-	8	Infinite	11
ABC thick	ABCtk	3	4	-	10	Infinite	17

Table 14. Traffic Characteristics used in Pavement Simulations.

Thickness type	Daily ESALs	Growth type	Growth Rate (%)	Speed (mph)
Thin	2000	Linear	0.4	60
Thick	6000	linear	0.4	60

Table 15. Material Characteristics used in Pavement Simulations.

Material	Poisson Ratio	Moduli (psi)	Source	Original Naming
Surface AC	0.30	-	This project	YB, WB, ZC, VC, UC, VD
Intermediate AC	0.30	-	RP2019-20	PI_RI19.0C
Base AC	0.30	-	RP2019-20	MO_RB25C
ABC	0.35	29008	Default	Aggregate_Base_A-1-a
Subgrade	0.40	10878	Default	Subgrade_A-6

The primary focus of the simulations was to evaluate the six surface mixtures that had been fully characterized by both the CF and SSR tests. To isolate the performance of these surface layers, all other material properties were kept constant across all simulations. The intermediate and base asphalt layers were selected from the NCSU database to represent an average performance, while default FlexPAVE™ materials were used for the unbound aggregate base and subgrade.

3.7.1. Rutting Predictions

Figure 18 shows the predicted change in surface layer rutting during the different pavement simulations, set to 0.5 inches as the maximum value of y-axis. Structures containing the YB mixture showed the most significant rutting, in some cases reaching depths equivalent to over 2 inches, two-thirds of the layer's thickness. Structures with the ZC mixture also exhibited pronounced rutting, with the surface layer alone contributing up to one-third of its thickness in rut depth. It is important to note that these values do not represent total pavement rutting, as contributions from the supporting layers are not accounted for in this analysis. Appendix E presents the contribution of each layer separately in full scale.

Structures with the VC and WB mixtures presented very similar rutting levels, while structures with the UC and VD mixtures were the best performers, with rut depths of less than one-tenth and one-twentieth of the surface layer thickness, respectively. The performance ranking of the mixtures in the simulation is highly consistent with the results from all laboratory rutting tests (APA, SSR, and IDT-HT). However, while all tests identified the YB mixture as a poor performer, its exceptionally high rutting potential was most clearly captured by the APA and SSR tests.

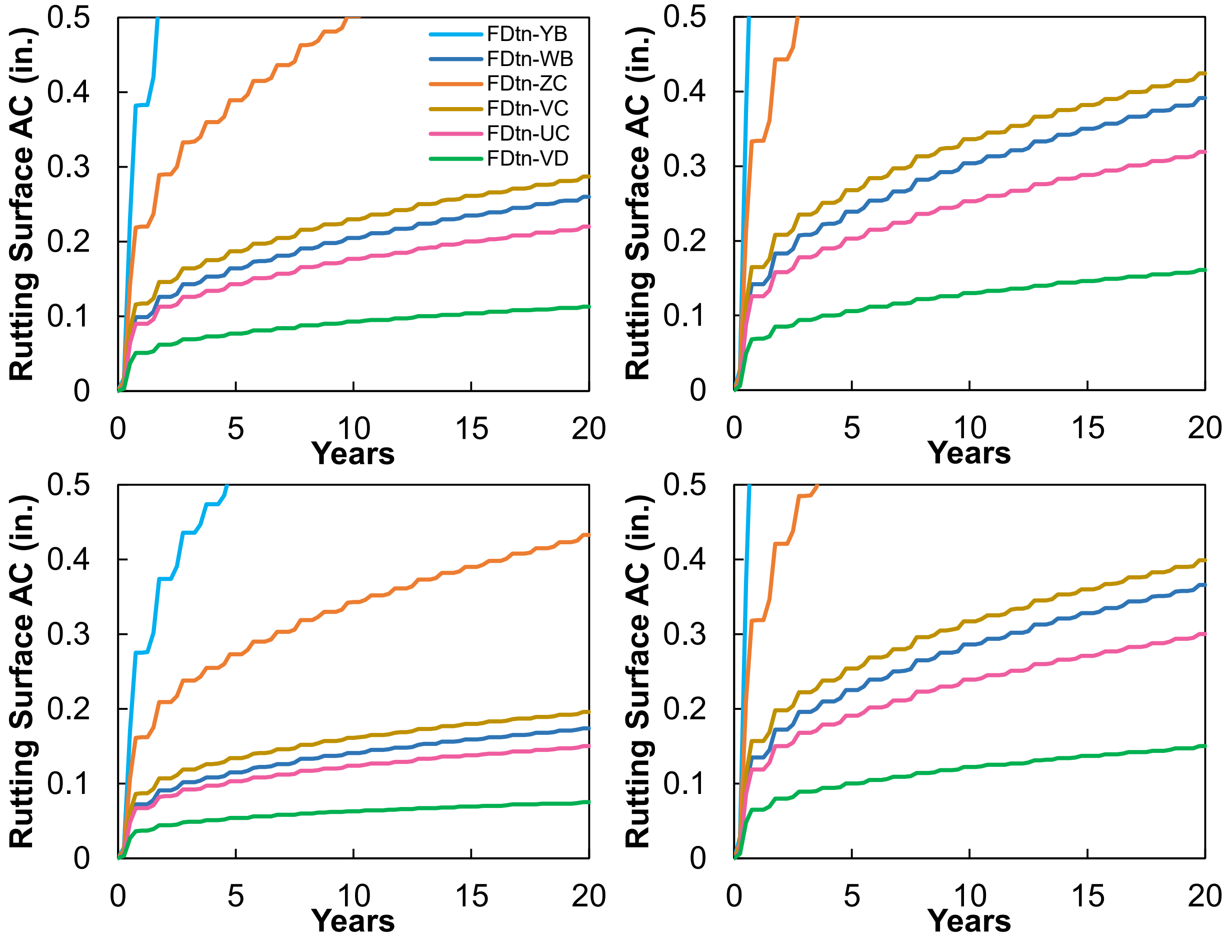


Figure 18. Rutting in the surface layer from FlexPAVE^{TM} simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

3.7.2. Fatigue Damage Predictions

For fatigue cracking, only the damage in the top 1/3rd of the pavement structure (referred to as the 'Top Damage') is shown since the bottom part of the damage is compounded by the intermediate or base asphalt (Saleh 2022). Figure 19 shows the expected damage over twenty years for the analyzed structures that contain both CF and SSR testing.

An analysis of these pavement simulations shows that structures containing the WB mixture had the worst long-term performance, which aligns with its poor results in the IDT-CT and CF cracking tests (discussed in more detail in Section 3.7.3). Although the WB mixture was in the group with a higher dynamic modulus, which can help minimize strain, its poor damage resistance appears to have outweighed this benefit. The top-performing pavements among those mixtures where both CF and SSR were performed were those containing the VC mixture, followed closely by the pavements with the YB mixture. This outcome was notable, as the YB mixture was a top performer in the IDT-CT test but had intermediate results in the CF test. Finally, a cluster of mixtures with average performance was observed, containing the VD, ZC, and UC mixtures.

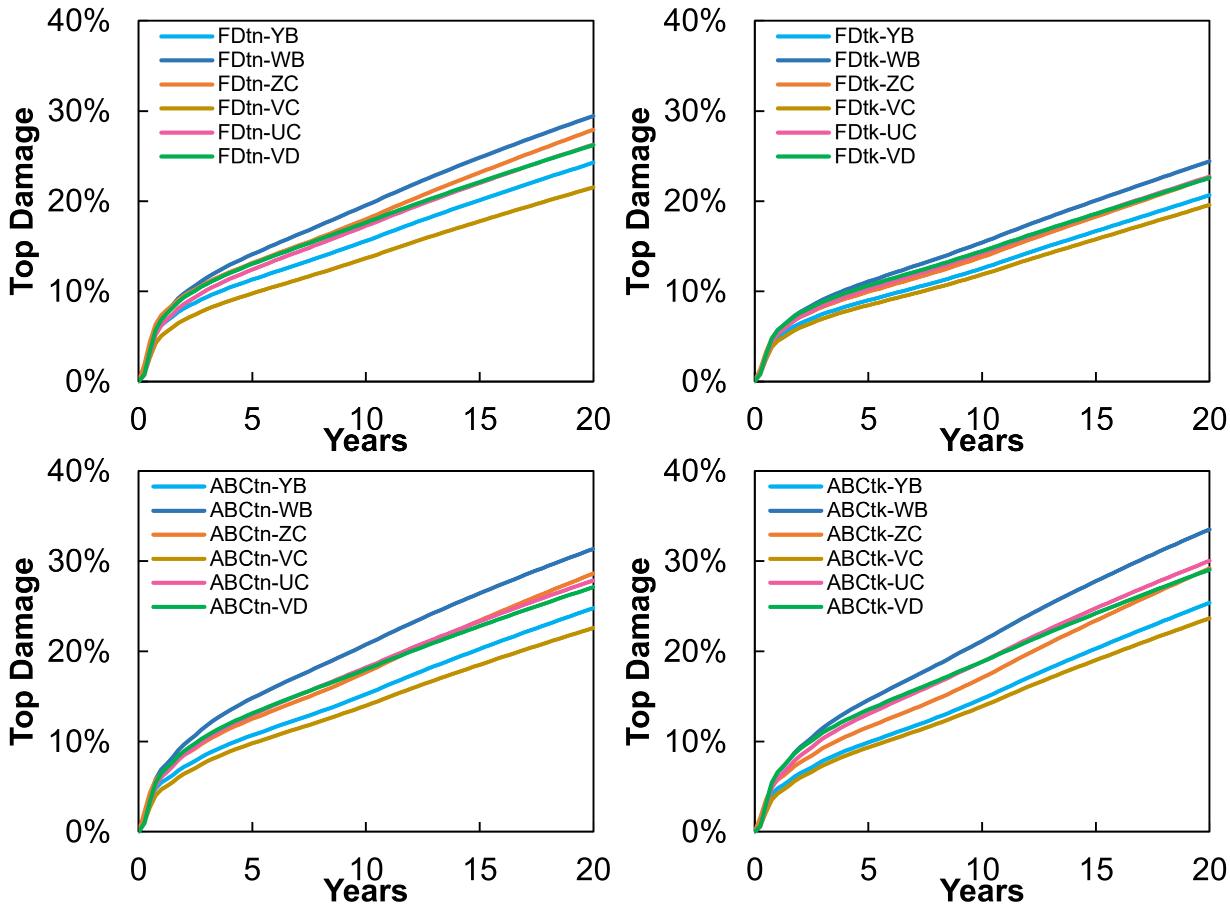


Figure 19. Top percent damage from FlexPAVE™ simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

Additional simulations were performed to evaluate fatigue damage on the mixtures that were tested for DM and CF, but had not been tested in the SSR. Simulations can be performed in FlexPAVE™ with only DM and CF properties because the rutting and cracking simulations do not interact. The results, shown in Figure 20 as percent Top Damage, indicate that this second group of mixtures had superior fatigue performance compared to the first. For example, the best performer from the first group (structure with VC mixture) had 29.0% damage in the FDtk structure at 20 years, while the worst performer from the second group (structure with HC mixture) had 28.4% damage in the FDtk structure at 20 years. This overview highlights the necessity of evaluating a large and diverse group of mixtures to avoid drawing potentially misleading conclusions from a limited sample set.

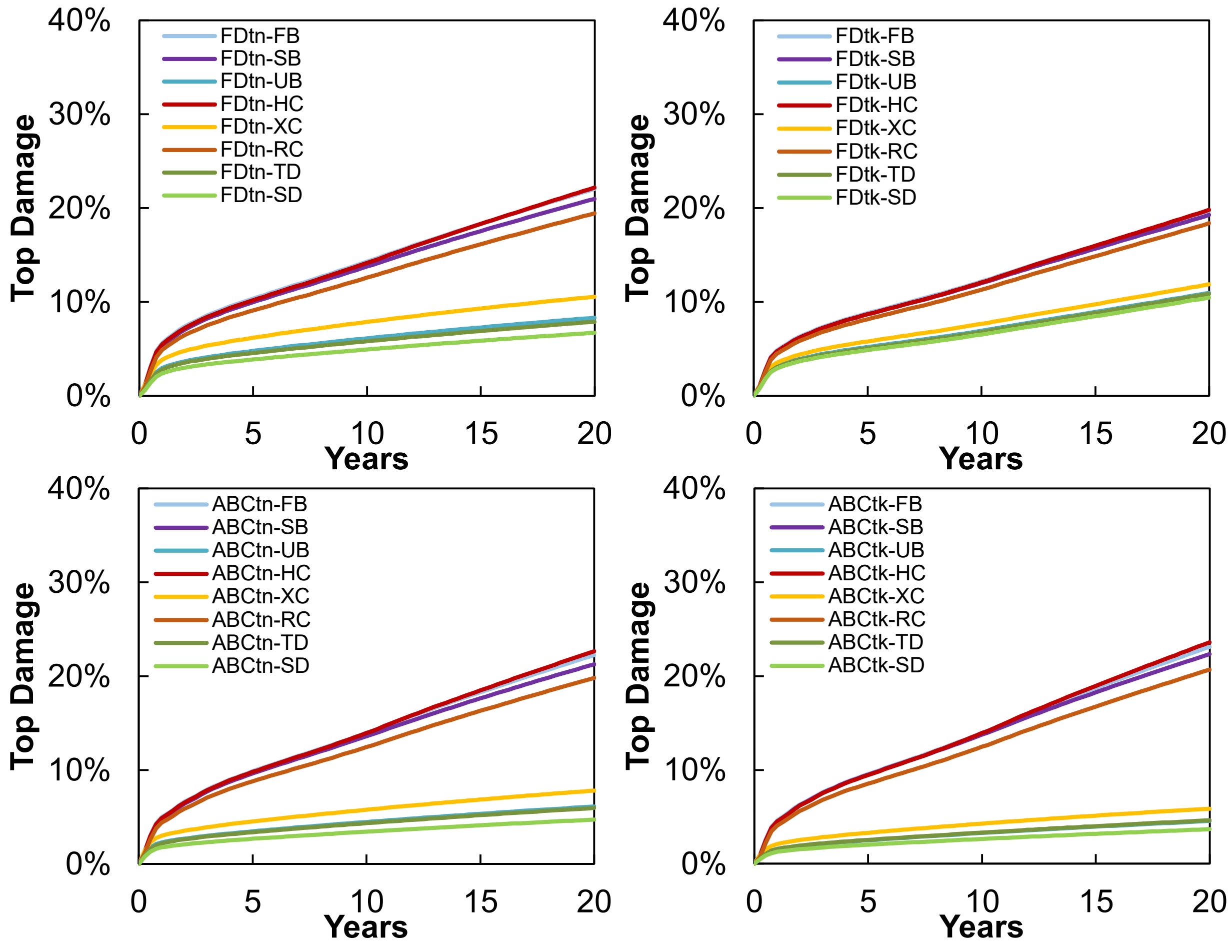


Figure 20. Top percent damage from FlexPAVE^{TM} additional simulations for the structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

Collective analysis of the data in Figure 19 and Figure 20 revealed four distinct groups, as demonstrated for the thick ABC pavement in Figure 21. Only the thick ABC pavement is shown, but the same groupings emerge for all pavement structures evaluated. As seen, the structures containing the WB mixture were the worst performers, followed by a fair performing group that included the UC, ZC, and VD mixtures. A third large cluster of mixtures with moderate performance included the YB, VC, HC, FB, SB, and RC mixtures. On the other hand, the best-performing group exhibited excellent durability over time, consisting of structures with the XC, TD, UB, and SD mixtures.

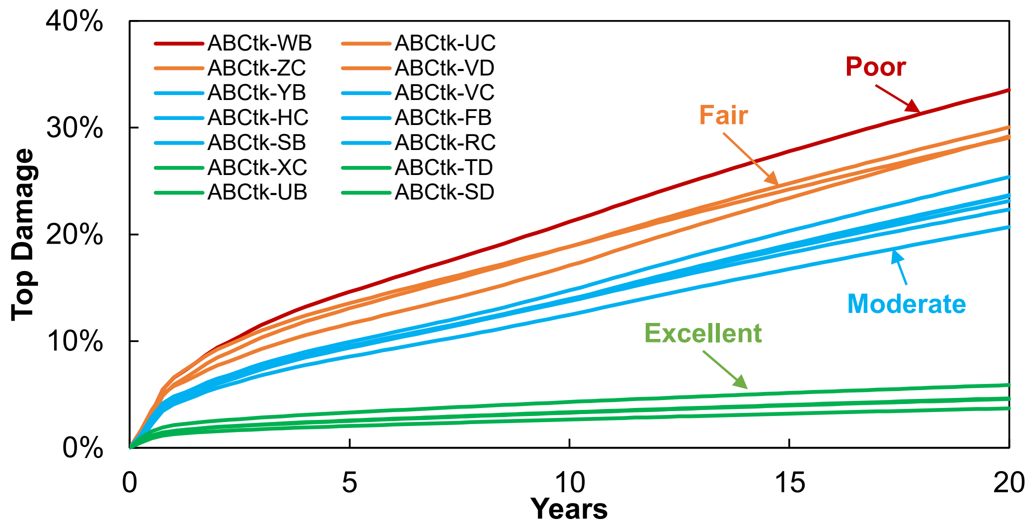


Figure 21. Comparison of pavement performance simulations showing grouping for ABCtk pavements.

3.7.3. Summary of Performance Evaluation

In total, 56 simulations were performed covering four structures and the evaluation of six mixtures from different locations in North Carolina for rutting and 14 mixtures for fatigue cracking. The performance conclusions are summarized in Table 16 for rutting and Table 17 for cracking. Qualitatively, the structures containing the surface mixtures were divided into excellent, good, moderate, fair, and poor performances as discussed above. For each test, the Tukey grouping was also categorically divided into either good (contained in the top-performing Tukey group), poor (contained in the worst-performing Tukey group), or moderate (neither good nor poor performing). For example, based on the results shown in Table C.1, mixtures with an average CT_{index} above 21.5 are considered good, mixtures with an average CT_{index} below 14.0 are conserved poor, and mixtures with average CT_{index} between 21.5 and 14.0 are considered moderate. Appendices C and D present the summary of Tukey groupings for BMD and AMPT tests, respectively.

Table 16. Summary of Rutting Ranking.

Surface AC	Surface Rutting at 20 Years (in.)				Performance			
	FDtn	FDtk	ABCtn	ABCtk	FlexPAVE^{TM}	APA	SSR	IDT-HT
VD	0.11	0.16	0.08	0.15	Excellent	Good	Good	Good
UC	0.22	0.32	0.15	0.30	Moderate	Moderate	Moderate	Moderate
WB	0.26	0.39	0.17	0.37	Moderate	Moderate	Moderate	-
VC	0.29	0.42	0.20	0.40	Moderate	Moderate	Moderate	Moderate
ZC	0.65	1.04	0.43	0.97	Fair	Poor	Poor	Poor
YB	1.40	2.39	0.90	2.24	Poor	Poor	Poor	Poor

The results highlight the important balance in mixture performance. For example, the YB mixture was a fair performer against cracking criteria but was the worst performer in rutting tests and simulations. On the other hand, the WB mixture showed moderate rutting resistance but delivered the worst performance in cracking tests and simulations. This demonstrates that over-optimizing for one property can negatively impact another. The ultimate goal is not to find a single "best" mixture, but to establish reliable criteria that improve overall performance and service life when multiple distresses are considered.

The rutting behavior from the simulations was in excellent agreement with all laboratory tests. For cracking, however, the findings were more complex as summarized in Table 17, which summarizes the categorical ratings from all the mixtures from FlexPAVE, IDT-CT, and CF. Starting with the RS9.5D mixtures, the pavement simulations showed that the SD and TD mixes delivered exceptional performance, while the VD mix was among the fair performers. This outcome was not clearly predicted by the laboratory index tests alone. In IDT-CT test, all three RS9.5D mixtures were grouped together with the lowest CT_{index} results. The CF test, however, was more discriminating. It identified the SD mixture as a top performer but suggested the TD mixture would perform worse than the VD mixture. The pavement simulations indicate that the higher dynamic modulus of the TD mixture likely reduced the strains within the structure, improving its performance enough to place it in the top tier. This result highlights the important interaction between a material's stiffness and its cracking resistance.

Table 17. Summary of Fatigue Damage and Durability Ranking.

Surface AC	Top Damage at 20 Years (%)				Performance		
	FDtn	**FDtk**	**ABCtn**	**ABCtk**	**FlexPAVE™**	**IDT-CT**	**CF**
SD	6.7	10.5	4.7	3.7	Excellent	Poor	Good
UB	8.3	11.0	6.1	4.6	Excellent	Moderate	Moderate
TD	7.9	10.9	6.0	4.6	Excellent	Poor	Moderate
XC	10.6	11.9	7.8	5.9	Excellent	Good	Poor
RC	19.4	18.4	19.8	20.7	Moderate	Moderate	Good
SB	21.0	19.3	21.3	22.3	Moderate	Moderate	Good
FB	22.0	19.7	22.3	23.1	Moderate	Moderate	Good
HC	22.2	19.8	22.7	23.6	Moderate	Good	Good
VC	21.5	19.6	22.6	23.7	Moderate	Moderate	Poor
YB	24.3	20.7	24.8	25.4	Moderate	Good	Moderate
VD	26.2	22.6	27.1	29.1	Fair	Poor	Moderate
ZC	27.9	22.6	28.6	29.2	Fair	Good	Poor
UC	26.3	22.7	27.9	30.1	Fair	Poor	Moderate
WB	29.5	24.4	31.4	33.5	Poor	Poor	Poor

For the mixtures categorized in the fair performance simulation group (VD, ZC, and UC), a comparison to lab data reveals a reasonable correlation. The UC mixture ranked as one of the worst performers in the IDT-CT test and as bottom-average in the CF test. The ZC mixture also performed poorly in the CF test; however, its high CT_{index} directly contradicted the fair pavement simulation finding.

Within the moderate performance simulation group (RC, SB, FB, HC, VC, and YB), IDT-CT correctly identified VC, HC, FB, SB, and RC mixtures as average performers. However, the second-highest CT_{index} for YB mixture was contradictory. CF test results classified YB and VC as low-average performers, yet it ranked FB, RC, SB, and HC in its top-performing group for S_{app} and D^R indexes. This discrepancy for the latter mixes may be attributed to their statistically lower dynamic modulus.

Finally, in the non-RS9.5D mix excellent performance simulation group (UB and XC), the laboratory correlations were again inconsistent. The XC mixture's top ranking in the IDT-CT aligned with the simulation, but its poor CF test result did not. On the other hand, the UB mixture's average-to-high ranking in both cracking tests did not fully predict its excellent final pavement performance.

While individual tests like the IDT-CT and CF are valuable for screening, no single test perfectly predicted the final pavement cracking performance from the FlexPAVE™ simulations. The different performance rankings between the lab tests and the final simulation emphasize that a combination of tests is likely necessary to fully characterize a mixture's potential. However, based on these findings, when the RS9.5D mixes are not considered, mixtures evaluated as poor in IDT-CT showed Poor to Fair performance in FlexPAVE™, and mixes shown as Good from IDT-CT were generally Excellent to Moderate in the FlexPAVE™ simulations.

4. IMPLEMENTATION IMPLICATIONS

4.1. Overview

Implementing performance testing requires significant investment from both contractors and agencies. This includes the cost of testing equipment (new and maintenance on existing ones), dedicated laboratory space, and routine training for technicians to ensure tests are performed correctly and consistently. Simpler tests like IDT-CT and IDT-HT can often be performed using the same loading frames that labs already possess for TSR purposes, making an easier the path to implementation.

As presented in the literature review (Appendix A), AASHTO PP 105 (2024a) outlines a staged approach for BMD implementation. The process begins with the selection of relevant performance tests, establishing preliminary thresholds, and the choice of an overall implementation approach (A, B, C, or D) to integrate performance tests with existing volumetric requirements.

A second stage is a trial period known as "shadowing", in which contractors submit mix designs for approval under the existing volumetric system while also conducting performance tests for informational purposes only. This process allows both the agency and industry to become familiar with the test procedures and to build a database of performance results. Following this, the third stage is the full implementation of mix design, where a proposed design should meet both performance thresholds and any applicable volumetric criteria.

Once the process is established for mix design, the focus shifts to production. A similar "shadow" period for quality assurance (QA) is initiated to gather data on plant-produced material, which may differ from lab-produced mix design material. This stage is important to gather production variability data, refining lot sizer and QA limits. Finally, if desired, the full implementation in QA can link performance results to pay adjustment factors, providing financial incentives for high-performing materials and penalties for materials that fail to meet requirements.

For NCDOT, implementing any proposed scenarios would require a review and revision of its Quality Management System (QMS). Specifically, Section 610, which relates to Asphalt Concrete Plant Mix Pavements, would need to be updated. The primary change would be the addition of IDT-CT performance criteria into Table 610-3: Mix Design Criteria. A draft example illustrating how this revision could look is presented in Appendix F. More detailed discussions of the implementation issues for mix design and QA are presented below.

4.2. Mix Design Specifications

Based on the review of state practices given in Appendix A, the simplest pathway to BMD implementation is via Approach A, where volumetric analysis is completed as normal, and test verification is done after ensuring that existing volumetric requirements are met. This approach is also consistent with how the NCDOT has currently implemented APA testing, thus, there is an existing precedent. However, within this approach, there are different pathways for distributing the additional work. Thus, to discuss the implications of integrating the IDT-CT into the current mix design process, five potential implementation scenarios were considered. These scenarios are differentiated primarily by who takes responsibility for the testing (the contractor or NCDOT) and what type of material should be shipped (compacted specimens, loose-mix, or component material). A summary of the scenarios is given in Table 18 below. Each of the five scenarios can be contrasted against the current state of the practice, referred to as the "Do-Nothing Scenario". In this scenario, the contractor produces specimens for both TSR and APA, tests for TSR, and ships APA samples to NCDOT for testing. The agency checks TSR and volumetric values for compliance, issuing the final job mix formula.

Table 18. Summary of Responsibilities from Alternative Mix Design Scenarios.

Test	Responsibility[a]	Scenario					
		D/N[b]	1	2	3	4	5
TSR	Produce loose-mix	C	C	C	C	C	C
	Reheat	-	-	-	-	-	-
	Fabricate specimens	C	C	C	C	C	C
	Test	C	C	C	C	C	C
	Compliance Check	A	A	A	A	A	A
APA	Produce loose-mix	C	C	C	C	C	A
	Reheat	-	-	-	-	A	-
	Fabricate specimens	C	C	C	C	A	A
	Test	A	A	A	A	A	A
	Compliance Check	-	-	-	-	-	-
IDT-CT	Produce loose-mix	-	C	C	C	C	A
	Reheat	-	-	-	A	A	-
	Fabricate Specimens	-	C	C	A	A	A
	Test	-	C	A	A	A	A
	Compliance Check	-	A	-	-	-	-

[a] C indicates contractor's responsibility, A means that the NCDOT would be responsible, - means that the step is not required under the given scenario.
[b] D/N refers to the Do-Nothing scenario.

The first two scenarios build upon the current baseline by incorporating the IDT-CT test into the mix design process. In Scenario 1, the contractor would be responsible for all aspects of the IDT-CT, including mixing, specimen fabrication, and testing. This approach places the majority of the resource burden on the contractor and may result in issues with respect to laboratory space, which was an issue that was highlighted as a concern in the industry survey (Section 1.3.1). While some contractors may already possess loading frame machines capable of performing the test, additional resources would still be required for personnel training, annual accreditation, equipment calibration, and maintenance. The additional resources to the contractor are considered High, as preparing and testing IDT-CT specimens is expected to add a full day to the contractor's workload (Zhou *et al.,* 2020). For the NCDOT, the additional resources are relatively Low and would consist of updating specifications, managing new compliance issues, dealing with disputes, and approving labs and periodic validation of labs conducting IDT-CT testing.

Scenario 2, on the other hand, would result in a more balanced workload. This scenario follows the current APA pattern, where the contractor produces the IDT-CT specimens and ships them to NCDOT for testing. This approach better divides the responsibilities but still requires the contractor to have additional oven and lab space for specimen fabrication. In this scenario, the demands for additional contractor resources are considered Moderate-to-High, as it would produce samples for both APA and IDT-CT, and ship them to the NCDOT. The TSR process remains the same. As for the NCDOT side, the responsibilities include conducting the APA and IDT-CT tests. The resources that would be needed for those tasks would include the same as in Scenario 1, plus training personnel, obtaining annual accreditation (if/when IDT-CT becomes an accredited test), equipment calibration, equipment maintenance, and the need to manage receiving and cataloging additional specimen testing. However, since the NCDOT already has a system in place to manage specimen receipt and testing, and since the personnel time required to perform a suite of IDT-CT testing is relatively low, the additional resources required from the NCDOT are considered Moderate.

A third alternative (Scenario 3) is envisioned to address some of the contractor lab space concerns and give the NCDOT greater control over specimen aging in the IDT-CT test. In this alternative, the contractor handles TSR and APA the same as is currently done. For IDT-CT, the contractor mixes the component materials, but does not proceed to compact specimens. Instead, the contractor would send enough loose-mixture to the NCDOT for its personnel to age and compact IDT-CT specimens. In this case, the contractor's additional resources are considered Low, as they would only be required to produce and ship loose-mix material for the IDT-CT test. This scenario shifts the specimen fabrication workload to the NCDOT and introduces one technical concern regarding the effects of reheating the loose mix, which may require further research to potentially develop adjustment factors (Diefenderfer *et al.* 2021a). Lastly, the additional resources on the agency would be considered Moderate-to-High, as the expected one-day turnaround time for IDT-CT production and testing would be added to the existing mix design approval timeline.

Scenario 4 expands the previous concept by having the contractor ship loose-mix material for both APA and IDT-CT tests. Compared against the Do-Nothing Scenario, the contractor's responsibilities actually decrease since there is no need to compact APA specimens. Consequently, though the agency workload increases, as it must now fabricate and test both APA and IDT-CT specimens. The tradeoff is that the NCDOT gains a higher degree of control over testing conditions. The impacts to the agency are considered to be greater than in Scenario 3, but only marginally so because the work to fabricate APA specimens could be done sequentially, especially if the different aging times used for APA and IDT-CT were maintained. Thus, the resource impact from Scenario 4 is considered to be High-to-Moderate. To offset some of these impacts, it may be possible to split some of the required IDT-CT compaction and testing to regional labs since the testing equipment required is much less specialized and expensive than the APA.

The final option, Scenario 5, grants the agency nearly full control over the mixture fabrication process. The contractor would only be required to ship batched component materials, with the agency responsible for all mixing, specimen fabrication, and performance testing. While this minimizes the contractor's duties substantially, it presents the most significant logistical and resource challenges for the NCDOT. This scenario would potentially increase the need for lab space, material storage, and a substantial increase in the time required for each mix design approval. Thus, the impact to the NCDOT would be High in this case.

A summary of the added resources to each side and control implications for each implementation scenario is presented in Table 19. The final column, control over testing conditions, rates the level of agency oversight, considering whether the agency performs or just checks the test results, controls sample fabrication, and short-term aging.

Table 19. Summary of Time, Cost and Control from Alternative Mix Design Scenarios per Mix Design.

Scenario	Contractor Added Resources	NCDOT Added Resources	Control Over Testing Conditions
Do-nothing scenario	None	None	+
Scenario 1	High	Low	++
Scenario 2	Moderate-to-High	Moderate	+++
Scenario 3	Low	Moderate-to-High	++++
Scenario 4	Reduction	High-to-Moderate	+++++
Scenario 5	Reduction	High	+++++++

4.3. Impacts on Quality Assurance

A primary challenge for implementing performance testing in quality assurance is the significant increase in turnaround time. Based on the agency survey, NCDOT's current turnaround for assurance series is one to three days. While the IDT-CT test itself is rapid, the entire process of sample fabrication from loose mix, conditioning, and testing can take a full day. Similarly, the APA test can take up two days to complete (Zhou *et al.*, 2020). This timeline is longer than the few hours required for volumetric tests, making lot-by-lot performance testing impractical under current procedures. Because of this timeline, there are few examples of states fully implementing performance testing for routine quality assurance. California and New Jersey, however, provide valuable models on how to approach this issue, and their practices are discussed in detail in Appendix A under the section *Integration of Performance Test into QA Practices*.

Also relevant to QA is the fact that performance tests can exhibit higher variability than volumetric tests, increasing the risk for both the agency (accepting poor material) and the contractor (having good material rejected) (Wang *et al*. 2023). That variability comes from sample fabrication and/or allowed gradation and binder content variability during production. Further research is needed to establish precise thresholds, but solutions include adopting relaxed QA thresholds compared to a mix design or implementing pay adjustment factors.

Four implementation alternative scenarios are identified from the literature and ongoing research for integrating the IDT-CT into the Quality Assurance process. All of these scenarios assume that IDT-CT testing has been adopted for mix design. The first scenario is a "do-nothing" baseline condition that represents the current state of practice. In this scenario, the agency keeps the usual 750-ton sampling frequency, measuring only volumetric properties to check for compliance with quality control values.

QA-Scenario 1 proposes implementing the IDT-CT for initial mix approval, with testing waived during routine production. This approach would allow the NCDOT to re-verify the performance of a Job Mix Formula (JMF) that has been inactive for an extended period, while preventing additional testing burdens during production. For this approach to be effective, the agency would need to determine whether the performance threshold for the re-verification should be the value from the original JMF approval or the established statewide threshold of 14. It is likely that an adjustment would need to be made if initial mix approval were done on the basis of plant-mixed materials.

The next two alternatives rely on Performance-Related Specifications (PRS), which are discussed in detail elsewhere (West *et al*., 2023). The first of these scenarios (QA-Scenario 2) would utilize a performance prediction model. This approach involves first calibrating a function to relate the mixture performance indices (CT_{index} and APA rut depth) to established Asphalt Quality Characteristics (AQCs) like asphalt binder content, gradation, and volumetrics. Then, in a second step, these AQCs are measured during production and used to predict performance index values. The initially calibrated relationship would allow the agency to predict how the IDT-CT would perform during production by monitoring routine volumetric results, without needing to run the performance test. In this model, the variation in volumetrics provides oversight for the predicted IDT-CT performance.

The second PRS-based scenario, QA-Scenario 3, involves tracking existing AQCs and using these values along with pre-defined deviations to trigger periodic performance testing. For instance, a high binder content could trigger a rutting test (APA or IDT-HT), while a low binder content could trigger a durability-related test (IDT-CT). This strategy focuses performance testing on lots that are most at-risk. Because the testing would be less frequent, this approach allows for the practical implementation of performance tests at any level of sophistication. There is some evidence that such an approach could be practical, as statistical analysis performed in this study and conducted in FHWA/NC 2023-03 showed strong correlations between gradation parameters and/or binder content and performance test results. While QA-Scenario 3 would likely be easier to implement, an upfront study would be needed to establish the AQC deviations that would be appropriate to trigger testing. Such a study would not be needed with QA-Scenario 2 since the relationship between AQCs and performance indices would be determined on a mixture-by-mixture basis.

QA-Scenario 4 would incorporate performance testing into routine quality assurance either in combination with existing AQC determination (partial performance-based specification) or in replacement of those AQCs (i.e., pure performance-based specification). This approach presents the most significant logistical and thus implementation challenge for North Carolina DOT, but could potentially have the highest payoff. To make this approach practical, strategies from other states could be adopted, such as redefining sampling rates for larger lot sizes. Such adjustments would involve making careful consideration of balance between risk for the contractor and the DOT and the resulting impacts on price. The payoff for QA-Scenario 4 could be greater freedom to the contractor, few potential points of contractual conflict (since several AQCs would be eliminated), and better overall certainty in the expected performance of asphalt materials.

4.4. Efficiency Opportunities

In view of making the process more practical, there are opportunities that could create efficiencies and or flexibilities in the process that could aid implementation. First, the implementation of a BMD approach where volumetric requirements are relaxed when test results show acceptable behaviors could provide contractors with more flexibility and develop buy-in. Testing-wise, the correlations shown in Figure 10 suggest that IDT-HT could be used as a screening tool for rutting in place of the more time-consuming APA test. This path may not mean that the APA is completely eliminated as it could still be used for verification or for cases showing marginal performance in the IDT-HT. This step would require establishing formal IDT-HT thresholds. It is noted that at least one other study has found a strong correlation between both tests (Boz *et al.,* 2025). A practical consideration is that the IDT-HT test requires an active operator, whereas the APA test can be run unattended. Also related to testing, the IDT-CT could potentially serve as a modified-dry-strength component of the TSR test, which could significantly reduce the time and sample preparation required for moisture sensitivity testing. Verification and validation studies would be required. In both of these cases, IDT-CT or IDT-HT implementation, more of the testing could possibly be shifted to regional labs. This approach would distribute the workload from the central lab and could be efficient, as regional labs likely already possess the necessary equipment for such testing.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

Based on the findings of this study, the following conclusions were reached:

- For performance testing, simpler monotonic tests, like IDT-CT, were favored by many states for their practicality and lower cost, but they are not suitable for pavement performance prediction.
- There existed a preference across agencies for using HWTT for rutting and IDT-CT for durability as performance testing.
- Contractors in North Carolina are receptive to the concept of BMD if it results in greater design freedom. There exist practical concerns regarding the cost, lab space, and training required for implementation.
- In North Carolina, APA testing was indicated to be the primary time bottleneck in the mix design approval process.
- The proposed protocol for aging mixtures at the compaction temperature for four hours prior to IDT-CT for RS9.5B and RS9.5C mixtures produced CT_{index} values lower than those observed from plant mix. Conversely, the 2-hour aging protocol prior to APA testing of laboratory-produced mixtures resulted in similar rut depths as plant-produced mixes.
- The standard short-term aging protocol (4 hours at compaction temperature) was too severe for polymer-modified mixes and resulted in artificially low cracking resistance values. A less harsh condition (4 hours at 143°C plus time to reach compaction temperature) yielded more reasonable and discriminating results, leading to its adoption for all subsequent RS9.5D mixture testing in the project. It is noted that both the 4 hours at compaction temperature and 4 hours at 143°C approach appear to create harsher aging conditions than plant production.
- The average and 15th percentile CT_{index} values for RS9.5B and RS9.5C mixtures in North Carolina were estimated to be 18.5 and 14.0, respectively. The RS9.5D mixtures showed an average CT_{index} of 11.2.
- North Carolina surface mixtures easily pass their respective APA rutting requirements. The average APA rut depth for RS9.5B, RS9.5C, and RS9.5D was 4.1 mm, 3.1 mm, and 1.0 mm, respectively.
- The APA test differentiated the performance of the RS9.5D mixtures from the other groups, but did not differentiate between the RS9.5B and RS9.5C mixtures. Rut depth was also found to correlate with effective asphalt content.
- The RS9.5B and RS9.5C mixtures had similar dynamic moduli across the temperatures and frequencies evaluated, however, the RS9.5D mixtures had statistically higher modulus at intermediate and high temperatures.
- No statistical difference in CF results was identified among the three mix types (RS9.5B, RS9.5C, and RS9.5D).
- The influence of the source region was not statistically significant for the IDT-CT and APA tests. However, for the DM and CF tests, the source region was found to be a statistically significant factor, though the specific regional differences were not consistent across all temperatures and metrics.

- The IDT-HT test correlated well with the APA ($R^2 = 0.99$) and the SSR test results ($R^2 = 0.95$), while also providing superior statistical differentiation between mixtures. The power-law relationship does suggest diminishing differentiability from IDT-HT at lower strength values.
- Pavement rutting performance simulations correlated strongly with all rutting-related laboratory test results (APA, SSR, IDT-HT). The YB mixture consistently ranked as the worst performer with the highest predicted rut depths, while the UC and VD mixtures were among the best performers, confirming the reliability of the lab tests to evaluate for rutting resistance.
- The correlation between laboratory tests and simulated damage performance was complex, and no single lab test perfectly predicted the final pavement performance rankings from FlexPAVE™.
- The IDT-CT test accurately separated the poor performing RS9.5B and RS9.5C mixes, but could not differentiate between good and moderately performing mixes based on the relationship to FlexPAVE™ predicted performance.
- The effect of asphalt binder source was either limited (differentiating only two of five sources in the IDT-CT), inconsistent (in the DM results), or not statistically significant (in the CF results). Thus, within the study limitations, the current practice of interchanging binder sources may not significantly impact asphalt mixture cracking or stiffness performance.

5.2. Recommendations

Based on the conclusions stated in the previous sub-section, this research makes the following recommendations:

- If the NCDOT chooses to adopt IDT-CT as part of its mix design criteria, it is recommended that the testing limit be set at 14 for RS9.5B and RS9.5C mix types, and aging be done for 4 hours at the compaction temperature prior to compaction. It is also recommended that the NCDOT invest resources in evaluating the efficacy of this value using data from field projects.
- This study evaluated mixtures at a fixed aging condition (4 hours at the compaction temperature), but also found that the CT_{index} values obtained after this condition were lower than those from PMLC specimens (based on two mixtures). Thus, if IDT-CT is implemented in QA processes, additional resources should be extended to more comprehensively evaluate the differences between lab-mixed and lab-compacted and PMLC specimens, as well as the impact of reheating PMLC specimens.
- Where applicable, preliminary IDT-HT strength criteria of approximately 150, 180, and 215 kPa for RS9.5B, RS9.5C, and RS9.5D mixtures can be used, but more work should be done to validate and verify these values.
- Investigate the feasibility of a modified TSR procedure where the IDT-CT serves as the dry-strength component, potentially reducing the time and sample preparation required for moisture sensitivity testing.
- Evaluate conducting the simpler performance tests (IDT-CT and IDT-HT) at the regional labs instead of only the central lab. This approach would distribute the workload and improve efficiency, as regional labs likely already possess the necessary loading frames for TSR testing.

- Conduct pilot “shadow” mix designs to evaluate the feasibility of implementing the IDT-CT and APA for mix design. During this period, performance tests would be run in parallel with existing volumetric measurements for informational purposes only. For this purpose, it is recommended that the shadow designs be performed using primarily Scenario 3 (Table 19) as this balances the resource burden on contractors while providing the agency with a high degree of control over the final performance test conditions. However, consideration should also be given to trying the other scenarios as well.
- Develop a practical training guide and certification program for technicians on the proper procedure for conducting the IDT-CT.
- Calibrate the proposed hybrid "trigger" QA method by evaluating how usual volumetric variability affects performance results. This would identify the key volumetric indicators (e.g., binder content, VMA) that best predict changes in IDT-CT and APA performance and establish the specific deviation thresholds that would trigger a performance test.

6. Implementation and Technology Transfer Plan

The Materials and Tests Unit of the NCDOT will be the primary users of the outcomes of this project. The experimental data and preliminary threshold values that this research has produced can be used by the NCDOT to improve their asphalt mix design process and may be used to improve quality assurance practices. This could lead to improved asphalt pavement performance and lower long-term costs for North Carolina roadways. The findings of this research will be communicated to the NCDOT in the form of this report, the appendices and supplementary information, and a closeout meeting with the project panel.

For follow-up activities, the research team believes that the NCDOT could consider allocating resources for the following activities:

- Expand the number of lab-mixed, lab-compacted mixtures tested to get better coverage of contractors and regions;
- Evaluate more plant-mixed, lab compacted mixtures to better understand the effects of plant production and reheating on the results;
- Develop training resources for NCDOT personnel on performing IDT-CT testing; and
- Evaluate the viability of using IDT-HT as a more efficient testing method than APA, and evaluate the viability of using the testing currently used for determining the tensile strength ratio for IDT-CT purposes.

7. References

AASHTO R *30-19: Standard Method of Test for Mixture Conditioning of Hot Mix Asphalt (HMA)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2021.

AASHTO *T 283-22: Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage*. American Association of State Highway and Transportation Officials, Washington, D.C., 2022a.

AASHTO M 320-23: Standard Specification for Performance-Graded Asphalt Binder. American Association of State Highway and Transportation Officials, Washington, D.C., 2023a.

AASHTO M 332-23: Standard Specification for Performance-Graded Asphalt Binder Using Multiple Stress Creep Recovery (MSCR) Test. American Association of State Highway and Transportation Officials, Washington, D.C., 2023b.

AASHTO T 240-23: Standard Method of Test for Effect of Heat and Air on a Moving Film of Asphalt Binder (Rolling Thin-Film Oven Test). American Association of State Highway and Transportation Officials, Washington, D.C., 2023c.

AASHTO *T 324-23: Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023d.

AASHTO *T 340-23: Standard Method of Test for Determining Rutting Susceptibility of Hot Mix Asphalt (HMA) Using the Asphalt Pavement Analyzer (APA)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023e.

AASHTO T 350-19: Standard Method of Test for Multiple Stress Creep Recovery (MSCR) Test of Asphalt Binder Using a Dynamic Shear Rheometer (DSR). American Association of State Highway and Transportation Officials, Washington, D.C., 2023f.

AASHTO *TP 132-23: Determining the Dynamic Modulus for Asphalt Mixtures Using Small Specimens in the Asphalt Mixture Performance Tester (AMPT)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023g.

AASHTO *TP 134-22: Stress Sweep Rutting (SSR) Test Using Asphalt Mixture Performance Tester (AMPT)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023h.

AASHTO *PP 105-24: Standard Practice for Balanced Design of Asphalt Mixtures*. American Association of State Highway and Transportation Officials, Washington, D.C., 2024a.

AASHTO T 315-24: Standard Method of Test for Determining the Rheological Properties of Asphalt Binder Using a Dynamic Shear Rheometer (DSR). American Association of State Highway and Transportation Officials, Washington, D.C., 2024b.

AASHTO *T 411-24: Determining the Damage Characteristic Curve and Failure Criterion Using Small Specimens in the Asphalt Mixture Performance Tester (AMPT) Cyclic Fatigue Test*. American Association of State Highway and Transportation Officials, Washington, D.C., 2024c.

AASHTO *T 166-24: Bulk Specific Gravity (Gmb) of Compacted Asphalt Mixtures Using Saturated Surface-Dry Specimens.* American Association of State Highway and Transportation Officials, Washington, D.C., 2024d.

Asphalt Institute *Mix Design Methods for Asphalt*. 6th ed. MS-02, Asphalt Institute, Lexington, KY, 1997.

Bowers, B.F. and Diefenderfer, S.D. *Technical Memorandum: Benchmarking of Asphalt Mixtures - Summary of Findings and Preliminary Recommendations*. Virginia Transportation Research Council, Charlottesville, VA, 2018.

Boz, I., Habbouche, J., Diefenderfer, S., and Bilgic, Y. Precision Estimates and Statements for Performance Indices from the Indirect Tensile Cracking Test at Intermediate Temperature. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2022, pp. 225-241. 340–351.

Boz, I., Habbouche, J., Diefenderfer, S.D., Coffey, G., Seitllari, A. and Ozbulut, O.E. Evaluating the Rutting Potential of Asphalt Mixtures with Simple and Practical Tests. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2679, No. 1, 2025, pp. 1006-1021.

Castorena, C. and Costa, R. *Literature Review for NCDOT RP 2024-14: Recycled Binder Availability Considerations for Mixing and Recycled Binder Replacement Percentage Specifications*. North Carolina Department of Transportation. 2024.

Cominsky, R.J., Huber, G.A., Kennedy, T.W., and Anderson, M. *The Superpave Mix Design Manual for New Construction and Overlays*. SHRP-A-407, Strategic Highway Research Program, Washington, D.C., 1994.

Diefenderfer, S.D., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: Phase I: Initial Roadmap Development and Specification Verification*. FHWA/VTRC 21-R15, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2021a.

Habbouche, J., Boz, I., and Diefenderfer, S.D. *Round Robin Testing Program for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase I*. FHWA/VTRC 23-R3, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2021a.

Jeong, J., Underwood, B.S., and Kim, Y.R. Cracking Performance Predictions Using Index-Volumetrics Relationships with Direct Tension Cyclic Fatigue Test and Illinois Flexibility Index Test (I-FIT). *Construction and Building Materials*, Vol. 315, 2022.

Kim, M., Mohammad, L.N., and Elseifi, M.A. Characterization of Fracture Properties of Asphalt Mixtures as Measured by Semicircular Bend Test and Indirect Tension Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2296, 2012, pp. 115–124.

McDaniel, R.S., Leahy, R.B., Huber, G.A., Moulthrop, J.S., and Ferragut, T. *The Superpave Mix Design System: Anatomy of a Research Program*. NCHRP Web-Only Document 186, National Cooperative Highway Research Program, Transportation Research Board, 2012.

NAPA. *Development of Marshall Procedures for Designing Asphalt Paving Mixtures*. Information Series 84, National Asphalt Pavement Association, Lanham, MD, 1982.

NCAT. *Quality Assurance (QA) Aspects of Performance Related Specifications (PRS)*. TFRS-01 Interim Report, National Center for Asphalt Technology, Auburn, AL, 2021.

NCHRP. Project Announcement for Guide for Implementing Performance Specifications [online]. *National Cooperative Highway Research Program*. Available from: https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4751 [Accessed 4 September 2022].

Ozer, H., Al-Qadi, I.L., Singhvi, P., Bausano, J., Carvalho, R., Li, X., and Gibson, N. Prediction of Pavement Fatigue Cracking at an Accelerated Testing Section Using Asphalt Mixture Performance Tests. *International Journal of Pavement Engineering*, Vol. 19, No. 3, 2018, pp. 264–278.

Saleh, N.F. *Quality Assurance Aspects of Performance-Related Specifications and Oxidative Aging of Asphalt Pavements*, Ph.D. Dissertation, North Carolina State University, Raleigh, NC, 2022.

Transportation Research Board (TRB). *Glossary of Transportation Construction Quality Assurance Terms: Seventh Edition*. Transportation Research Circular E-C235, Transportation Research Board, Washington, D.C., 2018.

Underwood, B.S., Castorena, C., Matini, N., Isied, M., Goenaga, B., and Kuchiishi, K. *Calibration of Structural Layer Coefficients for North Carolina Asphalt Pavements*. Final Report FHWA/NC/2019-20. North Carolina Department of Transportation. Raleigh, NC, 2021.

Wang, Y. D., Liu, J., and Liu, J. Integrating Quality Assurance in Balance Mix designs for Durable Asphalt Mixtures: State-of-the-Art Literature Review. *Journal of Transportation Engineering, Part B: Pavements*, Vol. 149, No. 1, 2023, 03122004.

West, R.C, Timm, D., Powell, B., Tran, N., Yin, F., Bowers, B., Rodenzo, C., Leiva. F., Vargas, A., Gu, F., Moraes, R., Nakhaei, M. *Phase VII (2018-2021) NCAT Test Track Findings.* NCAT Report 21-03, National Center for Asphalt Technology, Auburn, AL, 2021a.

West, R., Kim, Y. R., Saleh, N., Musselman, J., DeCarlo, K., Tran, N., Underwood, S., Yin, F., Pine, B., and Huber, G. *TFRS 01: Quality Assurance (QA) Aspects of Performance Related Specifications (PRS)*, Transportation Research Board, Washington, D.C. 2023.

Witczak, M.W., Kaloush, K., Pellinen, T., El-Basyouny, M., and Von-Quintus, H.L. *Simple Performance Test for Superpave Mix Design*. NCHRP Report 465, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2002.

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

Zhou, F., Im, S., Sun, L., and Scullion, T. Development of an IDEAL Cracking Test for Asphalt Mix Design and QC/QA. *Road Materials and Pavement Design*, 2017b.

APPENDIX A. DETAILED LITERATURE REVIEW

Introduction

Asphalt mixture is an engineered material that is subjected to different traffic and climatic conditions throughout their service life. The different load scenarios due to traffic, climate, or both simultaneously, in combination with changes in the material's properties with time, make asphalt mixtures susceptible to distresses such as fatigue cracking, thermal cracking, rutting, raveling, and others, affecting the functional and structural performance of the asphalt pavement. Further, different methodologies have been developed in the past to design asphalt mixtures that yield satisfactory pavement performance, such as the Hveem, Marshall, and Superpave methods (NAPA 1982, Asphalt Institute 1997). In general, these mix design methodologies seek sufficient asphalt binder to be added to the mixture to ensure durability without compromising the stability of the mix to withstand traffic loads.

The Marshall mix design is the most widely used mix design method worldwide because of its simplicity and relatively low cost. The detailed procedures vary from state to state but typically the methodology uses volumetric criteria and results from the Marshall stability test to select the optimum asphalt content. The main disadvantages associated with this method are related to the compaction method and the lack of relationship between the Marshall stability and pavement performance. Similar principles are followed with the Hveem mix design method, but the compaction method, measures of performance, optimum asphalt content criteria, and other considerations differ. More details can be found in Asphalt Institute (1997).

The Superpave mix design method was developed during the Strategic Highway Research Program (SHRP) as a response to the awareness that pavement materials were not providing satisfactory performance (McDaniel *et al.* 2012). Considering the shortcomings of the empirically based Marshall and Hveem test methods, Contract A-003A *Performance-Related Testing and Measuring of Asphalt-Aggregate Interaction and Mixtures* of SHPR was aimed at developing reliable and reproducible test methods that could be used to characterize asphalt mixtures in terms of fundamental engineering properties. The SHRP efforts resulted in a hierarchical mix design system intended to encompass material characterization, volumetric mix design, and performance testing analysis depending on the design traffic level (Cominsky *et al.* 1994). Level 1 mix design consisted of a volumetric design with strict attention to the selection of the asphalt binder and aggregates. Levels 2 and 3 also considered volumetrics in addition to performance tests and predictions that would allow for the mix to be optimized with respect to one more distress. However, Levels 2 and 3 were, by and large, never implemented presumably due to complicated test protocols, equipment costs, and the general belief that volumetrics and binder performance could ensure adequate pavement performance (McDaniel *et al.* 2012, Diefenderfer and Bowers 2019).

In the early 2000s, evidence of non-adequate pavement performance, the incorporation of recycled asphalt materials into asphalt mixtures, and other factors renewed the interest in integrating performance tests in mix design protocols (also known as performance mix design) and multiple research efforts have been conducted to allow for a successful implementation. In NCHRP Project 09-19 Witczak *et al.* (2002), simplified performance test (SPT) methods addressing permanent deformation, fatigue cracking, and low-temperature cracking were proposed as a final stage in the

Superpave volumetric mix design; however, the test protocols and attention to specific distress have varied across state agencies throughout the recent years.

In the early implementation of the Superpave mix design, more attention was given to improving rutting resistance and many state agencies incorporated rutting test requirements The adjustments made to the mix designs to meet rutting specifications affected the durability of the mix. Further, many agencies have indicated that cracking and raveling have become the main distresses controlling the service lives of asphalt pavements (West *et al*. 2018a). This observation has driven national attention to seek new approaches that balance multiple modes of distress. This 'new' approach is often called balance mix design (BMD) and has been defined by the FHWA BMD Task Force as:

> "*Asphalt mix design using performance tests on appropriately conditioned specimens that address multiple modes of distresses taking into consideration mix aging, traffic, climate and location within the pavement structure*"

Several test methods of varying complexities, analysis procedures, and costs have been developed in recent years to address rutting and fatigue cracking. This document presents a comprehensive review of the different test methods that have been proposed for BMD implementation, integration of performance tests into QA practices, expected equipment costs for a testing laboratory, as well as current state efforts for integrating performance tests into mix design.

BMD Performance Tests

Test methods for BMD generally fall into one of two categories; 1) tests used to evaluate the durability and 2) tests used to evaluate the stability of an asphalt mixture. For durability tests, the focus is primarily on the ability of the mixture to resist fatigue cracking and for stability tests the focus in on rutting resistance. The tests vary according to the precise method of test, analysis procedures, historical background and development, and cases where the method has been used. The following sections describe the tests according to these four characteristics.

Durability Test Methods

The durability of asphalt mixtures has been a concern since the development of early mix design methodologies. Hveem (1943) defined durability as the ability of asphalt materials to retain their original properties. A similar definition was given later by Finn (1967), who defined durability as the long-term resistance to the effects of aging. In general, the durability of asphalt mixtures has been historically associated with the material's resistance to cracking and raveling due to traffic and environmental loads.

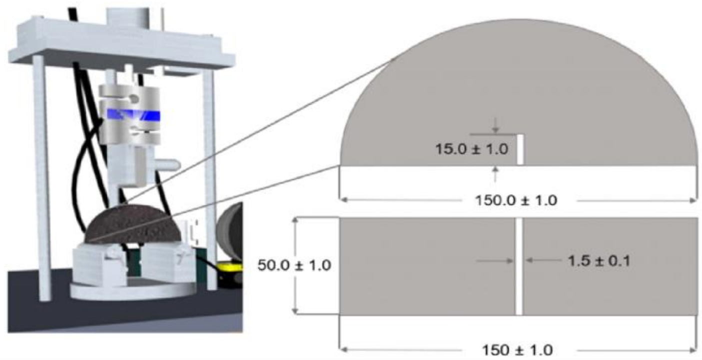
Performance test methods with different levels of complexity have been developed and proposed to characterize the cracking resistance of asphalt mixtures. The test methods can be categorized into two main categories depending on the nature of the load application; monotonic tests and repeated load tests. Monotonic tests are test methods that use a constant rate of loading and are generally based on fracture mechanics principles. Performance test methods such as the IDT-CT (also referred to as just IDT-CT), I-FIT, SCB-LTRC, and others fall in this category. Repeated load tests, including the Uniaxial Cyclic Fatigue Test, Bending Beam Fatigue Test, Overlay Test, and others, are cyclic tests that attempt to simulate the process of fatigue damage accumulation due to repeated traffic.

In the next sections, different durability test methods proposed for BMD implementation are reviewed. The review of each test method includes the method of calculation and description, the historical development, and a summary of studies using the test method. The summary of studies using each test method is mainly focused on the sensitivity of the test to asphalt mixture component properties or proportions (asphalt binder content, additives, recycled materials), air voids and aging, as suggested by FHWA-A-HIF-19-103 (Hajj *et al.* 2019).

I-FIT

Method and Calculation Description

The Illinois Flexibility Index Test (I-FIT) is a monotonic performance test used to determine the cracking resistance properties of asphalt mixtures based on Mode 1 fracture mechanics principles. The test is conducted on a 50-mm thick semi-circular bending (SCB) specimen with a 15-mm deep and 1.5-mm wide notch using a loading rate of 50 mm/min at a temperature of 25°C (AASHTO TP 124). The I-FIT setup and specimen geometry are shown in Figure A.1.



(a) (b)

Figure A.1. Overview of I-FIT test; (a) instrumentation and (b) specimen geometry. (Rivera-Pérez *et al.* 2021).

One of the material parameters determined by I-FIT is the fracture energy, G_f, which represents the energy dissipated through crack propagation. This parameter is calculated as the area under the load-displacement curve normalized by the ligament area (the product of the ligament length and thickness of the specimen) and is a function of the material's strength (defined by the peak load P_{max}) and ductility (defined by the maximum displacement) (Al-Qadi *et al.* 2015, Batioja-Alvarez *et al.* 2019). The fracture of energy is used to define multiple performance indicators such as the Flexibility Index (*FI*), Cracking Resistance Index (*CRI*), and the Balanced Cracking Index (*BCI*) as summarized in Table A.1. The development of the test and motivations for each cracking index will be discussed in the subsequent section.

Table A.1. Summary of Performance Indicators from I-FIT.

Index	Equation[a]	Significance	Reference
Flexibility Index, *FI*	$FI = \left(\frac{G_f}{\lvert m \rvert} A \right)$	A higher *FI* corresponds to higher flexibility and less brittle behavior.	(Al-Qadi *et al.* 2015)
Cracking Resistance Index, *CRI*	$CRI = \frac{G_f}{P_{\max}}$	A higher CRI indicates a better asphalt mixture in terms of cracking resistance.	(Kaseer *et al.* 2018)
Balanced Cracking Index, *BCI*	$BCI = \frac{G_f}{100} L_{75}$	A higher *BCI* indicates a higher cracking resistance.	(Majidifard *et al.* 2021)

[a] m = post-peak slope; A = Calibration coefficient for unit conversion and field aging shift. Defined as 0.01 for the mixtures evaluated in the ICT R27-128 project; and L_{75} = Displacement corresponding to the 75% of P_{max}.

Historical Development

The I-FIT protocol was developed under the ICT R27-128 project to evaluate the cracking resistance of asphalt mixtures with high contents of recycled materials (Al-Qadi *et al.* 2015, Ozer *et al.* 2016). Originally, the main output from the test was the *FI* of the asphalt mixture. The *FI* was developed to describe the fundamental fracture process and overall pattern of load-displacement curves and its underlying theory is inspired by the rate of crack growth definition by Bazant and Pratt (1988) for concrete materials (Al-Qadi *et al.* 2015). Three different versions of the *FI* were explored in the study and the selection of the final form was done based on the empirical correlations between the index and the approximate crack velocity. It considered the condition that as the asphalt mixture becomes more brittle (and more prone to cracking) the speed of crack propagation increases. The final form of the *FI* is calculated using the fracture energy normalized by the slope at the inflection point or post-peak, as shown in Table 1. In the earlier works, the *FI* was able to discriminate the cracking performance of lab- and plant-produced asphalt mixtures and field cores with varied mix design variables e.g., binder grade, RBR, etc. (Ozer *et al.* 2016). Further studies by Ling *et al.* using mix designs commonly used in Wisconsin showed that the *FI* was effective in discriminating between mix design factors and laboratory aging (Ling *et al.* 2017). Similar results were also found by Nemati *et al.* (2020).

A subsequent study by Kaseer *et al.* highlighted that the *FI* has limitations that can complicate its calculation and further implementation by highway agencies. The main issues are; *i*) the difficulty of determining the inflection point in some test results, *ii*) the variability associated with the test, and *iii*) the characterization of brittle mixtures (Kaseer *et al.* 2018). Consequently, the *CRI* was proposed as a surrogate to provide a greater distinction between asphalt mixtures with less variability and easier calculation. As shown in Table 1, the *CRI* is defined by the fracture energy normalized by the peak load. Majidifard *et al.* argued that the use of *CRI* is problematic due to the inherent difficulty in characterizing the cracking resistance in polymer-modified mixtures (or others) that have high peak loads (Majidifard *et al.* 2021). In their work, the *BCI* was proposed and calculated as the product between the fracture energy and the load-line displacement corresponding to the time when 75% of the peak load in the post-peak region is reached.

Brief Summary of Studies using the Test Method

The I-FIT has received considerable attention due to its capacity to rank asphalt mixture cracking performance. The topics studied in recent years include the effect of air voids and binder properties (Kim *et al.* 2012, Ling *et al.* 2017, Batioja-Alvarez *et al.* 2019, Ali *et al.* 2020, Nemati *et al.* 2020, Ishaq and Giustozzi 2021, Jeong *et al.* 2022), the effect of specimen geometry and test configuration (Rivera-Pérez *et al.* 2021), test thresholds (Ali *et al.* 2020), and oven-aging procedures (Lemke *et al.* 2018, Al-Qadi *et al.* 2019). Table A.2 presents a summary of the effects of different factors on the I-FIT-derived indices. In general, the literature suggests that the expected trends of different factors on the cracking indices, except for the effect of air voids, which are discussed below.

Batioja-Alvarez *et al.* investigated the effect of air voids on the cracking resistance of plant-mixed, field-compacted specimens using the *FI* (Batioja-Alvarez *et al.* 2019). In general, a positive correlation was found between air voids and *FI,* i.e., higher air void content produces higher *FI* values. In subsequent studies, Jeong *et al.* found a similar trend using the *FI, CRI,* and *BCI* and plant-mixed, lab-compacted mixtures (Jeong *et al.* 2022). The literature indicates that the indices from I-FIT suggest a counter-intuitive trend of cracking performance with air voids. Majidifard *et al.* examined a particular asphalt mixture where this trend was observed and found that the increase of air voids increased the fracture energy, L_{75}, and decreased the peak load and post-peak load (Majidifard *et al.* 2021). This inconsistent trend led to the development of equations to correct the sensitivity of *FI* with air voids. Barry proposed an equation based on the normalization of the fracture energy and post-peak slope using 7 percent as the reference air void content (Barry 2016).

Table A.2. Summary of Findings using the I-FIT Method.

Topic	Findings	Reference
Air voids	*FI, CRI,* and *BCI* increase with the increase in air void content.	(Kaseer *et al.* 2018, Batioja-Alvarez *et al.* 2019, Majidifard *et al.* 2021, Rivera-Pérez *et al.* 2021, Jeong *et al.* 2022)
Binder content	*FI, CRI,* and *BCI* increase with the increase in binder content.	(Ling *et al.* 2017, Kaseer *et al.* 2018, Majidifard *et al.* 2021)
Softer binder	*FI, CRI,* and *BCI* increase with the use of a softer binder.	(Kaseer *et al.* 2018, Ali *et al.* 2020, Majidifard *et al.* 2021)
Crumb rubber modificatio n	*FI, CRI,* and *BCI* decrease when crumb rubber is incorporated into the asphalt mixture.	(Majidifard *et al.* 2021, Rath *et al.* 2022)
Oxidative aging	*FI* and *CRI* decrease with oxidative aging.	(Kaseer *et al.* 2018, Ling and Buchanan 2022)
Recycled materials	*FI* and *CRI* decrease with the incorporation of recycled materials into the asphalt mixture.	(Bahia *et al.* 2016, Fakhri and Ahmadi 2017, Kaseer *et al.* 2018, Zaumanis *et al.* 2019)
Loading rate	*FI* shows a dependency on the loading rate.	(Rivera-Pérez *et al.* 2021)

The adjustment on the individual parameters yielded a non-linear relationship between *FI* and air voids (*AV*), as shown in Equation (2). Similarly, Kaseer *et al.* proposed Equation (3) to adjust *FI* using 7 percent as the reference air void content (Kaseer *et al.* 2018).

$$FI_{AV-Corrrected} = FI \times \frac{0.0651}{AV - AV^2} \quad (2)$$

$$FI_{AV-Corrrected} = FI \times \frac{7.0}{AV} \tag{3}$$

Ali *et al.* studied the effect of different binder PG grades on the *FI* of specimens with 4.0% air voids. It was found that the *FI* decreased with the increased high-temperature PG of the binder, except for the mixture with the PG 76-22 binder where a higher *FI* was obtained compared to the mixture with the PG 70-22 binder (Ali *et al.* 2022). Similarly, Ishaq and Giustozzi conducted a study to investigate the effect of binder properties on the *FI* calculated using specimens with 5.0% air voids (Ishaq and Giustozzi 2021). The study included a neat binder and binders modified with amino-wax additive, plastomers, plastomers-fibers, elastomers, and elastomers-fibers. It was found that the mixture containing only elastomers exhibited the highest *FI,* followed by the mixture with elastomers and fibers, amino-wax additive, neat binder, plastomers, and plastomer-fibers.

SCB-LTRC

<u>Method and Calculation Description</u>
The SCB test method developed by the Louisiana Transportation Research Center (LTRC), referred as SCB-LTRC, is a fracture mechanics-based test used to evaluate the intermediate temperature cracking resistance. The test procedure is described in ASTM D8044. The test is conducted at the climatic intermediate temperature using a loading rate of 0.5 mm/min on 57-mm thick semi-circular notched specimens. This protocol includes the evaluation of three different notch depths (3-mm width), 25 mm, 32 mm, and 28 mm, to determine the critical strain energy release rate as per Equation (4).

$$J_c = -\left(\frac{1}{b}\right)\frac{dU}{da} \tag{4}$$

Where J_c is the critical strain energy release rate, *a* is the notch depth, *b* is the sample thickness, *U* is the strain energy to failure and *dU/da* is the change of strain energy with notch depth. The latter is calculated by fitting a regression line that represents the relationship between strain energy and notch depth. Theoretically, only two notch depths need to be tested to calculate *dU/da*; however, three notch depths are specified to improve the accuracy of the calculation and account for measurement and theoretical uncertainty. A higher J_c corresponds to generally better cracking resistance.

<u>Historical Development</u>
The SCB-LTRC was developed by the Louisiana Transportation Research Center, however, the underlying principles of the SCB-LTRC test originated from the work of Rice (1968) who developed the *J*-integral concept to evaluate fracture characteristics of elastoplastic materials. Mull *et al.* utilized the concept of J_c to evaluate the fracture resistance of crumb rubber asphalt mixtures (Mull *et al.* 2002). In their study, the SCB configuration from gyratory-compacted cylindrical specimens was employed instead of the three-point bend beam configuration that was previously used by others to characterize the J_c of asphalt mixtures.

The relative ease of sample preparation and testing procedure motivated researchers at the LTRC to conduct in-depth evaluations of the test method. These efforts ultimately resulted in the incorporation of the SCB test into Louisiana's BMD approach. Some of those research efforts will be briefly summarized in the subsequent section.

Brief Summary of Studies using the Test Method

As mentioned earlier, most of the research studies conducted using the SCB-LTRC method were conducted by the LTRC in their effort to address the lack of a rapid, simple, and practical cracking performance test. However, studies outside Louisiana's research groups have also been realized to validate the test protocol. Table A.3 presents a summary of findings using the SCB-LTRC test method.

Table A.3. Summary of Findings using the SCB-LTRC Method.

Topic	**Findings**	**Reference**
Oxidative aging	Oxidative aging induces an adverse effect on cracking resistance using J_c.	(Kim *et al.* 2012, Cooper *et al.* 2015)
Recycled materials	J_c increases with the addition of RAP in WMA mixtures; however, the data showed high variability, thus complicating the interpretation.	(Zaremotekhases *et al.* 2022)
	The mixture containing RAP/RAS exhibited a higher J_c than the reference virgin mixture.	(Zhou *et al*. 2017a)
	The incorporation of 5% RAS did not substantially change the cracking resistance of the mix compared to the reference condition.	(Cooper *et al.* 2015)
	Asphalt mixtures containing no recycled asphalt pavement exhibited improved J_c.	(Cooper *et al.* 2014)
	Mixed trends are found concerning the effect of RAP on J_c, the results seem to be mix dependent.	(Kim *et al.* 2012)
Asphalt content	J_c increases with the increase of asphalt binder content.	(Zhou *et al*. 2017a)
Asphalt binder/ mixture mod.	J_c decreases with the use of WMA additives in asphalt mixtures.	(Kim *et al.* 2012, Zaremotekhases *et al.* 2022)
	J_c decreases with increasing REOB content.	(You *et al.* 2018)
	J_c is higher in mixtures containing SBS-modified binder than in mixes without SBS modification.	(Kim *et al*. 2012, Cooper *et al*. 2015, Zhou *et al*. 2017a)
	An engineered rejuvenating additive improved the cracking resistance of an asphalt mixture with 40 percent RAP.	(Elseifi *et al.* 2012)

IDT-CT

Method and Calculation Description

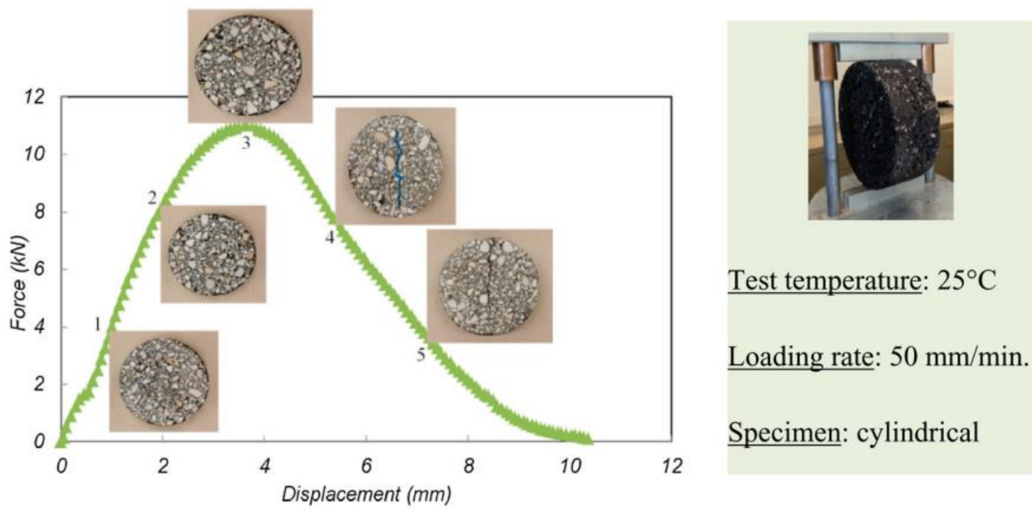
The IDT-CT is a monotonic performance test to try and estimate the cracking resistance of asphalt mixtures without cutting or notching the test specimens. The test is conducted at an intermediate temperature using a loading rate of 50 mm/min. Test specimens for mixtures with a NMAS of 19 mm or small are 62-mm thick and for mixtures with a NMAS larger than 19 mm, the specimens are 95 mm thick (ASTM D8225). The cracking tolerance index, or CT_{index}, is calculated using parameters based on the load-displacement curve and specimen dimensions, as shown by Equation (5).

$$CT_{Index} = \frac{t}{62} \times \frac{l_{75}}{D} \times \frac{G_f}{|m_{75}|} \times 10^6 \quad (5)$$

Where t is the specimen thickness, D is the specimen diameter, l_{75} is the displacement at 75% of the peak load after the peak, G_f is the failure energy and m_{75} is the slope at 75% peak load after the peak.

Historical Development

The IDT-CT protocol was developed under the NCHRP IDEA Project 195, which looked to develop a cracking test for asphalt mix design, quality control, and quality assurance. In general, the test is similar to the traditional Indirect Tensile Strength Test; however, the interpretation of the results differs in terms of the rationale used to define the cracking resistance index. The CT_{index} is derived from the load-displacement curve and its form is inspired by Paris' law and the definition of crack growth by Bazant and Prat in 1988 (Bažant and Prat 1988, Zhou *et al*. 2017b). This test has gained considerable attention in recent years among Departments of Transportation (DOTs) and practitioners given its simplicity, practicality, efficiency, repeatability, relatively low-cost equipment, and purportedly good correlation with field cracking performance (Diefenderfer and Bowers 2019, Seitllari *et al.* 2022). Figure A.2 shows a general overview of the test and typical results.



(a) (b)

Figure A.2. IDT-CT overview; (a) load-displacement curve and (b) test setup and conditions (Zhou *et al*. 2017b).

Brief Summary of Studies using the Test Method

Since the development of the IDT-CT and CT_{index}, several research efforts have been conducted to evaluate the ability of the index to discriminate the performance of asphalt mixture with different mix design characteristics, e.g., asphalt binder type, air voids, etc. This section will span a review of studies where the sensitivity of the CT_{index} to multiple mix design factors is evaluated. Table A.4 summarizes many of the relevant studies.

Chowdhury *et al.* evaluated the sensitivity of the CT_{index} to the asphalt binder characteristics using two virgin binders, a binder modified with Elvaloy and PPA (PBM), a binder modified with crumb rubber (CRMB), and a binder modified with a warm mix additive (SMB). The IDT-CT was conducted at temperatures of 15, 25, and 35°C and it was found that the PBM and CRMB mixtures exhibited CT_{index} values significantly higher than the other mixtures at all temperatures

(Chowdhury *et al.* 2022). Wu *et al.* conducted a study evaluating the effect of three commercial polymer additives in China on the cracking resistance of asphalt mixtures measured by the IDT-CT at multiple aging levels (Wu *et al.* 2021). The results suggested that polymer modification increases the CT_{index} of the base material, which was attributed to the higher ductility and lower crack propagation speed obtained through the binder modification. However, the CT_{index} values decreased progressively with long-term aging time. After 15 days of long-term aging (as per AASHTO R 30), all the mixtures seemed to have similar CT_{index} and the authors attributed this result to degradation of the additive during the conditioning period. Similar findings were reported by with SBS-modified binders (Ali *et al.* 2022).

Table A.4. Summary of Findings using the IDT-CT Method.

Topic	Findings	Reference
Air voids	No statistical difference was found in CT_{index} with air voids from 2.5% to 5.5%.	(Mivehchi *et al.* 2022)
	CT_{index} increases with the increase of air voids.	(Zhou *et al*. 2017b)
Binder content	CT_{index} increases with an increase in binder content.	(Ling and Buchanan 2022, Mivehchi *et al.* 2022)
Softer binder	CT_{index} increases with the use of softer binder.	(Zhou *et al*. 2017b)
Crumb rubber mod.	CT_{index} is not able to capture the favorable effects of crumb rubber incorporation into the asphalt mixture.	(Rath *et al.* 2022)
	CT_{index} increases with the incorporation of crumb rubber into the asphalt mixture.	(Chowdhury *et al.* 2022)
Polymer mod.	The expected bump in CT_{index} due to high polymer modification is not observed.	(Habbouche and Diefenderfer 2021)
	Polymer modification has a favorable effect on the CT_{index}.	(Zhou *et al*. 2017b)
Oxidate aging	Oxidative aging has an adverse effect on the CT_{index} of the asphalt mixtures evaluated.	(Zhou *et al*. 2017b)
Recycled materials	The incorporation of recycled materials decreases the CT_{index} but only marginal reductions are found when the contents of recycled materials are relatively low.	(Zhou *et al*. 2017b, Roja *et al*. 2021, Ling and Buchanan 2022, Mivehchi *et al*. 2022, Zalghout *et al*. 2022)
Loading rate	There is no significant effect of the loading rates evaluated in the study on the CT_{index}.	(Habbouche *et al*. 2022a)

Although multiple studies suggest that the IDT-CT can capture the effects of polymer modification, Habbouche and Diefenderfer presented a comparison of CT_{index} of polymer-modified (PMA) and high polymer-modified (HP) asphalt mixtures and found that the IDT-CT could not detect the expected impact of high polymer modification in improving the cracking performance of surface mixtures (Habbouche and Diefenderfer 2021). A study by Rath *et al.* concluded that the IDT-CT is not able to capture the favorable effect of ground tire rubber modification on the cracking resistance of asphalt mixtures, where a distinctive fracture behavior is observed (Rath *et al.* 2022).

Zhou *et al.* investigated the sensitivity of the CT_{index} to discriminate the cracking performance of asphalt mixtures containing RAP and RAS. They produced two mixes, one with 20 percent RAP and the other with 15 percent RAP and 5 percent RAS (Zhou *et al*. 2017b). The results from the IDT-CT test indicated a significant drop in the CT_{index} values when compared to the control mixture without recycled asphalt materials. This trend agrees with a more recent work by Ling and Buchanan, where it was found that the CT_{index} decreases roughly by 20 units per 1 percent increase of recycled binder replacement (Ling and Buchanan 2022). Zalghout *et al.* indicated no statistically significant changes in CT_{index} when adding 15 and 20 percent RAP to asphalt mixtures designed using the Maricopa Association of Governments Guidelines (Zalghout *et al.* 2022).

Cantabro

Method and Calculation Description

The Cantabro test is an asphalt mixture durability test procedure conducted using the Los Angeles abrasion testing machine as per AASHTO TP 108. The test is conducted at intermediate temperatures and for the test, asphalt specimens are placed into the abrasion drum without steel spheres and subjected to 300 revolutions at 30-33 rpm. The mass loss (*ML*) of the specimen during the agitation is used as an indicator of durability and it is expressed as a percentage of the mass of the specimen before the test, Equation (6).

$$ML = (A - B) \times 100\% / A \quad (6)$$

Where *A* and *B* are the mass of the specimen before and after the test respectively.

Historical Development

The Cantabro test has been used worldwide as a performance assessment for durability and to identify the minimum asphalt binder content for open-graded friction courses (OGFC) or porous friction courses (PFC) (Cooley *et al.* 2009, Arámbula-Mercado *et al.* 2016). The literature review suggests that the Cantabro protocol has been gaining increased attention to evaluate durability in other types of asphalt mixtures e.g., dense grade mixtures, cold recycled mixtures, etc., due to its simplicity, cost, and efficiency (Arámbula-Mercado *et al.* 2020).

Brief Summary of Studies using the Test Method

Several research efforts using the Cantabro test for dense graded mixtures have been conducted by researchers at Mississippi State University (MSU) and collaborators since 2009. An exhaustive review of the literature until 2017 is presented by Cox *et al.* and a summary of the most relevant findings is presented in Table A.5 (Cox *et al.* 2017).

Table A.5. Summary of Findings using the Cantabro Method.

Topic	Findings	Reference
Air voids	*ML* increases with an increase in air voids. However, the trend is not clear for mixtures with 19.0 NMAS.	(Doyle and Howard 2011, 2016)
Aging protocol	*ML* significantly increases with aging either by AASHTO R 30 (5 days at 85°C) or by Mississippi's OGFC aging protocol (7 days at 64°C) with respect to un-conditioned samples at comparable air void levels.	(Doyle and Howard 2016)
Recycled materials	*ML* tends to increase with the increase of recycled materials content.	(Celauro *et al.* 2010, Doyle and Howard 2016)
Asphalt content	*ML* decreases with increasing asphalt content.	(Doyle and Howard 2016, Vila-Cortavitarte *et al.* 2018)
Asphalt mod.	*ML* increases with the addition of 18% crumb rubber and 15% TB rubber and decreases with the addition of 4.5% SBS, 8% HDPE, and 24% Gilsonite with respect to the control condition.	(Zhou *et al*. 2021a)
	Mixtures containing high polymer (HP) binders had lower *ML* than mixes containing conventional SBS polymer-modified binders.	(Habbouche *et al*. 2022b)
	ML increased with the addition of recycled polystyrene.	(Vila-Cortavitarte *et al.* 2018)
Fiber	*ML* decreases with the addition of fibers in mixtures containing 70% of recycled asphalt concrete.	(Su and Hachiya 2008)

Uniaxial Cyclic Fatigue (CF)

Method and Calculation Description

The CF test is a repeated load performance test that is used to determine the cracking resistance properties of asphalt mixtures. It is most commonly coupled with the simplified viscoelastic continuum damage theory (S-VECD). The test is conducted using an Asphalt Mixture Performance Tester (AMPT) on cylindrical specimens. AASHTO T 411 is typically used when the NMAS of the mixture is 19 mm or smaller. For larger NMAS mixtures, AASHTO T 400 is used. The T 411 protocol uses a 38-mm diameter by 110-mm tall specimen while T 400 uses a 100 mm diameter by 130-mm tall specimen. The basic test protocol is the same in both test methods and consists of a repeated, tension only actuator controlled sinusoidal displacement at 10 Hz. The test temperature is generally around 21°C, but varies according to the performance grade of the binder used in the mixture. It should be noted that there exist other uniaxial test protocols that use similar principles and setups, such as the methods developed by the University of Nottingham, and Arizona State University, among others. However, this review will focus exclusively on the methods and development of the test protocol specified in AASHTO T 400 and T 411. Two key test outcomes are obtained from the CF tests, the damage characteristic curve, named the the material integrity (*C)* versus damage (*S*) curve, and the pseudo-energy-based failure criterion, D^R, calculated as indicated by Equations (7) and (8).

$$C = 1 - C_{11} S^{C_{12}} \tag{7}$$

$$D^R = \frac{Sum(1-C)}{N_f} \tag{8}$$

Where C_{11} and C_{12} are fitting coefficients of the power model, P is the number of specimens, *Sum*(1-*C*) is the integral area below the curve of (1-*C*) versus cycle number until the failure cycle, and N_f, is the number of cycles to failure. The results of the CF test allow for the calculation of the S_{app} parameter, Equation (9), which measures the amount of fatigue damage the material can tolerate considering the effect of the material's modulus and toughness (Castorena *et al.* 2021, Wang *et al.* 2022). A higher S_{app} value indicates higher fatigue cracking resistance.

$$S_{app} = 1000^{\frac{\alpha}{2}-1} \frac{a_{T(S_{app})}^{\frac{1}{\alpha+1}} \left(\frac{D^R}{C_{11}}\right)^{\frac{1}{C_{12}}}}{\left|E^*{}_{LVE,S_{app}}\right|^{\frac{\alpha}{4}}} \tag{9}$$

Where α is the damage growth rate, a_T is the time-temperature shift factor at a given temperature and $E^*{}_{LVE}$ is the average representative dynamic modulus.

Historical Development

The development of the CF test protocol originates from the work of Kim *et al.* (1997) and Daniel and Kim (2000). Kim and his colleagues used Schapery's work potential theory to study the constitutive behavior of asphalt concrete with damage in uniaxial tension mode. It was found that the constitutive model was able to successfully predict damage growth of the asphalt concrete under monotonic loading. Later, Daniel and Kim used the continuum damage-based constitutive model developed by Kim *et al.* to develop a prediction methodology and test procedure to characterize fatigue of asphalt concrete subjected to various loading and environmental conditions (Kim *et al.* 1997). Under the auspices of the NCHRP 09-19 project, a direct-tension uniaxial fatigue test method using 75-mm diameter and 150-mm height specimens and on-specimens LDTVs was developed. It was found that a single characteristic curve can be produced from cyclic and monotonic loading, i.e., the damage characteristics of the asphalt concrete are independent of the mode of loading. Underwood *et al.* developed the so-called simplified viscoelastic continuum damage model building on the work of Daniel and Kim to address certain shortcomings of the previous formulations and to allow for rapid characterization and/or prediction of the material's fatigue response (Underwood *et al.* 2012). For a more rigorous review of the S-VECD development, the reader is referred to previous work (Underwood *et al.* 2010, 2012).

Several efforts have been also made to develop failure criteria for asphalt mixtures under fatigue loading. Sabouri and Kim developed an energy-based failure criterion, referred as the average dissipated pseudo energy rate criterion or G^R, to overcome the shortcomings of previous failure criteria that were either arbitrary or not fundamentally based, e.g., reduction of the material's modulus to 50 percent of its initial value, or that depend on the mode of loading (Sabouri and Kim 2014). In their study, it was found that the G^R criterion was unique for a given mixture regardless of the temperature (if viscoplasticity effects were not encountered) and strain amplitude. However, Wang and Kim noted that the G^R model is significantly impacted by test variability because the model parameters are obtained using linear regression in a double logarithmic scale (Wang and Kim 2019). This fact and other limitations of the G^R criterion led to the development of D^R or average reduction in pseudo stiffness up to failure.

Brief Summary of Studies using the Test Method

Several research efforts have been conducted to evaluate the fatigue characteristics of asphalt mixtures using the CF test, however, the number of studies is relatively lower compared to the other durability test methods reviewed herein. It is speculated that this is probably because the equipment and personnel training required to conduct the CF test are more specialized and less accessible compared to other test methods. Table A.6 presents a summary of research studies using the test method for fatigue evaluation of asphalt mixtures.

Table A.6. Summary of Findings using the Uniaxial CF Method.

Topic	Findings	Reference
Air voids	S_{app} decreases with an increase in air void content for 9.5-mm and 19.0-mm NMAS asphalt mixtures.	(Jeong *et al.* 2022)
Oxidative aging	S_{app} and D^R decrease with the increase of long-term oven aging time.	(Saleh *et al.* 2020, Wang *et al.* 2022)
	For different types of mixtures evaluated, DR decreases with the increase in oven aging time.	(Zhang *et al.* 2019)
Recycled materials	S_{app} and D^R decrease with the increase in RAP content. The effect of using a softer binder with RAP is captured in the S_{app} parameter.	(Wang *et al.* 2022)
	In general, the use of RAP decrease fatigue performance by means of G^R, however, the addition of RAP up to 15% does not seem to significantly affect fatigue performance.	(Sabouri and Kim 2014)
Asphalt content	S_{app} increases with binder content.	(Wang *et al.* 2022)
Asphalt binder/ mixture mod.	The utilization of WMA additives seems to deteriorate fatigue resistance by means of G^R.	(Kim *et al.* 2017)
	The use of a WMA chemical additive increased the tolerance to fatigue damage with respect to the control HMA using S_{app}.	(Spadoni *et al*. 2022a)
	The incorporation of recycled plastic with graphene and plastomeric compounds yielded lower S_{app} values compared to the reference SBS-modified asphalt mixture.	(Spadoni *et al*. 2022b)

Bending Beam Fatigue Test

Method and Calculation Description

The BBF test is a repeated load performance test used to characterize the cracking resistance of asphalt mixtures (AASHTO TP 321). The test consists in subjecting an asphalt mixture beam specimen of 50 mm x 63 mm x 380 mm to controlled sinusoidal displacement at a frequency of 10 Hz. The test setup and specimen dimensions are presented in Figure A.3. The displacements are applied at two third points to produce a uniformly distributed moment at the mid-span of the beam. The strain selected is constant throughout the test and is generally between 250 to 750 microstrains for conventional asphalt mixtures. The number of cycles to failure is defined as the point at which the stiffness, *S,* multiplied by the number of cycles, *N,* is a maximum. Typically, the generic interpretation of fatigue is used but is worth mentioning that there exist analytical approaches to interpret BBF results using VECD theory.

Historical Development

The standard BBF procedure (also called the four-point bending test or third-point loading test) was based on the work conducted by researchers at the University of California Berkeley and at the Asphalt Institute. For the University of California, the test specimens consisted of beams of 38.1 mm x 38.1 mm x 381 mm, while The Asphalt Institute used larger specimens of 76.2 mm x 76.2 mm x 381 mm. In the early development of the test protocol, a pulsating loading of 0.1 seconds and a frequency of loading of 100 repetitions per minute were applied (Tangella *et al.* 1990). The current specimen size (AASHTO TP 321) and several modifications to the test equipment were proposed during the SHRP program. Some of the specific changes included improvements in the linear and torsional bearing to maintain zero moments at the beam ends, automation of temperature and test control, data acquisition, data reduction, among other significant modifications. Further, sinusoidal loading could be now applied at up to 25 Hz, with temperatures ranging between -10°C to 40°C. This set of modifications improved the repeatability of the test and reduced the overall testing time (Monosmith 1994).

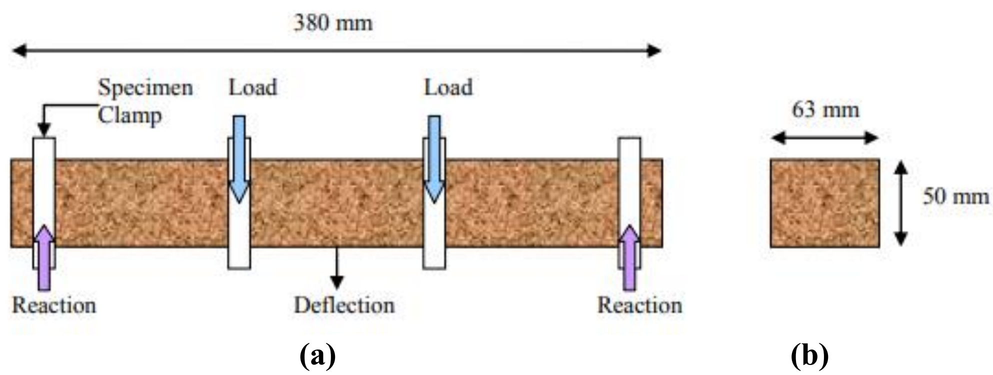


Figure A.3. BBF overview; (a) test configuration and (b) specimen cross section (Kim *et al.* 2011).

Brief Summary of Studies using the Test Method

A summary of relevant findings using the BBF protocol is presented in Table A.7. The reader is referred to the expanded fatigue test program of SHRP-404 for other results that support the relationship between mix properties and laboratory fatigue response (Monosmith 1994).

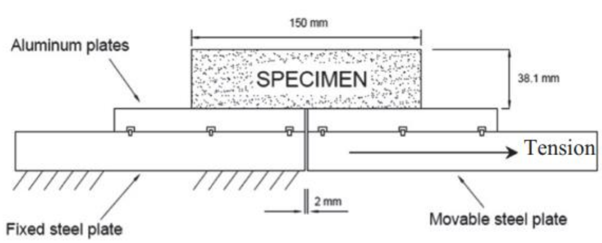
Overlay Test

Method and Calculation Description

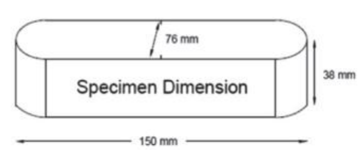
The Overlay Test (OT) is a repeated displacement-controlled performance test developed by the Texas transportation institute and is used primarily to evaluate reflective cracking resistance of asphalt mixture overlays. However, there have also been applications in the area of BMD and pavement performance simulations. The test procedure is described in the Texas Department of Transportation (TxDOT) procedure Tex-248-F. The protocol consists of gluing the test specimen to an apparatus that has a fixed and movable steel plate attached to a ram. The specimen geometry and apparatus setup are shown in Figure A.4. The ram moves to induce a triangular loading pattern that seeks to simulate the opening and closing of cracks at a frequency of 0.1 Hz and temperature of 25°C. The number of cycles to failure, N_f, is defined when a 93% reduction of the maximum applied load occurs. It should be noted that although there are devices specifically developed for the OT test, the test can be also conducted in an AMPT.

Table A.7. Summary of Findings using the BBF Method.

Topic	Findings	Reference
Air voids	An increase in air voids from 4 to 6 percent reduced the fatigue life of the asphalt mixture by 4.2 to 5.2 depending on the binder grade.	(Fakhri and Kheiry 2009)
Oxidative aging	Oxidative aging decreased the fatigue lives of asphalt mixtures.	(Islam and Tarefder 2015, Izadi *et al.* 2018)
Recycled materials	In general, fatigue life decreases with the increase of RAP and/or RAS content, independently of the base binder used in the mixture and incorporation of WMA technology.	(Kim *et al*. 2018a, Ding *et al*. 2019)
	The effects negative of RAP on the fatigue life of high-modulus asphalt mixtures were only clear when the RAP content is 60 percent.	(Ma *et al.* 2015)
Asphalt content	In general, increasing binder content increases the fatigue life of asphalt mixtures. It was found that a 1 percent increase in binder content would increase the fatigue life of asphalt mixtures between 10 and 350 percent.	(Fakhri and Kheiry 2009)
Asphalt binder/ mixture modification	The utilization of SBS-modified asphalt increased the fatigue lives up to three times more than non-polymer modified mixtures.	(Fakhri *et al.* 2013)
	The use of WMA additives in addition to rejuvenators can increase the fatigue life of high-content recycled asphalt mixtures and achieve better performance than conventional HMA.	(Kim *et al*. 2018a, Mirhosseini *et al*. 2020)
	The use of a softer binder improves the fatigue cracking resistance of recycled asphalt mixtures; however, the advantages of a softer binder were not observed with 20% RAS.	(Kim *et al*. 2018a)



(a)



(b)

Figure A.4. OT overview; (a) test configuration and (b) specimen dimensions (Walubita *et al.* 2013).

Historical Development

The OT concept was first introduced in the late 1970s by German and Lytton (1979) using long beams. The test setup and analysis theory have been modified several times since the introduction of the OT concept. The specimen geometry was modified by Zhou and Scullion (2004) in an attempt to develop a practical protocol for routine pavement characterization and design. Further, two loading schemes (Procedure A and B) were presented. Procedure A (one-phase loading) consisted of a cyclic triangular loading with a constant maximum displacement of 2 mm while Procedure B (two-phase loading) consisted of a constant displacement loading for 30 seconds to determine the relaxation modulus curve of the asphalt mixture before subjecting the specimen to the cyclic loading. Procedure A was recommended for routine evaluation of reflective cracking.

In Zhou and Scullion's work, Paris' law was used to determine fracture parameters and crack growth of asphalt mixtures. Later, Zhou *et al.* proposed a new fatigue cracking prediction approach considering cracking initiation and propagation based on OT results, where the crack initiation is determined based on the tensile strain at the bottom of the asphalt layer (Zhou *et al.* 2007). In this analysis, asphalt mixtures were assumed to be quasi-elastic materials.

Koohi *et al.* developed an alternative method to analyze OT results using principles of viscoelastic fracture mechanics and finite element simulations (Koohi *et al.* 2013). In their approach, the-phase loading pattern was changed to first include a non-destructive phase that includes 10 load cycles with an opening displacement of 0.05 mm. This phase was used to obtain the undamaged relaxation modulus of the asphalt mixture. It was followed by a destructive phase that includes 150 cycles with a maximum opening displacement of 0.653 mm. Finite element simulations were used to determine the crack growth function and the incremental pseudo *J*-integral with each load cycle. Later Gu *et al.* developed an alternative methodology based on the mechanical analysis of viscoelastic force equilibrium in the OT specimens (Gu *et al*. 2015a). In this effort, the destructive phase was modified to 200 load cycle with a maximum opening displacement of 0.32 mm. For more details on the fracture mechanic viscoelastic framework, the reader is referred to Gu et al's work (Gu *et al*. 2015a, 2015b). The OT setup has been also considered as a monotonic loading test to characterize fracture properties and cracking resistance of asphalt mixtures through conventional fracture mechanics parameters such as fracture energy, tensile strength, and others (Walubita *et al.* 2013, 2021, Lee *et al.* 2016).

Brief Summary of Studies using the Test Method

Since the development of the OT concept, several studies have been conducted to improve the OT setup and analysis theory. Some of these studies have been introduced in the previous section. Table A.8 presents a summary of research studies using the test method for cracking evaluation of asphalt mixtures.

Table A.8. Summary of Findings using the Overlay Test.

Topic	Findings	Reference
Oxidative aging	The OT cycles to failure decrease with increases in aging time in the Accelerated Pavement Weathering System.	(Fakhri and Kheiry 2009)
Recycled materials	In general, the OT cycles to failure decreases with the increase of recycled materials (RAP, RAS or both)	(Mogawer *et al.* 2011, Zhang *et al.* 2017)
Asphalt content	The OT cycles to failure increases with increases in asphalt content	(Li *et al.* 2014)
Asphalt binder/ mixture mod.	Rejuvenators increase the OT cycles to failure; however, the degree of improvement depends on the type of rejuvenator and dosage level.	(Im *et al.* 2014, Mohammadafzali *et al.* 2017)
	Asphalt mixes with PG 76-22 binder have superior performance in terms of OT cycles to failure.	(Li *et al.* 2014)
	Asphalt mixtures modified with rejuvenators exhibit a higher loss of cracking resistance with weathering time compared to the control mixture.	(Mohammadafzali *et al.* 2017)
	The use of waxed- and water-based WMA additives increases the cracking resistance of the asphalt mixture.	(Mogawer *et al*. 2011, Gu *et al*. 2015a)

Synthesis of Durability Test Methods

The previous section evaluated the tests' ability to capture the effect of asphalt mixture and other mix design parameters. The FHWA-HIF-19-103 (Hajj *et al.* 2019) publication recognizes other factors that are important to assess the appropriateness of test protocols for BMD implementation. These other factors include sample preparation, specimen conditioning and testing, training needs and applicability, equipment cost, repeatability, and field validation. Table A.9 (Adapted from Zhou *et al.* (2017a) presents a summary that identifies the aforementioned factors for each of the test methods evaluated herein.

The literature review suggests that most agree that the I-FIT, IDT-CT, and SCB-LTRC test methods are the most practical and easiest to adopt for BMD purposes. The following remarks are made:

- The I-FIT, IDT-CT, and SCB-LTRC are index tests that can be used to rank different materials and assess the acceptance of a mix design. However, it is worth mentioning that the results of the test are not amenable to conducting pavement performance predictions.
- The I-FIT and IDT-CT results do not have an expected trend regarding specimen air void content. While intuitively incorrect, this result may not be an issue and the indices from each test can be informative during the mix design process if the cracking evaluation is conducted with a fixed air void content. However, the test methods are non-ideal for QA processes that employ in-place density (Jeong *et al.* 2022). The trends are not clear for the SCB-LTRC.
- The I-FIT and IDT-CT use a loading rate of 50 mm/min and a similar interpretation of the test results. However, the SCB-LTRC test uses a loading rate of 0.5 mm/min. Asphalt mixtures are viscoelastic materials and therefore, the type of response in this test, i.e., development and growth of the fracture process zone, crack propagation, etc., is different than in the I-FIT and IDT-CT.

On the other hand, repeated load tests attempt to simulate the process of fatigue damage accumulation; therefore, the test results can be analyzed to perform pavement performance prediction, in addition to providing an index to rank or assess the acceptance of asphalt mixtures. The literature suggests that repeated load tests reviewed in this document (BBF, CF, and OT) are able to capture the effect of mix design properties on the fatigue performance of asphalt mixtures; however, some of the main challenges associated with the implementation of repeated load tests for BMD purposes are:

- Unlike monotonic tests, there is uncertainty in the time duration of the BBF, CF, and OT tests because the test results relied on the continuum degradation of the asphalt mixture, which depends on the stress or strain level selected and inherent damage resistance of the mixture. This can represent a challenge for routine characterization if multiple tests need to be conducted within a specific time frame.
- The preparation of BBF, CF, and OT specimens is lengthy due to coring, cutting, and gluing. Further, the training required to conduct the tests (setting up the equipment, mounting the specimen, etc.) and the data analysis are relatively more specialized than monotonic tests.
- The cost of the equipment for BBF, CF, or OT is significantly higher compared to the monotonic tests. This may represent a roadblock for state DOTs and contractors if multiple test machines are needed to supply the demand for testing.

Table A.9. Comparison of Different Cracking Test Methods (Adapted from Zhou *et al*. 2017a).

Test name	Test standard	Cross section geometry	Parameter (s)	Correlation to field performance	Typical test variability	Advantages	Drawbacks	Equipment cost[1]
IDT-CT	ASTM D8225	Circular; *D*= 150 mm *T*= 62 mm	CT_{index}	Fairly good correlation to No. of passes to first crack and reflective cracking rate Zhou *et al.* (2017b). Good indicator for resistance to top-down cracking (NCAT Test Track) West *et al*. (2021a).	Relatively high (COV < 40%) reported	No coring, cutting, and gluing is required. The test duration is less than 5 minutes.	The trends with air voids are not clear, which complicates the implementation for QA processes that use in-place density. Five test specimens are often necessary due to test variability.	<$10,000
SCB-LTRC	ASTM D8044	Semi-circular; *T*= 57 mm, *ND* = 25.4, 31.8 and 38.1 mm	J_c	Fair correlation between J_c and cracking rate Kim *et al.* (2012). The J_c was unable to discriminate mixes with significant cracking (NCAT Test Track) West *et al*. (2021a)	Medium variability (COV = 20-25%)	The analysis theory considers a fundamental fracture mechanics parameter.	At least twelve specimens are required for a full fatigue characterization of the mixture.	-
I-FIT	AASHTO TP 124	Semi-circular; *T*=50 mm *ND* = 15 mm	*FI, CRI* and *BCI*	Relatively good relationship between *FI* and ALF cycles, as reported by Ozer *et al*. (2018). Fair correlation to top-down cracking field performance (NCAT Test Track) West *et al*. (2021a).	Relatively high (COV < 50%)	The test duration is amenable for routine characterization	The test requires cut and notched specimens. A minimum of six replicates are needed given the test variability.	<$10,000
Cantabro	AASHTO TP 108	Circular; *D = 150 mm* *T = 110 – 120 mm*	*ML*	The ability of the test to rank field performance is not clear due to the variability of the test results	Medium variability (COV = 20-25%)	No cutting require; mix design specimens can be used as test specimens; relatively short test duration;	The test does not explicitly evaluate any fundamental fracture mechanics/continuum damage properties.	<$10,000
Uniaxial Cyclic Fatigue	AASHTO T 411 AASHTO T 400	Circular; *D* = 38 mm, 100 mm *T* = 110 mm, 130 mm (Dimensions depend on test standard)	S_{app}, D^R	Promising correlation to field performance (FHWA-ALF, MnROAD). Good correlations observed with top-down cracking (NCAT Test Track) West *et al*. (2021a)	Low variability reported by	Allows the prediction of fatigue performance at various conditions. Results can be combined with pavement analysis model. Literature review suggests that the cracking parameters follow the expected trends with mix design properties.	Fabrication of the test specimens requires coring, cutting and gluing. Requires specialized software to analyze the test results. Uncertainty in the test duration can represent a drawback for routine characterization.	>$100,000
Overlay Tester	Tex-248-F	Oblong *T* = 38 mm	N_f,	Promising correlation to field performance (FHWA-ALF). Good indicator of a mixture's resistance to top-down cracking (NCAT Test Track) West *et al*. (2021a).	Relatively high (COV = 30-50%)	Interpretation of the results is relatively simple.	Requires five sample due to inherent test variability. Cutting and gluing is necessary.	$40,000-$50,00
Bending Beam Fatigue	AASHTO TP 321	Rectangular *W*= 63 mm *D*= 50 mm *L*= 380 mm	N_f	Good capacity to rank asphalt mixture validate in SHRP-404.	High variability (COV > 50%)	The test mimics bottom-up fatigue cracking and can be used for pavement performance predictions.	Fabrication of specimens and test setup can be cumbersome. The test specimens require significant amount of material. Typical analysis framework requires multiple tests at different strain levels and temperatures to calibrate fatigue model.	>$100,000

[1] Costs are for the equipment used for the given test alone. As described in Section 2.4 some equipment can be used for several different tests. Final costs may be less than the amounts listed in this table because of these multipurpose options

Stability Test Methods (Rutting)

Stability of asphalt mixtures was previously determined in terms of Marshall Stability test, and Hveem Stabilometer. When Superpave methodology was envisioned, and further applied to current United States practice, designers noticed a lack of a mechanical "proof" parameter analog to stability, relying on conformance to the material specifications and volumetric mix criteria (Witczak *et al.* 2002). It has been noticed that the rutting distress reduced its occurrence in asphalt pavements due to the new methodology, even so, some states cautiously adopted a rutting limit in their local designs. When envisioning new BMD approaches, it is fundamental that both cracking and rutting potential are assessed in asphalt mixtures, and it changes are properly accounted and balanced.

The following sections describe different test methodologies discussed as BMD possible application. The items are divided in method description, historical background and studies using the test, covering from wheel-tracking "torture" tests, monotonic indirect tensile tests, and more fundamental dynamic shear tests.

Asphalt Pavement Analyzer (APA) Test

<u>Method and Calculation Description</u>
The Asphalt Pavement Analyzer (APA) test is a wheel loading test that directly measures the rut depth on-specimen, which serves as a performance index for rutting potential in service. The test consists by an aluminum wheel load with a linear pressured rubber hose that simulates the effect of a wheel tracking back and forth over the pavement, at a high temperature. The sample is placed inside a mold and kept "confined" while the test is running. After 8000 cycles, the accumulated rut depth on the sample is recorded.

AASHTO T 340 (2023b) recommends that six samples should be used, with a desired air void content of 7 ± 0.5% to represent the as-constructed (and thus worst case) scenario. For gyratory samples, either plant or laboratory produced, dimensions of 150 mm in diameter by 75 ± 2 mm in height, or beams measuring 125 ± 2 mm wide by 300 ± 2 mm long by 75 ± 2 mm tall. Roadway cores can also be tested, having 140 to 152 mm in diameter with a minimum height of 50 mm. Cores taller than 75 mm should be trimmed, and the uncut face of the core shall be used to test (AASHTO, 2023b).

The wheel load should be 690 ± 35 kPa (100 ± 5 lbf), while the hose pressure should be 445 ± 22 N (100 ± 5 psi). According to AASHTO T 340 (2023b), the test temperature should be the same as the high temperature Performance Grade (PG_H) for the location where the material will be used. Prior to testing, the specimen should be conditioned for not less than 6 hours and no more than 24 hours. Rut depth measurements can be taken manually or automatically at multiple locations on the specimen. Whichever method is adopted the average rut depth across all locations and for all specimens is recorded at the end of 8000 cycles (AASHTO, 2023b).

<u>Historical Development</u>
The APA evolved from the Georgia Loaded Wheel Tester (GLWT) (Cooley *et al.* 2000). Originally, developed in 1985 by the Georgia Department of Transportation (GDOT) and Georgia Institute of Technology, to be a simple testing procedure to evaluate rutting potential of a hot asphalt mix (Lai 1986). At that time, the United Stated relied on Marshall's procedure for its asphalt mixture designs and rutting was a major distress (Collins *et al.* 1996). The advantages of

the GLWT were that it could be run in a semi-automated mode that could test three beam samples at the same time (Collins *et al.* 1996).

The first study used a “loaded foot” kneading compactor to produce the samples, beams of 125 mm wide by 300 mm long by 75 mm thick dimensions. After pouring into a mold, a sliding rack was positioned over it, and a stationary kneading compaction would produce the beams (Lai 1986). A decade later, a rolling compaction machine was developed for the specimen fabrication, simulating the field compaction effort (Lai and Shami 1995). With the development of Superpave design method, the test also migrated to gyratory samples, with the specimen measuring 150 mm diameter by 75 mm height was found to be consistent with the given beam results (Collins *et al.* 1996). The test temperature has increased over time, from 35°C (Georgia’s average temperature at which rutting has been found to occur) (Lai 1986), to 40°C (Lai 1993, Collins *et al.* 1995), 50°C (Collins *et al.* 1995, 1996) and 60°C (Collins *et al.* 1996). For the pass/fail limit, GDOT set its limit at 7.5 mm in its first study (Lai 1986), while others, like Utah DOT Maryland DOT, considered 5.0 mm (Collins *et al.* 1996).

The APA follows the same basic procedure as the GLWT, using the load wheel back and forth over a sample to induce rutting, measuring the rut depth at 8000 cycles. However, it can be also used to test samples submerged in water, and either three beams or six gyratory samples can be tested simultaneously (Choubane *et al.* 2000, Cooley *et al.* 2000). The wheel load and hose pressure have also remained the same since the first GWLT (AASHTO, 2023b).

<u>Brief Summary of Studies using the Test Method</u>

The APA test is well-stablished as a rutting resistance “go/not go” test in the literature, with a solid test procedure. Recently, rather than debate its test method, much effort has been given to determine the effects of volumetrics, binder properties, RAP and additive additions on the test results, and its suitability to be incorporated in a Balanced Mix Design (BMD) program. Table A.10 summarizes some of the findings from literature regarding APA results.

Table A.10. Summary of Findings using the APA Method.

Topic	Temperature; Air voids	Findings	Reference
Test method	(40, 50, 60°C; 2.0 to 7.0%)	Higher air voids and higher test temperature decrease rutting resistance.	(Collins *et al.* 1996)
Binder effect	(58, 64, 70, 76°C; 7%)[a]	A lower J_{nr} increases the rutting resistance.	(Bernier *et al.* 2012)
Aggregate effect	(64°C; 4%) [a] (50°C; 6%)	Aggregates with rougher texture, better adhesion with binder, and better angularity increase rutting resistance.	(Kandhal and Mallick 2001, Hussan *et al.* 2019)
Volumetric effect	(50°C; 6%) (64°C; 4%)	Coarser gradation increases rutting resistance.	(Hussan *et al.* 2019, Park *et al.* 2022)
	(50°C; 6%)	Higher OBC and higher VMA decreases rutting resistance.	(Hussan *et al.* 2019)
RAP content	(58, 64, 70, 76°C; 7%)[a] (58°C; 7%)	Adding RAP increases rutting resistance.	(Bernier *et al.* 2012, Colbert and You 2012)
Rejuvenators	(64°C; 7 ± 1%)[a]	Adding rejuvenators can decrease rutting resistance.	(Song *et al.* 2018)
WMA	(64°C; 7 ± 0.5%)[a] (64°C; 7 ± 1%)[a]	Adding WMA agent can either increase or decrease rutting resistance.	(Zhao *et al.* 2013, Song *et al.* 2018)

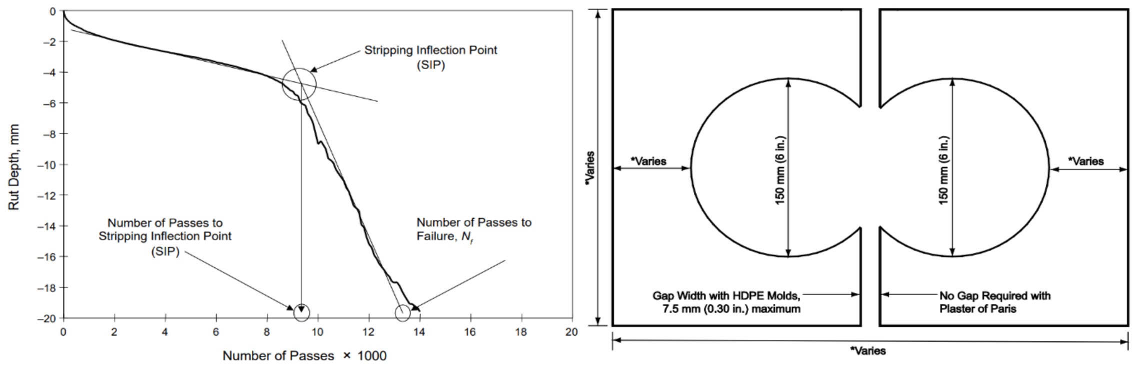
[a] Test performed at binder's Performance Grade High temperature of each mix

Hamburg Wheel Tracking Test (HWTT)

Method and Calculation Description

Currently the most used rutting test reported as Balanced Mix Design efforts is the Hamburg Wheel Tracking Test (HWTT). As with the previous test, it also uses a small-scale wheel load simulation over an asphalt mixture sample to directly evaluate rutting depth. In this case, a steel wheel directly contacts the surface at a high temperature. Testing can be performed on either saturated or non-saturated samples, with the former being more common in the U.S. When performed in a saturated mode, the test may also provide an indicator of moisture susceptibility. The rut depth can be recorded for each cycle and used to calculate additional indices such as the creep slope and the called "stripping inflection point" (*SIP*), see Figure A.5(a).

AASHTO T 324 (2023a) requires the steel wheel to have diameter of 203.2 ± 2 mm, and be 47 ± 0.5 mm wide. The load of the wheel is 703 ± 4.5 N, moving at the center of specimen, making 52 ± 2 passes per minute. Laboratory compacted beams (320 mm long, 260 mm wide) and gyratory specimens (150 mm diameter), also cores from field, can be used in the test. The thickness of specimen in any case should be at least twice the nominal maximum aggregate size, generally yielding 38 mm to 100 mm tall specimens. When using gyratory specimens, they must be cut to fit in the mounting system, see Figure A.5(b). A pair of gyratory samples of the same material should be tested simultaneously, while the reported results should be the average of two pairs. The air voids recommended for use in the testing are 7 ± 0.5% for gyratory specimens and 7 ± 1% for slabs samples.



(a) **(b)**

Figure A.5. HWTT overview; (a) typical results and (b) gyratory mounting (AASHTO 2023a).

After the test, the samples are pre-conditioned in the device's water bath, for at least 45 minutes, and not more than 60 minutes. During the test, an LVDT records the rut depth of at least each 20 passes of the wheel. The test terminates at either 20,000 passes count or at a pre-determined rut depth. The test temperature, as well as rut depth limits are not pre-defined in the AASHTO standard, giving liberty to each agency to establish their own specifications. Generally, the test is performed at 50°C, but some states use PG_H based temperatures, varying from 44°C to 56°C (Mohammad *et al*. 2016a).

Historical Development

In the 1970s, Helmut-Wind Incorporated of Hamburg proposed a specification requirement to measure combined rutting and stripping susceptibility (Mohammad *et al*. 2016a). Later, in 1990, a selected group representing AASHTO, FHWA, NAPA, SHRP, AI, and TRB engaged in a two-week tour over six different European countries, with the objective to pinpoint potential improvements to asphalt mixtures practices (AASHTO *et al.* 1990). One of the tests and equipment observed during that tour was a wheel-tracking device from Hamburg, Germany, which reportedly had an excellent field-performance correlation. The device was then assigned to Colorado Department of Transportation (CDOT) and FHWA's Turner-Fairbank Highway Research Center to demonstrate and evaluate (Aschenbrener 1995). In their studies, comparing the lab results with known field performance materials, several mixture characteristics were found to be well captured by the test, such as: clay content, dust to asphalt, dust coating aggregate, short term aging, aggregate quality, anti-stripping, and compaction temperature (Aschenbrener and Currier 1993, Aschenbrener *et al.* 1994, Stevenson and Aschenbrener 1994, Aschenbrener 1995). The first studies from CDOT to use the equipment inherited the French plate compactor (pneumatic tire) and kneading compactor to produce beam samples (Aschenbrener 1995). However, the method was swiftly adapted and validated to use Superpave Gyratory compacted specimen (Izzo and Tahmoressi 1999). More recently, Mohammad *et al.* (2016) documented the capabilities of Hamburg devices commercially available in the US, proposing several changes to AASHTO T 324's (2023a) requirements, such as wheel position waveform, frequency, speed, temperature control in water bath, measurement positioning, and results analysis, in order to unify the repeatability and accuracy of the equipment nationwide (Mohammad *et al*. 2016a).

Brief Summary of Studies using the Test Method

Among the literature, Hamburg Wheel-Tracking test have been used to asses both moisture susceptibility and rutting of asphalt materials. Because of this flexibility, using only the rut depth presented by the test can mislead in a rutting interpretation (Yin *et al.* 2014). The comparison between materials can be done regarding the total rut depth at the end of the test (TRD), at a certain number of passes, and using stripping inflection point (*SIP*). These classical approaches have their limitations, and alternatives have been proposed, as rutting-only, and stripping-only related index from the test results (Yin *et al*. 2014, 2020a, Wen *et al*. 2016).

Zhang *et al.* (2021) conducted a study comparing field cores from 50 pavements across US and tested using HWTT procedure. In general, the rut depth from the device overestimates what is observed in the field. A better relation was seen for modified binder mixtures, even though their results showed that the test results alone may not be able to ensure a good field rut ranking. Ranking analysis indicated that choosing a rut depth at a proper number of test cycles would be more appropriate than using the number at 20,000 passes. Besides that, selecting a test temperature based on binder properties, local climate or traffic level would be more useful than applying a single temperature. Finally, a proposed model allying HWTT results with pavement age, AADTT, and temperatures of the pavement presented a good prediction of field rutting (R^2=0.79), with more significant effect from the first three (Zhang *et al.* 2021). Other influences in HWTT results are related to impacts of volumetric, binder properties, and additives. Table A.11 summarizes some findings reported in literature.

Table A.11. Summary of Findings using HWTT Method.

Topic	Findings	Reference
Binder effect	Lower J_{nr} and higher $\|G^*\|/sin\delta$ yields smaller rut depths, but the correlation was not high (R^2=0.47 and 0.57)	(Salim *et al.* 2019)
Gradation effect	Coarser gradation decreases rut depth.	(Larrain and Tarefder 2016, Kim *et al*. 2018b)
	Coarser gradation in the coarser fraction and finer gradation in the fine fraction increases resistance. An aggregate skeleton with uniform gradation in each size shows a lower rutting resistance.	(Lv *et al.* 2020)
Volumetric effect	Excessive and inadequate binder content reduces rutting resistance.	(Lv *et al.* 2020)
RAP content	Mixtures with high RAP replacement ratio have lower rut depths.	(Safi *et al.* 2019)
WMA	Adding WMA agent can reduce the combined effect of rutting and stripping resistance,	(Zhao *et al.* 2013)

High Temperature IDT

Method and Calculation Description

The IDT high temperature (IDT-HT) test has been recently discussed as a simple alternative for rutting criteria. The advantages of this method are that it requires inexpensive equipment and the test can be finished in less than one minute. The samples are produced directly from the gyratory compactor at around 7% air voids, with no cutting, coring or gluing afterwards. To condition, it should be enveloped in plastic and placed in water for 2 hours, at 10°C lower than the 50%

reliability level 7-day maximum temperature. An indirect diametric load is applied using a constant rate of 50 mm/min. The results are described in terms of the peak stress strength reached during the test. Usually, Tensile Strength Ratio standard, AASHTO T 283 (2022a), is referred to as the test guidelines for this indirect tension test. Because of its simple index, an easy pass/fail criterion is reported as mixture criteria in most states that apply this test methodology, being well related to other rutting tests (Christensen and Bonaquist 2002). It is also purported to be promising especially for quality control/quality assurance testing, since it can also be used to test very thin cored specimen (Christensen *et al.* 2000).

Historical Development

The indirect tensile test idea was originally developed for portland cement concrete and known as the Brazilian test. The test was adapted during the 1960's to evaluate anisotropy and tensile strength of asphalt mixtures (Livneh and Shklarsky 1962). More recently, researchers have found that its results correlate well with the Mohr-Coulomb failure theory for shear strength and that it could be used to determine cohesion of asphalt mixtures (Christensen *et al.* 2004). Using a critical representative high temperature, and a monotonic loading condition that fits what is expected from field, it was found that the test could evaluate rutting resistance of asphalt materials and correlated well with other well-established tests and field data (Christensen *et al.* 2004). Christensen and Bonaquist (2002) further argued for the IDT-HT test based on the idea that the critical stress state for rutting would not occur directly under the center of tire (where hydrostatic like forces exist), but at the edge, where the confining stresses are similar to the failure in a typical IDT-HT.

The first studies suggested that a test done at loading rate about 3.75 to 7.50 mm/min, and 15 to 20°C below the critical temperature, would better mimic the stresses response at a critical high temperature (Christensen *et al.* 2000, 2004). Further studies improved the test methodology to a temperature 10°C below the critical 20-mm depth temperature, estimated from LTPPBind and used a loading rate of 50 mm/min, reducing drastically the testing time and still achieving good relations with expected trends (Mohseni *et al.* 2005, Christensen and Bonaquist 2007, AAT 2011). This new setup enabled the specimen setup and test to be completed in two minutes and eliminated the need to use a temperature-controlled chamber.

Brief Summary of Studies using the Test Method

As cited before, the displacement loading rate and specified temperature used for IDT-HT have changed over the years. Table A.12 synthetizes some findings over the literature using the test methodology. Informing which temperature and loading rate were used. Other rutting tests are more commonly and used as comparison for mixture properties.

Table A.12. Summary of Findings using the IDT-HT Method.

Topic	Temperature; Loading rate	Findings	Reference
Binder effect	(35°C; 3.75 mm/min)[a]; (44°C; 50 mm/min)[b]	Higher PG grade leads to higher peak stress result.	(Christensen *et al.* 2000, Bennert *et al.* 2018)
Gradation/ volumetric effect	(44°C; 50 mm/min)[b]	Test results are sensitive to volumetric, and composition effects	(Bennert *et al.* 2018)
Other tests	(35°C; 3.75 mm/min)[a]	Good correlation with RSCH (R^2=0.8); and ALF rutting results	(Christensen *et al.* 2000, Stuart *et al.* 2000)
	(44°C; 50 mm/min)[b]	Good correlation with APA at 64°C (R^2=0.8)	(Bennert *et al.* 2018)
	(49°C; 50 mm/min)[b]	Same ranking as APA, and Confined FN	(Meroni *et al.* 2021)
	(50.2°C; 50 mm/min)[b]	Reasonable correlation to HWTT (R^2=0.63)	(Yin *et al*. 2020b)

[a] Test performed at 20°C below 50-mm depth critical temperature, 50% reliability; [b] Test performed at 10°C below 20-mm depth from LTPPBind 3.1, 50% reliability.

IDEAL-RT

Method and Calculation Description

Similar to the IDT-HT and the previously described IDT-CT test procedure, IDEAL-RT has been proposed as a simple alternative for rutting characterization. Like IDT-HT, it is based on shear strength results, described as an RT_{Index}, expressed in Equation (10), where the shear strength is defined by Equation (11).

$$RT_{Index} = \left(6.618 \times 10^{-5}\right) \times \tau_f \tag{10}$$

$$\tau_f = 0.356 \frac{P_{\max}}{t \times w} \tag{11}$$

Where RT_{Index} is the rutting tolerance index, τ_f is the shear strength in Pascals, $P_{\max}$ is the maximum load in Newtons, t is the specimen thickness in meters, and w is the width of the upper loading strip (0.0191 m).

The test was designed specifically for BMD implementation to be a simple, fast, and inexpensive test that could be finished in less than one day. Differently from IDT-HT, the temperature is set to 50°C and a unique apparatus that induces shear rather than biaxial tension is used (Figure A.6). This supporting cradle has a concave surface, with radius curvature same as the specimen diameter of 150 mm. The specimen should be produced at a 7±0.5% air void and has the same dimensions as the IDT-CT (150 mm diameter and 62 mm height). Specimens can be conditioned using a water bath or temperature chamber and it is stated in the literature that no great difference in results to each other (Zhou *et al*. 2021b). The test is conducted using a 50 mm/min loading rate, and can relate both cohesion and friction angle from shear strength of asphalt mixes (Zhou 2021). More specific information can be seen in ASTM D8360.

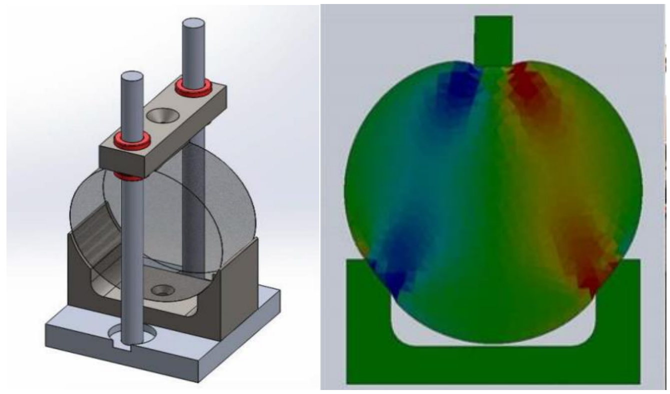


Figure A.6. IDEAL-RT Test apparatus and stresses (Yin *et al*. 2020b, Zhou 2021).

<u>Historical Development</u>

IDEAL-RT test was developed with quality assurance and control in mind. Regardless of how well a mix can be designed, if its quality is not well controlled during production, the field performance will be diminished (Zhou 2021). Some tests are great options for BMD, but may not be practical enough to add as a quality control check. Such a test should be practical enough to be done in less than one day, have a good correlation to field, and have simple testing and sample preparation. In view of these issues, IDEAL-RT was proposed as an alternative to Quality Control framework, with some case studies described in (Zhou *et al*. 2020, 2021b, Zhou 2021).

<u>Brief Summary of Studies using the Test Method</u>

Similar to the previous test, the main findings in the literature relate the IDEAL-RT to other consolidated rutting methodologies, well known for their good correlation to field performance. Yi *et al.* (2020) found a correlation of R^2=0.72 (excluding one outlier) for HWTT and IDEAL-RT results when comparing LMLC specimen to re-heated and non-re-heated PMLC (Yin *et al*. 2020b). Zhou (2021), using a varied set of 23 mixtures containing binder PGs of 58 to 70, 0 to 40% RAP content, 0 to 5% RAS content, and some with rejuvenators added, found a strong correlation with HWTT results with 95% and 98 confidence level (Zhou 2021).

Confined Flow Number (FN)

<u>Method and Calculation Description</u>

Moving to a more fundamental testing procedure, the Flow Number (FN) test uses a repeated loading to induce axial stresses in a cylindrical sample. The effects observed during the test are related to the densification process, even though when the sample is close to the failing point, some shear deformation might be visually seen.

The AASHTO T 378 (2022b) describes how to execute the Flow Number test, in either confined or unconfined conditions. At least three samples shall be cored and cut from gyratory compacted specimen, in order to achieve a 150 mm height and 100 mm diameter condition, with 7±0.5% air voids. The conditioning time for samples is four hours in air chamber at the test temperature. The test is compatible to AMPT. For confined testing, a membrane should encase the sample and for either confined or unconfined testing a greased double-latex friction reducer should be place on both ends of the specimen. The confining stress applied is 30 kPa, with a 600 kPa haversine load of 0.1 seconds pulse, followed by a 0.9 seconds rest. The test temperature is determined by LTPPBind 3.1, 20 mm below surface, with 50% reliability.

During the test, the sample accumulates permanent deformation strains through three stages: first the specimen consolidates (the primary region), second there is a stable growth of deformation (the secondary region), and lastly a rapid increase in deformation (the tertiary region). The number of cycles for the sample achieve the tertiary stage is defined as Flow Number. The standard specifies a fitting function, the Franken Model (1977), shown in Equation (12). The number of cycles needed to reach the flow number is material dependent and for highly rut resistant mixtures the number of cycles can be high. Usually, agencies prefer to set a limit-based criterion for different traffic conditions.

$$\varepsilon_p = AN^B + C\left(e^{D\times N} - 1\right) \quad (12)$$

Where ε_p is the permanent axial strain measured during the test, N is the number of cycles, and A, B, C, and D are fitting coefficients for each mixture.

Historical Development

The development of the FN test began out of a need to add to Superpave mix design and incorporate simple performance test methodologies. In 1999, NCHRP Project 9-19 evaluated several tests procedures, from triaxial dynamic modulus, shear tests, triaxial static creep, and triaxial repeated load test. The idea behind this effort was to find a test method that could accurately and reliably measure mixture characteristics over a diverse range of traffic and climate conditions (Witczak *et al.* 2002). From all the triaxial repeated loading tests, the highest ranked parameter with the best statistical measures was the FN. At first, the testing temperatures were 38°C and 54°C, so that results could be compared to dynamic modulus conditions. For confined state, the applied confining stress was 138 kPa, while the axial stresses varied from 827 to 965 kPa. For the test pulse and rest time, the same were kept since the beginning (Witczak *et al.* 2002). Later studies used different confinement stresses: 6.9 to 69 kPa, and axial stresses: 207, 523, and 827 kPa trying to achieve a representative stress state under the edge of the tire (Gibson *et al.* 2009, 2012). Also, calibration factors using the test method and different temperatures were implemented in MEPDG mechanistic-empirical, actual AASHTO PavementME. The procedure enables one to predict permanent deformation over time using lab-measured coefficients from materials and focused on modelling rutting using the secondary stage (Von-Quintus *et al.* 2012).

Brief Summary of Studies using the Test Method

The confined FN test is somewhat established in literature in terms of procedure. Some variations are found still in the selected stresses, temperature, depending on the standard applied. Some findings in literature using the test protocol are presented in Table A.13.

Table A.13. Summary of Findings using the Confined FN Method.

Topic	Temperature; Axial | Conf. Stress	Findings	Reference
Binder effect	(38°C, 54°C; 138, 207, 828 kPa | 69, 138 kPa)	FN increases with increased viscosity	(Rodezno *et al.* 2010)
	(55°C[a]; 600 kPa | 30 kPa)	FN increases with PG grade	(Islam *et al.* 2019)
	(50°C; 400 kPa | unconfined)	Good correlation between FN and $J_{nr3.2}$	(Salim *et al.* 2019)
Volumetric effect	(38°C, 54°C; 138, 207, 828 kPa | 69, 138 kPa)	FN sensitive to increased V_{be}, air voids, gradation, and asphalt content	(Rodezno *et al.* 2010)
	(55°C[a]; 600 kPa | 30 kPa)	FN increases with increased V_{ba}, air voids, VMA, VFA, and asphalt content, and N_{design}	(Islam *et al.* 2019)
Other tests	(50°C; 400 kPa | unconfined)	Good correlation between HWTT and FN (R^2=0.79)	(Salim *et al.* 2019)

Stress Sweep Rutting (SSR) Test

Method and Calculation Description

The Stress Sweep Rutting (SSR) Test follows the basic structure of the FN and applies repeated loading into confined specimen to obtain cumulative strain results. AASHTO TP 134 (2023d) describe its steps to test cored and cut 100 mm diameter and 150 mm tall cylindrical specimens using the AMPT device. The improvement from previous protocol is that a set of temperatures, axial stresses, and loading time are chosen in a way to enable rutting prediction over any given traffic and climate conditions, with the same number of specimens as FN.

Two temperatures are selected for testing, one low (between 17 and 32°C), and one high, depending on the project location. The test is separated into three sets of 200 cycles at 69 kPa confinement stress, reducing greatly the amount of testing time. For the high temperature, the axial stress increments for each 200 cycles are 689, 423, and 896 kPa, with 0.4 seconds of loading time and 3.6 seconds as rest, while for the low temperature, the stresses are 483, 689, and 896 kPa, with 0.4 seconds for loading time, and 1.6 seconds as rest. The first loading pulse from high temperature test are set as reference, while the other are shifted to it, forming a universal model for each material.

During the test, only permanent deformation is recorded, taking actuator load cell displacement. The calculations can be performed automatically using FlexMAT™ Rutting spreadsheet, or using the following Equations (13) - (15). An index is also associated with the SSR, the rutting strain index or RSI, so that users can leverage the test result without having to do full simulations with the FlexPAVE™ software. The idea behind the index is to compare the average permanent strain of an asphalt layer to its total thickness in a standard structure, over a given climate location (Ghanbari *et al.* 2022). A mixture with higher RSI is more likely to have lower rutting resistance. The calculations are automatically done using FlexMAT™ Rutting spreadsheet.

$$\varepsilon_{vp} = \frac{\varepsilon_0 \times N_{red}}{\left(N_1 + N_{red}\right)^{\beta}} \tag{13}$$

$$N_{red} = 10^{p_2} \times 10^{-0.877 \times D} \times N\left(\xi_p\right)^{p_1} \left(\frac{\sigma_v}{P_a}\right)^D \quad (14)$$

$$D = d_1 T + d_2 \quad (15)$$

Where the reference condition is the first set of high temperature, ε_{vp} is the viscoplastic strain; ε_0, N_1, and β are obtained through numerical optimization, N is the number of cycles for a certain loading condition, ξ_p is the reduced load time, T is the test temperature, σ_v is the vertical stress, P_a is the atmospheric pressure (101.325 kPa), d_1 and d_2 are the vertical stress shift-factor coefficient, and p_1 and p_2 are reduced load time shift factor coefficients.

Historical Development

Due to known strain accumulation response dependency from axial load level, history, pulse time, rest period, and temperature in confined state, a broader procedure was needed to simulate strain state found in field (Kandhal and Cooley 2003). The background of SSR interconnects with the first studies that defined the triaxial repeated load test. However, the focus of the early development work that led to the SSR was to describe the strain accumulation prior to failure. The development arc began with efforts during NCHRP 9-19 to understand the role of loading time and stress levels (Gibson 2006). Later, Yun (2008) found that the rest period had a significant role in affecting the amount of permanent strain that accumulated. Yun described this effect using a backstress concept while Subramanian (2011) used a rate hardening plasticity and viscoelastic relaxation concept. The primary offshoot of both of these efforts was a realization that instead of using multiple temperatures, and stress levels, a protocol that could interconnect these variables in less test effort was needed. A mention should be made regarding the Triaxial Stress Sweep (TSS), which preceded SSR and was developed following the Yun and Subramanian efforts, which used square loading pulses. It used a set of three temperatures, with increasing deviatoric stress amplitudes, different rest periods to widen the evaluated range, totalizing at least eight specimens with mounted LVDTs to model a single mixture (Choi and Kim 2013). After the testing protocol, the called shift model captures the effects of deviatoric stress, load time and temperature on accumulated strain of asphalt mixture, shifting any given condition to the reference 689 kPa as deviatoric stress amplitude using an incremental model described in the equations above (Choi and Kim 2014). Afterwards, studies worked to improve the overall practicality of the test. These enhancements included using only one high and one low temperature and removing the necessity of LVDTs. These changes reduced the number of samples to only two replicates for each experiment, and from 16 hours to only 6 hours of testing (Kim and Kim 2017).

Brief Summary of Studies using the Test Method

The test procedure was improved over time until its final version of SSR. In this process, it was validated using known rutting prediction tools (Kim and Kim 2017). Some literature found great correlation with expected trends, listed in the Table A.14. Since its capabilities to assess rutting are related to performance predicting, there are a number of studies using its results to benchmark mixtures, and to find an optimum volumetric characteristics and field density for a given mix (Wang *et al.* 2019, Jeong *et al.* 2020, 2021).

Table A.14. Summary of Findings using the SSR Method.

Topic	Findings	Reference
Binder effect	SSR results indicate better rutting resistance for higher PG	(Ghanbari *et al.* 2022)
	RSI increase with decreased binder content	(Veeraragavan *et al.* 2022)
	$J_{nr3.2}$ is well correlated with FlexPAVE predictions	(Barros *et al*. 2022a)
Volumetric effect	RSI decreases with coarser gradation, higher asphalt content, higher air voids	(Ghanbari *et al.* 2022)
	RSI decrease with increased NMAS	(Veeraragavan *et al.* 2022)
RAP addition	RSI decreases with RAP content increasing	(Ghanbari *et al.* 2022)
	Increasing RBR does not affect rutting that much	(Chkaiban *et al.* 2022)
Other tests	No apparent relationship between RSI and HWTT	(Chkaiban *et al.* 2022, Veeraragavan *et al.* 2022)
	No direct comparison between RSI and unconfined FN at 204 kPa	(Barros *et al*. 2022a, 2022b)

Synthesis of Stability Test Methods

There are several ways to assess an asphalt mixture's stability. To properly relate to what is expected from the material, a reliable field performance correlation should exist, but this method should also be simple enough to conduct that it can be implemented. That being said, the three main categories of rutting tests regarding BMD applications are described in sequence. Table A.15 shows an overview found in literature regarding the six tests presented above, with respect to costs of equipment, sample preparation, and testing time.

The first test type would be characterized as a torture test where a small-scale simulation of a wheel load over a sample is used. This type of method gives results that are fairly simple to understand. Samples can be tested either dry or on wet conditions, conceiving the possibility to also obtain moisture susceptibility inferences. The amount of rutting is dependent on the number of cycles applied and the overall rut resistance of the material. Often times these tests are performed to either critical rutting levels or until a preset number of cycles. These tests generally have (or are believed to have) strong correlation to field performance and so these types of tests are often used as a benchmark when generating new rutting tests methodologies.

The second type of stability methods are the indirect diametric monotonic loading tests. This simple methodology has been discussed often lately because of its simple costs for equipment, and sample preparation. The tests can take less than a day, or even six hours from sample collecting, preparation and testing, which is an acceptable interval for quality assurance and control protocols. The analysis is based on the maximum shear induced in the sample, or on the energy curve for displacement versus shear strength. From the two described tests, cohesion and friction angle are the main characteristics assessed. Also, promising correlations to wheel-loading tests have been shown.

Finally, the third type of test method focuses on mechanistic characterization and the use of test results for performance predictions. The repeated axial tests focus heavily on the stresses directly below the tire but can mimic both densification and shear phenomenon. Usually, there are performance prediction tools associated with the tests, which improve the benefit of the methodology to not only rank materials form their index, but to predict an expected rutting in field. The biggest shortcoming of these types of tests are the cumbersome sample preparation protocols and the total testing time, which can take more than a day. The equipment is also more expensive and the training requirements are greater.

Table A.15. Comparison of Different Rutting Test Methods (Adapted from Zhou *et al.* 2020).

Test name	Test Standard	Parameter | Criteria	Test Temperature	Specimen Preparation	Total Time	Equipment Costs[a]
APA	AASHTO T 340	Rut depth | Pass/fail	Location PG_H at 98% reliability	Four samples, no cutting, nor gluing	Two days	>$100,000
HWTT	AASHTO T 324	Rut depth, *SIP* | Pass/fail	50°C	Four samples, one cut per specimen	Two days	$50,000
IDT-HT	AASHTO T 283[b]	Peak Strength | Pass/fail	Location PG_H -10°C	Three samples, no coring/cutting	<One day	<$10,000
IDEAL-RT	ASTM D8360	RT_{Index} | Pass/fail	50°C	Three samples, no coring/cutting	<One day	<$10,000
Confined FN	AASHTO T 378	*FN* | Pass/fail or predictive model with more tests	Location PG_H with 50% reliability	Three samples, one coring per sample, two cuts per sample, membrane	Four days	$85,000
SSR	AASHTO TP 134	*RSI*, p_1, p_2, d_1, d_2 | Pass/fail or predictive model with FlexPAVE™	High and Low Temperature (based on PG_H)	Three samples, one coring per sample, two cuts per sample, membrane	Four days	$85,000

[a] Costs are for the equipment used for the given test alone. As described in Section 2.4 some equipment can be used for several different tests. Final costs may be less than the amounts listed in this table because of these multipurpose options

[b] IDT-HT uses an adaptation of AASHTO T 283 (2022a)

Integration of Performance Test into QA Practices

The previous sections covered the different test protocols that have been proposed for the design asphalt mixtures based on BMD principles. These protocols are orientated at improving the quality of the asphalt mixtures in terms of rutting and cracking performance. Although necessary, an adequate mix design alone is not a guarantee of satisfactory pavement performance and efforts must be done to integrate performance tests with specifications that assure that reasonable process and procedures are being followed in the production of the final engineered product. Depending on the quality characteristics used for acceptance, these specifications are often called quality assurance (QA) specifications or performance-related specifications (PRS).

The integration of performance test into QA practices can represent a challenge for state DOTs due to the adjustments and resources needed to conduct performance tests in a routine basis. Further, guidance is needed for determining minimum sampling frequency needed (and risks associated), and for sampling, fabrication and testing procedures that could be unknown to DOTs personnel. These challenges motivated the development of the ongoing TFRS 01: *Quality Assurance Aspects of Performance Related Specifications* and NCHRP 10-107: *Guide for Implementing Performance Specifications*. The general objective of TFRS 01 is to integrate PASSFlex^{TM} (system of tools based on fundamental engineering properties for PRS) within a QA system to ensure long-term asphalt performance and reduce life cycle costs (NCAT 2021). Specifically, TFRS 01 addresses:

(d) the use of the cyclic fatigue S_{app} and rutting *RSI* test parameters (the reader is referred to section *2.1.5* and *2.2.6* for more details), index thresholds, and acceptance limits in support of performance engineered mix design approaches and to facilitate further implementation of the tests and performance predictions
(e) material selection and mixture design changes that can impact the test results (cyclic fatigue, SSR, and their index parameters) and trends associated with owner agency specified performance thresholds, and
(f) the major elements of a QA system per TRB E-Circular 235 and associated buyer/seller and payment risks (TRB 2018).

In TFRS 01, three different sets of laboratory experiments are conducted to answer the following questions needed to develop guidance on integrating PASSFlex^{TM} suite in a QA framework.

(a) What are the differences between properties of lab-mixed, lab-compacted (as-designed) and plant-mixes, lab-compacted (as-constructed)?
(b) How does production variability affect the properties of the mixtures and resultant S_{app} and RSI and how these affect QA and PRS?
(c) How do constituent material changes affect S_{app} and RSI?
(d) How sensitive are S_{app} and RSI index parameters to changes in mixtures within the same category?

As mentioned before, TFRS 01 is an ongoing research project, and this literature review will be updated with the major findings and guidelines proposed therein. According to public information, NCHRP 10-107 researchers are developing a guide that will assist state DOTs with performance specification implementation including integrating tests into QA systems. The guide will "specifically address, but not be limited to, implementation activities such as using pilot projects and shadow specifications, establishing appropriate control and specification limits, gaining buy-in from agency and industry personnel, and managing risk." (NCHRP 2022).

Caltrans requires performance testing for both mix design and quality assurance on specific mix types: Type A HMA (maximum aggregate blend of 25% RAP) and Rubberized Hot Mix Asphalt-Gap-Graded (RHMA-G). In order to do so, the testing frequency for volumetric properties (air voids N_{design}, VMA, and dust proportion) was reduced to allow for the inclusion of the Hamburg Wheel-Tracking Test and TSR. Both tests are required once per 10000 tons, or once per project, and performance thresholds are the same for mix design and quality assurance.

New Jersey uses BMD Approach A and Approach B, where volumetric requirements are met first, but optimum binder content is selected based on cracking and rutting test results on four specialty mix types: binder rich intermediate course (BRIC), high performance thin overlay (HPTO), bridge deck waterproof surface course (BDWSC), and hot mix asphalt high RAP (HRAP) that contains more than 20% RAP by aggregate blend (NJDOT, 2019). For quality assurance, sampling should occur normally for volumetrics (every lot of 3500 tons, five random samples or once per 700 tons); however, for performance testing, the first sample comes from the test strip or first lot and is then sampled every lot for HPTO and HRAP HMA, and every second lot for BRIC and BDWSC.

The required performance testing include APA and Overlay Tester (OT) for HPTO, BRIC, and High RAP HMA mix types. For BDWSC mixtures, APA and BBF testing are required. Performance thresholds are slightly relaxed compared to mix design requirements for some tests (i.e., APA rut depth for HPTO is relaxed from 4 mm to 5 mm). If a lot fails to meet the requirements, NJDOT utilizes a pay adjustment table, which can go up to requiring removal and replacement in worst scenarios.

For both mix design and quality assurance, contractors should submit samples for volumetrics verification, samples at specified height for APA and OT testing, boxes of loose-mix to produce BBF specimens, and spare boxes of loose-mix material. The agency is in charge of producing the BBF samples, cutting the OT samples, and conduct all performance testing.

Integrating a similar system requires addressing several key challenges. NCDOT's current QA frequency (once per 750 tons) means that adding performance testing would significantly impact turnaround times. A potential solution is to review and possibly reduce sampling frequencies (lot sizes) for performance tests, similar to the approach used by other states.

Test Costs

One important factor for conducting BMD tests is the cost. The numbers showed here represents a sample of national territory sellers, from the start of 2023. None of the displayed values cover taxes, crating, shipping, discounts that may apply, installation, yearly calibrations, nor personnel training. With respect to personnel costs it is noted that as test complexity increases the cost of training would increase, but that resources exist for all of the tests discussed here and that examples of successful training regimes also exist for each of the tests.

In addition, the costs shown in this section do not include those associated with sample compaction (ovens, compactors, mixers, etc.) since these pieces are generally already available in labs that would be performing these tests. It is noted that most tests use a Superpave gyratory compactor and with the exception of AMPT based tests standard height gyratory compactors are sufficient. It should be noted that the Bending Beam Fatigue Test requires beam samples rather than cylinders and that while APA and HWTT can also be tested with beam specimen, but recent literature is predominantly using gyratory samples. Finally, for CF, only 38 mm diameter samples are covered in this document since the focus of BMD tests is surface and (in some cases) intermediate layer mixes with NMAS up to 19 mm (AASHTO 2024b).

Table A.16 summarizes the different steps that may be necessary to prepare test specimens after the samples have been compacted and which ones are used for the different test methods. It is important to mention that some equipment, like the cutting saws, can be used to prepare different test specimens. That being said, the cost in the last column shows the standalone price of purchasing the equipment for only the specific test. A combination of more than one of these tests can reduce significantly the values. In some cases, a range in values were found from the cost study and in these cases, the value shown is an average of the maximum and minimum range detailed below.

Table A.16. Summary of Additional Steps for Specimen Preparation Prices.

Test name	Coring	Cutting	Notching	Gluing	Membrane	Stand-Alone Price
I-FIT	-	yes	yes	-	-	$9,000
SCB	-	yes	yes	-	-	$9,000
IDT-CT	-	-	-	-	-	-
Cantabro	-	-	-	-	-	-
CF	yes	yes	-	yes	-	$10,615[a]
BBF	-	yes	-	-	-	$5,850[b]
OT	-	yes	-	yes	-	$8,000[a]
APA	-	-	-	-	-	-
HWTT	-	yes	-	-	-	$5,050
IDT-HT	-	-	-	-	-	-
IDEAL-RT	-	-	-	-	-	-
Confined FN	yes	yes	-	-	yes	$9,450
SSR	yes	yes	-	-	yes	$9,450

[a] Considering a kit of plates and gluing for three specimen
[b] Additional considerations should be made for beam sample compaction.

For coring, a commercially available machine costs around $4,000, with each coring bit being sold separately. Custom built setups can also be constructed at a substantially lower price if a machine shop is available to weld or bolt together an appropriate frame. In case of the CF specimen, the lab would need to have a 38 mm coring bit (about $315), while for confined FN and SSR, a 100 mm coring bit would be needed (about $415). For cutting, two types of equipment have been used. Larger capacity equipment ranges from $3,800 to $5,500 and requires a 16″ diameter saw blade (about $480). More expensive options are available that include automatic feed and sizing options. Smaller capacity saws are needed for more precise cuts on smaller samples and these generally cost between $3,000 and $5,000 and require a 10" saw blade (about $110). The larger capacity saw is recommended whenever multiple geometries may be used. The smaller capacity saw is recommended for notch-based tests and for cutting 38 mm diameter specimens. An addition consideration should be made regarding clamps to hold the specimen, with 4 in. or 6 in. clamps (about $350) and HWTT trimming jig ($400) to a more sophisticated clamp table (about $1,200), needed for OT and BBF precise cutting. These trimming jigs can be optionally be manufactured according to the laboratory needs.

Lastly, some more complex test might require gluing specimens to plates. Both CF and OT require plates, targets, and gluing jigs, which are often reported by sellers as optional items when a machine setup in bought. For CF, each specimen uses six targets (about $35) and two plates (about $330). These accessories are generally sold in a three-specimen set. The gluing jigs for targets and plates cost approximately $600 each. Each specimen needs to be kept in the gluing jib for about 30 minutes, but the technician time to glue the specimens is usually about five to 10 minutes and. users can perform other activities while gluing takes place. The glue itself can be bought in small packages at an average cost of about $5 per three specimens. The OT does not require targets, but more total glue is needed to attach specimens to the plates. Each OT specimen uses two plates ($320), and the gluing jig plus needed accessories for three specimen costs about $1,450. Since there is a bigger surface to be glued, a slow-bonding glue might be required, justifying the need of more than one jig. The glue itself can be bought in one-per-sample packages (about $14 per sample) or in bigger quantities.

Confined tests, like confined FN and SSR, require that a membrane cover the sides of specimen, while a double-greased membrane is placed on top and bottom of specimen, to ensure no friction with the testing plates. A kit with membranes can be bought in large quantity packages, and added to the grease, which is used in very small amounts, can cost about $7.40 per specimen. Membranes can be reused but careful training of testing technicians is needed to avoid testing mistakes.

With respect to testing equipment, there are three main categories: stand-alone versions, electromechanical loading-frame compatible tests, and AMPT compatible. This consideration should be carefully taken into account, since some tests can be performed in more than machine setup. For example, kits exist for some tests like the SCB, I-FIT, and IDT-CT/IDEAL-RT that can leverage modern era Marshall load frames. Also, at least one AMPT manufacturer also includes a range of kits for multiple tests (i.e., CF, SSR/FN, IDT based tests, OT, etc.). A summary of testing equipment is showed in Table A.17.

Some tests such the Cantabro, APA, and BBF require standalone machines while others can be stand alone or tested with a kit attached to an AMPT and/or an electromechanical load frame. For those tests that require a stand-alone machine, options range from an LA Abrasion machine (Cantabro) to a completely stand-alone machine (BBF, APA, HWTT). It is noted that the stand-alone option for the BBF, a four-point bending apparatus (starting from $31,200) and an additional temperature chamber (approximately $10,000) is needed. Universal Testing Machines can also support a four-point accessory to perform this test, and perform any AMPT test given the right accessories, but at higher prices. Many general-purpose testing machines can cost approximately $150,000 or more. For the OT there is a stand-alone machine available in market that costs about $55,000, but a kit is available for the AMPT at a cost of approximately $4,000. For the APA, there are two main machines available, one for two-simultaneously tests (the APA Jr.) and one for three simultaneous tests (the standard machine). The prices for the first range from $66,000 to $72,000, while the second can be about $125,000. HWTT setups usually involve machines that run two tests simultaneously and costs range from $50,000 to $63,000. Some vendors guarantee that both tests can be performed in their equipment; however, only limited evidence was found of researchers actually doing so. The described prices do not include any shipping, installation, training, nor any calibration that may be done yearly.

Table A.17. Summary of Testing Equipment Prices.

Test name	Stand-Alone Machine	EM Load-Frame Compatible	Additional Accessories	AMPT Compatible	Additional Accessories
I-FIT	N/A	yes	$700	yes	$1,000
SCB	N/A	yes	$700	yes	$1,000
IDT-CT	N/A	yes	$550	yes	$600
Cantabro	$8,000	no	-	no	-
CF	N/A	no	-	yes	$6,800
BBF	>$32,200	no	-	no	-
OT	$55,000	no	-	yes	$4,000
APA	>$66,000	no	-	no	-
HWTT	>$50,000	no	-	no	-
IDT-HT	N/A	yes	$550	yes	$600
IDEAL-RT	N/A	yes	$900	yes	$900
Confined FN	N/A	no	-	yes	$305
SSR	N/A	no	-	yes	$305

Due to the recent increased attention given to BMD by state agencies and academia, there are several electromechanical loading frames available for conducting the tests listed in this literature review. A given laboratory should give a closer look to what technologies are associated with the product, what is the minimum loading capacity to apply a specific test, and what is the type of data acquisition. Overall, there is a large span in prices ranging from $4,150 for single purpose machines to $12,000 for machines capable of doing all I-FIT, SCB, IDT-CT, IDT-HT, and IDEAL-RT tests. An external temperature chamber can be used for more efficient testing, but some agencies and labs also use water baths for this purpose. Additionally, each test might require additional accessories to be performed. Sometimes these fixtures are included in the cost of the machine. For SCB and I-FIT, a semi-circular three-point loading jig can cost about $700; for IDT-CT and IDT-HT, an IDT loading jig can cost about $550; and the IDEAL-RT jig can cost about $900 (which requires an IDT loading jig).

Finally, a universal testing machine like AMPT is capable of doing nine out of the thirteen cited tests, but comes at a substantially higher up front cost and maintenance cost. Recently, due to improvement in technology and controlling systems, a reduction in maintenance and machine prices has been noted. Even so, AMPT prices start from $72,000 (not considering any accessory, installation, shipping, training, and calibration). Like with the electromechanical machines, each test should require its own accessory kit. For SCB and I-FIT, the three-point jig can cost about $1,000; the IDT jig can cost about $600 for IDT-CT and IDT-HT; and the IDEAL-RT jig can be about $900.

The CF requires a general-purpose load frame or an AMPT. With an AMPT a set of spacers related to specimen height of 110 mm is needed to ensure that the machine actuator will have sufficient gauge length (about $780). Additionally, a set of six clamps and three LVDTs to read on-specimen stresses and strains are needed. These accessories can cost about $5,300. For the OT, the LVDT is not mandatory, since the actuator can be used to estimate stresses, nonetheless, the loading frame kit for this test can cost about $4,000. Finally, for confined FN and SSR, the actuator can be also used to estimate material response, so only the set of confining plates with drainage pipe is required, that costs about $305.

Integration of Performance Tests into Mix Design

In 2015, the former Federal Highway Administration (FHWA) Expert Task Group (ETG) on Mixtures and Construction formed a BMD Task Force and defined BMD as "an asphalt mix design using performance tests on appropriately conditioned specimens that address multiple modes of distress taking into consideration mix aging, traffic, climate, and location within the pavement structure". Some reasons were highlighted for using the BMD approach: evaluating the designed mixtures relative to performance rather than volumetric properties alone, addressing the increased use of recycled materials, and evaluating additives and modifiers that are not directly considered with the traditional volumetric mix design (FHWA 2016).

Originally, the Superpave mix design system included different requirements depending on the traffic. Lower traffic roads were to be designed based on volumetric properties and higher traffic roads were to be initially designed based on volumetric properties, but then adjusted based on performance testing. Some performance tests were proposed by the Strategic Highway Research Program (SHRP) but were not implemented except for a few projects because they were not considered practical for routine use. In the early years of Superpave implementation, rutting was the primary focus, and it was partially addressed by the aggregate and binder grade requirements. In addition, many state highway agencies added rutting test requirements to their specifications. This reduced pavement rutting issues but ultimately led to cracking becoming the primary form of distress. Therefore, increasing concerns regarding cracking and durability issues motivated the asphalt pavement industry to use BMD as a new approach for design and production acceptance of asphalt mixtures (NAPA 2021).

In 2020, the FHWA conducted virtual visits to some State DOTs that were early adopters of BMD to learn more about the details and implementation efforts of BMD. Successful practices reported from these visits were gathered and integrated into an overall BMD implementation process as part of mix design approval and quality assurance. Listed below are the eight major tasks that were suggested for successful BMD implementation (FHWA 2022, Hajj *et al.* 2022).

1. Understanding the "why" and benefits of BMD
2. Overall planning
3. Selecting performance tests
4. Performance testing equipment: acquiring, managing resources, training, and evaluating
5. Establishing baseline data
6. Specifications and program development
7. Training, certification, and accreditation
8. Initial implementation into engineering practice

These eight tasks along with their sub-tasks and brief descriptions are shown in Table A.18 below. More details and guidance on executing these tasks can be found in the tech brief and mentioned paper (FHWA 2022, Hajj *et al.* 2022).

Table A.18. Eight Potential Tasks for BMD Implementation (FHWA 2022).

	Task	Sub-Task	Description
1	Motivations and Benefits	–	Understand why the agency wants to pursue BMD and what the potential benefits would be
2	Overall Planning	2.1	Identification of Champions
		2.2	Establishing Stakeholders Collaboration
		2.3	Doing Homework
		2.4	Establishing Goals
		2.5	Mapping the Tasks
		2.6	Identifying Available External Technical Information and Support
		2.7	Developing an Implementation Timeline
3	Selecting Performance Tests	3.1	Identifying Primary Modes of Distress
		3.2	Identifying and Assessing Performance Test Appropriateness
		3.3	Validating the Performance Tests
4	Performance Testing Equipment: Acquiring, Managing Resources, Training, and Evaluating	4.1	Acquiring Equipment
		4.2	Managing Resources
		4.3	Conducting Initial Training
		4.4	Evaluating Performance Tests
		4.5	Conducting Inter-Laboratory Studies
5	Establishing Baseline Data	5.1	Reviewing Historical Data & Information Management System
		5.2	Conducting Benchmarking studies
		5.3	Conducting Shadow Projects
		5.4	Analyzing Production Data
		5.5	Determining How to Adjust Asphalt Mixtures Containing Local Materials
6	Specifications and Program Development	6.1	Sampling and Testing Plans
		6.2	Pay Adjustment Factors (If Part of the Goals)
		6.3	Developing Pilot Specifications and Policies
		6.4	Conducting Pilot Projects
		6.5	Final Analysis and Specification Revisions
7	Training, Certifications, and Accreditations	7.1	Developing and/or Updating Training and Certification Programs
		7.2	Establishing or Updating Laboratory Accreditation Program Requirements
8	Initial Implementation	–	–

Review of AASHTO PP 105

The formal AASHTO process for balanced mixture design is outlined in AASHTO PP 105. This standard highlights four different approaches (A through D) for balanced mix design (AASHTO 2024a).

Approach A: Volumetric Design with Performance Verification - Approach A starts with obtaining the optimum asphalt content following the standard volumetric mix design method. Selected performance tests are then performed on the optimum asphalt content design to assess its resistance to rutting, cracking, and moisture damage. The design is accepted as a BMD and production can start if the mixture passes the selected performance tests criteria; otherwise, the whole mix design procedure is repeated using other materials or different mix proportions until all volumetric and performance criteria are met.

Approach B: Volumetric Design with Performance Optimization - Approach B is similar to Approach A with the exception of it using multiple asphalt contents. In other words, the mix design process is performed, and the optimum asphalt content is obtained; however, the selected performance tests are conducted on the optimum asphalt content and two or more additional contents. The asphalt content that meets all performance tests criteria is selected. In case none of the mixtures satisfies the criteria, the entire mix design process is repeated as in approach A.

Approach C: Performance-Modified Volumetric Mix Design - Approach C begins with the current volumetric mix designs methods to obtain an initial design. The designed mixture is then subjected to performance testing and the initial mixture components such as aggregate proportions and binder content are adjusted to satisfy performance tests. The final design has to meet performance testing criteria but not necessarily all volumetric criteria.

Approach D: Performance Design - Approach D is primarily focused on performance analysis. The mixture components are established and adjusted following performance analysis. Volumetric properties may be checked after the designed mixture meets laboratory performance tests criteria.

State Efforts Integrating BMD

According to the National Asphalt Pavement Association (NAPA 2020), 14 State DOTs have already begun the implementation of BMD and 16 other State DOTs are in the process of pre-implementation. Table A.19 shows the BMD approach and the tests selected by the agencies that have started the implementation of BMD. In this table, the states and state agencies are listed in the first two columns, followed by the implemented BMD approach(s) according to AASHTO PP 105 (2024a) and the selected rutting and cracking test(s). The last column states whether these tests are performed for production acceptance. This column may state that testing is required for production acceptance in terms of rutting performance only (such as Illinois DOT), or both rutting and cracking performance (such as Louisiana DOT). Some state agencies do not require any performance testing during production (such as Oklahoma DOT), and some may require testing for informational purposes such that results would not lead to any payment adjustment or halting of production as these decisions are still made based on the resulting volumetrics (such as New York State DOT).

Table A.19. BMD Implementation for Different States (NAPA 2020).

State	Agency	Approach	Rutting Test	Cracking Test	Performance Testing for Production Acceptance
Alabama	ALDOT	D	IDT-HT	IDT-CT	Yes, "Pass/Fail"
California	Caltrans	C	FN, HWTT	BBF, I-FIT	Yes, HWTT for "Pass/Fail"
Illinois	IDOT	A	HWTT	I-FIT	Yes, HWTT for "Pass/Fail"
Louisiana	LaDOTD	A	HWTT	SCB-LTRC	Yes, "Pass/Fail"
Missouri	MoDOT	B	HWTT	IDT-CT	Yes, HWTT for "Pass/Fail", IDT-CT for Pay Adjustment
New Jersey	NJDOT	A & B	APA	OT, BBF	Yes, "Pass/Fail" or Pay Adjustment
New York	NYSDOT	A	IDT-HT	I-FIT, IDT-CT	No, Informational only
Oklahoma	ODOT	C	HWTT	IDT-CT	No
Pennsylvania	PennDOT	C	HWTT	IDT-CT	N/A
Tennessee	TDOT	D	HWTT	IDT-CT	TBD
Texas	TxDOT	A	HWTT	OT, IDT-CT	Yes, "Pass/Fail"
Virginia	VDOT	A & D	APA	Cantabro, IDT-CT	Yes, "Pass/Fail"
Vermont	Vtrans	A	HWTT	I-FIT	Yes, PWL
Wisconsin	WisDOT	A	HWTT	IDT-CT	No, informational testing

Alabama

The Alabama Department of Transportation (ALDOT) currently uses HWTT to evaluate the rutting resistance of all SMA mixes, high traffic Wearing Surface Layers, and Upper Binder Layer mixes (ALDOT 2022a). For ALDOT, mixes using PG 67-22 and PG 76-22 binder must show 10 mm rutting or less at 10,000 and 20,000 cycles, respectively. The ALDOT's most recent special provision of BMD for local roads requires the asphalt contractors to use the High Temperature Indirect Tensile Strength (IDT-HT) per ALDOT 458 and the Alabama Cracking Test (AL-CT) per ALDOT 459 for the evaluation of rutting and cracking resistance, respectively (ALDOT 2022b). The AL-CT test is similar to the IDT-CT described in ASTM D8225. The main difference is the test specimen height, where specimens tested following AL-CT are to have a height of 95 mm instead of 62 mm if the mixture NMAS is 37.5 mm or larger, while in ASTM D8225, the 95 mm height is required for mixtures with NMAS of 25 mm or larger. A minimum IDT-HT strength of 20 psi is required for mix design approval. The performance test criteria for CT_{index} from the AL-CT test are 55, 83, and 110 for low (A/B ESAL Range), medium (C/D ESAL Range) and high traffic (E ESAL Range), as shown in Table 20 (ALDOT 2020). These criteria are required for mix design approval only, as production acceptance is based on asphalt binder content and air voids. However, if two consecutive performance test results fall below the minimum IDT-HT of 20 psi or the CT_{index} limits shown in Table A.20, production will not be accepted until performance tests criteria are satisfied. Note that these special provisions apply only to the trial sections for BMD validation research projects.

Table A.20. ALDOT IDT-CT Criteria (ALDOT 2020).

Minimum Allowable CT_{index}		
ESAL Range	**Design CT_{index}**	**Production CT_{index}**
A/B	55	50
C/D	83	75
E	110	100

California

The California Department of Transportation (Caltrans) has implemented a framework for BMD mixtures that incorporates performance-based specifications and its mechanistic empirical design program, CalME. These mixtures are usually designed and placed on roads with very high traffic volume. For Caltrans, the BMD specifications are applied to plant-produced mixtures, and the performance testing procedure includes the repeated simple shear test, BBF test, and HWTT. The specification criteria for durability assessment were selected based on repetitions to 5% permanent deformation shear strain for the repeated simple shear test and 50% loss of flexural stiffness at 20°C and a test frequency of 10 Hz for the BBF test (Harvey *et al.* 2014). The HWTT is required for production acceptance with the limits shown in Table A.21 (Caltrans 2022).

Table A.21. Caltrans HWTT Thresholds (Caltrans 2022).

Binder Grade	**HWTT minimum number of passes at:**	
	0.5-inch Rut Depth	***SIP***
PG 58	10,000	Report Only
PG64	15,000	
PG 70	20,000	
PG 76 or higher	25,000	

Georgia

The Georgia DOT (GDOT) currently uses Superpave mix design method and requires HWTT for rutting susceptibility and moisture damage assessment with the criteria shown in Table A.22 (GDOT 2021). The specifications do not require any specific tests or limits for fatigue testing other than stating that BBF or any other approved fatigue test can be used for fatigue performance testing if required. However, a recent BMD research study sponsored by GDOT was conducted to benchmark the cracking resistance of plant-produced mixtures being produced in Georgia using the CT_{index}. obtained from the IDT-CT test (Sala *et al.* 2022). The research study results were analyzed and CT_{index} thresholds were proposed for future implementation in GDOT specifications for asphalt mix design approval and acceptance testing. These thresholds were proposed for three mixture types and for two aging conditions classified as reheated mixtures (RH) and critically aged mixtures (CA), using three approaches: minimum value, 25th percentile, and average values. RH mixtures undergo short-term aging for 4 hours at 135°C and CA mixtures undergo long-term aging for 8 hours at 135°C after short-term aging. The recommended minimum CT_{index} values are shown in Table A.23.

Table A.22. GDOT HWTT Criteria (GDOT 2021).

Binder Grade	Mix Type	Number of Passes	Maximum Rut Depth	*SIP*
PG 64-22 and PG 67-22	4.75 mm, 9.5 mm SP Type I, and 9.5 mm SP Type II	15,000	12.5 mm	> 15,000
PG 64-22 and PG 67-22	12.5 mm SP, 19 mm SP, and 25 mm SP	20,000	12.5 mm	> 20,000
PG 76-22	All Mix types	20,000	12.5 mm	> 20,000

Table A.23. GDOT Proposed IDT-CT Criteria (Sala *et al.* 2022).

Mix Type	Minimum CT_{index}					
	Minimum Value		25th Percentile		Average Value	
	RH	CA	RH	CA	RH	CA
Stone Matric Asphalt	75	35	115	65	150	90
Surface	30	20	40	10	55	25
Base/Intermediate	15	10	20	20	30	15

Illinois

The Illinoi DOT requires HWTT, I-FIT, and a modified version of the tensile strength ratio test to evaluate rutting, fatigue, and moisture susceptibility, respectively. In 2015, initial cracking thresholds were recommended based on the correlation between I-FIT tests conducted for varying asphalt mix types and field cracking performance (Al-Qadi *et al.* 2015). In 2019, a long-term aging protocol for implementation of the I-FIT was published (Al-Qadi *et al.* 2019). This effort recommended aging compacted specimens in forced-draft ovens at 95°C for three days. For acceptance, laboratory-produced laboratory-compacted specimens must have a mean *FI* for unaged and oven-aged specimens greater than 8.0 and 5.0, respectively. In addition, plant-produced laboratory-compacted specimens must have a mean *FI* for unaged and oven-aged specimens greater than 8.0 and 4.0, respectively. Contractors were also provided an optional approach to use oven aging for one day at 95°C to screen for problematic mixtures; a minimum *FI* of 6 should be met for this approach (Al-Qadi *et al.* 2019). The thresholds adopted by the Illinois DOT for HWTT and I-FIT tests are shown in Table A.24 and Table A.25, respectively (IDOT 2022). During production, if HWTT or I-FIT testing fail to meet the criteria, the contractor shall resample for testing to be repeated by the DOT and production is continued as long as the other mix criteria are met. If the second set of HWTT or I-FIT tests fail, no additional mixture shall be produced until both tests yield passing results. The HWTT is required to meet the criteria for high ESAL mixtures only.

Table A.24. Illinois Modified AASHTO T 324 Requirements (IDOT 2022).

PG Grade	Minimum Number of Wheel Passes for a 12.5 mm Rut Depth
PG 58-xx (or lower)	5,000
PG 64-xx	7,500
PG 70-xx	15,000
PG 76-xx (or higher)	20,000

Table A.25. Illinois Modified AASHTO T 393 Requirements (IDOT 2022).

Mixture	Short Term Aging, Minimum *FI*	Long Term Aging, Minimum *FI*
HMA	8	5
SMA	16	10
IL-4.75	12	-

Iowa

The Iowa DOT follows the conventional Superpave volumetric approach in designing most of its asphalt mixtures. However, rutting resistance must be evaluated by the contractor or a third-party mix design laboratory using the HWTT for mixtures designed for very-high-volume traffic and/or produced using a certain aggregate mineralogy. The HWTT testing temperature is a function of the asphalt binder high temperature PG. The current specifications require the *SIP* to be at a minimum of 10,000 cycles and 14,000 cycles for plant produced mixtures with traffic designation Standard (S), and High (H) or Very High (V), respectively. Special types of asphalt mixtures may require additional performance testing and acceptance criteria. The Iowa DOT is also currently considering the addition of the disc-shaped compact tension test to evaluate mixture resistance to thermal cracking in their BMD efforts (West *et al*. 2018a).

Louisiana

The Louisiana DOT has implemented BMD Approach A using conventional volumetric criteria along with HWTT (Louisiana refers to this as loaded wheel tracking (LWT), but their specifications indicate that the test is run according to AASHTO T 324) and SCB tests to evaluate rutting and intermediate temperature cracking, respectively. The roadway acceptance test sampling consists of collecting five random cores from five sublots, totaling 25 random cores. Some of these cores are subjected to LWT and SCB testing, while others undergo density measurement and verification (Mohammad *et al*. 2016b). The current rutting specifications require conducting the HWTT test at 50°C and mixtures are accepted if the rut depth at 20,000 cycles is lower than 6 mm for mixtures containing polymer and crumb rubber modified asphalt binders (Level 2 traffic) and lower than 10 mm for mixtures containing unmodified binders (Level 1 traffic). For cracking, the current specifications require the SCB fracture energy (SCB-LTRC) at 25°C to be greater than 0.5 kJ/m^2 and 0.6 kJ/m2 for Level 1 and 2 traffic, respectively (Cooper *et al.* 2016). Efforts related to assessing the changes in test parameters from different specimen types (mix design vs. plant produced vs. field cores), developing an accelerated aging protocol, and employing the SCB test into quality control are ongoing.

Maine

In 2019, the Maine DOT initiated a research study to evaluate the cracking and rutting resistance of asphalt mixtures using several performance tests, including the HWTT, IDT-CT, CF test, and SSR. The study attempted to correlate rutting and cracking tests and set thresholds for HWTT and IDT-CT for future field-produced BMD mixtures. The HWTT was conducted at 45 or 48°C for mixtures with unmodified or modified binders, respectively, whereas the IDT-CT was conducted at 25°C. Both tests were conducted at an air-void level of 7.0 ± 0.5%. The study could not confirm the existence of a correlation between the HWTT rut depth and the RSI from the SSR test, nor between the CT_{index} from the IDT-CT and Sapp from CF. Suggested thresholds were a maximum of 12.5 mm rut depth at 20,000 cycles for the HWTT and a minimum CT_{index} of 150 (Veeraragavan *et al.* 2022).

Missouri
The Missouri DOT sponsored a 2020 study aiming to implement BMD. The HWTT was selected for the evaluation of rutting resistance, and three cracking tests, DCT, I-FIT, and IDT-CT were studied for cracking resistance. Testing was done at different temperatures and laboratory tests were compared to field performance data obtained from MoDOT's pavement condition surveys to establish the links necessary to determine performance tests thresholds and to calibrate the specification. It was found that the three cracking tests relate well to field cracking performance with the DCT test controlling low temperature and block cracking and the I-FIT and IDT-CT tests controlling fatigue and reflective cracking. Thresholds were recommended and presented in four tables based on traffic and position in pavement for non-SMA mainline and shoulder mixtures where cracking tests would be conducted at -12°C and rutting testing at 50°C. An example of the recommended thresholds for the highest criticality surface layer for a mainline mixture requires a CT_{index} of 150 and allows up to a 6 mm rut depth at 20,000 passes, with the *SIP* being at a minimum of 15,000 passes (Buttlar *et al.* 2020).

Nebraska
The Nebraska DOT sponsored a research study entitled "Feasibility and Implementation of Balanced Mix Design in Nebraska" (Nsengiyumva *et al.* 2020). The study investigated the SCB test for cracking and developed a test named "Gyratory Stability" for rutting as possible tests for future BMD implementation. The SCB test method was developed by examining critical testing variables such as the minimum number of replicates, specimen thickness, notch length, loading rate, and the testing temperature. Following the determination of these variables based on repeatability and practicality, the effect of different testing fixtures on the test results were investigated. The study recommended limit values for some of the mentioned test variables. A preliminary limit of 6 was suggested for *FI* and the IDT-CT was recommended for future evaluation. For rutting susceptibility, it was concluded that the FN was impractical and time consuming despite all its advantages. Thus, the researchers developed a simpler more practical test method and named it Gyratory Stability test. This test is comprised of a disc-shaped specimen loaded using the Marshall Stability test fixture. Similar to the Marshall Stability test, the results of this test are the G-stability and G-flow, which reflect the maximum load and the displacement at that load, respectively. This test was compared with the FN test and a good correlation (R^2=0.92) was found between the (stability/flow) ratio from the G-Stability test and the (flow number/flow strain) ratio from the FN test. Using the correlation developed between these two tests and fixing the G-flow value as the average of the tested specimen results, traffic dependent G-stability limits were recommended as shown in Table A.26.

Table A.26. Nebraska DOT Recommended G-stability Criteria (Nsengiyumva *et al.* 2020).

Traffic (MESALs)	Minimum Flow Number[a]	G-stability (kN)
<3	-	-
3 to < 10	53	5.55
10 to <30	190	17.24
> 30	740	64.17

[a] Recommended criteria from NCHRP report 673, page 142 (AAT 2011)

New Jersey
The New Jersey DOT (NJDOT) currently uses BMD Approach A on several types of asphalt mixtures including high reclaimed asphalt pavement (HRAP), high-performance thin overlay

(HPTO), binder-rich intermediate course (BRIC), bituminous-rich base course (BRBC), and bridge deck waterproofing surface course (BDWSC) mixtures. The performance-testing matrix at both the mix design and plant-production stages include APA testing at 64°C, tensile strength ratio and OT testing at 25°C, and BBF testing at 15°C (Bennert *et al.* 2021). The current NJDOT criteria for APA, OT, and BBF tests for the different mixture types are shown in Table A.27.

Table A.27. New Jersey DOT Performance Tests Criteria (Bennert 2022).

Mix Type	Performance Test				Air Void (%)
	Rutting		Cracking		
	Test	Criteria	Test	Criteria	
HPTO	APA at 64°C	< 4.0 mm[a]	OT at 25°C	> 600 cycles	5.0±0.5%
BRIC		< 6.0 mm[a]		> 700 cycles	3.5±0.5%
BDWSC		< 3.0 mm[a]	BBF at 15°C	>100,000 cycles[b]	1.5±0.5%
BRBC		< 5.0 mm[a]		100M cycles[c]	5.5±0.5%
HRAP, PG 64E-22		< 4.0 mm[a]	OT at 25°C	>275 cycles[d], > 150 cycles[e]	6.0±0.5%
HRAP, PG 64S-22		< 7.0 mm[a]		>200 cycles[d], > 100 cycles[e]	6.0±0.5%

[a] at 8000 cycles, [b] 1500 με and 10 Hz, [c] per NCHRP 9-38, [d] Surface Mixtures, [e] Intermediate/Base Mixtures

With the shift to using quicker and simpler tests such as the IDT-CT and IDT-HT to evaluate cracking and rutting, respectively, tentative thresholds for IDT-HT strength and the CT_{index} at intermediate temperature have been determined and are undergoing further evaluation for possible implementation (Bennert *et al.* 2021). Through informal discussion, the proposed IDT-CT and IDT-HT criteria for NJDOT were obtained and are shown in Table A.28 below.

Table A.28. NJDOT Proposed IDT-CT and IDT-HT Criteria (Bennert 2022).

Mixture Type	Layer	Binder Grade	Minimum CT_{index}	Minimum IDT-HT Strength (psi)
HRAP	Surface	PG64E-22	190	34
		PG64S-22	170	25
	Intermediate /Base	PG64E-22	150	34
		PG64S-22	130	25
BRIC	All	All	350	27
HPTO	All	All	350	34

New York

A recent study sponsored by New York State DOT evaluated the performance of eleven of the state's approved asphalt mixtures to set a baseline for the existing mixture performance (Bennert *et al.* 2022). The mixtures were evaluated for fatigue resistance using IDT-CT, OT, and SCB-FI at 6.0 ± 0.5% air voids, and Overlay Tester at 6.0 ± 0.5% air voids. Rutting susceptibility was evaluated using the APA, HWTT, and IDT-HT, all conducted at 6.0 ± 0.5% air voids. Mixture optimum asphalt content was varied by -0.5%, +0.5%, and 1% and testing was conducted again. NJDOT test criteria were used to evaluate the mixtures performance, except for IDT-HT and IDT-CT. For these two tests, criteria were established by comparing their results to the remaining tests. This comparison was done by plotting their results against the other tests and using correlations to obtain the criteria. Correlations resulted in chosen limits of a minimum IDT-HT of 30 psi at 44°C and minimum CT_{index} of 135 at 25°C. Chosen criteria are shown in Table A.29. Results showed that six of the eleven evaluated mixtures had poor fatigue cracking performance at the volumetric optimum asphalt content. Two methods were considered for ensuring satisfactory fatigue performance, the first is by using the asphalt content that ensures good fatigue performance, and the second being using that same asphalt content plus the production allowable limit of 0.4% to ensure better performance. Comparing the air voids levels at these performance-based asphalt contents resulted in an average design air void level of 4%. However, there was high variability in the data in terms of standard deviation and range, which led to the conclusion that volumetric design alone is not enough to ensure good performance.

Table A.29. Performance Test Criteria for the Performance Tests used in the NYSDOT BMD Study (Bennert *et al.* 2022).

Rutting Test and Criterion	Fatigue Cracking and Criterion
APA < 4.0 mm rutting at 8,000 cycles, 64°C	OT > 250 cycles ay 93% Load Reduction and 25°C
HWTT < 12.5 mm rutting at 20,000 cycles, 50°C	CT_{index} > 135 at 25°C
IDT-HT > 30 psi at 44°C	*FI* > 8.0 at 25°C

Ohio

A two-phase research study was undertaken by the Ohio DOT (ODOT) to select a method for characterizing the durability of asphalt mixtures containing RAP and RAS. After the completion of Phase 1, the I-FIT was selected to be the better test for ODOT's needs (Rodezno *et al.* 2018). However, with the emergence of the IDT-CT, ODOT pursued Phase 2 of the study to compare I-FIT and IDT-CT and to make an informed decision as to the viability of the IDT-CT (Abbas *et al.* 2021). The study concluded that the two tests can be used as surrogates to each other and recommended selecting the IDT-CT test since it had several advantages compared to the I-FIT. A standard IDT-CT test method was developed to be used by ODOT and its contractors in mix design approval and QC/QA with the proposed CT_{index} criteria shown in Table A.30 below.

Currently, ODOT requires IDT-CT testing for all Type A Superpave mix design and uses APA testing for mixtures with more than 15% fine aggregates and that do not meet the fine aggregate angularity criteria (ODOT 2022a). The IDT-CT results are to be reported from testing conducted at 25°C and 7.0 ± 0.5% air voids (ODOT 2022b). The APA is conducted at 49°C and 7.0 ± 0.5% air voids with a maximum allowable rut depths of 5 mm for PG58-28 and PG64-22 mixtures, and 3 mm for all other mixtures (ODOT 2022a, 2022c). In addition, ODOT incorporates APA and BBF test criteria in its specification for bridge deck waterproofing HMA. For this type of mixture, the APA is run at 64°C for specimens with 4.0 ± 0.1% air voids with a maximum allowed rut depth of 4 mm, and the BBF is run on 4.0 ± 1.0% air voids specimens with a frequency of 10 Hz, and a minimum number of cycles to failure of 100,000 cycles at a strain level of 1,500 macrostrain (ODOT 2014).

Table A.30. ODOT Proposed CT_{index} Limits (Abbas *et al.* 2021).

Mix Type	Minimum CT_{index}
Item 442 (Superpave) 12.5 mm (Surface)	80
Item 442 (Superpave) 19 mm (Intermediate)	60
Item 441 (Marshall) Type 1 Surface Mixes	80
Item 441 (Marshall) Type 1 Intermediate Mixes	80
Item 441 (Marshall) Type 2 Intermediate Mixes	60
Item 302 (Marshall) Mixes	60

Oklahoma

The Oklahoma DOT (ODOT) currently requires HWTT for rutting susceptibility for its Superpave, SMA, and rich intermediate layer (RIL) mixtures. The criterion for HWTT number of passes to 12.5 mm rut depth at 50°C is a function of the binder grade and is shown in Table A.31 (ODOT 2019). The ODOT sponsored a research study to select a durability test along with the HWTT for BMD implementation (Cross and Li 2019). In this study, the I-FIT was selected initially as the durability test for implementation by ODOT. However, the study later looked into performing IDT-CT testing and compared the results of the two tests concluding that they were not highly correlated. It was recommended that ODOT focus on the IDT-CT test for future BMD implementation due to its simplicity and due to the high variability in the I-FIT results. As an initial recommendation, a minimum CT_{index} of 80 was discussed as the criterion for short-term aged specimens, with a possibility of dropping the binder grade in case of failure to meet this criterion occurs.

Table A.31. ODOT HWTT Criteria (ODOT 2019).

Binder Grade	Minimum Number of Passes to 12.5 mm Rut Depth
PG 58	10,000
PG 64	10,000
PG 70	15,000
PG 76	20,000

Oregon

Previous research efforts established a performance-based BMD framework that proposed the use of the I-FIT with typical *FI* values ranging from 9 to 14 for production mixtures. Recent efforts developed a long-term aging protocol consisting of aging mixtures at 95°C for 24 hours to simulate 3-5 years of in-field aging. Subsequently, the *FI* threshold was refined to a minimum of 6 for Level 3 mixtures (1-10 million ESALs on rural highways and 1-3 million ESALs on urban highways) and 8 for Level 4 mixtures (>10 million ESALs on rural highways and >3 million ESALs on urban highways). The HWTT was used and a rut depth threshold of 3 mm for Level 3 mixtures and 2.5 mm for Level 4 mixtures after 20,000 passes was recommended (Coleri *et al.* 2020, Sreedhar *et al.* 2021).

Texas

The Texas DOT currently uses BMD Approach A for premium asphalt mixtures such as porous friction courses, stone matrix asphalt, thin overlay mixtures, and hot in-place recycling of asphalt concrete surfaces. The HWTT and OT are required to evaluate mixture resistance to rutting/moisture damage and reflection/bottom-up cracking, respectively. Initially, an optimum binder content (OBC) is determined using Superpave volumetric mix design criteria. The HWTT and OT are then used to evaluate specimens at three binder contents (OBC, OBC + 0.5%, and OBC + 1.0%), and the final OBC is selected to satisfy the requirements of both tests. The minimum number of HWTT passes at a test temperature of 50°C to 12.5 mm rut depth are 10,000 cycles, 15,000 cycles, and 20,000 cycles for mixtures produced with high temperature binder PG of 64°C and lower, 70°C, and 76°C and higher, respectively. The OT requirements include a minimum critical fracture energy of 1 in-lb/in^2 and a maximum crack propagation rate of 0.45. It is also required that a correlation between IDT-CT and OT be established to decide on a limit for the IDT-CT to be used instead of the OT (TxDOT 2019). In a 2020 study, a quality control/quality assurance (QC/QA) acceptance protocol was recommended using practical performance tests such as the IDT at high and intermediate temperatures. This process includes sampling produced loose mixtures and compacting performance test specimens to an air-void level of 7 ± 0.5%. after conditioning for 2 hours at 135°C. A minimum IDT shear strength of 1.02 MPa at 50°C and a minimum CT_{index} at 25°C of 105 were recommended as production acceptance criteria (Zhou *et al.* 2020).

Utah

The Utah DOT requires HWTT and IDT-CT for assessing the rutting and cracking susceptibility of its Superpave mixtures. The HWTT is conducted at an air-void level of 7.0 ± 1.0% and at a temperature of 46°C, 50°C, or 54°C for mixtures with binders of high temperature PGs of 58°C, 64°C, or 70°C, respectively. Their criteria is to have less than 10 mm rut depth at 20,000 cycles for N_{des} of 75 gyrations or greater, or at 10,000 cycles for N_{des} less than 75 gyrations (UDOT 2021, 2022a). The IDT-CT is required to be performed and submitted along with the mix design. The test is conducted at an air-void level of 7 ± 0.5% and at a temperature of 25°C (UDOT 2022b).

Virginia

In 2017, an initial effort was undertaken by researchers at the Virginia Transportation Research Council (VTRC) to benchmark the performance of several asphalt mixtures (Bowers and Diefenderfer 2018). The Cantabro mass loss test, the IDT-CT, and the APA rut test were recommended for use in the BMD method for VDOT for assessing their asphalt mixtures in terms of durability, cracking resistance, and rutting potential, respectively. Moreover, initial performance threshold criteria were developed for the selected tests where Cantabro mass loss was limited to a maximum of 7.5%, APA rut depth to a maximum of 8.0 mm, and the CT_{index} to a minimum of 70 (Bowers and Diefenderfer 2018). In 2021, the developed performance-based specifications were assessed and verified prior to full implementation in Virginia. This study validated that the IDT-CT and APA rut tests selected for use in the BMD method are in agreement with fundamental performance tests. Based on results from the mixtures tested in this study, the performance criteria previously established were shown to be reasonable based on additional mixture testing (Diefenderfer *et al.* 2021a). Following this, a special provision was released by the Virginia DOT for designing surface mixtures with BMD. The provision requires conducting the APA and the IDT-CT tests for evaluating rutting and cracking performance of their BMD mixtures, respectively. The APA is run at an air-void level of 7.0 ± 0.5% and at a temperature of 64°C, with a maximum allowed rut depth of 8 mm at 8000 passes. The IDT-CT is run at an air-void level of 7.0 ± 0.5% and at a temperature of 25°C, with a minimum required CT_{index} of 70 (Diefenderfer *et al.* 2021b).

Wisconsin

In 2017, the Wisconsin DOT implemented the regressed air void approach aiming to solve the durability issues encountered when following the Superpave mix design approach (West *et al.*

2018b). This method led to improvements in cracking resistance without compromising rutting resistance, but still had significant limitations and hindered innovation which led to unacceptable field performance. Hence, the Wisconsin DOT conducted a research study in 2021 focusing on the implementation of BMD into their design specifications. In this study, the HWTT, IDT-CT, and DCT tests were used to evaluate mixture rutting and moisture resistance, intermediate-temperature cracking resistance, and low-temperature cracking resistance, respectively. Based on the test results and findings, it was concluded that the WisDOT should continue using the regressed air voids approach for low traffic mixes. On the other hand, it was recommended that BMD Approach C be implemented for medium and high traffic mixes, as well as SMA mixes. Benchmarking efforts were undertaken and a preliminary performance test criteria were suggested for the different mixes as shown in Table A.32 (West *et al.* 2021b). In this table, the parameters obtained from the HWTT differ from traditional rut depth and *SIP*. The primary HWTT data analysis used in this study follows the method by Yin *et al.* (2014), which changes the HWTT curve into a steady-state (corrected) rut depth portion for the evaluation of rutting resistance and a post-stripping portion for the evaluation of moisture susceptibility, as illustrated in Figure A.7 below. The stripping number (*SN*) parameter in this analysis represents the number of passes at which stripping occurs in the mixture and is determined as the inflection point of the rut depth curve and is typically much lower than the *SIP*. The parameters obtained from this method are the SN and the corrected rut depth at 20,000 passes (CRD_{20k}). Lower CRD_{20k} and higher *SN* values are desired for better rutting resistance and moisture resistance, respectively (Yin *et al*. 2014, 2020a).

Table A.32. WisDOT Recommended Performance Tests Thresholds (West *et al.* 2021b).

Traffic Level	HWTT[a]		IDT-CT[b]	DCT[b]
	CRD_{20k} (mm)	SN (passes)	CT_{index}	G_f (J/m²)
SMA Mix	≤ 6.0	≥ 2,000	≥ 80	≥ 400
HT Mix				
MT Mix	≤ 7.0		≥ 40	≥ 300
LT Mix	≤ 8.0			

[a] conducted on short-term aged specimens
[b] conducted on long-term aged specimens

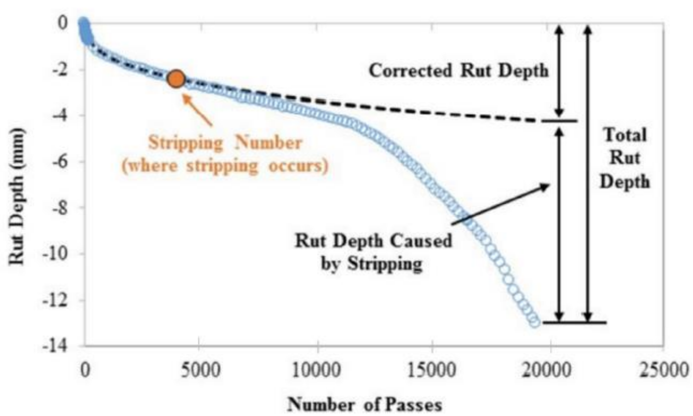


Figure A.7. Alternative HWTT data analysis based on CRD_{20k} and *SN* (West *et al.* 2021b).

Summary of Tests and Threshold Limits

In summary, the most common tests used for rutting and cracking susceptibility are the HWTT (followed by APA) and the IDT-CT (followed by SCB-type tests), respectively. Out of the 18 reviewed state DOTs, 14 are using or considering the use of the HWTT, and 11 are using or considering the use of the IDT-CT in their BMD specifications. The temperatures at which these two most common tests are conducted are almost always consistent, being typically 50°C for HWTT and 25°C for the IDT-CT. The most common mean air-void level at which all rutting and cracking performance tests are conducted is 7.0%. A summary of the rutting and cracking tests used by the reviewed states are shown in Table A.33 and Table A.34, respectively. In these tables, the tests used are listed along with the temperature at which they are conducted, air-void level, and set or recommended criteria. Testing conditions and criteria used or recommended by the states differ for each test and can be affected by many variables such as the binder PG, aging level, traffic, modifier usage, mix type, and number of design gyrations. To account for that and have an overall summary, the tables often contain a range of temperatures, air-void levels, or criteria.

Table A.33. State DOTs BMD Rutting Tests Summary.

State	Test	Temp.	AV	Criteria
Alabama	HWTT	50°C	7±0.5%	Passes to 10-mm rut depth ≥ 10,000-20,000 [a]
	IDT-HT	50°C	7±0.5%	IDT-HT strength ≥ 20 psi
California	HWTT	45-55°C[a]	6.5-7±0.5%[b]	Passes to 12.5-mm rut depth ≥ 10,000-25,000[a]
Georgia	HWTT	50°C	7±1%	Passes to 12.5-mm rut depth and to *SIP* > 15,000-20,000[a,b]
Illinois	HWTT	50°C	7±1%	Passes to 12.5-mm rut depth ≥ 5,000-20,000[a]
Louisiana	HWTT	50°C	7±0.5%	Rut depth at 20,000 cycles < 6-10 mm[c]
Maine	HWTT	45 or 48°C[d]	7±0.5%	Rut depth at 20,000 cycles < 12.5 mm
Nebraska	G-Stability	54°C	7±0.5%	G-Stability ≥ 5.55-64.17 kN[c]
New Jersey	APA	64°C	1.5-6±0.5%[b]	Rut depth at 8,000 cycles < 3-7 mm[b]
New York	APA	64°C	6±0.5%	Rut depth at 8,000 cycles < 4 mm
	HWTT	50°C	6±0.5%	Rut depth at 20,000 cycles < 12.5 mm
	IDT-HT	44°C	6±0.5%	IDT-HT strength > 30 psi
Ohio	APA	49 or 64°C[b]	4±0.1% or 7±0.5%[b]	Rut depth at 8,000 cycles < 3-5 mm[a,b]
Oklahoma	HWTT	50°C	7±1%	Passes to 12.5-mm rut depth ≥ 10,000-20,000[a]
Oregon	HWTT	50°C	7±1%	Rut depth at 20,000 cycles ≤ 2.5-3.0 mm[b]
Texas	HWTT	50°C	7±0.5%	Passes to 12.5-mm rut depth ≥ 10,000-20,000[a]
	IDT-HT	50°C	7±0.5%	IDT shear strength ≥ 1.02 MPa
Utah	HWTT	46-54°C[a]	7±1%	Passes to 10-mm rut depth ≥ 10,000-20,000[e]
Virginia	APA	64°C	7±0.5%	Rut depth at 8,000 cycles ≤ 8 mm
Wisconsin	HWTT	46°C	7±0.5%	CRD_{20k} ≤ 6-8 mm[b], SN ≥ 2,000

[a] Parameter or temperature is binder PG grade dependent, [b] Parameter is mix type dependent, [c] Parameter is traffic dependent, [d] Temperature varies depending on a polymer modified or non-polymer modified binder, [e] Parameter is gyration dependent

Table A.34. State DOTs BMD Cracking Tests Summary.

State	Test	Temp.	AV	Criteria
Alabama	IDT-CT	25°C	7±0.5%	Design $CT_{index} \geq$ 55-110 Production $CT_{index} \geq$ 50-100[a]
California	BBF	20°C	6±0.5%	50% Loss of flexural stiffness at 10 Hz
Georgia	IDT-CT	25°C	7±0.5%	CT_{index} using Average Value Method ≥ 15-150[b,c]
Illinois	I-FIT	25°C	7±1%	$FI \geq$ 5.0-16[b,c]
Louisiana	SCB-LTRC	25°C	7±0.5%	$J_c >$ 0.5-0.6 kJ/m^2 [a]
Maine	IDT-CT	25°C	7±0.5%	$CT_{index} \geq 150$
Nebraska	SCB	23°C	4±0.5%	$FI \geq 6.0$
New Jersey	OT	25°C	3.5-6±0.5%[c]	Cycles to 93% Load Reduction > 100-700[c]
	BBF	15°C	1.5-5.5±0.5%[c]	>100,000 cycles at 1500 microstrain, 10 Hz
New York	OT	25°C	6±0.5%	Cycles to 93% Load Reduction > 250
	IDT-CT	25°C	6±0.5%	$CT_{index} > 135$
	I-FIT	25°C	6±0.5%	$FI \geq 8.0$
Ohio	IDT-CT	25°C	7±0.5%	$CT_{index} \geq$ 60-80[c]
	BBF	NR[d]	4±1.0%	>100,000 cycles at 1500 microstrain, 10 Hz
Oklahoma	IDT-CT	25°C	7±0.5%	$CT_{index} \geq 80$
	I-FIT	25°C	7±0.5%	$FI \geq 8.0$
Oregon	I-FIT	25°C	7±1%	$FI \geq$ 6-8.0[c]
Texas	OT	25°C	7±1%	Critical fracture energy ≥ 1 in-lb/in^2 and crack propagation rate ≤ 0.45
	IDT-CT	25°C	7±0.5%	$CT_{index} \geq 105$
Utah	IDT-CT	25°C	7±0.5%	Not Assigned
Virginia	IDT-CT	25°C	7±0.5%	$CT_{index} \geq 70$
Wisconsin	IDT-CT	25°C	7±0.5%	$CT_{index} \geq$ 40-80[c]

[a] Parameter is traffic dependent, [b] Parameter is age level dependent, [c] Parameter is mix type dependent, [d] Not reported

Knowledge Gaps

The literature review indicates that have been many efforts to study different rutting and fatigue cracking protocols for integration into BMD. However, there are aspects related to the selection of test methods and implementation of BMD that are still not fully elucidated. Some of these aspects are:

- There is no clear indication as to which is the most adequate rutting and cracking test to be adopted by state agencies when considering the tradeoff between simplicity, cost, degree of fundamental characterization, and any unique aspects that exist in that state.
- The studies evaluating different test methods mainly focus on one category at a time. However, there is a need to evaluate cracking and rutting test protocols simultaneously to identify compatible methods and potential optimization of the BMD process when similar

specimen geometries and test equipment are used for both cracking and rutting characterization.

- BMD test methods can provide an indication of the material's performance; however, most of the proposed protocols for rutting and cracking do not have a mechanically sound methodology to link test results with fundamental engineering properties, which is essential to the practical integration of mixture design and pavement design.
- There is no explicit indication about which of the AASHTO PP 105 (2024b) approaches is the most efficient at improving overall performance of asphalt concrete mixtures. A and B are seen as more conservative, keeping volumetric-based design while adding performance verification or optimization. On the other hand, approaches C and D can give the contractors more freedom in terms of achieving the best performance of a selected test, and consistency over production.
- There is, as of yet, no agreed upon best practice for integrating performance testing into quality assurance. However, two ongoing projects are examining some of the key issues (NCHRP 10-107 and TFRS01). The findings from these studies are not yet public, but in general both agree that the simplicity of testing and turnaround time are critical issues to consider.

There are no set universal performance thresholds for BMD tests as the interpretation of results may change for different locations and mixtures, and thus it is necessary to determine and verify adequate thresholds for a certain agency.

References

AASHTO *T 283-22: Standard Method of Test for Resistance of Compacted Asphalt Mixtures to Moisture-Induced Damage*. American Association of State Highway and Transportation Officials, Washington, D.C., 2022a.

AASHTO *T 378-22: Determining the Dynamic Modulus and Flow Number for Asphalt Mixtures Using the Asphalt Mixture Performance Tester (AMPT)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2022b.

AASHTO *T 324-23: Standard Method of Test for Hamburg Wheel-Track Testing of Compacted Asphalt Mixtures*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023a.

AASHTO *T 340-23: Standard Method of Test for Determining Rutting Susceptibility of Hot Mix Asphalt (HMA) Using the Asphalt Pavement Analyzer (APA)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023b.

AASHTO *TP 132-23: Determining the Dynamic Modulus for Asphalt Mixtures Using Small Specimens in the Asphalt Mixture Performance Tester (AMPT)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023c.

AASHTO *TP 134-22: Stress Sweep Rutting (SSR) Test Using Asphalt Mixture Performance Tester (AMPT)*. American Association of State Highway and Transportation Officials, Washington, D.C., 2023d.

AASHTO *PP 105-24: Standard Practice for Balanced Design of Asphalt Mixtures*. American Association of State Highway and Transportation Officials, Washington, D.C., 2024a.

AASHTO *T 411-24: Determining the Damage Characteristic Curve and Failure Criterion Using Small Specimens in the Asphalt Mixture Performance Tester (AMPT) Cyclic Fatigue Test*. American Association of State Highway and Transportation Officials, Washington, D.C., 2024b.

AASHTO, FHWA, NAPA, SHRP, AI, and TRB *European Asphalt Study Tour*. Washington, D.C., 1990.

Advanced Asphalt Technologies (AAT). *A Manual for Design of Hot Mix Asphalt with Commentary*, NCHRP Report 673, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2011.

Abbas, A.R., Nazzal, M., Quasem, T., Mansour, M., and Husain, S.F. *Crack Resistance and Durability of Ohio DOT Asphalt Mixtures Using I-FIT & IDEAL-CT: Phase 2*. FHWA/OH-2022-01, Ohio Department of Transportation, Columbus, OH, 2021.

ALDOT *Special Provision No. 18-0806(2): Balanced Asphalt Mix Design for Local Roads*. Alabama Department of Transportation, Geneva County, AL, 2020.

ALDOT *Standard Specifications for Highway Construction*. Alabama Department of Transportation, Henry County, AL, 2022a.

ALDOT *Special Provision No. 22-PS0080: Asphalt Pavement - Balanced Mix Design ESAL Range C/D Benchmarking Sections*. Alabama Department of Transportation, Montgomery, AL, 2022b.

Ali, A., Kabir, S.F., Al-Badr, B., Alfalah, A., Xie, Z., Decarlo, C., Elshaer, M., and Mehta, Y. Laboratory Performance of Dense Graded Asphalt Mixtures Prepared Using Highly Polymer Modified Binders Containing Corn Oil as Softening Agent. *Construction and Building Materials*, Vol. 345, 2022.

Ali, U.M., Al-Qadi, I.L., and Ozer, H. Flexibility Index Threshold Optimization for Various Asphalt Concrete Mixes and Climatic Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2674, No. 1, 2020, pp. 104–112.

Al-Qadi, I.L., Ozer, H., Lambros, J., El-Khatib, A., Singhvi, P., Khan, T., Rivera-Perez, J., and Doll, B. *Testing Protocols to Ensure Performance of High Asphalt Binder Replacement Mixes Using RAP and RAS*. FHWA-ICT-15-017, Illinois Department of Transportation, Springfield, IL, 2015.

Al-Qadi, I.L., Ozer, H., Zhu, Z., Singhvi, P., Mohamed, U., Sawalha, A.M., Francisco, A., Luque, E., Jesus, J., Mainieri, G., and Zehr, T.G. *Development of Long-Term Aging Protocol for Implementation of the Illinois Flexibility Index Test (I-FIT)*. FHWA-ICT-19-009, Illinois Department of Transportation, Springfield, IL, 2019.

Arámbula-Mercado, E., Chavarro-Muñoz, S.J., Hu, S., and Moseley, H. Performance of Hot and Cold Recycled Mixtures with High Reclaimed Asphalt Pavement Content. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2674, No. 9, 2020, pp. 714–726.

Arámbula-Mercado, E., Hill, R.A., Caro, S., Manrique, L., Park, E.S., and Fernando, E. *Understanding Mechanism of Raveling to Extend Open Graded Friction Course (OGFC) Service Life*. Florida Department of Transportation, Tallahassee, FL, 2016.

Aschenbrener, T. Evaluation of Hamburg Wheel-Tracking Device to Predict Moisture Damage in Hot-Mix Asphalt. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1492, 1995, pp. 193–201.

Aschenbrener, T. and Currier, G. *Influence of Testing Variables on the Results from the Hamburg Wheel-Tracking Device*. CDOT-DTD-R-93-22, Colorado Department of Transportation, Denver, CO, 1993.

Aschenbrener, T., Terrel, R.L., and Zamora, R.A. *Comparison of the Hamburg Wheel-Tracking Device and the Environmental Conditioning System to Pavements of Known Stripping Performance*. CDOT-DTD-R-94-1, Colorado Department of Transportation/Federal Highway Administration, Denver, CO, 1994.

Asphalt Institute *Mix Design Methods for Asphalt*. 6th ed. MS-02, Asphalt Institute, Lexington, KY, 1997.

Bahia, H.U., Teymourpour, P., Swiertz, D., Ling, C., Varma, R., Mandal, T., Chaturabong, P., Lyngdal, E., and Hanz, A. *Analysis and Feasibility of Asphalt Pavement Performance-Based Specifications for WisDOT*. WisDOT No. 0092-15-04, Wisconsin Department of Transportation, Madison, WI, 2016.

Barros, L.M., Nascimento, L.A.H., Aragão, F.T.S., Underwood, B.S., and Pivetta, F.C. Characterization of the Permanent Deformation of Asphalt Mixtures Based on Indexes and on Pavement Structural Performance. *Construction and Building Materials*, Vol. 326, 2022a, pp. 126555.

Barros, L.M., Nascimento, L.A.H., and Aragão, F.T.S. Effects of Binder and Aggregate Properties on the Permanent Deformation of Asphalt Mixtures by Means of Uniaxial and Triaxial Tests. *Construction and Building Materials*, Vol. 332, 2022b, pp. 127346.

Barry, M.K. *An Analysis of Impact Factors on the Illinois Flexibility Index Test*. M.S. Thesis. University of Illinois at Urbana-Champaign, Urbana, Illinois, 2016.

Batioja-Alvarez, D., Lee, J., and Haddock, J.E. Understanding the Illinois Flexibility Index Test (I-FIT) using Indiana Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 6, 2019, pp. 337–346.

Bažant, Z.P. and Prat, P.C. Effect of Temperature and Humidity on Fracture Energy. *ACI Materials Journal*, Vol. 85, No. No. 4, 1988, pp. 262–271.

Bennert, T. Personal Communication Over Email, December 27, 2022.

Bennert, T., Haas, E., and Wass, E. Indirect Tensile Test (IDT) to Determine Asphalt Mixture Performance Indicators During Quality Control Testing in New Jersey. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, No. 28, 2018, pp. 394–403.

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

Bennert, T., Haas, E., Wass, E., Tulanowski, D., and Zavery, Z. Statewide Assessment of Balanced Mixture Design for New York State's Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 11, 2022, pp. 340–351.

Bernier, A., Zofka, A., and Yut, I. Laboratory Evaluation of Rutting Susceptibility of Polymer-Modified Asphalt Mixtures Containing Recycled Pavements. *Construction and Building Materials*, Vol. 31, 2012, pp. 58–66.

Bowers, B.F. and Diefenderfer, S.D. *Technical Memorandum: Benchmarking of Asphalt Mixtures - Summary of Findings and Preliminary Recommendations*. Virginia Transportation Research Council, Charlottesville, VA, 2018.

Buttlar, W.G., Urra-Contreras, L., Jahangiri, B., Rath, P., and Majidifard, H. *Support for Balanced Asphalt Mixture Design Specification Development in Missouri*. Missouri Department of Transportation. CMR 20-010. Missouri Department of Transportation, Jefferson City, MO, 2020.

Caltrans *Standard Specifications*. California State Transportation Agency, Department of Transportation, 2022.

Castorena, C., Underwood, B.S., Kim, Y.R., Lee, K., Tran, N., and Taylor, A.J. *Ruggedness and Interlaboratory Studies for Asphalt Mixture Performance Tester (AMPT) Cyclic Fatigue Test: Phase I Report*. FHWA-HRT-21-057, Federal Highway Administration, McLean, VA, 2021.

Celauro, C., Bernardo, C., and Gabriele, B. Production of Innovative, Recycled and High-Performance Asphalt for Road Pavements. *Resources, Conservation and Recycling*, Vol. 54, No. 6, 2010, pp. 337–347.

Chkaiban, R., Hajj, E.Y., and Hand, A.J.T. Influence of Balanced Mix Design Approaches on Pavement Design Making Through an Illustrative Example. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 10, 2022, pp. 495–506.

Choi, Y.T. and Kim, Y.R. Development of Calibration Testing Protocol for Permanent Deformation Model of Asphalt Concrete. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2373, 2013, pp. 34–43.

Choi, Y.T. and Kim, Y.R. Implementation and Verification of a Mechanistic Permanent Deformation Model (Shift Model) to Predict Rut Depths of Asphalt Pavement. *Road Materials and Pavement Design*, Vol. 15, No. sup1, 2014, pp. 195–218.

Choubane, B., Page, G.C., and Musselman, J.A. Suitability of Asphalt Pavement Analyzer for Predicting Pavement Rutting. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1723, 2000, pp. 107–115.

Chowdhury, P.S., Noojilla, S.L.A., and Reddy, M.A. Evaluation of Fatigue Characteristics of Asphalt Mixtures Using Cracking Tolerance Index (CTIndex). *Construction and Building Materials*, Vol. 342, 2022.

Christensen, D.W. and Bonaquist, R. Use of Strength Tests for Evaluating the Rut Resistance of Asphalt Concrete. *Asphalt Paving Technology: Journal of the Association of Asphalt Paving Technologists*, Vol. 71, 2002, pp. 692–711.

Christensen, D.W. and Bonaquist, R. *Using the Indirect Tensile Test to Evaluate Rut Resistance in Developing Hot-Mix Asphalt Designs: Practical Approaches to Hot-Mix Asphalt Mix Design and Production Quality Control Testing*. Transportation Research Circular C-124, Transportation Research Board, Washington, D.C., 2007.

Christensen, D.W., Bonaquist, R., Anderson, D.A., and Gokhale, S. *Indirect Tension Strength as a Simple Performance Test*. Transportation Research Circular E-C068, Transportation Research Board, Washington, D.C., 2004.

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

Colbert, B. and You, Z. The Determination of Mechanical Performance of Laboratory Produced Hot Mix Asphalt Mixtures using Controlled RAP and Virgin Aggregate Size Fractions. *Construction and Building Materials*, Vol. 26, No. 1, 2012, pp. 655–662.

Coleri, E., Sreedhar, S., and Obaid, I.A. *Development of a Balanced Mix Design Method in Oregon*. FHWA-OR-RD-21-03, Oregon Department of Transportation, Corvallis, OR, 2020.

Collins, R., Shami, H., and Lai, J.S. Use of Georgia Loaded Wheel Tester to Evaluate Rutting of Asphalt Samples Prepared by Superpave Gyratory Compactor. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1545, 1996, pp. 161–168.

Collins, R., Watson, D., and Campbell, B. Development and Use of Georgia Loaded Wheel Tester. *Transportation Research Board*, Vol. 1492, 1995, pp. 202–207.

Cominsky, R.J., Huber, G.A., Kennedy, T.W., and Anderson, M. *The Superpave Mix Design Manual for New Construction and Overlays*. SHRP-A-407, Strategic Highway Research Program, Washington, D.C., 1994.

Cooley, L.A., Brumfield, J.W., Mallick, R.B., Mogawer, W.S., Partl, M.N., Poulikakos, L.D., and Hicks, G. *Construction and Maintenance Practices for Permeable Friction Courses*. NCHRP Report 640, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2009.

Cooley, L.A., Kandhal, P.S., Buchanan, M.S., Fee, F., and Epps, A. *Loaded Wheel Testers in the United States: State of the Practice*. Transportation Research Circular E-C016, Transportation Research Board, Washington, D.C., 2000.

Cooper, S.B., King, W., and Kabir, S. *Testing and Analysis of LWT and SCB Properties of Asphalt Concrete Mixtures*. FHWA/LA/536, Louisiana Department of Transportation and Development, Baton Rouge, LA, 2016.

Cooper, S.B., Mohammad, L.N., and Elseifi, M.A. Laboratory Performance of Asphalt Mixtures Containing Recycled Asphalt Shingles. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2445, 2014, pp. 94–102.

Cooper, S.B., Mohammad, L.N., Elseifi, M.A., and Medeiros, M.S. Effect of Recycling Agents on the Laboratory Performance of Asphalt Mixtures Containing Recycled Asphalt Shingles. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2506, 2015, pp. 54–61.

Cox, B.C., Smith, B.T., Howard, I.L., and James, R.S. State of Knowledge for Cantabro Testing of Dense Graded Asphalt. *Journal of Materials in Civil Engineering*, Vol. 29, No. 10, 2017.

Cross, S.A. and Li, J.Q. *Implement Balanced Asphalt Mix Design in Oklahoma*. FHWA-OK-19-01, Oklahoma Department of Transportation, Oklahoma City, OK, 2019.

Daniel, J.S. and Kim, Y.R. Development of a Simplified Fatigue Test and Analysis Procedure Using a Viscoelastic, Continuum Damage Model. *Asphalt Paving Technology: Journal of the Association of Asphalt Paving Technologists*, 2000, pp. 619–650.

Diefenderfer, S.D. and Bowers, B.F. Initial Approach to Performance (Balanced) Mix Design: The Virginia Experience. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 2, 2019, pp. 335–345.

Diefenderfer, S.D., Boz, I., and Habbouche, J. *Balanced Mix Design for Surface Asphalt Mixtures: Phase I: Initial Roadmap Development and Specification Verification*. FHWA/VTRC 21-R15, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2021a.

Diefenderfer, S.D., Boz, I., and Habbouche, J. *Balanced Mix Design for Asphalt Surface Mixtures: 2019 Field Trials*. FHWA/VTRC 21-R21, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2021b.

Ding, X., Chen, L., Ma, T., Ma, H., Gu, L., Chen, T., and Ma, Y. Laboratory Investigation of the Recycled Asphalt Concrete with Stable Crumb Rubber Asphalt Binder. *Construction and Building Materials*, Vol. 203, 2019, pp. 552–557.

Doyle, J.D. and Howard, I.L. Evaluation of the Cantabro Durability Test for Dense Graded Asphalt. *In*: Han, J. and Alzamora, D.E. (eds) *Geo-Frontiers*. 2011.

Doyle, J.D. and Howard, I.L. Characterization of Dense-Graded Asphalt with the Cantabro Test. *Journal of Testing and Evaluation*, Vol. 44, No. 1, 2016, pp. 77–88.

Elseifi, M.A., Mohammad, L.N., Ying, H., and Cooper, S.B. Modeling and Evaluation of the Cracking Resistance of Asphalt Mixtures Using the Semi-Circular Bending Test at Intermediate Temperatures. *Road Materials and Pavement Design*, Vol. 13, No. sup1, 2012, pp. 124–139.

Fakhri, M. and Ahmadi, A. Evaluation of Fracture Resistance of Asphalt Mixes Involving Steel Slag and RAP: Susceptibility to Aging Level and Freeze and Thaw Cycles. *Construction and Building Materials*, Vol. 157, 2017, pp. 748–756.

Fakhri, M., Hassani, K., and Ghanizadeh, A.R. Impact of Loading Frequency on the Fatigue behavior of SBS Modified Asphalt Mixtures. *Procedia - Social and Behavioral Sciences*, Vol. 104, 2013, pp. 69–78.

Fakhri, M. and Kheiry, P.T. Fatigue Characterization of Hot Mix Asphalt Using Four-Point Bending Beam Test. *LCMT 2009 8th Annual International Conference*. Liverpool, United Kingdom, 2009.

FHWA. *Asphalt Mixture ETG Meeting Technical Report*. Federal Highway Administration, Salt Lake City, UT, 2016.

FHWA. *Balanced Asphalt Mix Design: Eight Tasks for Implementation*. FHWA-HIF-22-048, Federal Highway Administration, 2022.

Finn, F. *Factors Involved in the Design of Asphaltic Pavement Surfaces*. NCHRP Report 39, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 1967.

Francken, L. Permanent Deformation Law of Bituminous Road Mixes in Repeated Triaxial Compression. *Proceedings of 4th International Conference on Structural Design of Asphalt Pavements*. Ann Arbor, MI, 1977.

GDOT *Standard Specifications Construction of Transportation Systems*. Georgia Department of Transportation, Atlanta, GA, 2021.

Germann, F.P. and Lytton, R.L. *Methodology for Predicting the Reflective Cracking Life of Asphalt Concrete Overlays*. FHWA/TX-79/09+207-5, Texas Department of Transportation, Austin, TX, 1979.

Ghanbari, A., Underwood, B.S., and Kim, Y.R. Development of a Rutting Index Parameter Based on the Stress Sweep Rutting Test and Permanent Deformation Shift Model. *International Journal of Pavement Engineering*, Vol. 23, No. 2, 2022, pp. 387–399.

Gibson, N., Qi, X., Shenoy, A., Al-Khateeb, G., Kutay, M.E., Andriescu, A., Stuart, K., Youtcheff, J., and Harman, T. *Performance Testing for Superpave and Structural Validation*. FHWA-HRT-11-045, Federal Highway Administration, McLean, VA, 2012.

Gibson, N.H. *A Comprehensive Model for the Compressive Behavior of Asphalt Concrete*. Ph.D. Dissertation, University of Maryland, College Park, MD, 2006.

Gibson, N.H., Kutay, M.E., Keramat, D., and Youtcheff, J. Multiaxial Strain Response of Asphalt Concrete Measured During Flow Number Performance Test. *Asphalt Paving Technology: Asphalt Paving Technology*, Vol. 78, 2009, pp. 25–65.

Gu, F., Luo, X., Zhang, Y., and Lytton, R.L. Using Overlay Test to Evaluate Fracture Properties of Field-Aged Asphalt Concrete. *Construction and Building Materials*, Vol. 101, 2015b, pp. 1059–1068.

Gu, F., Zhang, Y., Luo, X., Luo, R., and Lytton, R.L. Improved Methodology to Evaluate Fracture Properties of Warm-Mix Asphalt Using Overlay Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2506, 2015a, pp. 8–18.

Habbouche, J., Boz, I., Diefenderfer, B.K., and Kim, S. Multi-Level Laboratory Performance Evaluation of Conventional and High Polymer-Modified Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2022b, pp. 297–314.

Habbouche, J., Boz, I., and Diefenderfer, S.D. *Interlaboratory Study for the Indirect Tensile Cracking Test at Intermediate Temperature: Phase II*. FHWA/VTRC 23-R3, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2022a.

Habbouche, J. and Diefenderfer, B.K. *Laboratory and Field Performance Evaluation of Pavement Sections with High Polymer-Modified Asphalt Overlays*. FHWA/VTRC 21-R16, Virginia Department of Transportation/Federal Highway Administration, Richmond, VA, 2021.

Hajj, E.Y., Aschenbrener, T., and Nener-Plante, D. Examples of Successful Practices with State Implementation of Balanced Design of Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2022, pp. 44–66.

Hajj, E.Y., Hand, A.J.T., Chkaiban, R., and Aschenbrener, T.B. *Index-Based Tests for Performance Engineering Mixture Design for Asphalt Pavements*. Report No. FHWA-HIF-19-103, Federal Highway Administration, Washington, D.C., 2019.

Harvey, J.T., Wu, R., Signore, J., Basheer, I., Holikatti, S., Vacura, P., and Holland, T.J. *Performance-Based Specifications: California Experience to Date*. Transportation Research Circular E-C189, Transportation Research Board, Washington, D.C., 2014.

Hussan, S., Kamal, M.A., Hafeez, I., and Ahmad, N. Comparing and Correlating Various Laboratory Rutting Performance Tests. *International Journal of Pavement Engineering*, Vol. 20, No. 10, 2019, pp. 1239–1249.

Hveem, F. Quality Tests for Asphalts - A Progress Report. *Proceedings of the Association of Asphalt Paving Technologist*, Vol. 15, 1943, pp. 111–152.

IDOT *Standard Specifications for Road and Bridge Construction*. Illinois Department of Transportation, Springfield, IL, 2022.

Im, S., Zhou, F., Lee, R., and Scullion, T. Impacts of Rejuvenators on Performance and Engineering Properties of Asphalt Mixtures Containing Recycled Materials. *Construction and Building Materials*, Vol. 53, 2014, pp. 596–603.

Ishaq, M.A. and Giustozzi, F. Correlation Between Rheological Fatigue Tests on Bitumen and Various Cracking Tests on Asphalt Mixtures. *Materials*, Vol. 14, No. 24, 2021.

Islam, M.R., Kalevela, S.A., and Nesselhauf, S.K. Sensitivity of the Flow Number to Mix Factors of Hot-Mix Asphalt. *Infrastructures*, Vol. 4, No. 2, 2019, pp. 34.

Islam, M.R. and Tarefder, R.A. Study of Asphalt Aging through Beam Fatigue Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2505, 2015, pp. 115–120.

Izadi, A., Motamedi, M., Alimi, R., and Nafar, M. Effect of Aging Conditions on the Fatigue Behavior of Hot and Warm Mix Asphalt. *Construction and Building Materials*, Vol. 188, 2018, pp. 119–129.

Izzo, R.P. and Tahmoressi, M. Use of the Hamburg Wheel-Tracking Device for Evaluating Moisture Susceptibility of Hot-Mix Asphalt. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1681, 1999, pp. 76–85.

Jeong, J., Underwood, B.S., and Kim, Y.R. Rutting Performance Prediction Using Index-Volumetrics Relationships with Stress Sweep Rutting Test and Hamburg Wheel-Track Test. *Construction and Building Materials*, Vol. 295, 2021, pp. 123664.

Jeong, J., Underwood, B.S., and Kim, Y.R. Cracking Performance Predictions Using Index-Volumetrics Relationships with Direct Tension Cyclic Fatigue Test and Illinois Flexibility Index Test (I-FIT). *Construction and Building Materials*, Vol. 315, 2022.

Jeong, J., Wang, Y.D., Ghanbari, A., Nash, C., Nener-Plante, D., Underwood, B.S., and Kim, Y.R. Pavement Performance Predictions Using Performance-Volumetric Relationship and Evaluation of Construction Variability: Example of MaineDOT Shadow Project for the Development of Performance-Related Specifications. *Construction and Building Materials*, Vol. 263, 2020, pp. 120150.

Kandhal, P.S. and Cooley, L.A. *Accelerated Laboratory Rutting Tests: Evaluation of the Asphalt Pavement Analyzer*. NCHRP Report 508, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2003.

Kandhal, P.S. and Mallick, R.B. 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.

Kaseer, F., Yin, F., Arámbula-Mercado, E., Epps Martin, A., Daniel, J.S., and Salari, S. Development of an Index to Evaluate the Cracking Potential of Asphalt Mixtures Using the Semi-Circular Bending Test. *Construction and Building Materials*, Vol. 167, 2018, pp. 286–298.

Kim, D. and Kim, Y.R. Development of Stress Sweep Rutting (SSR) Test for Permanent Deformation Characterization of Asphalt Mixture. *Construction and Building Materials*, Vol. 154, 2017, pp. 373–383.

Kim, D., Norouzi, A., Kass, S., Liske, T., and Kim, Y.R. Mechanistic Performance Evaluation of Pavement Sections Containing RAP and WMA Additives in Manitoba. *Construction and Building Materials*, Vol. 133, 2017, pp. 39–50.

Kim, M., Mohammad, L.N., and Elseifi, M.A. Characterization of Fracture Properties of Asphalt Mixtures as Measured by Semicircular Bend Test and Indirect Tension Test. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2296, 2012, pp. 115–124.

Kim, M., Mohammad, L.N., Jordan, T., and Cooper, S.B. Fatigue Performance of Asphalt Mixture Containing Recycled Materials and Warm-Mix Technologies Under Accelerated Loading and Four Point Bending Beam Test. *Journal of Cleaner Production*, Vol. 192, 2018a, pp. 656–664.

Kim, S., Shen, J., and Jeong, M.M. Effects of Aggregate Size on the Rutting and Stripping Resistance of Recycled Asphalt Mixtures. *Journal of Materials in Civil Engineering*, Vol. 30, No. 2, 2018b.

Kim, Y.R., Jadoun, F.M., Former, T.H., and Muthadi, N. *Local Calibration of the MEPDG for Flexible Pavement Design*. FHWA\NC\2007-07, North Carolina Department of Transportation, Raleigh, NC, 2011.

Kim, Y.R., Lee, H.-J., and Little, D.N. Fatigue Characterization of Asphalt Concrete Using Viscoelasticity and Continuum Damage Theory. *Asphalt Paving Technology: Journal of the Association of Asphalt Pavement Technologist*, Vol. 66, 1997, pp. 520–551.

Koohi, Y., Luo, R., Lytton, R.L., and Scullion, T. New Methodology to Find the Healing and Fracture Properties of Asphalt Mixes Using Overlay Tester. *Journal of Materials in Civil Engineering*, Vol. 25, No. 10, 2013, pp. 1386–1393.

Lai, J.S. *Development of a Simplified Test Method to Predict Rutting Characteristics of Asphalt Mixes*. GDOT Project 8503, Georgia Department of Transportation, Atlanta, GA, 1986.

Lai, J.S. Results of Round-Robin Test Program to Evaluate Rutting of Asphalt Mixes Using Loaded Wheel Tester. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1417, 1993, pp. 127–134.

Lai, J.S. and Shami, H. Development of Rolling Compaction Machine for Preparation of Asphalt Beam Samples. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1492, 1995, pp. 18–25.

Larrain, M.M.M. and Tarefder, R.A. Effects of Asphalt Concrete Gradation, Air Voids, and Test Temperatures on Rutting Susceptibility by Using the Hamburg Wheel Tracking Device (HWTD). *In*: Chen, D., Lee, J., and Steyn, W.J. (eds) *Geo-China 2016*. American Society of Civil Engineers, Reston, VA, 2016, pp. 83–90.

Lee, S.I., Faruk, A.N.M., and Walubita, L.F. Comparison of Fracture Cracking Parameters from Monotonic Loading Tests: Indirect Tension and Monotonic Overlay Tester Tests. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2576, 2016, pp. 19–27.

Lemke, Z., Sadek, H., Swiertz, D., Reichelt, S., and Bahia, H.U. Effects of Reheating Procedure and Oven Type on Performance Testing Results of Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, No. 26, 2018, pp. 124–133.

Li, J., Oh, J., Naik, B., Simate, G.S., and Walubita, L.F. Laboratory Characterization of Cracking-Resistance Potential of Asphalt Mixes Using Overlay Tester. *Construction and Building Materials*, Vol. 70, 2014, pp. 130–140.

Ling, C. and Buchanan, S. A Sensitivity Study to Evaluate Intermediate Temperature Cracking Tests for Vermont Asphalt Mixes. *Journal of Transportation Engineering, Part B: Pavements*, Vol. 148, No. 2, 2022.

Ling, C., Swiertz, D., Mandal, T., Teymourpour, P., and Bahia, H.U. Sensitivity of the Illinois Flexibility Index Test to Mixture Design Factors. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2631, 2017, pp. 153–159.

Livneh, M. and Shklarsky, E. The Splitting Test for Determination of Bituminous Concrete Strength. *Proceedings of the Association of Asphalt Paving Technologists*, Vol. 31, 1962, pp. 457–476.

Lv, Q., Huang, W., Zheng, M., Sadek, H., Zhang, Y., and Yan, C. Influence of Gradation on Asphalt Mix Rutting Resistance Measured by Hamburg Wheel Tracking Test. *Construction and Building Materials*, Vol. 238, 2020, pp. 117674.

Ma, T., Wang, H., Huang, X., Wang, Z., and Xiao, F. Laboratory Performance Characteristics of High Modulus Asphalt Mixture with High-Content RAP. *Construction and Building Materials*, Vol. 101, 2015, pp. 975–982.

Majidifard, H., Jahangiri, B., Rath, P., and Buttlar, W.G. Development of a Balanced Cracking Index for Asphalt Mixtures Tested in Semi-Circular Bending with Load-LLD Measurements. *Measurement: Journal of the International Measurement Confederation*, Vol. 173, 2021.

McDaniel, R.S., Leahy, R.B., Huber, G.A., Moulthrop, J.S., and Ferragut, T. *The Superpave Mix Design System: Anatomy of a Research Program*. NCHRP Web-Only Document 186, National Cooperative Highway Research Program, Transportation Research Board, 2012.

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

Mirhosseini, F.A., Tahami, A., Hoff, I., Dessouky, S., Kavussi, A., Fuentes, L., and Walubita, L.F. Performance Characterization of Warm-Mix Asphalt Containing High Reclaimed-Asphalt

Pavement with Bio-Oil Rejuvenator. *Journal of Materials in Civil Engineering*, Vol. 32, No. 12, 2020.

Mivehchi, M., Wen, H., Wen, Y., and Wang, L. Study of Measures to Design Asphalt Mixes Including High Percentages of Recycled Asphalt Pavement and Recycled Asphalt Shingles. *Transportation Research Record: Journal of the Transportation Research Board*, 2022, pp. 036119812211089.

Mogawer, W.S., Austerman, A.J., Bonaquist, R., and Roussel, M. Performance Characteristics of Thin-Lift Overlay Mixtures: High Reclaimed Asphalt Pavement Content, Recycled Asphalt Shingles, and Warm-Mix Asphalt Technology. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2208, 2011, pp. 17–25.

Mohammad, L.N., Elseifi, M.A., Raghavendra, A., and Mengqiu, Y. *Hamburg Wheel-Track Test Equipment Requirements and Improvements to AASHTO T 324*. NCHRP Web-Only Document 219, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2016a.

Mohammad, L.N., Kim, M., and Challa, H. *Development of Performance-Based Specifications for Louisiana Asphalt Mixtures*. Louisiana Transportation Research Center. FHWA/LA.14/558, Louisiana Department of Transportation and Development, Baton Rouge, LA, 2016b.

Mohammadafzali, M., Ali, H., Musselman, J.A., Sholar, G.A., and Massahi, A. The Effect of Aging on the Cracking Resistance of Recycled Asphalt. *Advances in Civil Engineering*, Vol. 2017, 2017.

Mohseni, A., Carpenter, S., and D'Angelo, J. Development of Superpave High-Temperature Performance Grade (PG) Based on Rutting Damage. *Asphalt Paving Technology: Journal of the Association of Asphalt Paving Technologists*, Vol. 74, 2005, pp. 197–254.

Monosmith, C. *Fatigue Response of Asphalt-Aggregate Mixes*. SHRP Project A-404, Strategic Highway Research Program, 1994.

Mull, M.A., Stuart, K., and Yehia, A. Fracture Resistance Characterization of Chemically Modified Crumb Rubber Asphalt Pavement. *Journal of Material Science*, 2002, pp. 557–566.

NAPA. *Development of Marshall Procedures for Designing Asphalt Paving Mixtures*. Information Series 84, National Asphalt Pavement Association, Lanham, MD, 1982.

NAPA. BMD Implementation Efforts [online]. *National Asphalt Pavement Association*. Available from: https://www.asphaltpavement.org/expertise/engineering/resources/bmd-resource-guide/implementation-efforts [Accessed 1 December 2022].

NAPA. *Balanced Mix Design Resource Guide*. Informational Series 143, National Asphalt Pavement Association, Greenbelt, MD, 2021.

NCAT. *Quality Assurance (QA) Aspects of Performance Related Specifications (PRS)*. TFRS-01 Interim Report, National Center for Asphalt Technology, Auburn, AL, 2021.

NCHRP. Project Announcement for Guide for Implementing Performance Specifications [online]. *National Cooperative Highway Research Program*. Available from: https://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=4751 [Accessed 4 September 2022].

Nemati, R., Dave, E.V., and Sias, J.E. Statistical Evaluation of the Effects of Mix Design Properties on Performance Indices of Asphalt Mixtures. *Journal of Testing and Evaluation*, Vol. 48, No. 1, 2020.

Nsengiyumva, G., Kim, Y.-R., and Hu, J. *Feasibility and Implementation of Balanced Mix Design in Nebraska*. Nebraska Department of Transportation Research Reports. SPR-P1(19) M080, Nebraska Department of Transportation, Lincoln, NE, 2020.

ODOT *Supplement Specification 856: Bridge Deck Waterproofing Asphalt Surface Course*. Ohio Department of Transportation, Columbus, OH, 2014.

ODOT *Standard Specifications for Highway Construction*. Oklahoma Department of Transportation, Oklahoma City, OK, 2019.

ODOT *Supplemental Specification 800: Revisions to the 2019 Construction & Material Specifications*. Ohio Department of Transportation, Columbus, OH, 2022a.

ODOT *Supplement 1033: Determining the Cracking Resistance of Asphalt Mixtures Using the Indirect Tensile Asphalt Cracking Test*. Ohio Department of Transportation, Columbus, OH, 2022b.

ODOT *Supplement 1057: Method for Testing Asphalt Mixture Rutting Using the Asphalt Pavement Analyzer*. Ohio Department of Transportation, Columbus, OH, 2022c.

Ozer, H., Al-Qadi, I.L., Lambros, J., El-Khatib, A., Singhvi, P., and Doll, B. Development of the Fracture-Based Flexibility Index for Asphalt Concrete Cracking Potential Using Modified Semi-Circle Bending Test Parameters. *Construction and Building Materials*, Vol. 115, 2016, pp. 390–401.

Ozer, H., Al-Qadi, I.L., Singhvi, P., Bausano, J., Carvalho, R., Li, X., and Gibson, N. Prediction of Pavement Fatigue Cracking at an Accelerated Testing Section Using Asphalt Mixture Performance Tests. *International Journal of Pavement Engineering*, Vol. 19, No. 3, 2018, pp. 264–278.

Park, B., Cocconcelli, C., and Chun, S. Gradation Characteristics-Based Interlayer Mixture Design Method for Enhanced Rutting Resistance of Asphalt Pavements. *Canadian Journal of Civil Engineering*, Vol. 49, No. 4, 2022, pp. 607–616.

Rath, P., Gettu, N., Chen, S., and Buttlar, W.G. Investigation of Cracking Mechanisms in Rubber-Modified Asphalt through Fracture Testing of Mastic Specimens. *Road Materials and Pavement Design*, Vol. 23, No. 7, 2022, pp. 1544–1563.

Rice, J.R. A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks. *Journal of Applied Mechanics*, Vol. 35, 1968, pp. 379–386.

Rivera-Pérez, J., Ozer, H., Lambros, J., and Al-Qadi, I.L. Illinois Flexibility Index Test: Effect of Specimen Geometry and Test Configuration on the Asphalt Concrete Damage Zone. *Journal of Transportation Engineering, Part B: Pavements*, Vol. 147, No. 1, 2021, pp. 04020085.

Rodezno, M.C., Taylor, A.J., and Gu, F. *Crack Resistance and Durability of RAS Asphalt Mixtures-Phase 1*. FHWA/OH-2018-15, Ohio Department of Transportation, Columbus, OH, 2018.

Rodezno, M.C., Kaloush, K.E., and Corrigan, M.R. Development of a Flow Number Predictive Model. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2181, 2010, pp. 79–87.

Roja, K.L., Masad, E., and Mogawer, W. Performance and Blending Evaluation of Asphalt Mixtures Containing Reclaimed Asphalt Pavement. *Road Materials and Pavement Design*, Vol. 22, No. 11, 2021, pp. 2441–2457.

Sabouri, M. and Kim, Y.R. Development of a Failure Criterion for Asphalt Mixtures Under Different Modes of Fatigue Loading. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2447, 2014, pp. 117–125.

Safi, F.R., Al-Qadi, I.L., Hossain, K., and Ozer, H. Total Recycled Asphalt Mixes: Characteristics and Field Performance. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 12, 2019, pp. 149–162.

Sala, V.D., Tran, N., Yin, F., and Bowers, B.F. Evaluating Impact of Corrected Optimum Asphalt Content and Benchmarking Cracking Resistance of Georgia Mixtures for Balanced Mix Design Implementation. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2022, pp. 13–29.

Saleh, N.F., Mocelin, D.M., Castorena, C., and Kim, Y.R. Modeling Changes in Asphalt Mixture Properties with RAP Content. *In*: Anupam, K., Papagiannakis, A.T., Bhasin, A., and Little, D. (eds) *Advances in Materials and Pavement Performance Prediction II*, 2020.

Salim, R., Gundla, A., Zalghout, A., Underwood, B.S., and Kaloush, K.E. Relationship between Asphalt Binder Parameters and Asphalt Mixture Rutting. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 6, 2019, pp. 431–446.

Seitllari, A., Boz, I., Habbouche, J., and Diefenderfer, S.D. Assessment of Cracking Performance Indices of Asphalt Mixtures at Intermediate Temperatures. *International Journal of Pavement Engineering*, Vol. 23, No. 1, 2022, pp. 70–79.

Song, W., Huang, B., and Shu, X. Influence of Warm-Mix Asphalt Technology and Rejuvenator on Performance of Asphalt Mixtures Containing 50% Reclaimed Asphalt Pavement. *Journal of Cleaner Production*, Vol. 192, 2018, pp. 191–198.

Spadoni, S., Ingrassia, L.P., Mariani, E., Cardone, F., and Canestrari, F. Long-Term Performance Assessment of a Warm Recycled Motorway Pavement. *Case Studies in Construction Materials*, Vol. 17, 2022a.

Spadoni, S., Ingrassia, L.P., Mocelin, D., Kim, Y.R., and Canestrari, F. Comparison of Asphalt Mixtures Containing Polymeric Compounds and Polymer-Modified Bitumen Based on the VECD Theory. *Construction and Building Materials*, Vol. 349, 2022b.

Sreedhar, S., Coleri, E., Obaid, I.A., and Kumar, V. Development of a Balanced Mix Design Method in Oregon to Improve Long-Term Pavement Performance. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2675, No. 12, 2021, pp. 1121–1137.

Stevenson, J.D. and Aschenbrener, T. *Comparison of Test Results from Laboratory and Field Compacted Samples*. CDOT-DTD-R-94-3, Colorado Department of Transportation, Denver, CO, 1994.

Stuart, K.D., Mogawer, W.S., and Romero, P. *Validation of Asphalt Binder and Mixture Tests That Measure Rutting Susceptibility*. FHWA-RD-99-204, Federal Highway Administration, McLean, VA, 2000.

Su, K. and Hachiya, Y. Examination of Fiber-Added Recycled Asphalt Concretes for Surface Course in Airport Pavements. *First International Conference on Transportation Infrastructure*, Beijing, [Online]. Available from: https://www.researchgate.net/profile/Yoshitaka-Hachiya/publication/265104546_Examination_of_Fiber_-Added_Recycled_Asphalt_Concretes_for_Surface_Course_in_Airport_Pavements/links/543fb4330cf2fd72f99cf9a7/Examination-of-Fiber-Added-Recycled-Asphalt-Concretes-for-Surface-Course-in-Airport-Pavements.pdf [Accessed 29, January 2023] 2008.

Subramanian, V. *A Viscoplastic Model with Rate-Dependent Hardening for Asphalt Concrete in Compression*. Ph.D. Dissertation, North Carolina State University, Raleigh, NC, 2011.

Tangella, S.C.S.R., Craus, J., Deacon, J.A., and Monismith, C.L. *Summary Report on Fatigue Response of Asphalt Mixtures*. SHRP Project A-003-A, Strategic Highway Research Program, 1990.

Transportation Research Board (TRB). *Glossary of Transportation Construction Quality Assurance Terms: Seventh Edition*. Transportation Research Circular E-C235, Transportation Research Board, Washington, D.C., 2018.

TxDOT *Special Specification 3074: Superpave Mixtures – Balanced Mix Design*. Texas Department of Transportation, Austin, TX, 2019.

UDOT *Materials Manual of Instruction Section 990: Method of Test for Hamburg Wheel-Track Testing of Compacted Hot-Mix Asphalt (HMA)*. Utah Department of Transportation, Salt Lake City, UT, 2021.

UDOT *2023 Standard Specifications for Road and Bridge Construction*. Utah Department of Transportation, Salt Lake City, UT, 2022a.

UDOT *Materials Manual of Instruction Section 963: Superpave Mix Design Verification with the IDEAL Test*. Utah Department of Transportation, Salt Lake City, UT, 2022b.

Underwood, B.S., Baek, C., and Kim, Y.R. Simplified Viscoelastic Continuum Damage Model as Platform for Asphalt Concrete Fatigue Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2296, 2012, pp. 36–45.

Underwood, B.S., Kim, Y.R., and Guddati, M.N. Improved Calculation Method of Damage Parameter in Viscoelastic Continuum Damage Model. *International Journal of Pavement Engineering*, Vol. 11, No. 6, 2010, pp. 459–476.

Veeraragavan, R.K., Nener-Plante, D., Myers, L., Nash, C., and Tran, N.H. Balanced Mix Design Benchmarking of Field-Produced Asphalt Mixtures in Maine, U.S. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 5, 2022, pp. 263–276.

Vila-Cortavitarte, M., Lastra-González, P., Calzada-Pérez, M.Á., and Indacoechea-Vega, I. Analysis of the Influence of Using Recycled Polystyrene as a Substitute for Bitumen in the

Behaviour of Asphalt Concrete Mixtures. *Journal of Cleaner Production*, Vol. 170, 2018, pp. 1279–1287.

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

Walubita, L.F., Faruk, A.N.M., Alvarez, A.E., and Scullion, T. The Overlay Tester (OT): Using the Fracture Energy Index Concept to Analyze the OT Monotonic Loading Test Data. *Construction and Building Materials*, Vol. 40, 2013, pp. 802–811.

Walubita, L.F., Fuentes, L., Lee, S.I., Guerrero, O., Mahmoud, E., Naik, B., and Simate, G.S. Correlations and Preliminary Validation of the Laboratory Monotonic Overlay Test (OT) Data to Reflective Cracking Performance of In-service Field Highway Sections. *Construction and Building Materials*, Vol. 267, 2021.

Wang, Y.D., Ghanbari, A., Underwood, B.S., and Kim, Y.R. Development of a Performance-Volumetric Relationship for Asphalt Mixtures. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2673, No. 6, 2019, pp. 416–430.

Wang, Y.D. and Kim, Y.R. Development of a Pseudo Strain Energy-Based Fatigue Failure Criterion for Asphalt Mixtures. *International Journal of Pavement Engineering*, Vol. 20, No. 10, 2019, pp. 1182–1192.

Wang, Y.D., Underwood, B.S., and Kim, Y.R. Development of a Fatigue Index Parameter, Sapp, for Asphalt Mixes Using Viscoelastic Continuum Damage Theory. *International Journal of Pavement Engineering*, Vol. 23, No. 2, 2022, pp. 438–452.

Wen, H., Wu, S., Mohammad, L.N., Zhang, W., Shen, S., and Faheem, A. Long-Term Field Rutting and Moisture Susceptibility Performance of Warm-Mix Asphalt Pavement. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2575, 2016, pp. 103–112.

West, R.C., Rodezno, M.C., Leiva, F., and Taylor, A.J. *Regressing Air Voids for Balanced HMA Mix Design*. WHRP 0092-16-06, Wisconsin Department of Transportation, Madison, WI, 2018b.

West, R.C., Rodezno, M.C., Leiva, F., and Yin, F. *Development of a Framework for Balanced Mix Design*. NCHRP 20-07/Task 406, National Center for Asphalt Technology, Auburn, AL, 2018a.

West, R.C, Timm, D., Powell, B., Tran, N., Yin, F., Bowers, B., Rodenzo, C., Leiva. F., Vargas, A., Gu, F., Moraes, R., Nakhaei, M. *Phase VII (2018-2021) NCAT Test Track Findings.* NCAT Report 21-03, National Center for Asphalt Technology, Auburn, AL, 2021a.

West, R.C., Yin, F., Rodezno, M.C., and Taylor, A.J. *Balanced Mixture Design Implementation Support*. WHRP 0092-20-04, Wisconsin Department of Transportation, Madison, WI, 2021b.Witczak, M.W., Kaloush, K., Pellinen, T., El-Basyouny, M., and Von-Quintus, H.L. *Simple Performance Test for Superpave Mix Design*. NCHRP Report 465, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2002.

Wu, B., Luo, C., Pei, Z., Xia, J., Chen, C., and Kang, A. Effect of Different Polymer Modifiers on the Long-Term Rutting and Cracking Resistance of Asphalt Mixtures. *Materials*, Vol. 14, No. 12, 2021, pp. 3359.

Yin, F., Arámbula-Mercado, E., Lytton, R., Martin, A.E., and Cucalon, L.G. 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.

Yin, F., Chen, C., West, R.C., Martin, A.E., and Arámbula-Mercado, E. Determining the Relationship among Hamburg Wheel-tracking Test Parameters and Correlation to Field Performance of Asphalt Pavements. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2674, No. 4, 2020a, pp. 281–291.

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

You, T., Shi, Y., Mohammad, L.N., and Cooper, S.B. Laboratory Performance of Asphalt Mixtures Containing Re-Refined Engine Oil Bottoms Modified Asphalt Binders. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, No. 28, 2018, pp. 88–95.

Yun, T. *Development of a Viscoplastic Constitutive Model Using a Rate-Dependent Yield Criterion for HMA in Compression*. Ph.D. Dissertation, North Carolina State University, Raleigh, NC, 2008.

Zalghout, A., Castro, S., Karam, J., and Kaloush, K. Laboratory and Field Evaluation of Plant Produced Asphalt Mixtures Containing RAP in Hot Climate: A Case Study from Phoenix, Arizona. *Construction and Building Materials*, Vol. 351, 2022, pp. 128322.

Zaremotekhases, F., Sadek, H., Hassan, M., and Berryman, C. Impact of Warm-Mix Asphalt Technologies and High Reclaimed Asphalt Pavement Content on the Performance of Alternative Asphalt Mixtures. *Construction and Building Materials*, Vol. 319, 2022.

Zaumanis, M., Arraigada, M., Wyss, S.A., Zeyer, K., Cavalli, M.C., and Poulikakos, L.D. Performance-Based Design of 100% Recycled Hot-Mix Asphalt and Validation Using Traffic Load Simulator. *Journal of Cleaner Production*, Vol. 237, 2019.

Zhang, J., Simate, G.S., Hu, X., Souliman, M., and Walubita, L.F. Impact of Recycled Asphalt Materials on Asphalt Binder Properties and Rutting and Cracking Performance of Plant-produced Mixtures. *Construction and Building Materials*, Vol. 155, 2017, pp. 654–663.

Zhang, R., Daniel, J.S., and Dave, E.V. Evaluation of Viscoelastic Properties and Cracking Behaviour of Asphalt Mixtures with Laboratory Aging. *In*: *Poulikakos, L.D., Cannone Falchetto, A., Wistuba, M.P., Hofko, B., Porot, L., and Di Benedetto, H. (eds) RILEM 252-CMB Symposium. RILEM 252-CMB 2018. RILEM Bookseries Vol. 20. Springer International Publishing*, 2019, pp. 33–38.

Zhang, W., Chen, X., Shen, S., Mohammad, L.N., Cui, B., Wu, S., and Khan, A.R. Investigation of Field Rut Depth of Asphalt Pavements Using Hamburg Wheel Tracking Test. *Journal of Transportation Engineering, Part B: Pavements*, Vol. 147, No. 1, 2021, pp. 04020091.

Zhao, S., Huang, B., Shu, X., and Woods, M. Comparative Evaluation of Warm Mix Asphalt Containing High Percentages of Reclaimed Asphalt Pavement. *Construction and Building Materials*, Vol. 44, 2013, pp. 92–100.

Zhou, F. *Implementation of the IDEAL Cracking Test for Asphalt Mix Design and QC/QA*. NCHRP 20-44 (16), National Cooperative Highway Research Program, Transportation Research Board, 2021.

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

Zhou, F., Hu, S., Scullion, T., Mikhail, M., and Walubita, L.F. A Balanced HMA Mix Design Procedure for Overlays. *Asphalt Paving Technology: Journal of the Association of Asphalt Paving Technologists*, Vol. 76, 2007, pp. 823–850.

Zhou, F., Im, S., Hu, S., Newcomb, D., and Scullion, T. Selection and Preliminary Evaluation of Laboratory Cracking Tests for Routine Asphalt Mix Designs. *Road Materials and Pavement Design*, Vol. 18, 2017a, pp. 62–86.

Zhou, F., Im, S., Sun, L., and Scullion, T. Development of an IDEAL Cracking Test for Asphalt Mix Design and QC/QA. *Road Materials and Pavement Design*, 2017b.

Zhou, F. and Scullion, T. *Upgraded Overlay Tester and Its Application to Characterize Reflection Cracking Resistance of Asphalt Mixtures*. FHWA/TX-04/0-4467-1, Texas Department of Transportation, Austin, TX, 2004.

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

Zhou, L., Huang, W., Zhang, Y., Lv, Q., and Sun, L. Mechanical Evaluation and Mechanism Analysis of the Stripping Resistance and Healing Performance of Modified Asphalt-Basalt Aggregate Combinations. *Construction and Building Materials*, Vol. 273, 2021a.

APPENDIX B. DETAILED ANALYSIS OF CONTRACTOR SURVEY

Table B.1. Contractor Questions Related to Demographic and General Information.

Contractor	A	B	C	D	E	F
Size of company (tons per year)	1,000,000 in VA 20,000 in NC	1,600,000	600,000	1,000,000	350,000	300,000
Number of plants	9 in VA 0 in NC	15, Building the 16th	3	10	3	1 in NC, 3 in SC
Larger plant (ton per year)	300,000	500,000	300,000	Not sure	150,000	400,000 to 500,000
Smallest plant (ton per year)	80,000	20,000	70,000	<100,000	60,000	200,000 to 300,000
Area of actuation	Division 1	Divisions 1, 2, 3, 4, 5, 8 and 7	Divisions 3, 5, 6, and 8	Divisions 11, 12, 13	Divisions 7, 8 and 9	Division 10
Know about BMD?	Yes, from VA experience	Not much	Yes, from seminars and word of mouth	Not much	Yes, from seminars	Yes, from seminars and classes
General opinion	Need to allow freedom to design by BMD specs. Currently, there are too many restrictions from agencies.	Not much	Would like to get away from volumetrics. Curious of how performance testing would work	Not sure	Not against it	Performance over rules makes more sense. Relearning asphalt mixes and testing may be an issue
Issues if additional testing added to QC?	More work, twice as many samples, for volumetrics and performance.	Not sure	Equipment, lab space	Can cause rushed procedures	More equipment, more work during mix design	Room for more equipment, more training, more time.

Table B.2. Contractor Questions Related to Operational Issues.

Contractor	A	B	C	D	E	F
Number of design per year per mix type	4 to 6 designs per plant, covering all design types	This year ~ 150 but 50 in a typical year due to a new plant and increase RAP	Not every year, 4 per year	10 per year	6 or 7	Based on material changes, about 6 a year
What triggers new mix design?	New sources, and trend lines	New RAP spec, change in quarries, new plants, sand sources	Change of material: quarry and increase in RAP	New sources, replacement designs	Percentage or recycle mix they don't have (low RAP or virgin)	Material changes, production trend lines, quarry changes
How much time is spent on mix designs?	One week per design	Full time design tech	Week to week, and half per design	3 days to a week	3 weeks total, a week each	One week per design
Mix design steps	Target point based on gravities, past experience, similar mixes	Based on past experience	Based on previous design experience, start at lower increments	Target point based on experience. They have done trial gradations before	Use what works. Consistent quarries, don't need to change much	Update gradations averages. Adjust targets based on experience. Middle increments, then high and low.
Perform verification on APA before submitting a design?	Yes, have been doing for VA	No	No	No	No	No
How often APA have failed?	Never	No	Years ago	No	No	No
Does NCDOT ever request APA samples with redesign?	No	On big changes (binder and material delete)	No	No	No	No
How long to run full series QC?	3 hours	< 3 hours	3.5 hours	3 hours	1.5 hours	2.5 hours
Do all labs have a test press for design?	No, have only two	No, have only five	No, have only one	No, have only two	One at design lab, and one in another plant	Yes
Do you have a central design lab?	Can do design at all plants but test at only two	Yes. They are using 2 other labs this year sue to volume	Yes	Yes	Yes	Yes
Do you submit any mix design? If so, what percentage?	Yes, Trimat did 4	P401's only	P401 and OGFC	Fed jobs, and occasional designs due to time crunch	Only P401, one a year	No

Table B.3. Contractor Questions Related to Training.

Contractor	A	B	C	D	E	F
No. of mix tech.	0 in NC, Five in VA	Six	One	One	One	One
Number of level 1 or 2 techs	Two	Level 1: three, Level 2: sixteen	Level 1: three, Level 2: two	Level 1: six, Level 2: three	Level 1: two, Level 2: two	Level 1: two, Level 2: one
How do you train for mix design certification?	Train stone properties, then test during production, then mix design testing	Class of 8-10 days in lab and classroom, plus hands on in lab	Hands on in the lab	Hands on in the lab	During winter, trains on mixing and batching, QC manager does calculations	Physical side, then paperwork and blends. Begin with past mixes that worked, then tweak
How do you train for QC?	Same as mix design. Hands on with senior techs	Read specs, then hands on (2-3 weeks, then training on paperwork). Rotate with senior techs	Hands on in the lab, NCDOT On-the-Job Training Program	Hands on in the lab	Hands on with senior tech. Time depends on how fast they learn, plus class scheduling	Daily moistures, gradations. Then show by example then technical items. Paperwork during production. QMS manual review
How do you train for TSR?	Allows techs to make the pills and explain test, then send all TSR's to anti strip supplier	Mostly mix design techs, or senior techs (4-5 people)	Only the senior tech does it	Only for level 1, run by level 2 people	Only the senior tech does it	Only the senior tech does it
How often train new personnel?	7-9 people over 3 years	See previous answer	Yearly	Mostly long term employees. Only one hire last year and half	As they are hired	Every year when hired
How often train existing personnel?	Yearly meetings for refreshing	Done during production. Would like to set a training session	Active training daily, NCDOT training	Check on people in labs. Leave level 2 alone, new employees daily, experienced weekly	Assessed internally by QC manager and lead tech, and by NCDOT	Internally through assessments with NCDOT. Would like to set a yearly training session
How many hours train for QC?	Average three weeks	Two to six months	One year	Four to five weeks	Six months	At least a month
How many hours for mix design?	One year for testing experience, and one season for mix design	Two to three years	A year with continuous designs, otherwise five years	Three seasons	Six months of dedicated mix designs	One year

Table B.4. Contractor Questions Related to General Practices.

Contractor	A	B	C	D	E	F
Overall strategy for mix design	Keep what works for each division. Sometimes change materials due to price	Maximize RAP content, minimize asphalt binder content	Management controls materials and may switch sometimes	Maximize RAP without fractionating, use #9's because they have surplus	Maximize RAP, but mostly based on experience with materials	Cheapest mix possible, Maximize RAP and RAS
Number of active JMFs by mix type	4 for NC, 25 for VA	300 to 500 approved, 200 used	5 per plant, 15 total	200 total	2 active per mix type, one with RAS and one without	2 active per mix type, one with RAS and one without
Are there any 9.5B and 9.5C identical?	Yes	No	Yes, with different asphalt contents	No	Not anymore since 2018	Not anymore since 2018
Factors considered when developing a mix design	Cost of aggregates, required mixes	Cost of asphalt binder, RAP, aggregates, sand, number of bins, be able to run multiple mixes, tank availability	Aggregates availability, sand sources, cost of materials, design something that looks good and durable	Try to keep gradation and specs in the middle if range. Does not take asphalt content into account	Excess RAP, anti-strip, no material changes, back up design	Overall cost and quarry availability
Account for differences between lab mixes and plant mixes?	Add baghouse fines to design (1.0% to 1.5%)	Add baghouse fines to design, adjust gradations	Add baghouse fines to design (1% works) Designs are true to plant	Redo and adjust or get new material and asphalt binder	Mostly gradation changes or gravities	Modifies RAS and RAP procedures during mixing to get more realistic AC target
Why would you switch JMF during production?	Trend lines going out. May change to different state or private mixes	Trend lines going out, or running out of materials	Trend lines going out, or running out of materials	Material changes	Issues on mixes, or seasonal changes (no RAS on Fall/Winter)	Trend lines going out, but usually at end of the day

APPENDIX C. DETAILED SUMMARY OF BMD EXPERIMENTS

Plant F (RS9.5B) BMD Summary

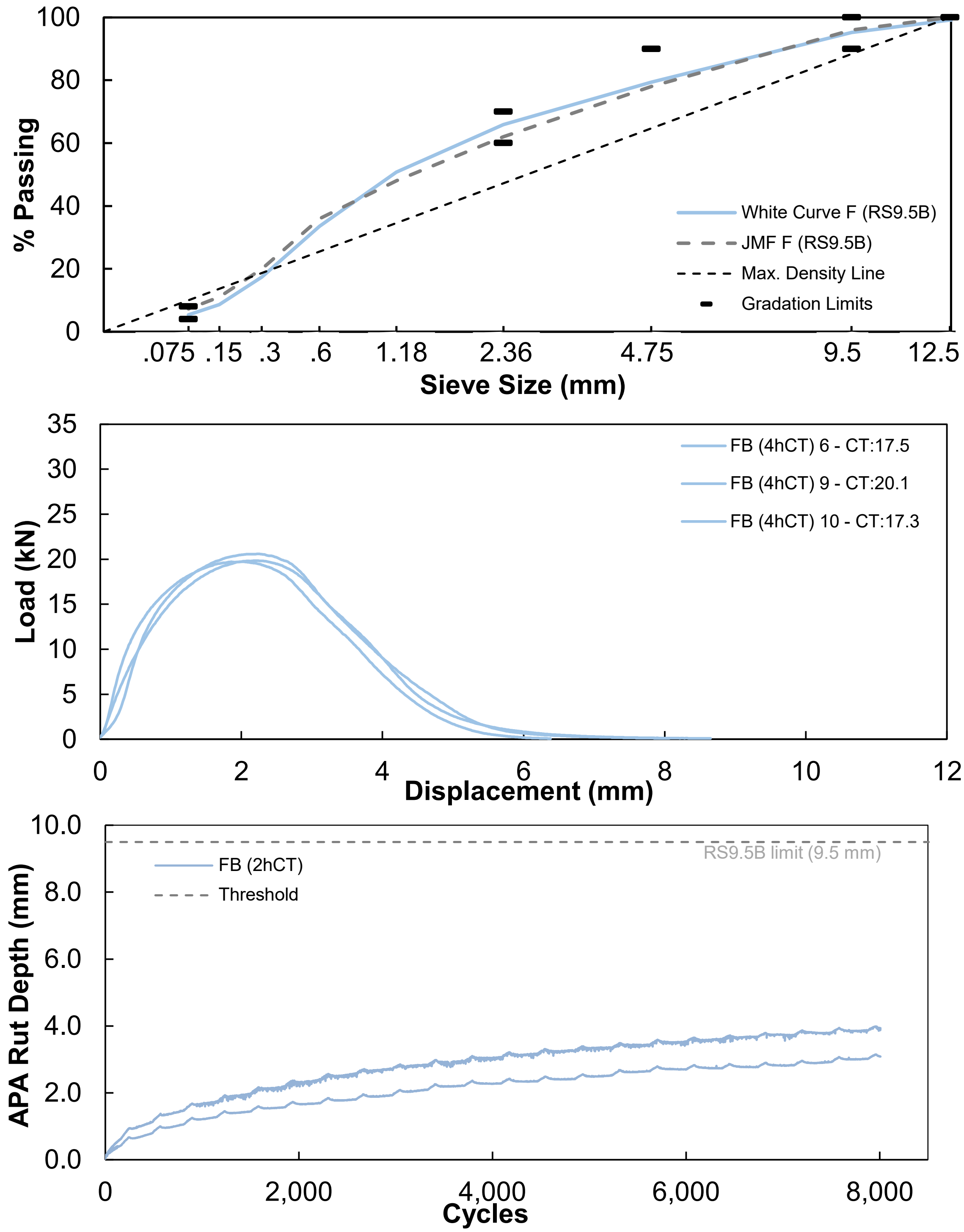


Figure C.1. Plant F (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant Y (RS9.5B) BMD Summary

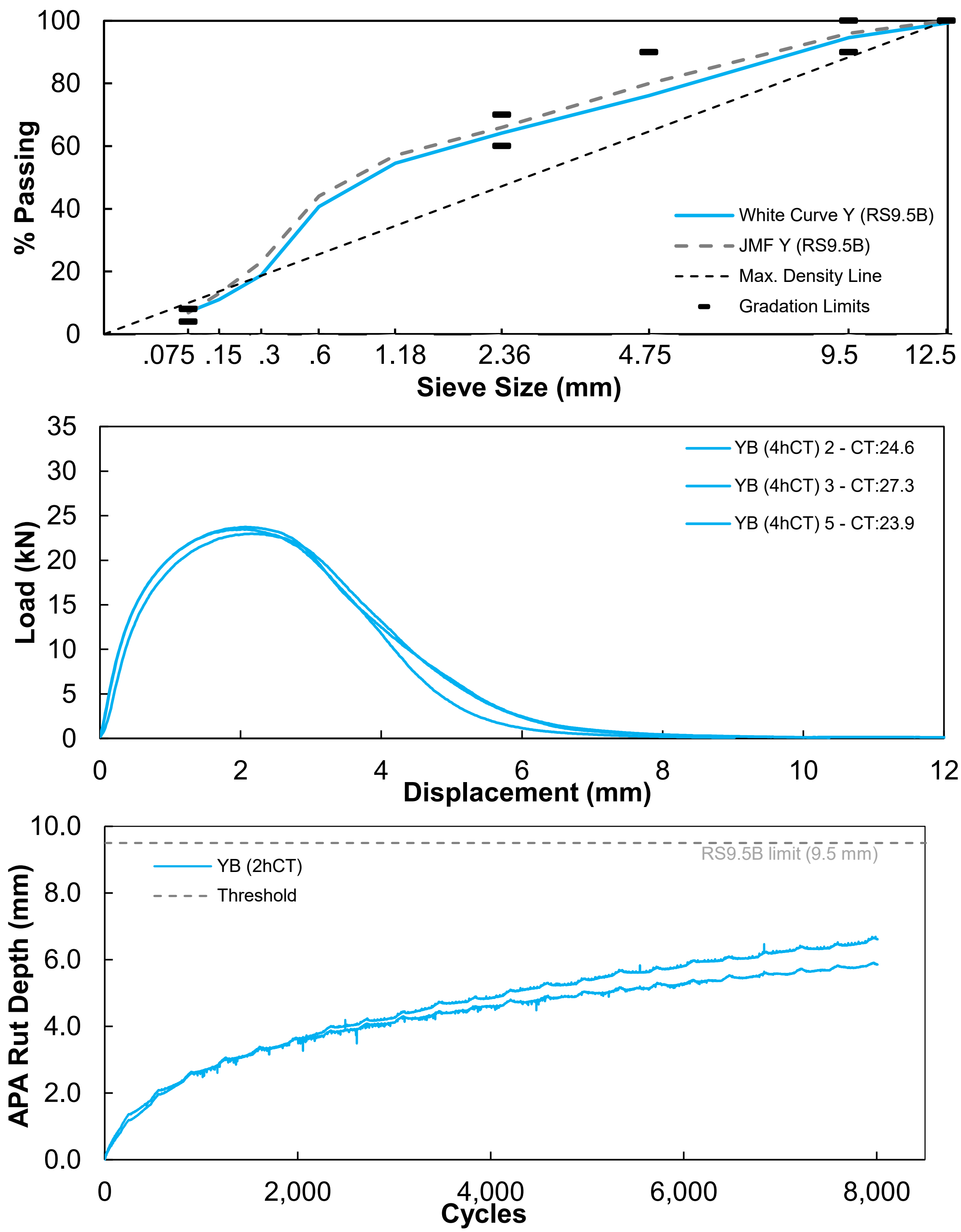


Figure C.2. Plant Y (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant W (RS9.5B) BMD Summary

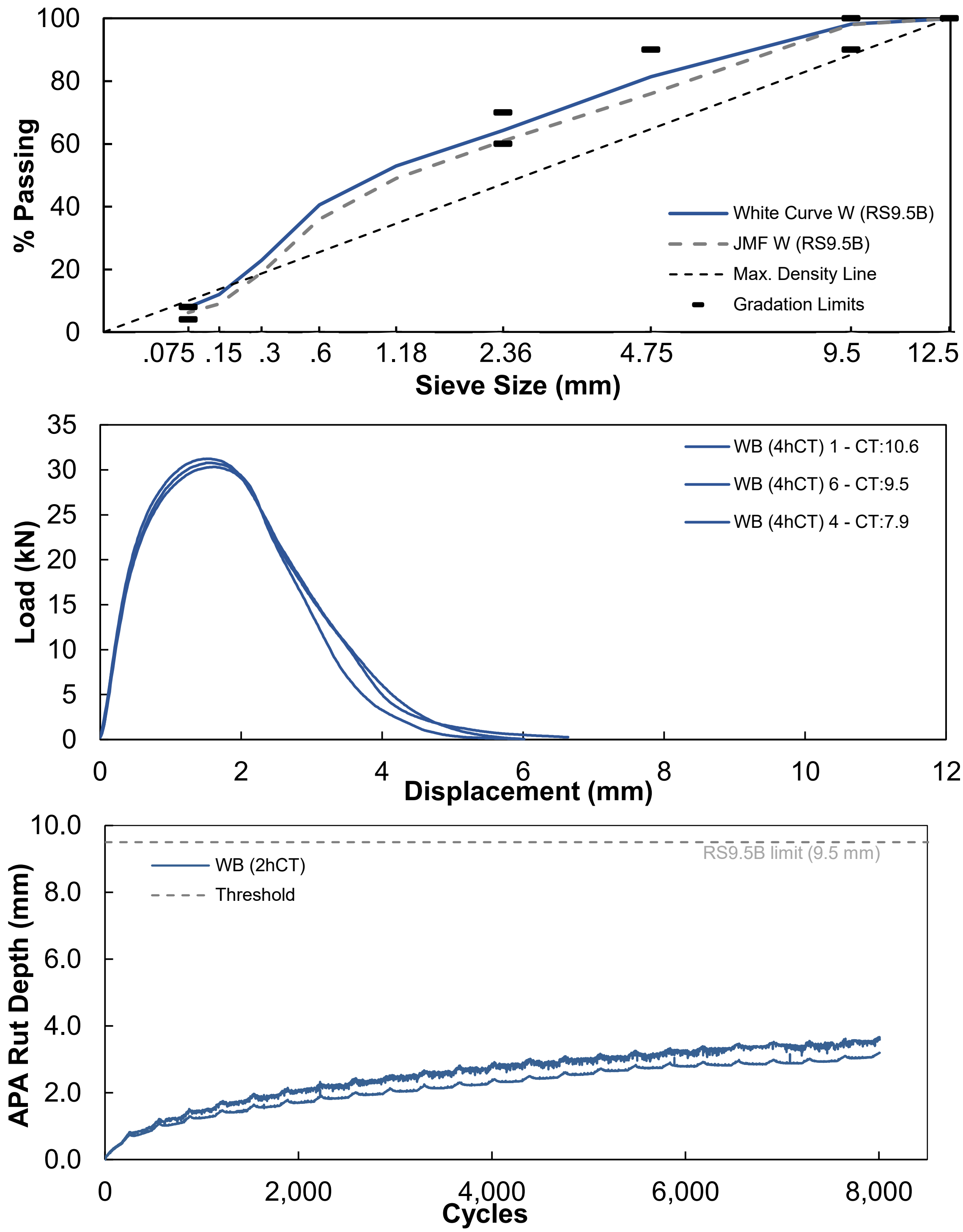


Figure C.3. Plant W (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant S (RS9.5B) BMD Summary

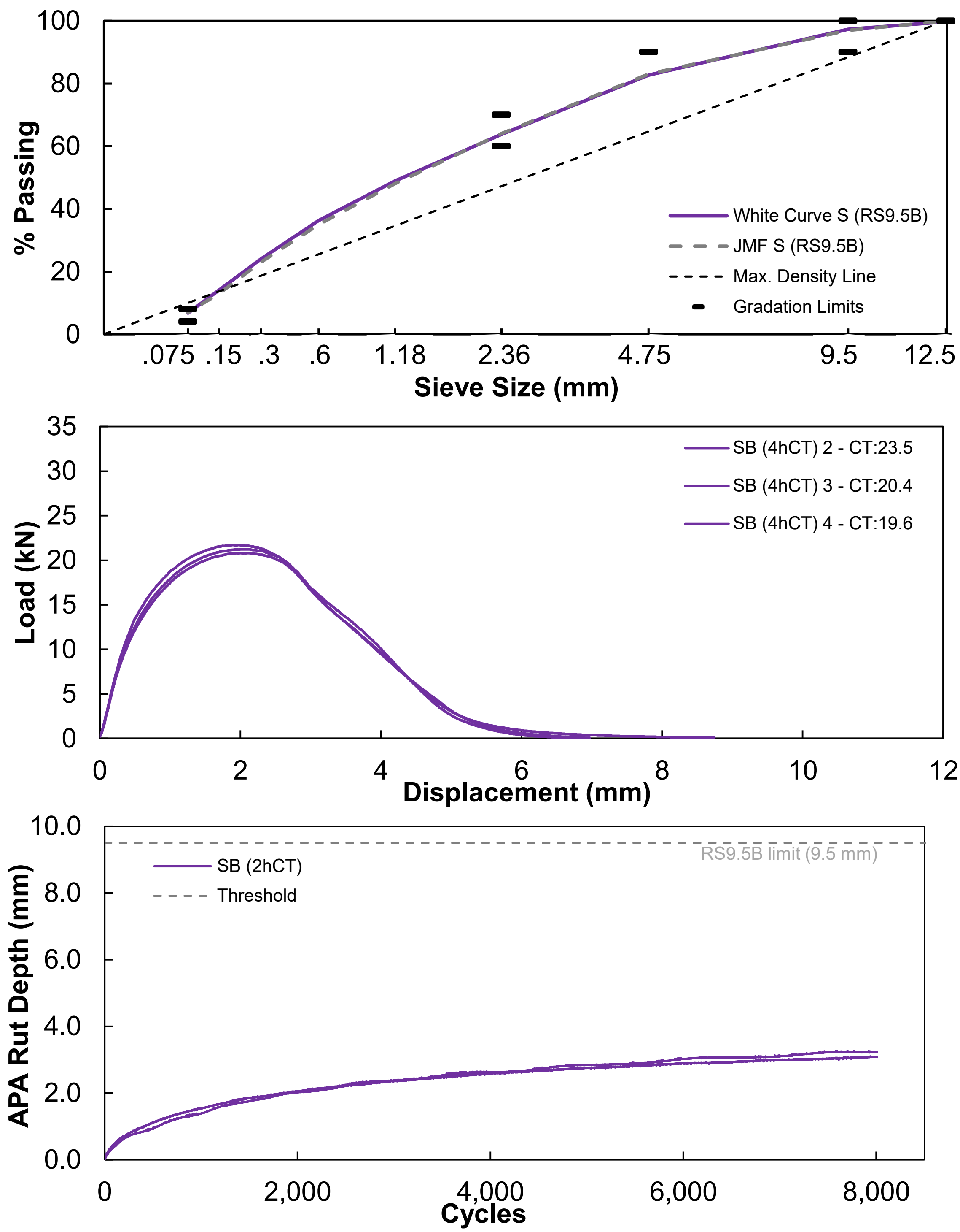


Figure C.4. Plant S (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant U (RS9.5B) BMD Summary

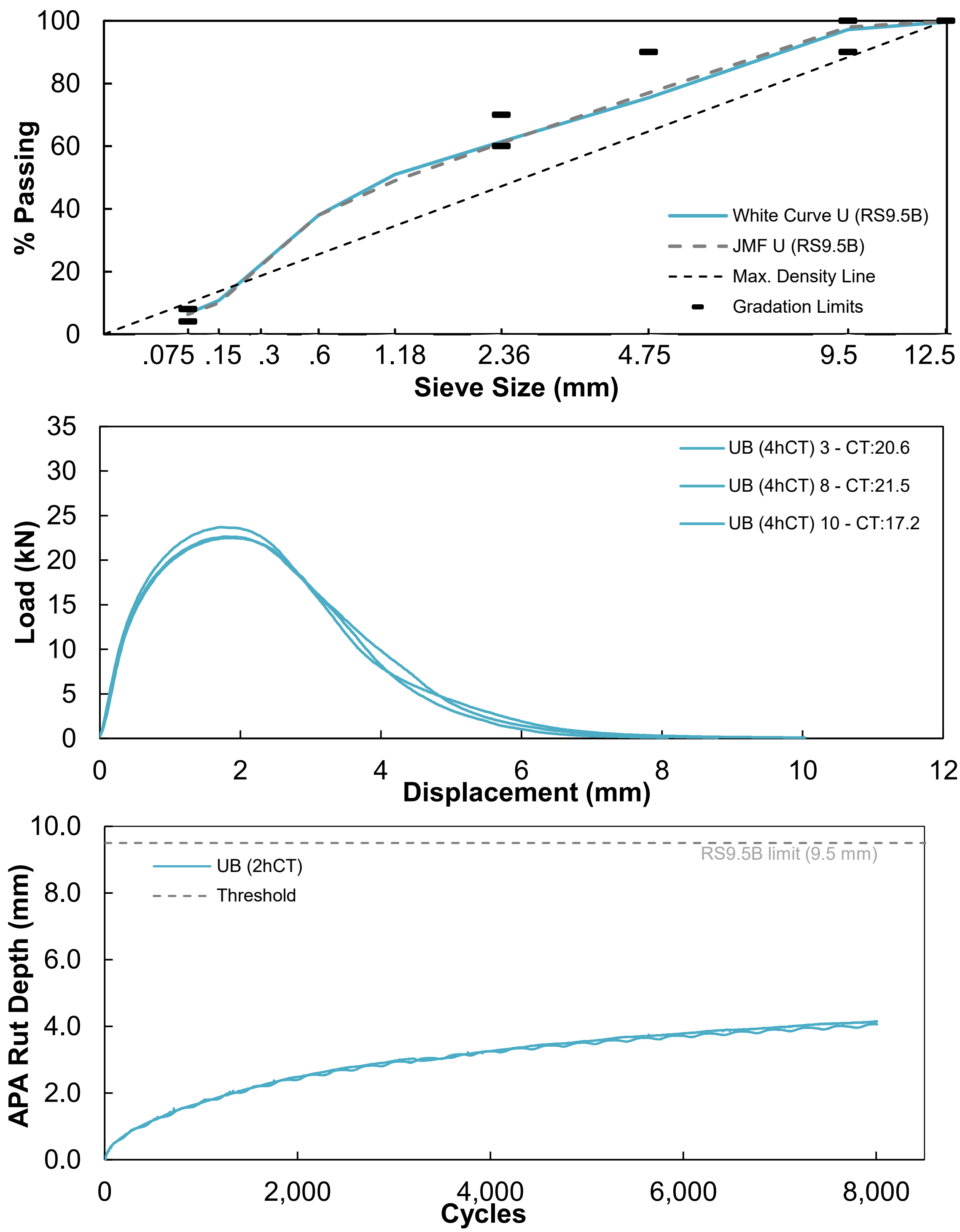


Figure C.5. Plant U (RS9.5B) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant H (RS9.5C) BMD Summary

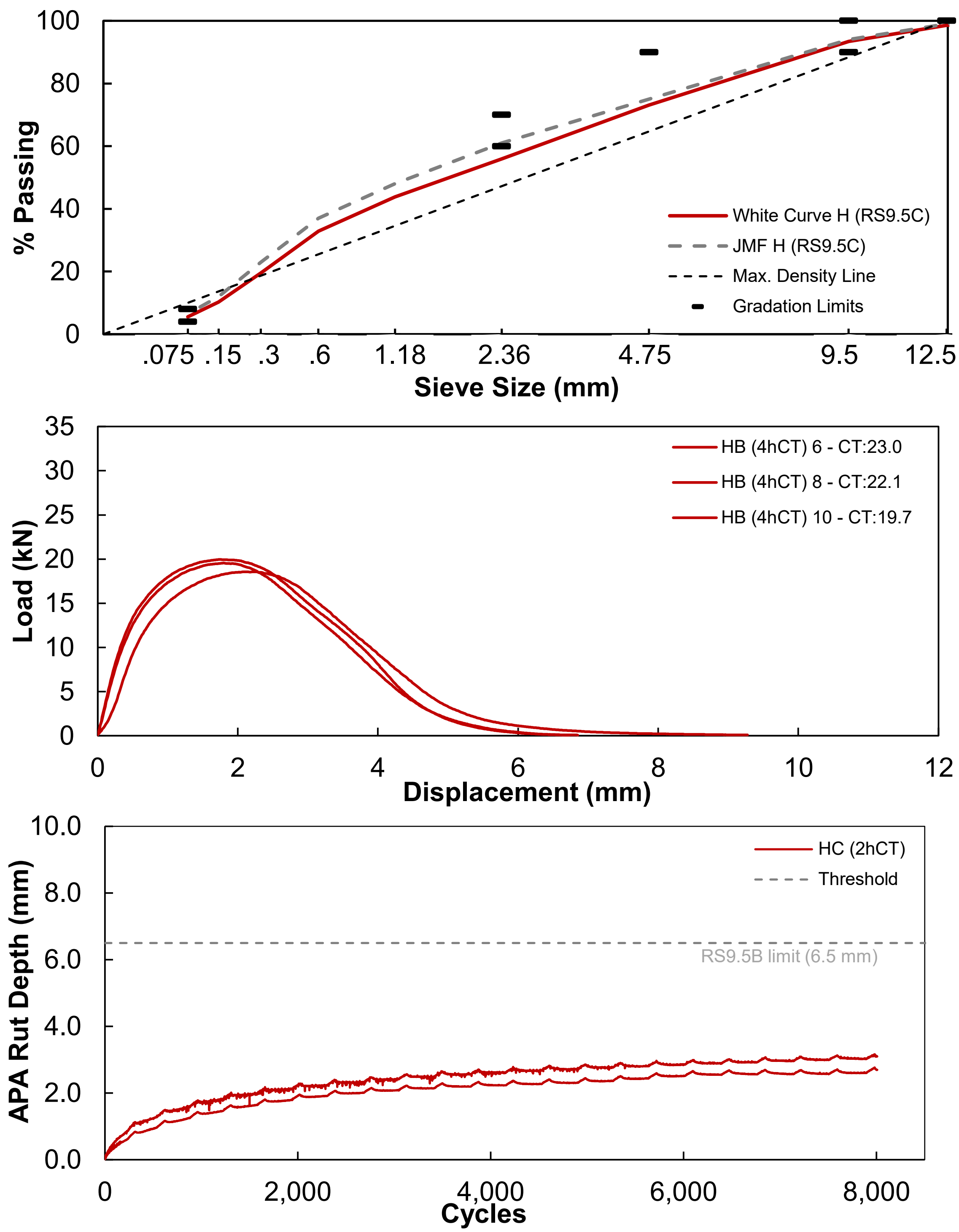


Figure C.6. Plant H (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant Z (RS9.5C) BMD Summary

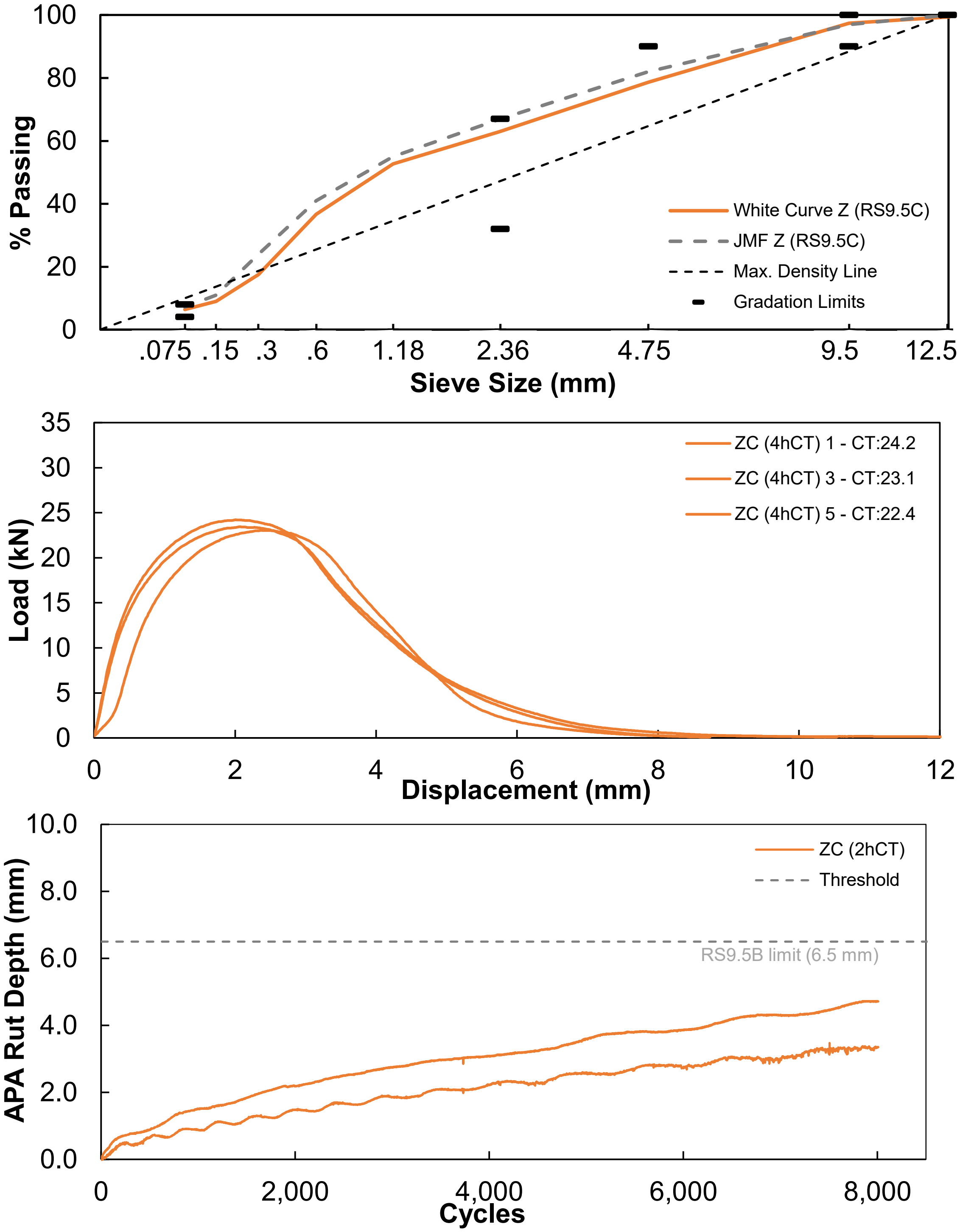


Figure C.7. Plant Z (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant V (RS9.5C) BMD Summary

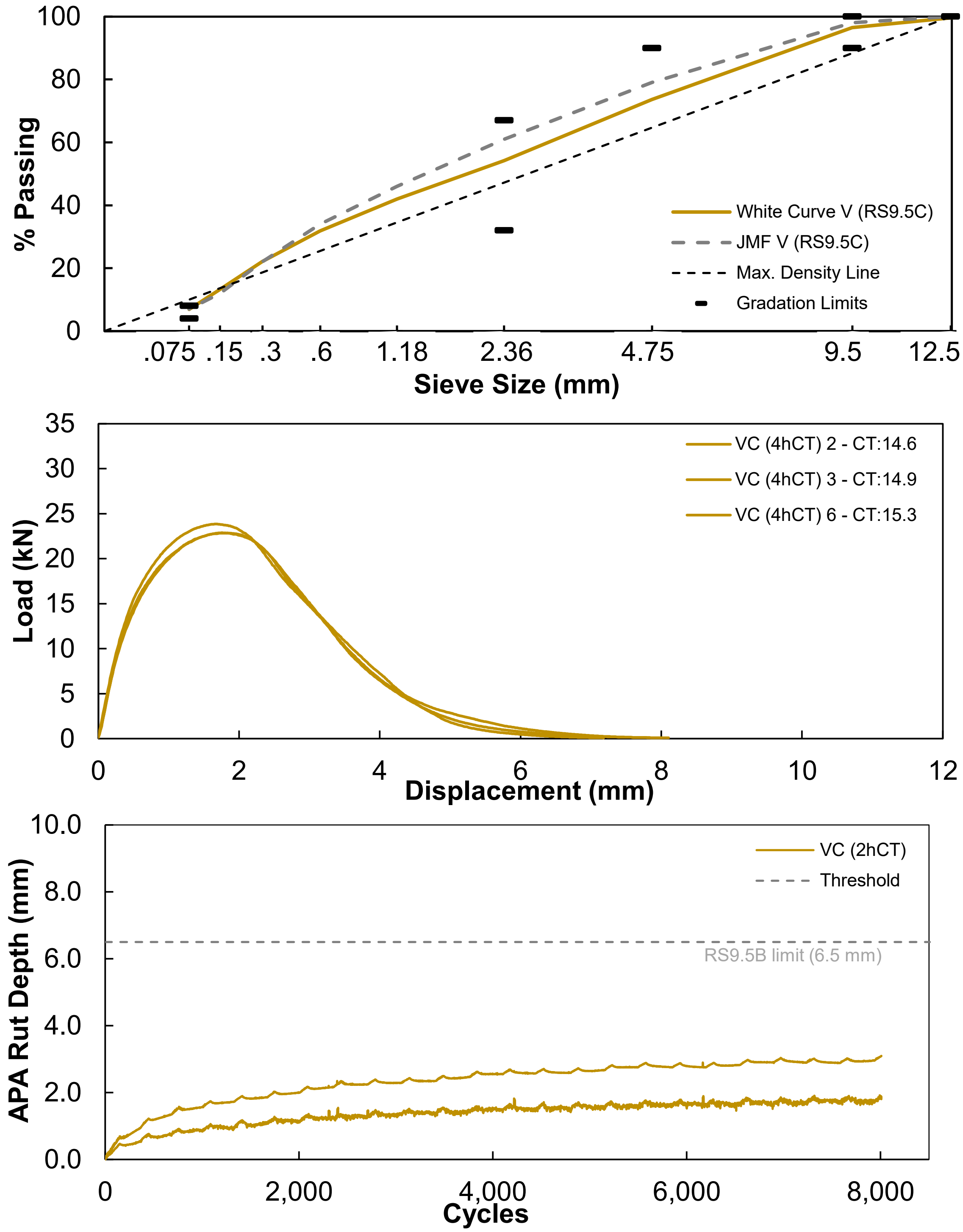


Figure C.8. Plant V (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant X (RS9.5C) BMD Summary

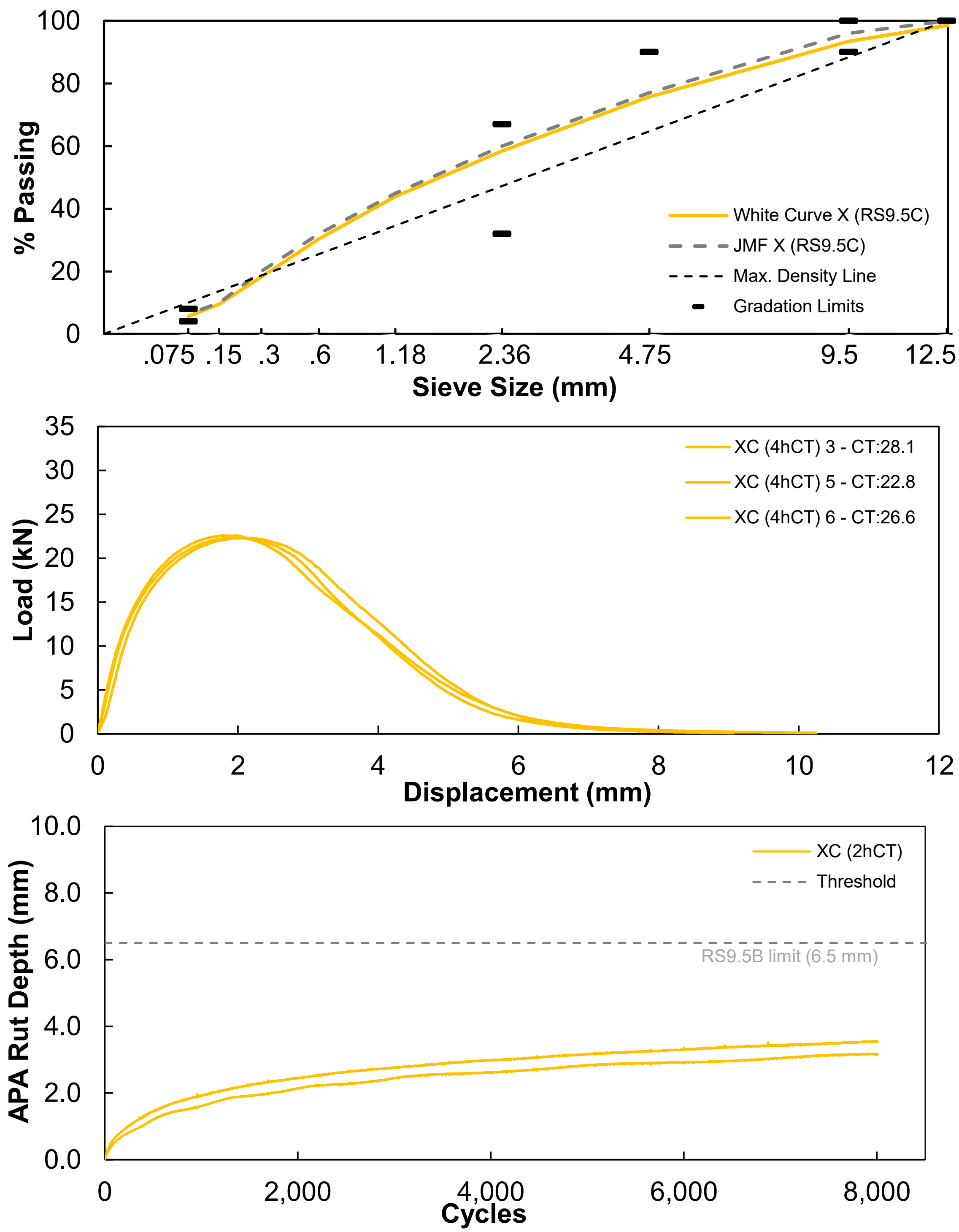


Figure C.9. Plant X (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant U (RS9.5C) BMD Summary

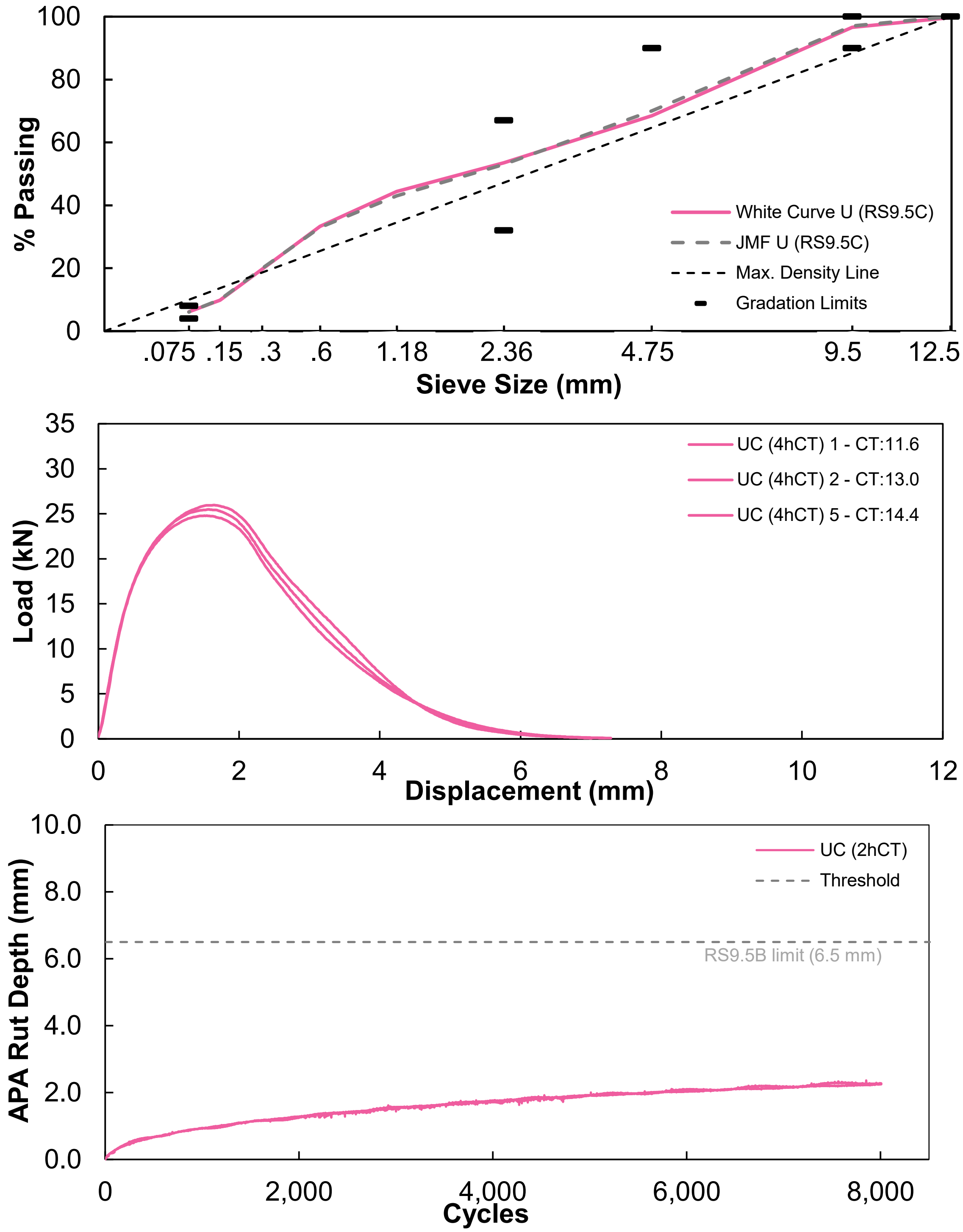
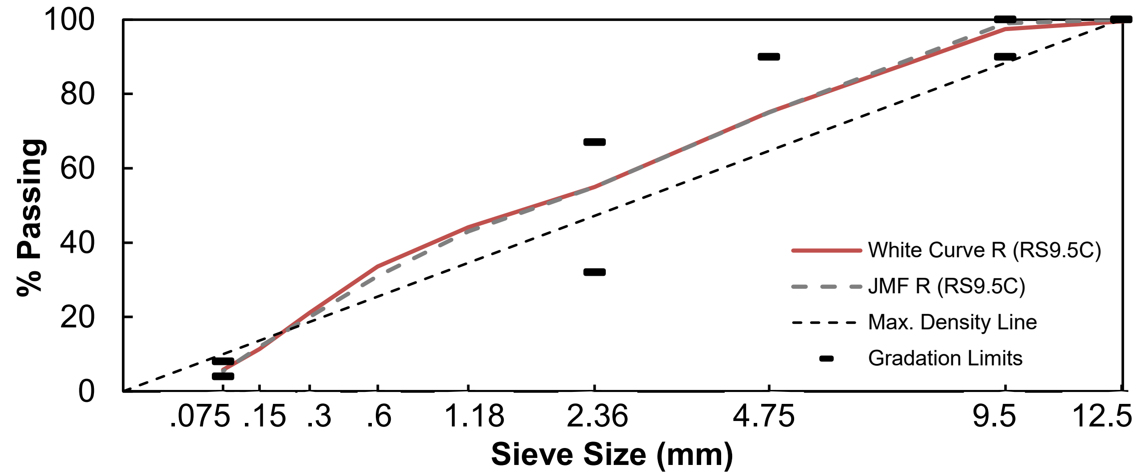
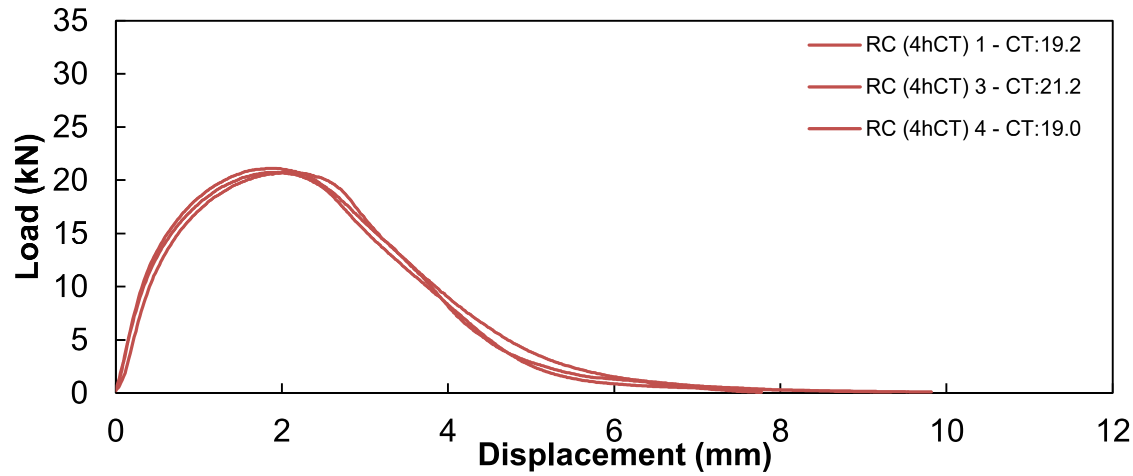


Figure C.10. Plant U (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant R (RS9.5C) BMD Summary





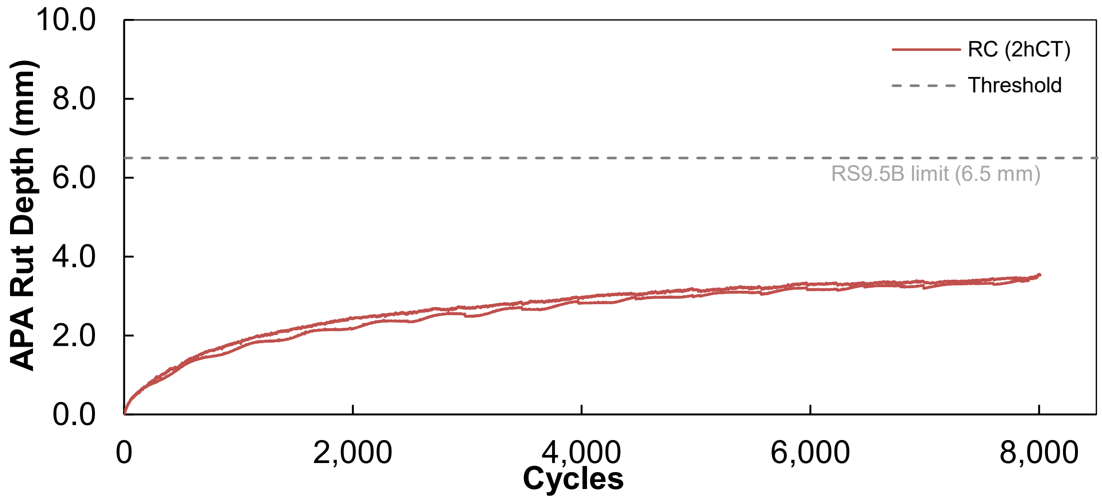


Figure C.11. Plant R (RS9.5C) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant V (RS9.5D) BMD Summary

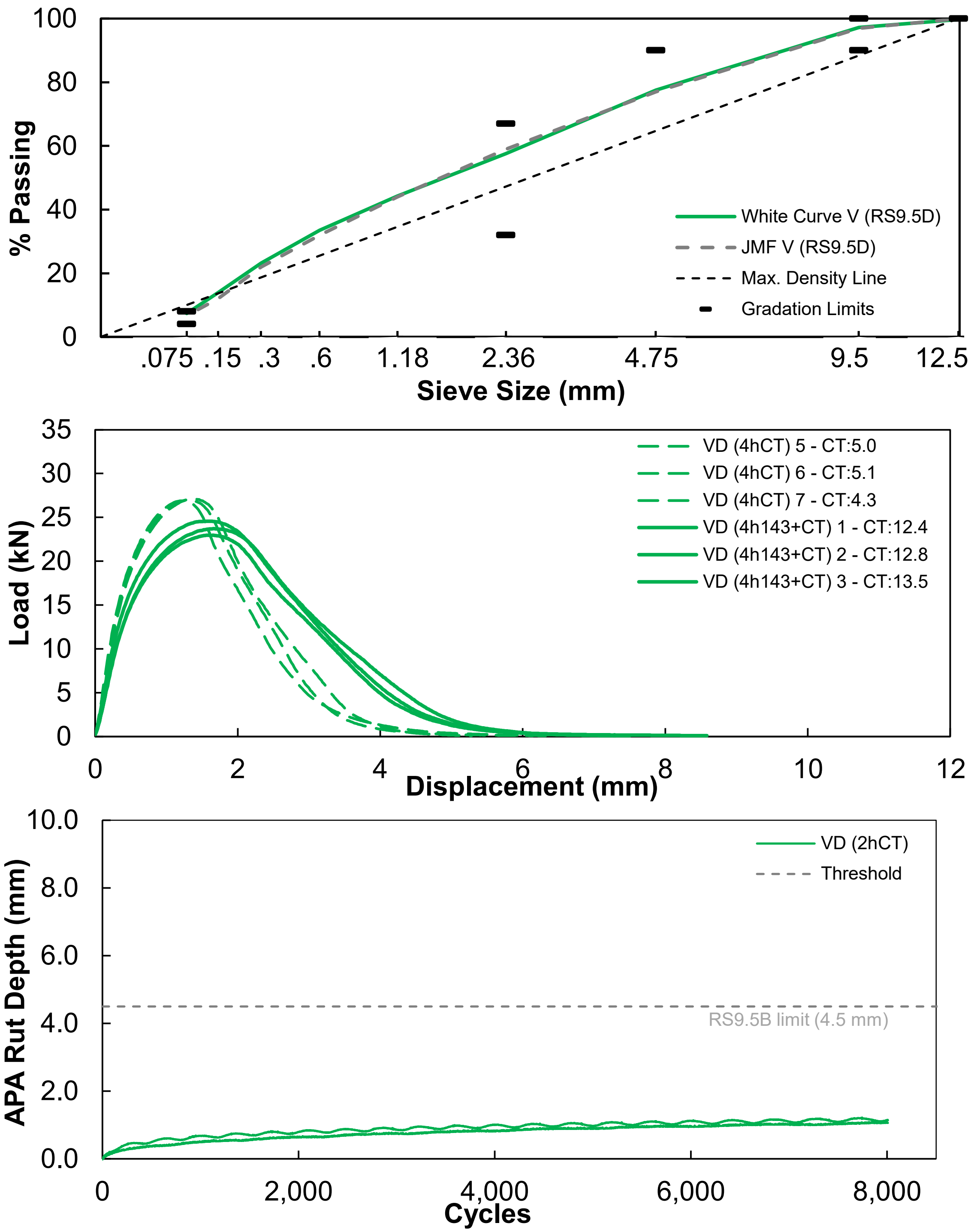


Figure C.12. Plant V (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant T (RS9.5D) BMD Summary

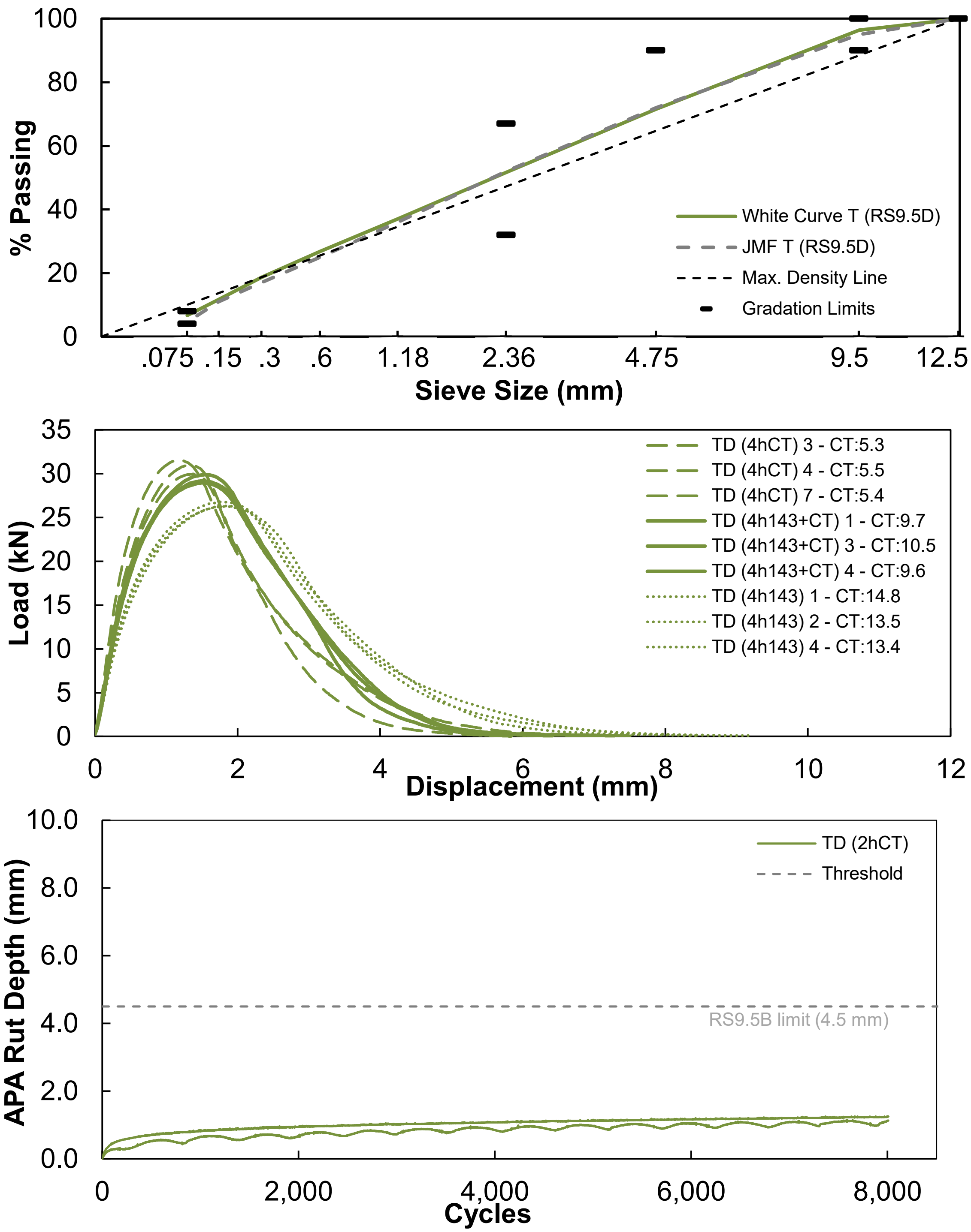


Figure C.13. Plant T (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

Plant S (RS9.5D) BMD Summary

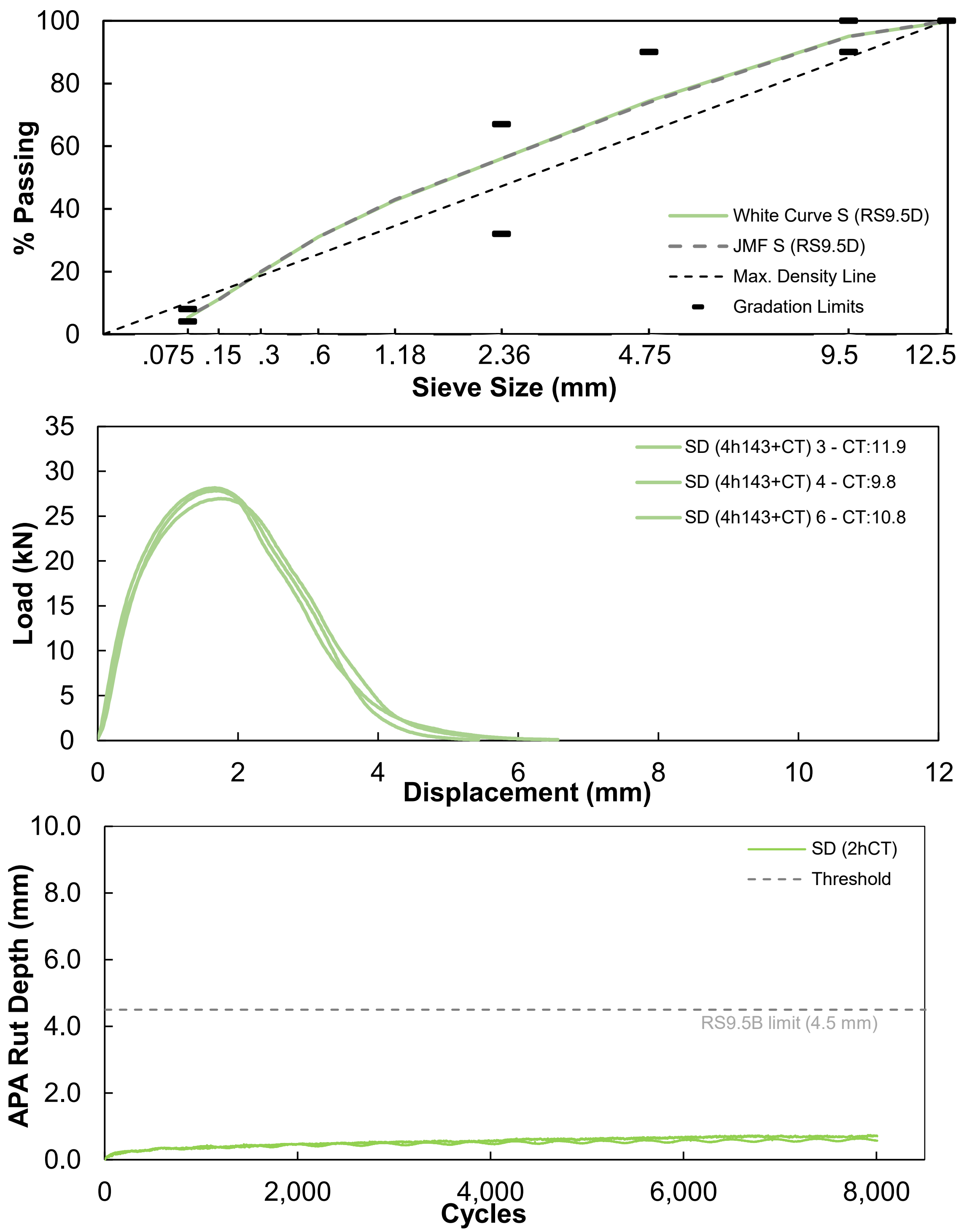


Figure C.14. Plant S (RS9.5D) BMD summary: (top) Gradation, (middle) IDT-CT load *vs.* displacement curve, (bottom) APA rut depth curve.

IDT-CT Statistical Results

Table C.1. Analysis of Variance for IDT-CT on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Individual Mixtures	0.0707, 0.5682	**<0.0001**

Tukey HSD Grouping	Mean	Group Letters						
XC	25.81	A						
YB	25.26	A	B					
ZC	23.24	A	B	C				
HC	21.59	A	B	C	D			
SB	21.16		B	C	D			
RC	19.79			C	D			
UB	19.77			C	D			
FB	18.32				D	E		
VC	14.92					E	F	
UC	13.00						F	G
VD	12.89						F	G
SD	10.82						F	G
TD	9.96							G
WB	9.32							G

Table C.2. Analysis of Variance for IDT-CT on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F
Mix Type	**0.0210**, **0.0022**	**<0.0001**

Tukey HSD Grouping	Mean	Group Letters	
RS9.5C	19.72	A	
RS9.5B	18.76	A	
RS9.5D	11.22		B

Table C.3. Analysis of Variance for IDT-CT on Mix Type (Excluding WB and RS9.5D).

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Mix Type	0.0915, 0.1511	0.3822

Table C.4. Mix Type and Source Region Mixed Model for IDT-CT.

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Mix Type	0.0210, 0.0022	0.0010
Source Region	0.2067, 0.2478	0.1955

Tukey HSD Grouping	Mean	Group Letters
Not performed for Source Region		

APA Statistical Results

Table C.5. Analysis of Variance for APA on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F
Individual Mixtures	1.000, **0.0252**	**<0.0001**
Tukey HSD Grouping	**Mean**	**Group Letters**
YB	6.24	A
UB	4.11	B
ZC	4.03	B
RC	3.55	B C
FB	3.49	B C
WB	3.42	B C
XC	3.36	B C
SB	3.16	B C
HC	2.91	B C D
VC	2.47	B C D E
UC	2.26	C D E F
TD	1.24	D E F
VD	1.11	E F
SD	0.65	F

Table C.6. Analysis of Variance for APA on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F
Mix Type	0.1159, **0.0107**	**<0.0001**
Tukey HSD Grouping	**Mean**	**Group Letters**
RS9.5B	19.72	A
RS9.5C	18.76	B
RS9.5D	11.22	C

Table C.7. Analysis of Variance for APA on Mix Type (Excluding YB and RS9.5D).

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Mix Type	0.4019, 0.1537	0.1526

Table C.8. Mix Type and Source Region Mixed Model for APA.

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F
Mix Type	0.1159, 0.0107	<0.0001
Source Region	0.5904, 0.8095	0.0586
Tukey HSD Grouping	**Mean**	**Group Letters**
Not performed for Source Region		

Table C.9. APA Rut Depths Difference to Reported in the Job Mix Formula.

Mixture	NCSU APA Rut Depth (mm)	JMF APA Rut Depth (mm)	Difference from JMF (mm)	Percent Difference from JMF
FB	3.49	5.90	-2.4	-41%
YB	6.24	3.50	+2.7	+78%
WB	3.42	5.40	-2.0	-37%
SB	3.16	3.60	-0.4	-12%
UB	4.11	2.00	+2.1	+106%
HC	2.91	3.70	-0.8	-21%
ZC	4.03	5.10	-1.1	-21%
VC	2.47	3.20	-0.7	-23%
XC	3.36	5.10	-1.7	-34%
UC	2.26	2.10	+0.2	+8%
RC	3.55	3.70	-0.2	-4%
TD	1.24	1.10	+0.1	+13%
VD	1.11	0.70	+0.4	+59%
SD	0.65	0.90	-0.3	-28%

APPENDIX D. DETAILED AMPT EXPERIMENTAL RESULTS

Plant F (RS9.5B) AMPT Summary

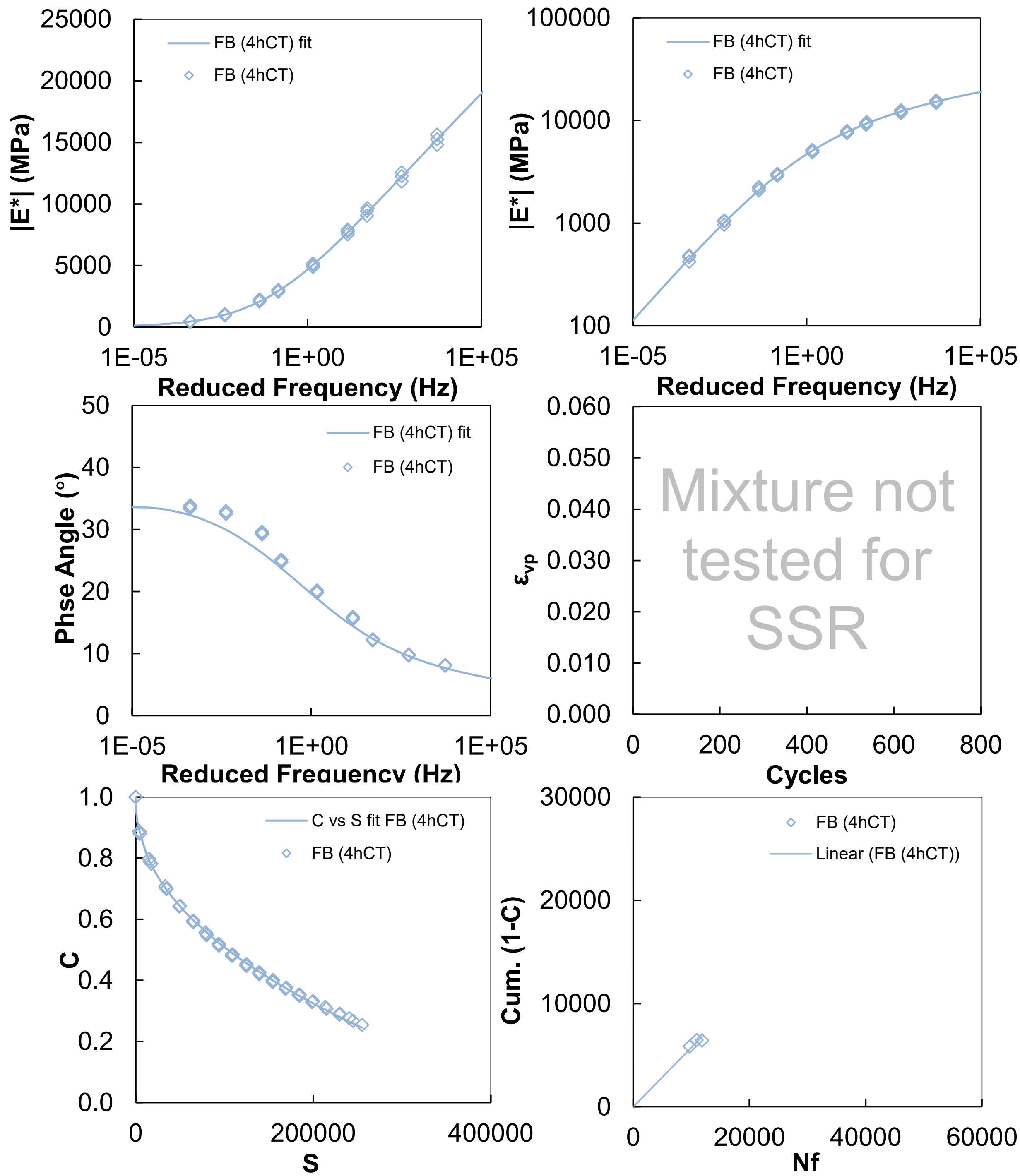


Figure D.1. Plant F (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant Y (RS9.5B) AMPT Summary

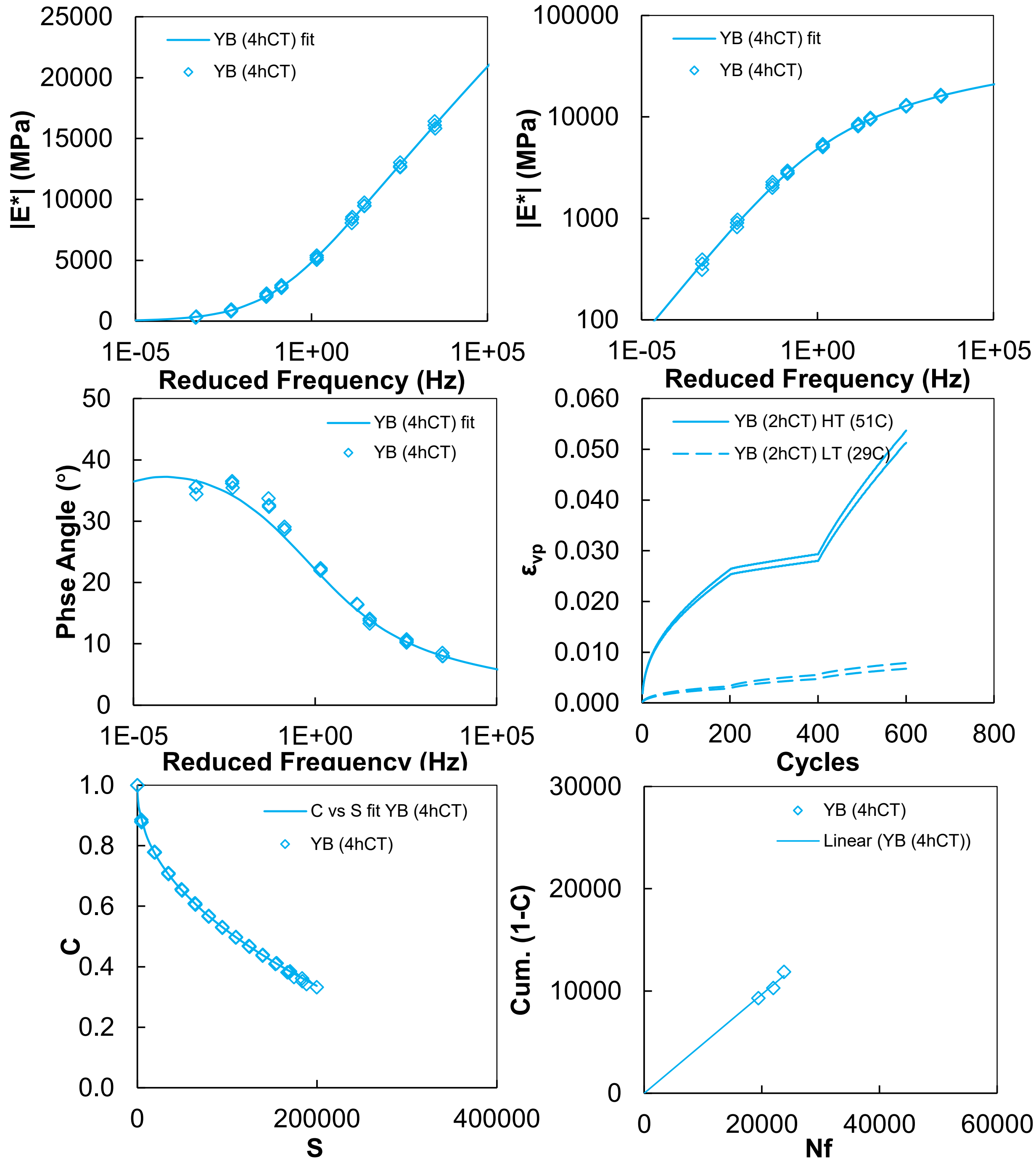


Figure D.2. Plant F (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant W (RS9.5B) AMPT Summary

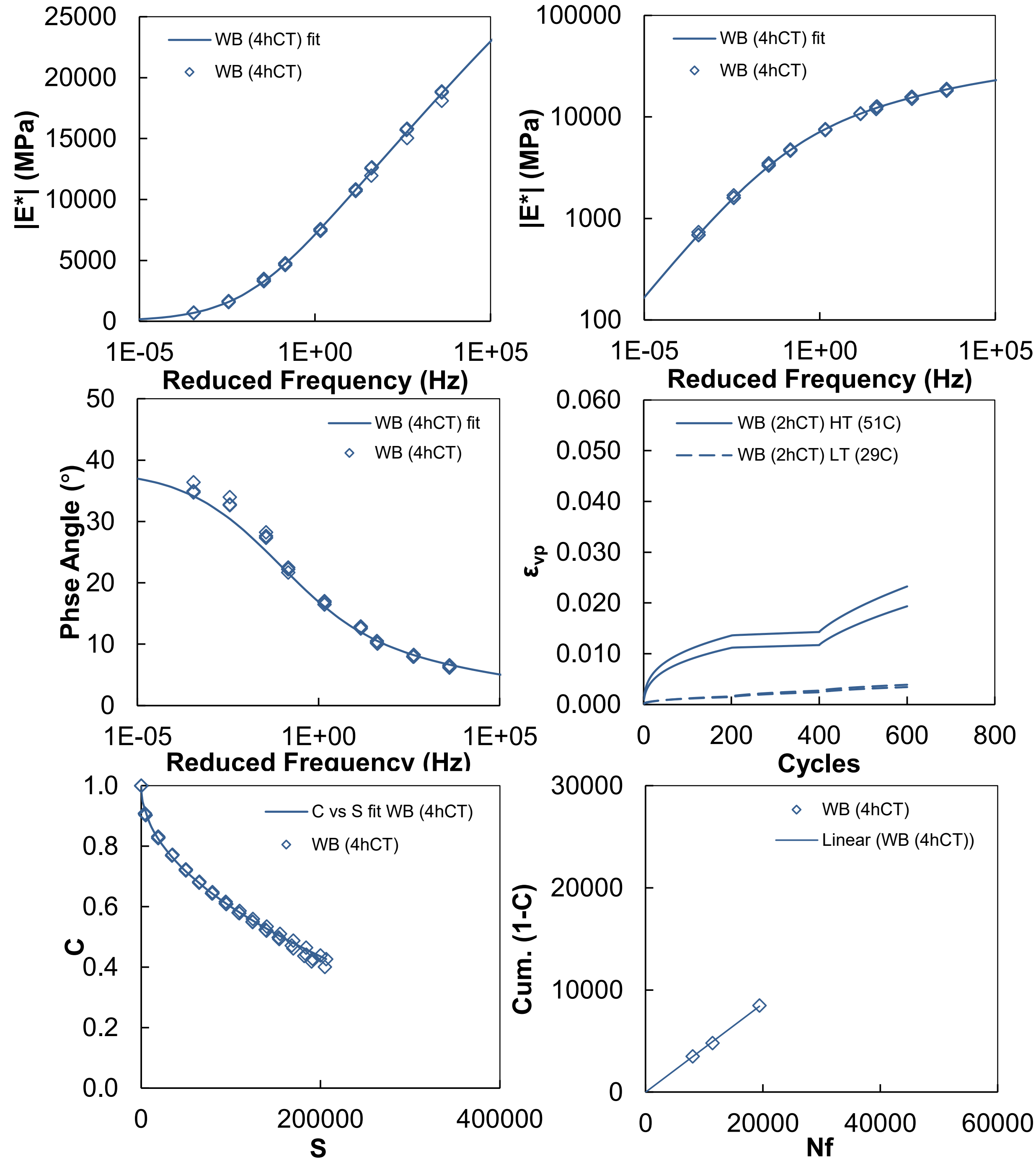


Figure D.3. Plant W (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant S (RS9.5B) AMPT Summary

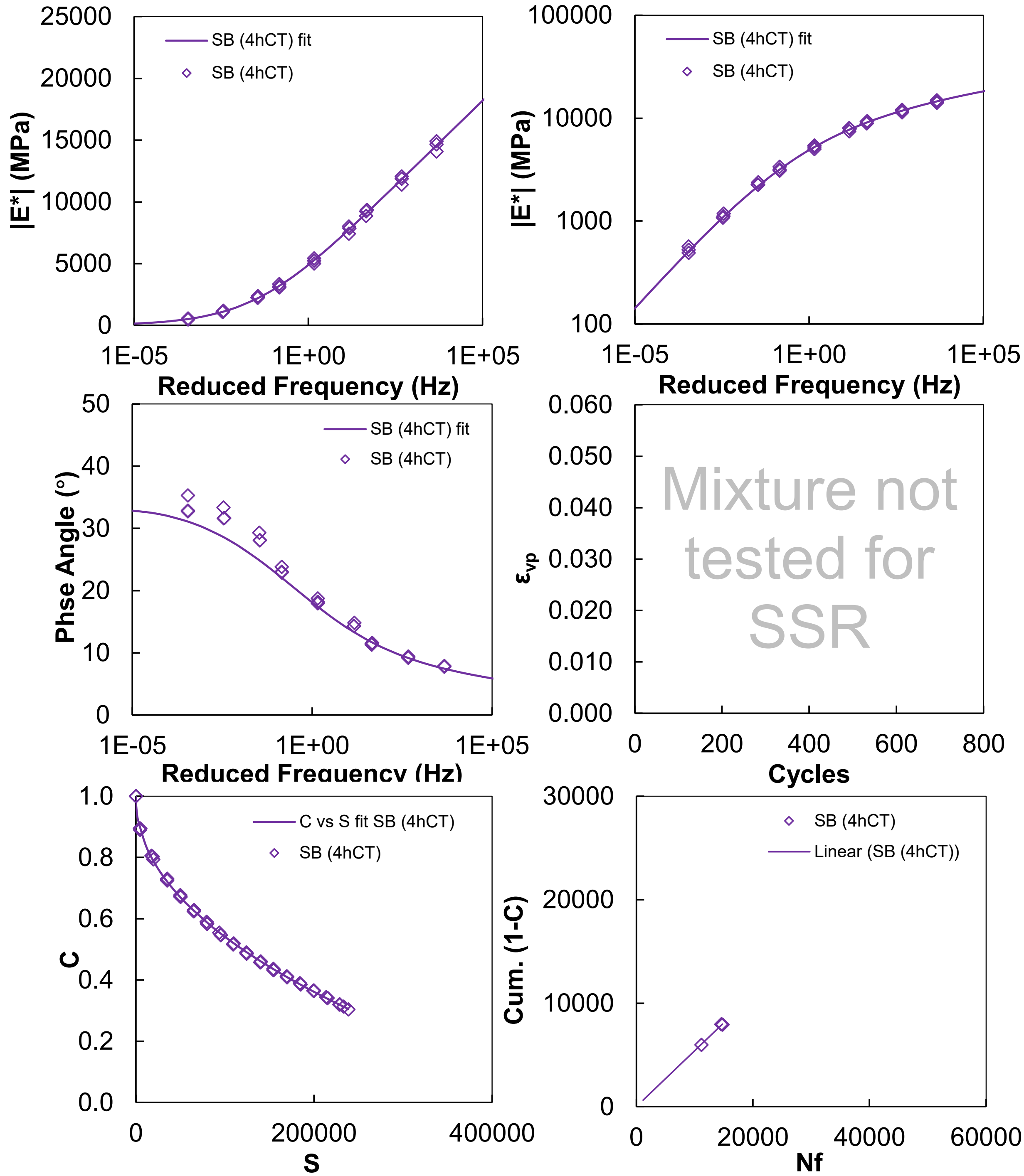


Figure D.4. Plant S (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant U (RS9.5B) AMPT Summary

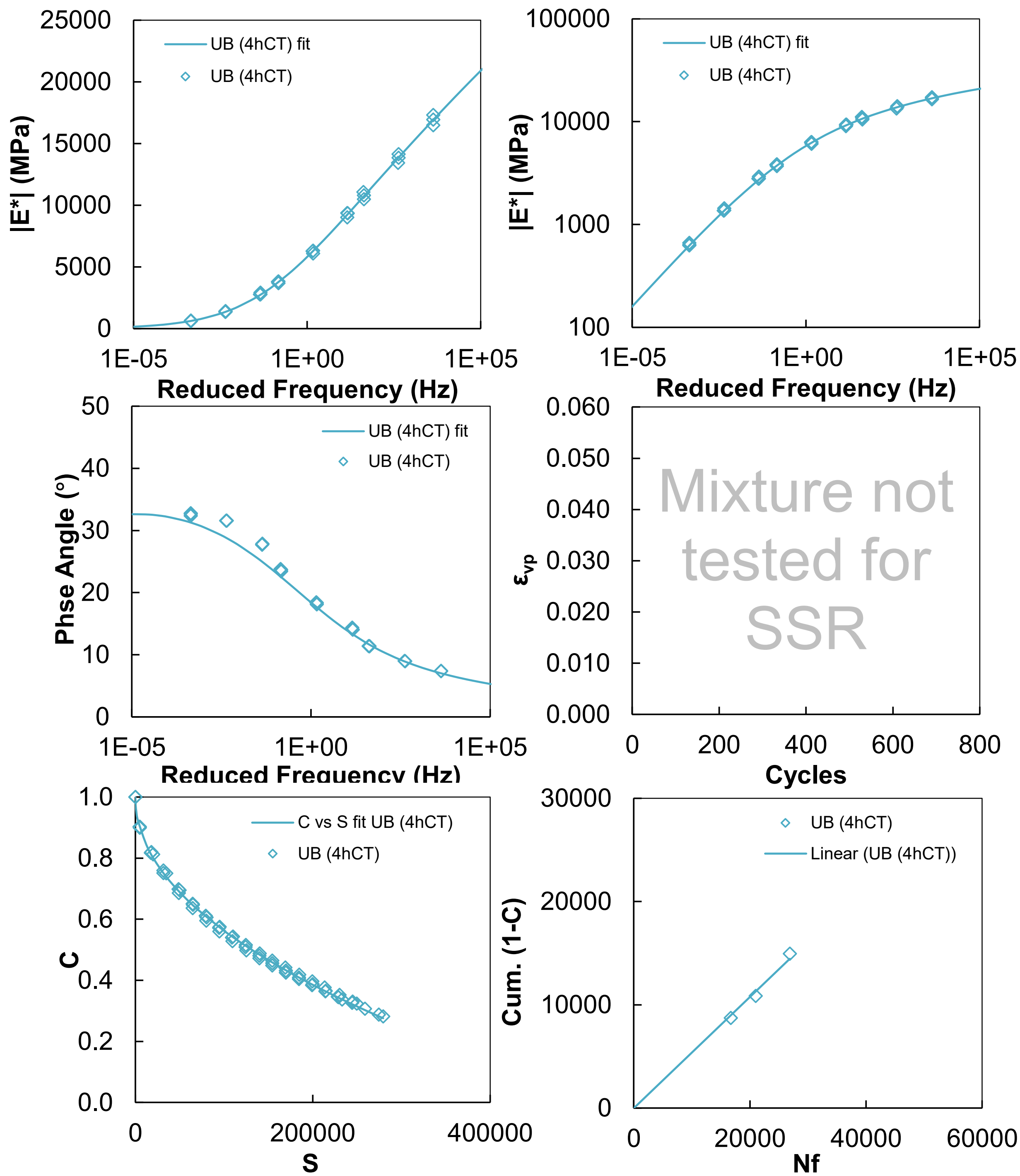


Figure D.5. Plant U (RS9.5B) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant H (RS9.5C) AMPT Summary

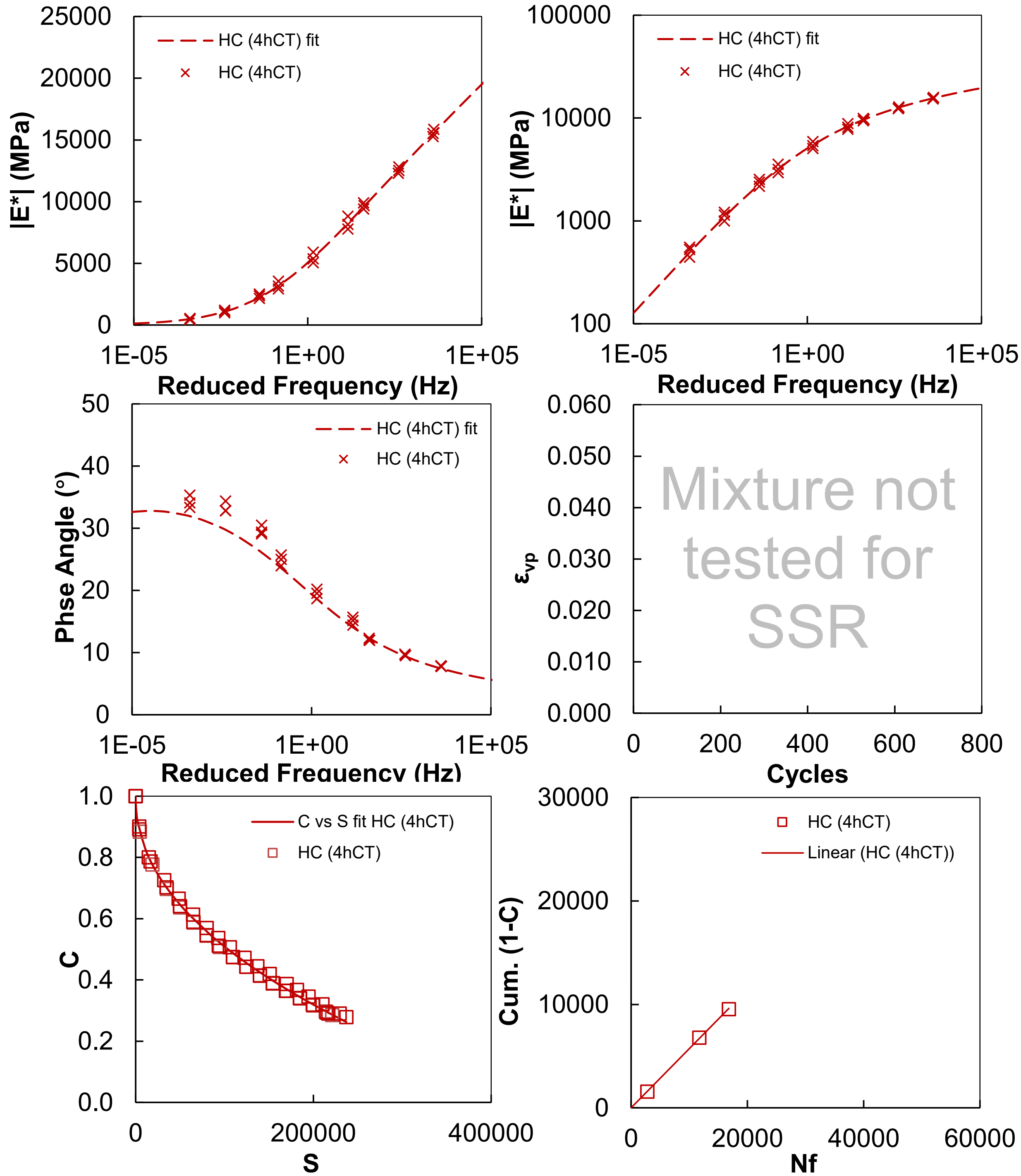


Figure D.6. Plant H (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant Z (RS9.5C) AMPT Summary

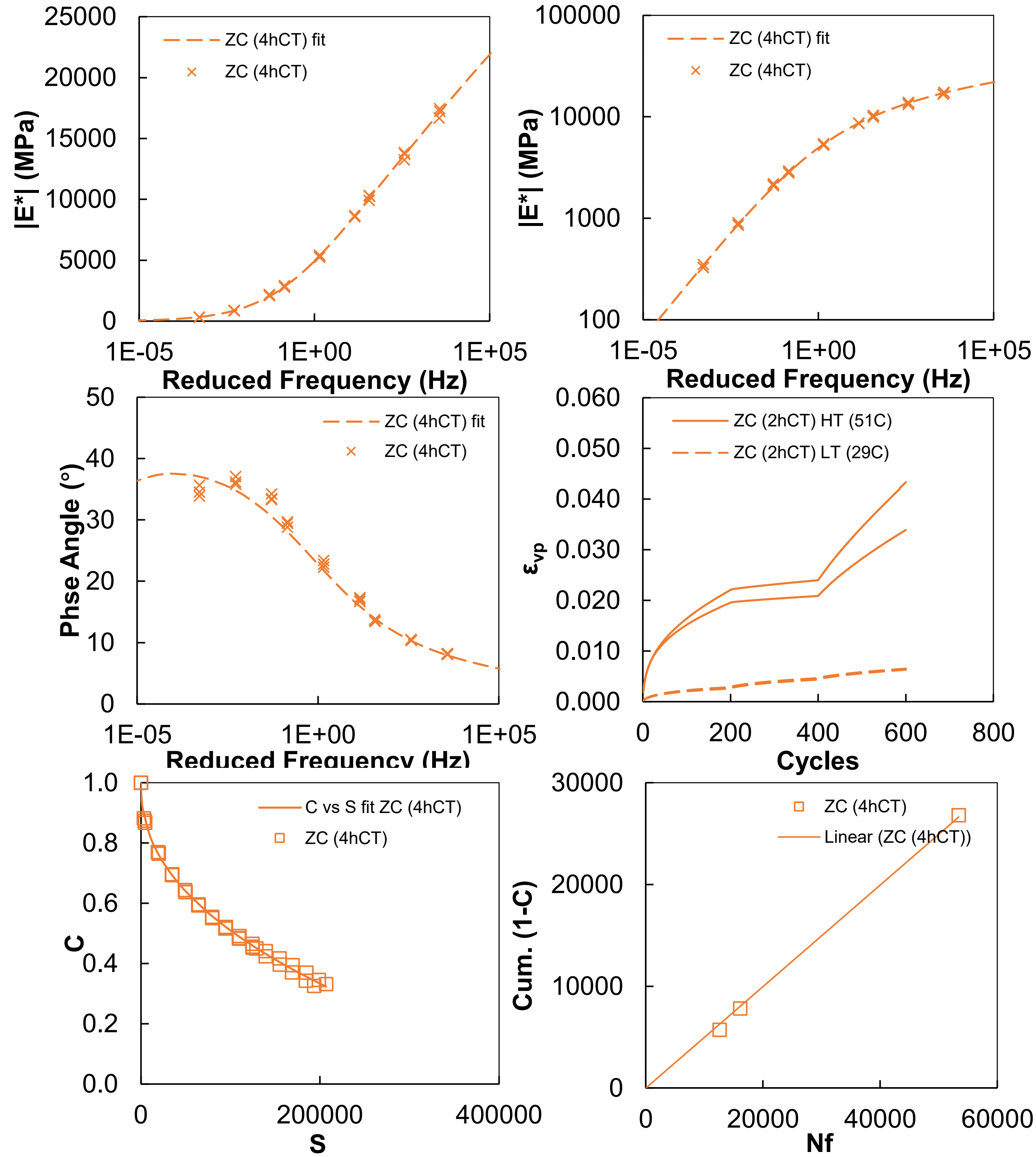


Figure D.7. Plant Z (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant V (RS9.5C) AMPT Summary

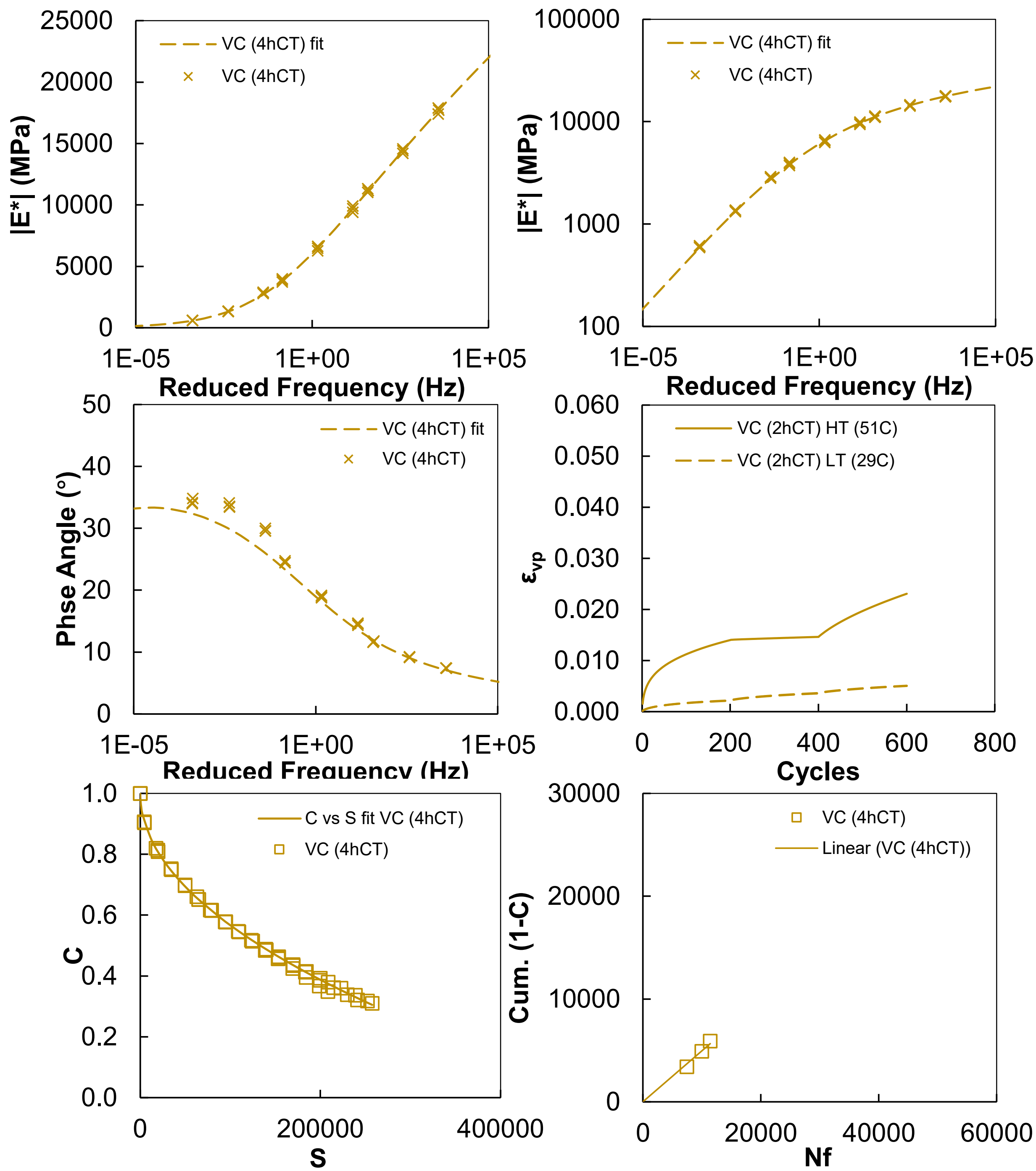


Figure D.8. Plant V (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant X (RS9.5C) AMPT Summary

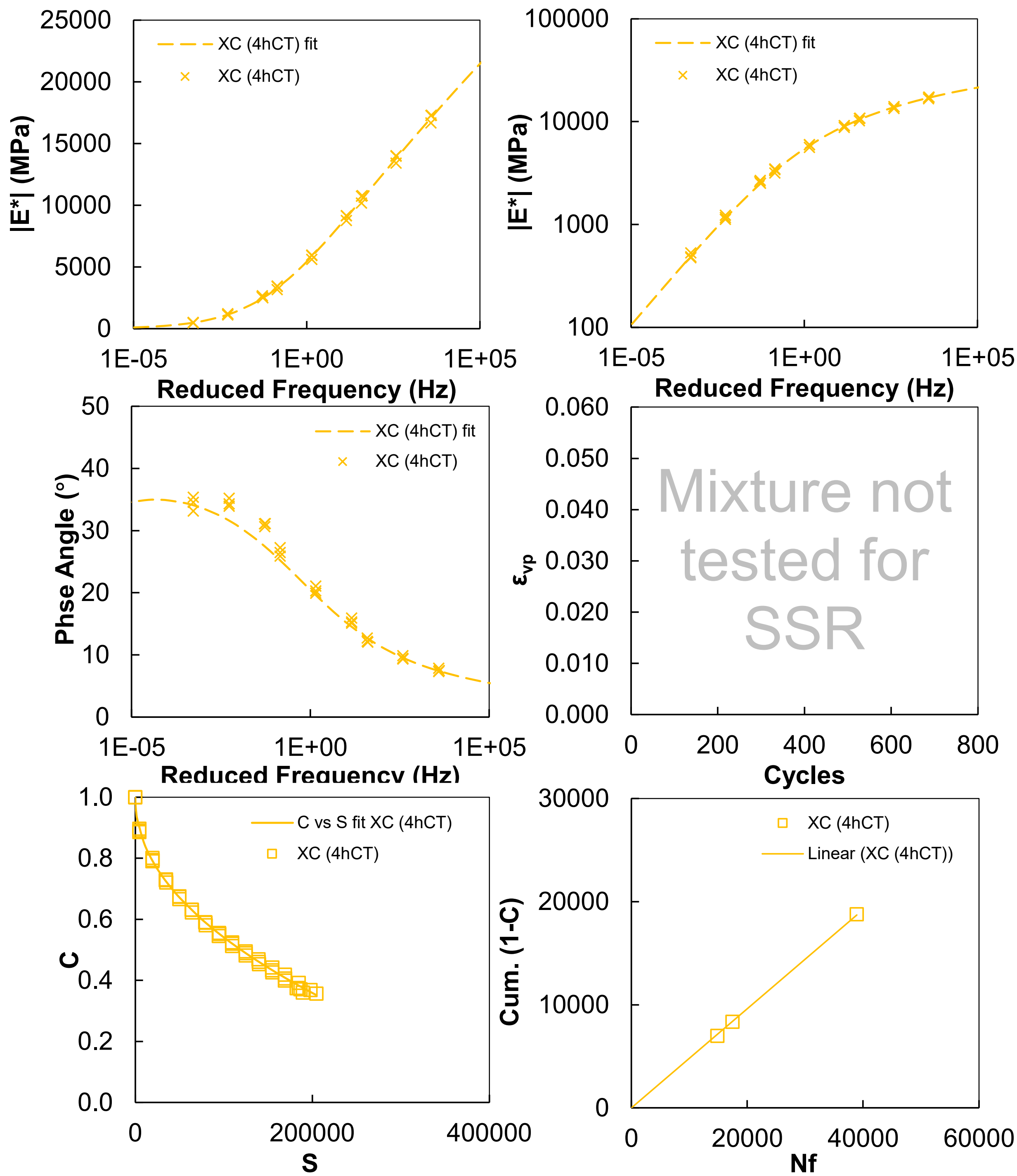


Figure D.9. Plant X (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant U (RS9.5C) AMPT Summary

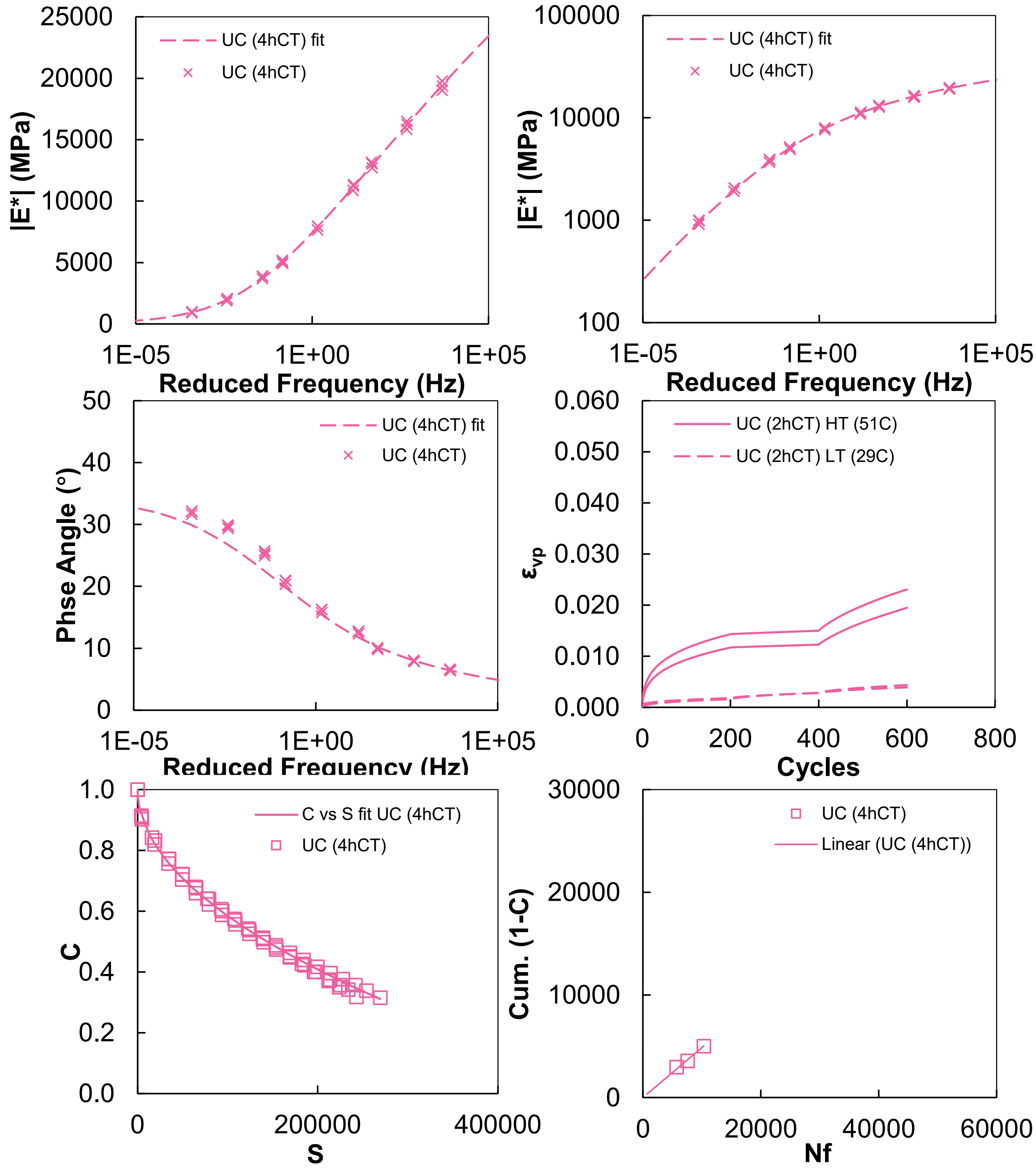


Figure D.10. Plant U (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant R (RS9.5C) AMPT Summary

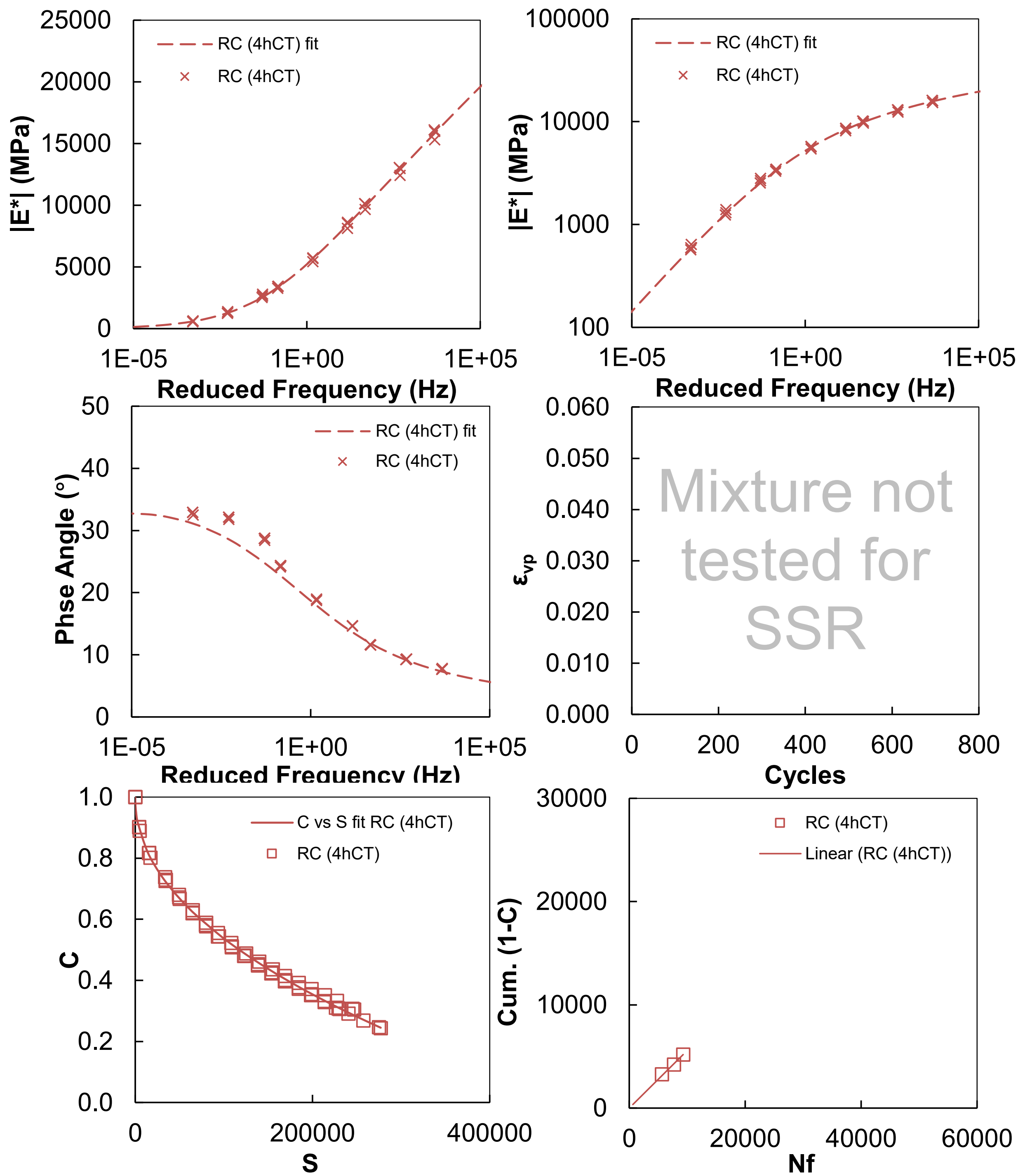


Figure D.11. Plant R (RS9.5C) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant V (RS9.5D) AMPT Summary

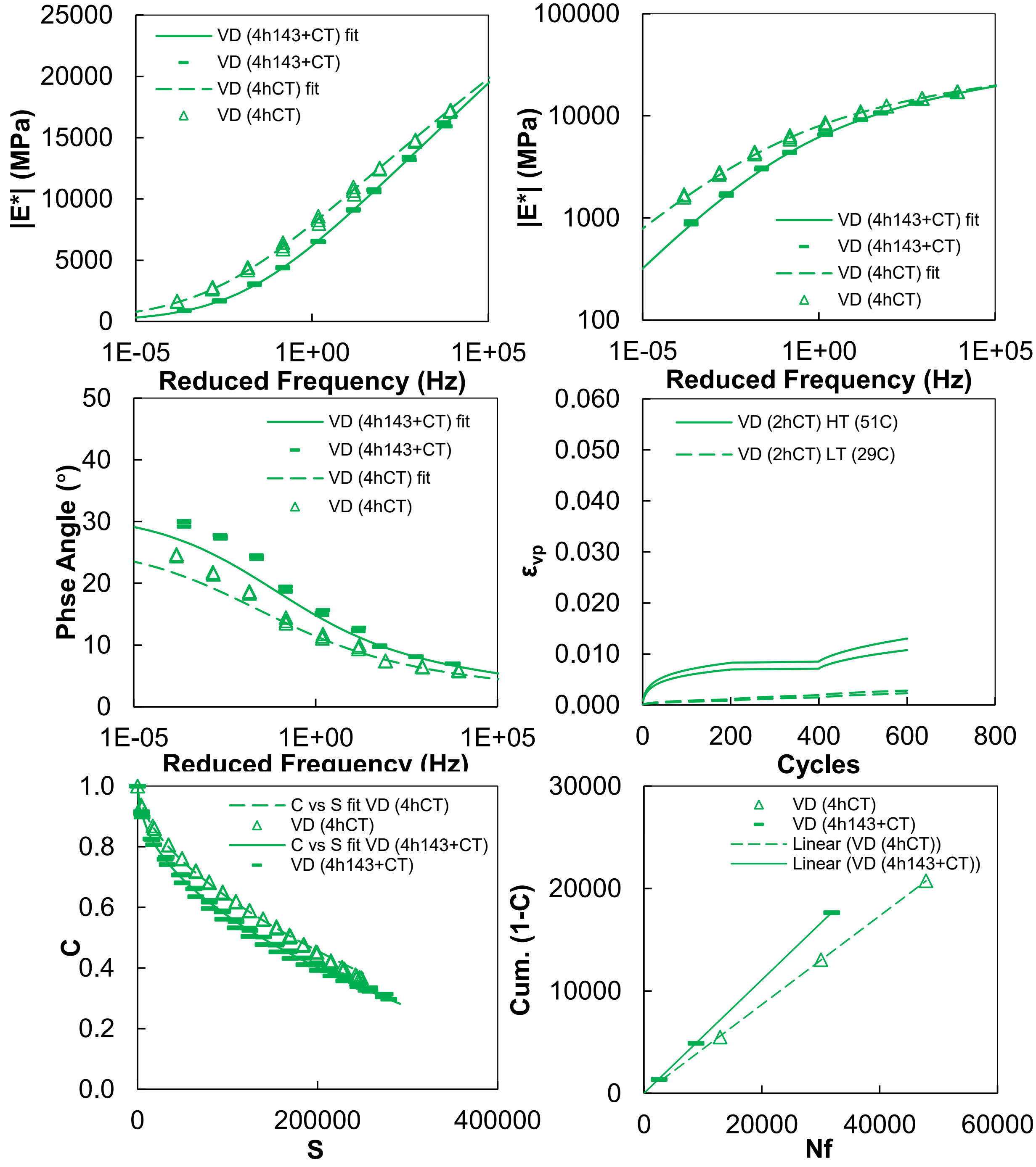


Figure D.12. Plant V (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (middle-right) Accumulated permanent strain curve, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant T (RS9.5D) AMPT Summary

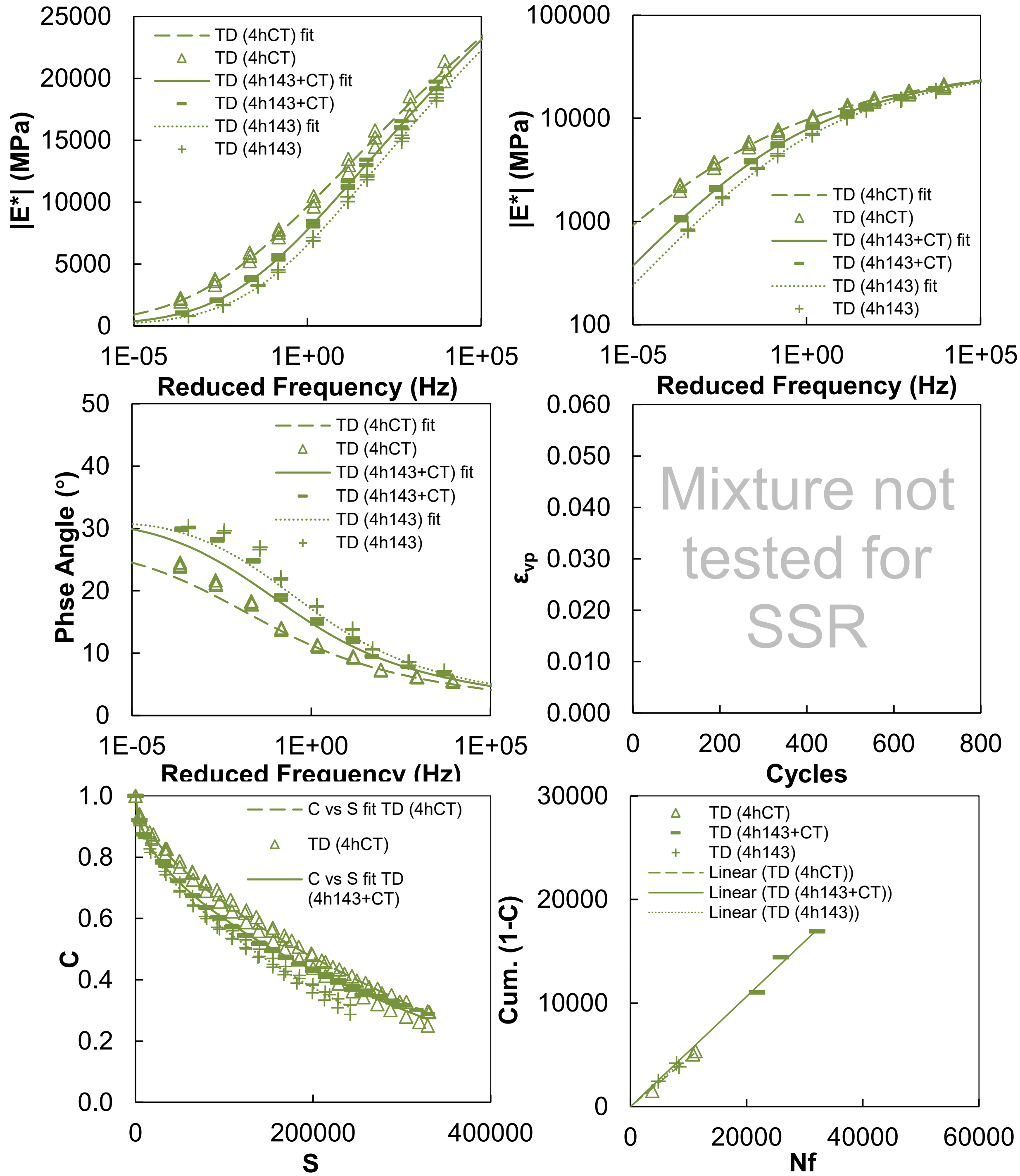


Figure D.13. Plant T (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

Plant S (RS9.5D) AMPT Summary

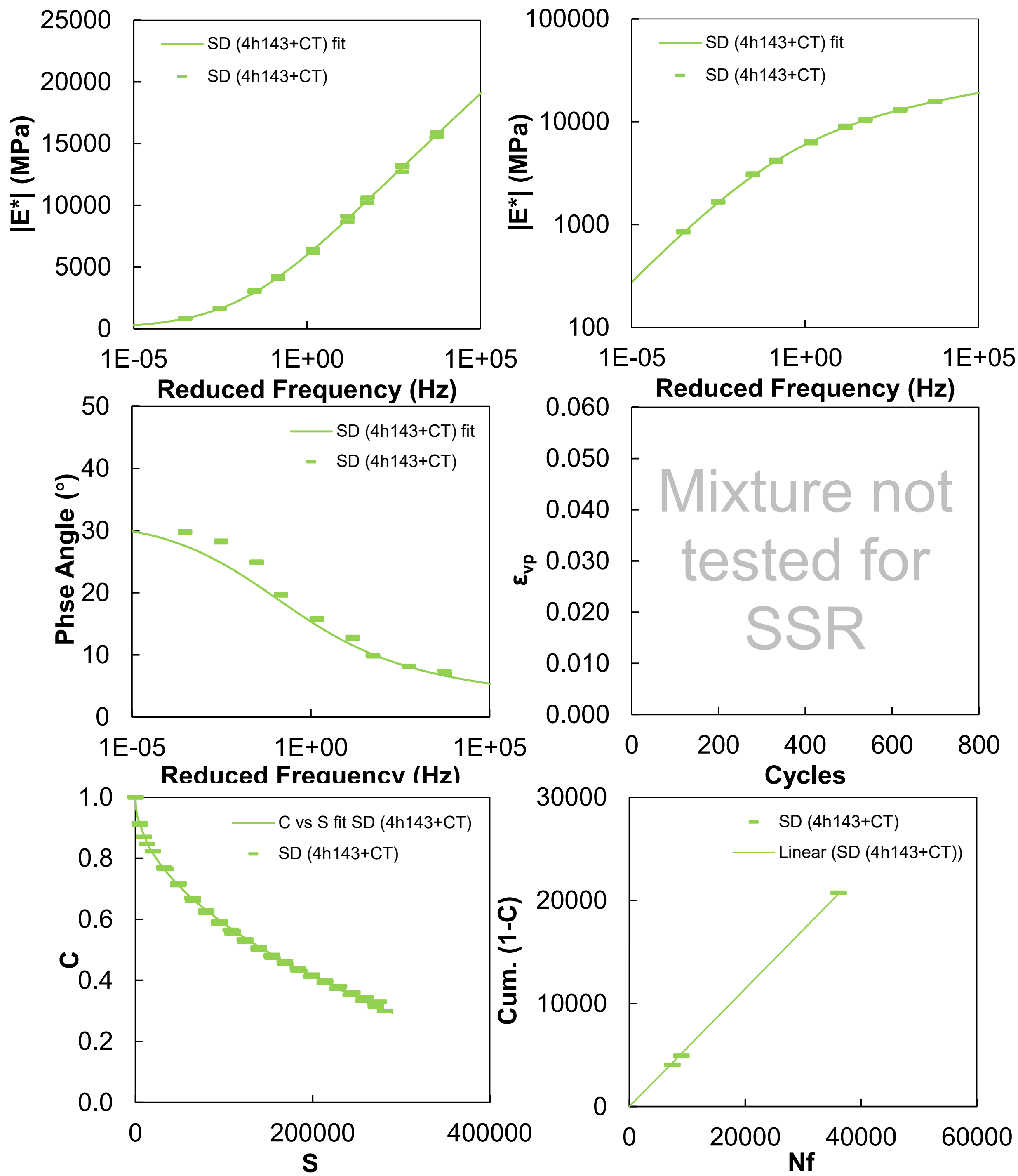


Figure D.14. Plant S (RS9.5D) AMPT summary: (top-left) Dynamic Modulus in semi-log space, (top-right) Dynamic modulus in log-log space, (middle-left) Phase angle, (bottom-left) Integrity (C) vs. accumulated damage (S), (bottom-right) D^R curve.

DM Statistical Results

Table D.1. Analysis of Variance for DM (4°C, 1Hz) on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F						
Individual Mixtures	0.8160, 0.9960	**<0.0001**						
Tukey HSD Grouping	**Mean**	**Group Letters**						
TD	16206.0	A						
UC	16194.0	A						
WB	15533.0	A						
VC	14404.0		B					
UB	13823.0		B	C				
XC	13816.0		B	C				
ZC	13605.0		B	C	D			
VD	13301.0			C	D	E		
SD	13029.0			C	D	E	F	
RC	12835.0				D	E	F	
YB	12807.0				D	E	F	
HC	12555.0					E	F	G
FB	12225.0						F	G
SB	11780.0							G

Table D.2. Analysis of Variance for DM (20°C, 1Hz) on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F					
Individual Mixtures	0.1233, 0.5339	**<0.0001**					
Tukey HSD Grouping	**Mean**	**Group Letters**					
TD	8234.3	A					
UC	7859.0	A	B				
WB	7519.3		B				
VD	6541.0			C			
VC	6469.0			C			
SD	6350.0			C	D		
UB	6244.3			C	D		
XC	5850.0				D	E	
RC	5621.7					E	F
HC	5451.3					E	F
ZC	5348.0					E	F
SB	5229.7						F
YB	5223.0						F
FB	5042.7						F

Table D.3. Analysis of Variance for DM (40°C, 1Hz) on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F					
Individual Mixtures	0.3279, 0.8299	**<0.0001**					
Tukey HSD Grouping	**Mean**	**Group Letters**					
TD	2097.3	A					
UC	2006.0	A					
VD	1690.3		B				
SD	1662.0		B				
WB	1624.0		B				
UB	1390.3			C			
VC	1348.7			C	D		
RC	1312.3			C	D		
XC	1174.3				D	E	
SB	1128.0					E	
HC	1115.6					E	
FB	1022.6					E	F
YB	897.3						F
ZC	870.8						F

Table D.4. Analysis of Variance for DM (4°C, 1Hz) on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Mix Type	0.4065, 0.7791	0.2158

Table D.5. Analysis of Variance for DM (20°C, 1Hz) on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F	
Mix Type	0.7846, 0.9606	**0.0139**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
RS9.5D	7041.8	A	
RS9.5C	6099.8		B
RS9.5B	5851.8		B

Table D.6. Analysis of Variance for DM (40°C, 1Hz) on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F	
Mix Type	0.6472, 0.2306	**<0.0001**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
RS9.5D	1816.6	A	
RS9.5C	1304.6		B
RS9.5B	1212.4		B

Table D.7. Mix Type and Source Region Mixed Model for DM (4°C, 1Hz).

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F	
Mix Type	0.4065, 0.7791	0.2164	
Source Region	0.1590, 0.2471	**0.0185**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
Piedmont	14185	A	
Coastal Plains	13930	A	B
Mountains	12306		B
Not performed for Source Region			

Table D.8. Mix Type and Source Region Mixed Model for DM (20°C, 1Hz).

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F	
Mix Type	0.7846, 0.9606	**0.0188**	
Source Region	0.6074, 0.5054	**0.0324**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
Piedmont	6710.1	A	
Coastal Plains	6137.7	A	B
Mountains	5580.1		B

Table D.9. Mix Type and Source Region Mixed Model for DM (40°C, 1Hz).

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F	
Mix Type	0.6472, 0.2306	**0.0008**	
Source Region	0.2507, 0.6565	**0.0072**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
Piedmont	1600.2	A	
Mountains	1302.2	A	B
Coastal Plains	1270.1		B

CF Statistical Results

Table D.10. Analysis of Variance for S_{app} on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F				
Individual Mixtures	0.1562, 0.4186	<0.0001				
Tukey HSD Grouping	**Mean**	**Group Letters**				
FB	25.1	A				
SD	24.3	A	B			
RC	23.4	A	B	C		
SB	23.2	A	B	C	D	
HC	21.9	A	B	C	D	E
VD	21.8	A	B	C	D	E
UB	21.3	A	B	C	D	E
YB	19.2		B	C	D	E
TD	18.7			C	D	E
WB	18.4			C	D	E
VC	18.1				D	E
ZC	17.9					E
UC	17.8					E
XC	17.5					E

Table D.11. Analysis of Variance for D^R on Individual Mixtures.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F				
Individual Mixtures	0.1127, 0.5488	<0.0001				
Tukey HSD Grouping	**Mean**	**Group Letters**				
FB	0.578	A				
HC	0.571	A				
RC	0.560	A				
SD	0.555	A				
VD	0.546	A	B			
SB	0.539	A	B	C		
UB	0.532	A	B	C	D	
TD	0.530	A	B	C	D	
UC	0.492		B	C	D	
VC	0.489		B	C	D	E
YB	0.483			C	D	E
ZC	0.481			C	D	E
XC	0.477				D	E
WB	0.431					E

Table D.12. Analysis of Variance for S_{app} on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F
Mix Type	0.8336, 0.9722	0.0919

Table D.13. Analysis of Variance for D^R on Mix Type.

Attribute	Equal Variance (Levene and Barlett) Prob>F	Difference in Means Welch's ANOVA (α=0.05) Prob>F
Mix Type	**0.0114**, **0.0203**	**0.0237**
Tukey HSD Grouping	**Mean**	**Group Letters**
RS9.5D	0.544	A
RS9.5B	0.513	A
RS9.5C	0.512	A

Table D.14. Mix Type and Source region Mixed Model for S_{app}.

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F	
Mix Type	0.8336, 0.9722	0.2186	
Source Region	0.1322, **0.0408**	**0.0483**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
Mountains	23.3	A	
Piedmont	20.9	A	B
Coastal Plains	19.6		B

Table D.15. Mix Type and Source Region Mixed Model for D^R.

Attribute Model	Equal Variance (Levene and Barlett) Prob>F	Difference in Means ANOVA (α=0.05) Prob>F	
Mix Type	**0.0114**, **0.0203**	0.6605	
Source Region	0.0680, **0.0128**	**0.0182**	
Tukey HSD Grouping	**Mean**	**Group Letters**	
Mountains	0.543	A	B
Piedmont	0.534	A	
Coastal Plains	0.492		B

APPENDIX E. DETAILED PERFORMANCE EVALUATION SUMMARY

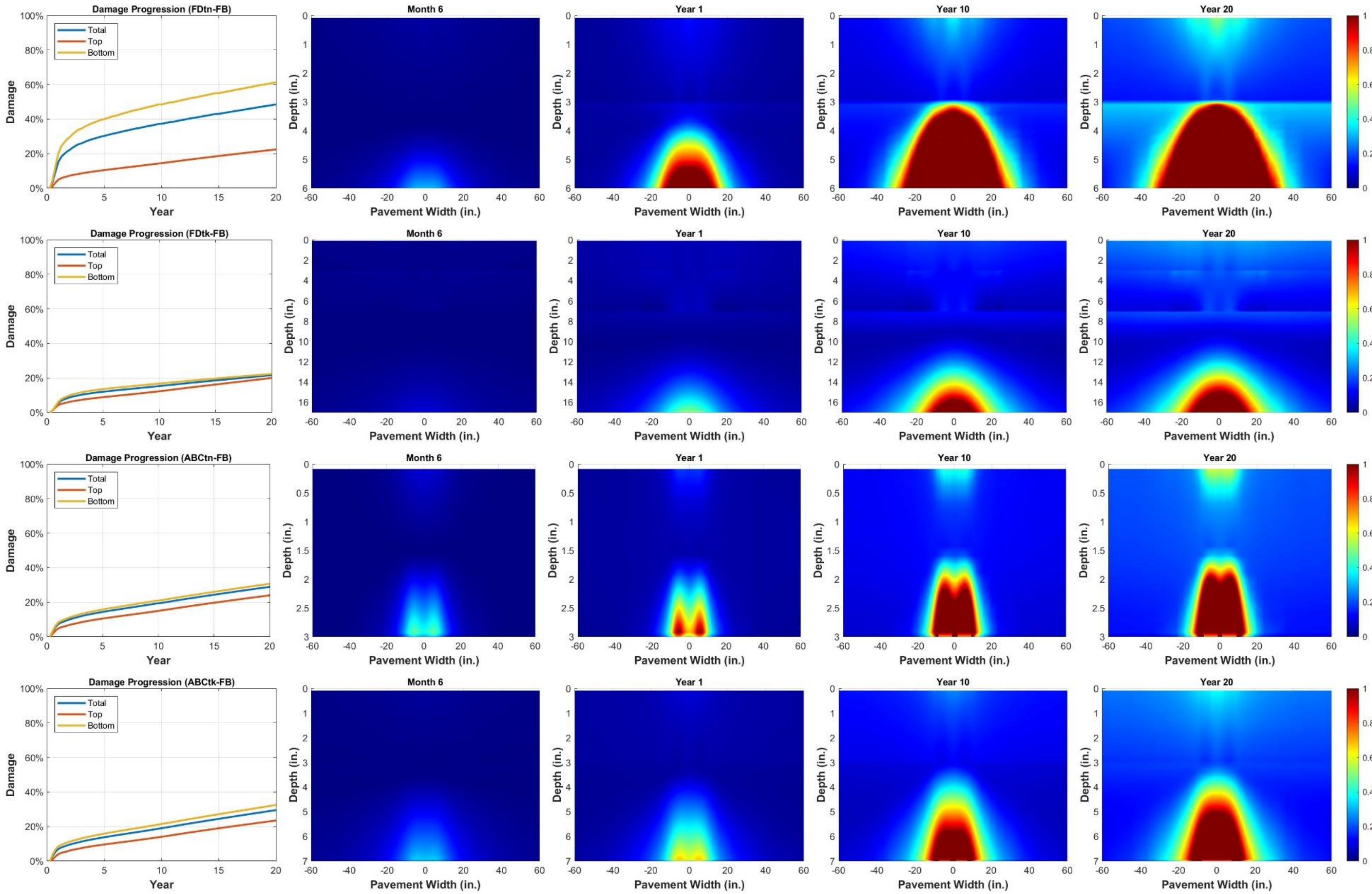


Figure E.1. Damage contours FB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

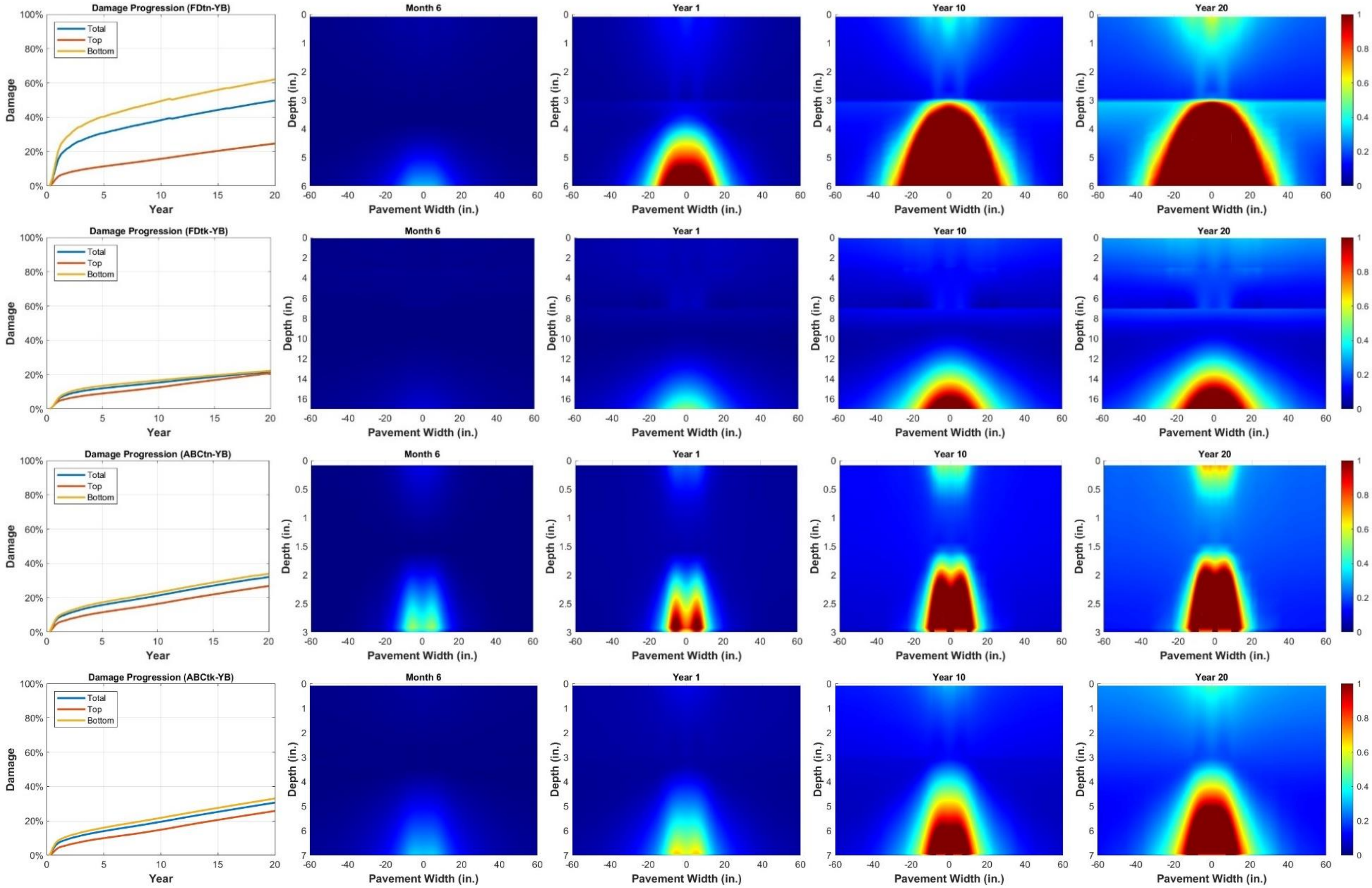


Figure E.2. Damage contours YB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

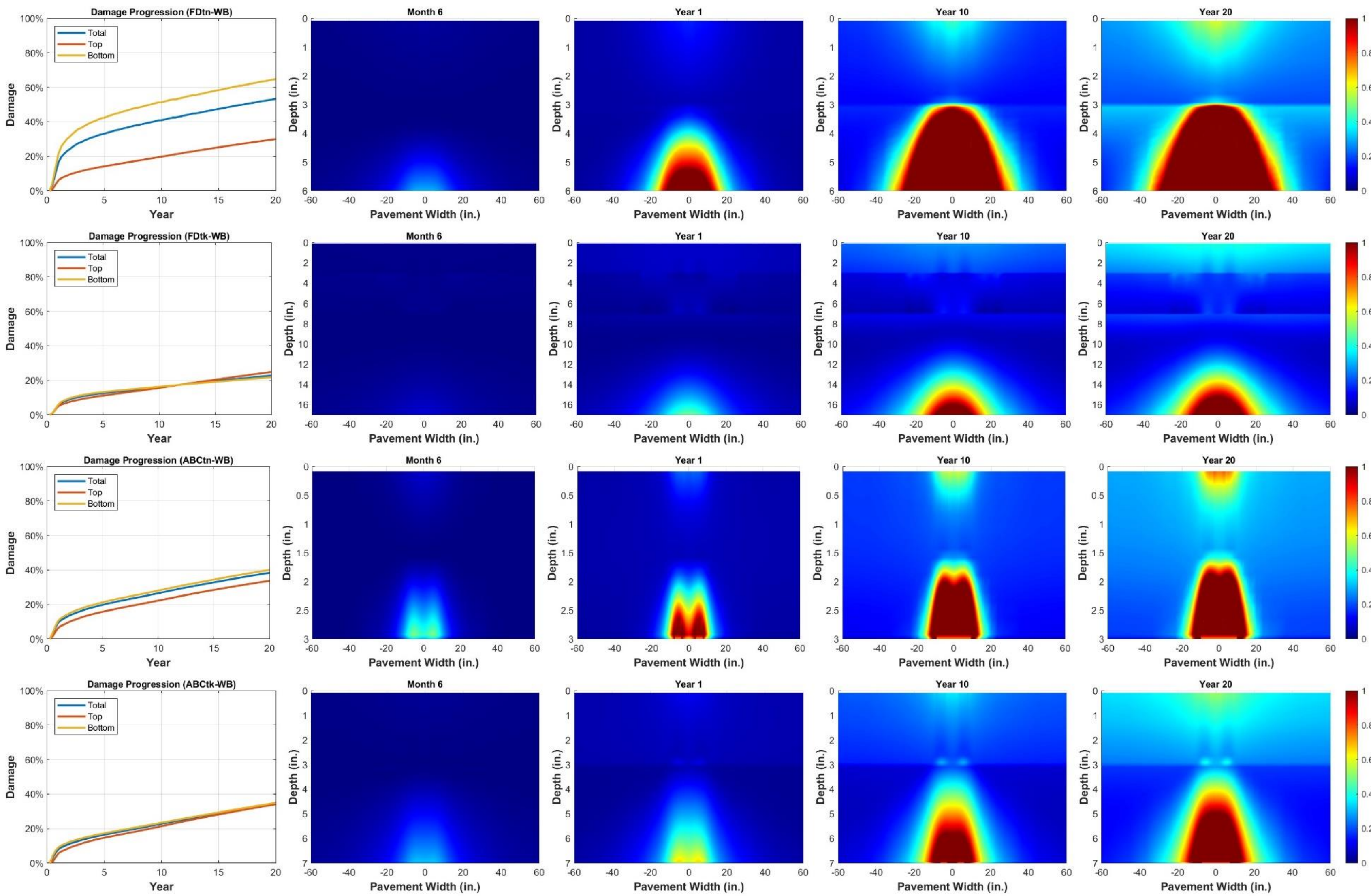


Figure E.3. Damage contours WB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

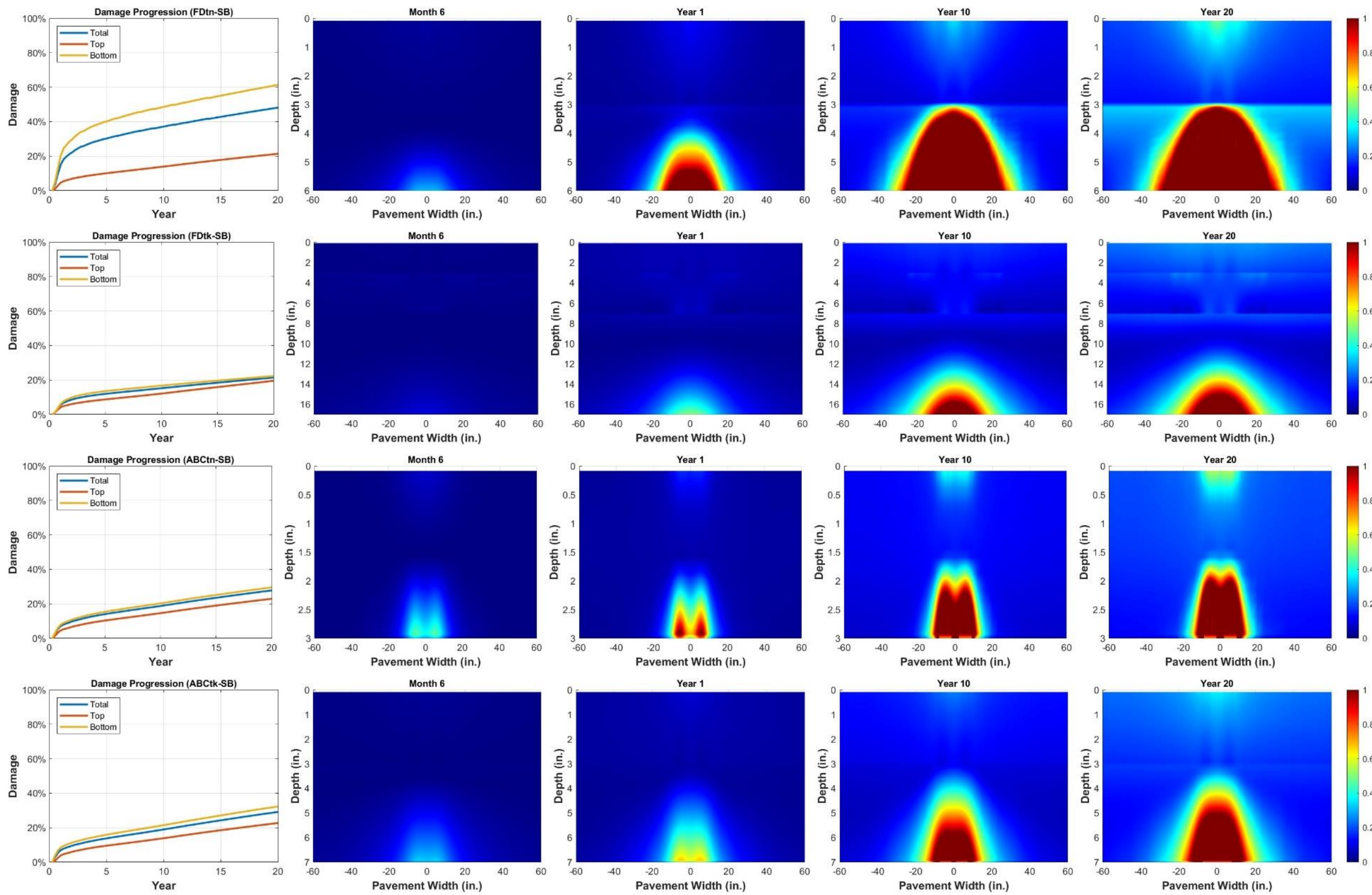


Figure E.4. Damage contours SB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

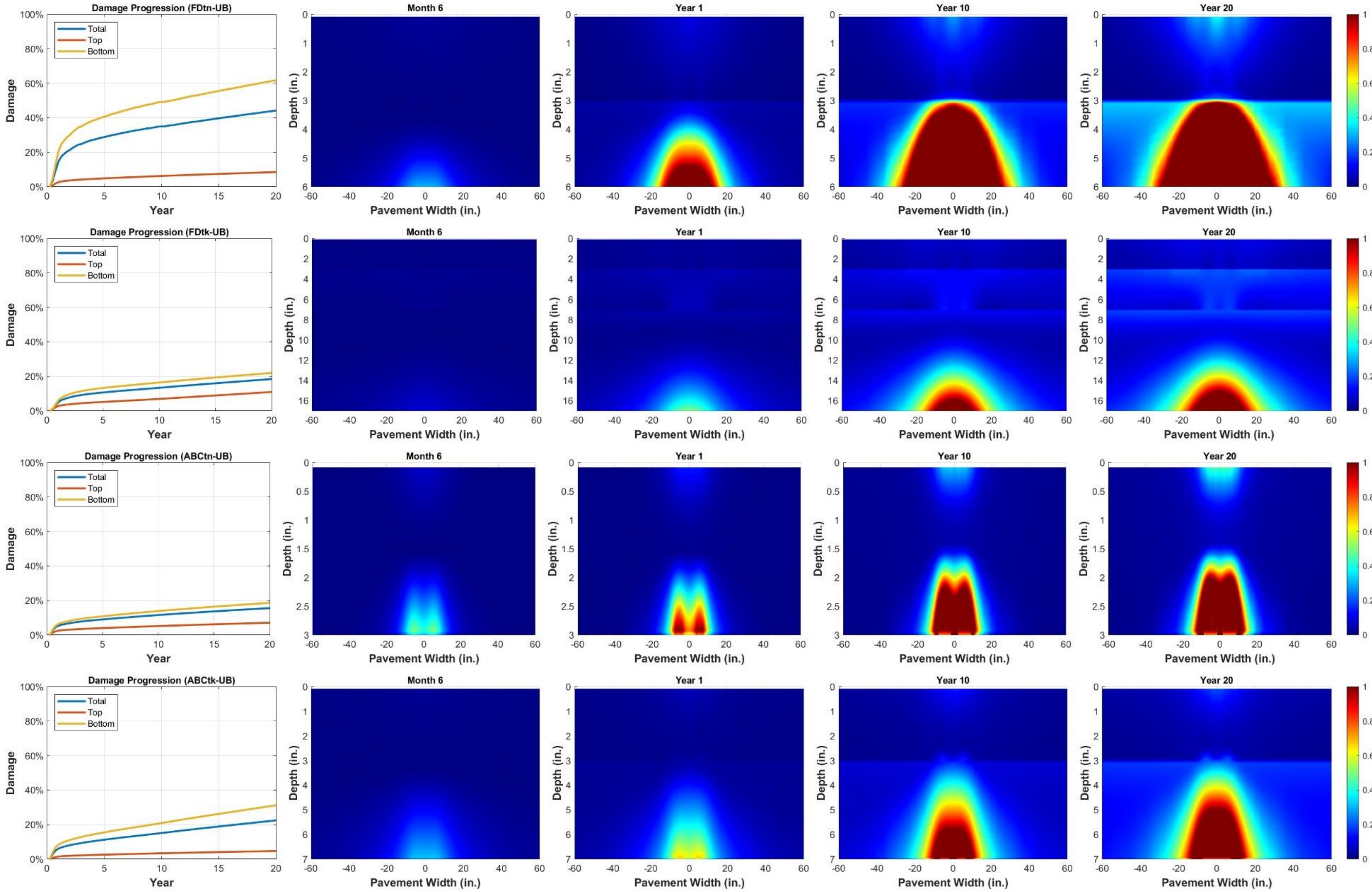


Figure E.5. Damage contours UB mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

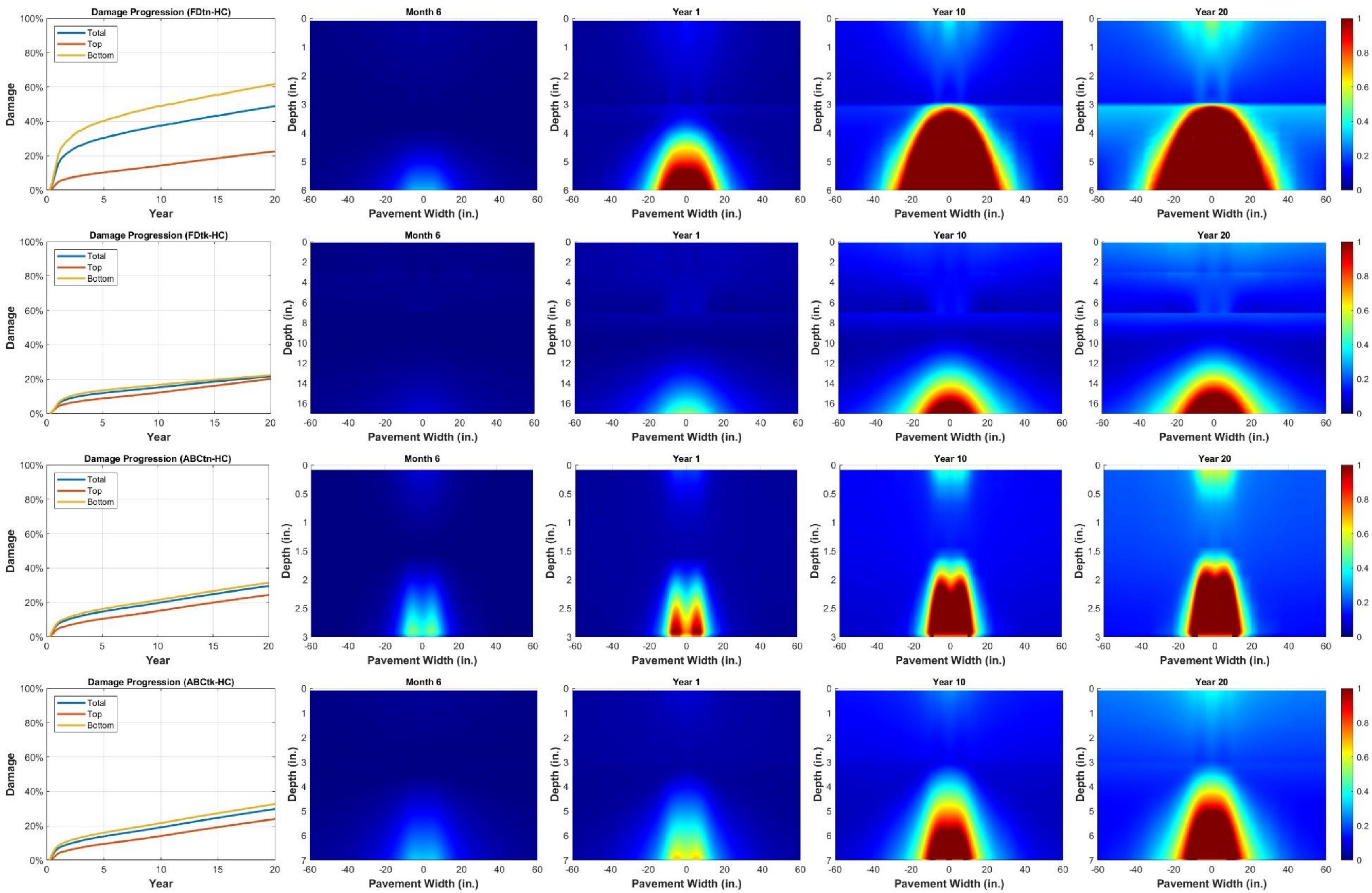


Figure E.6. Damage contours HC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

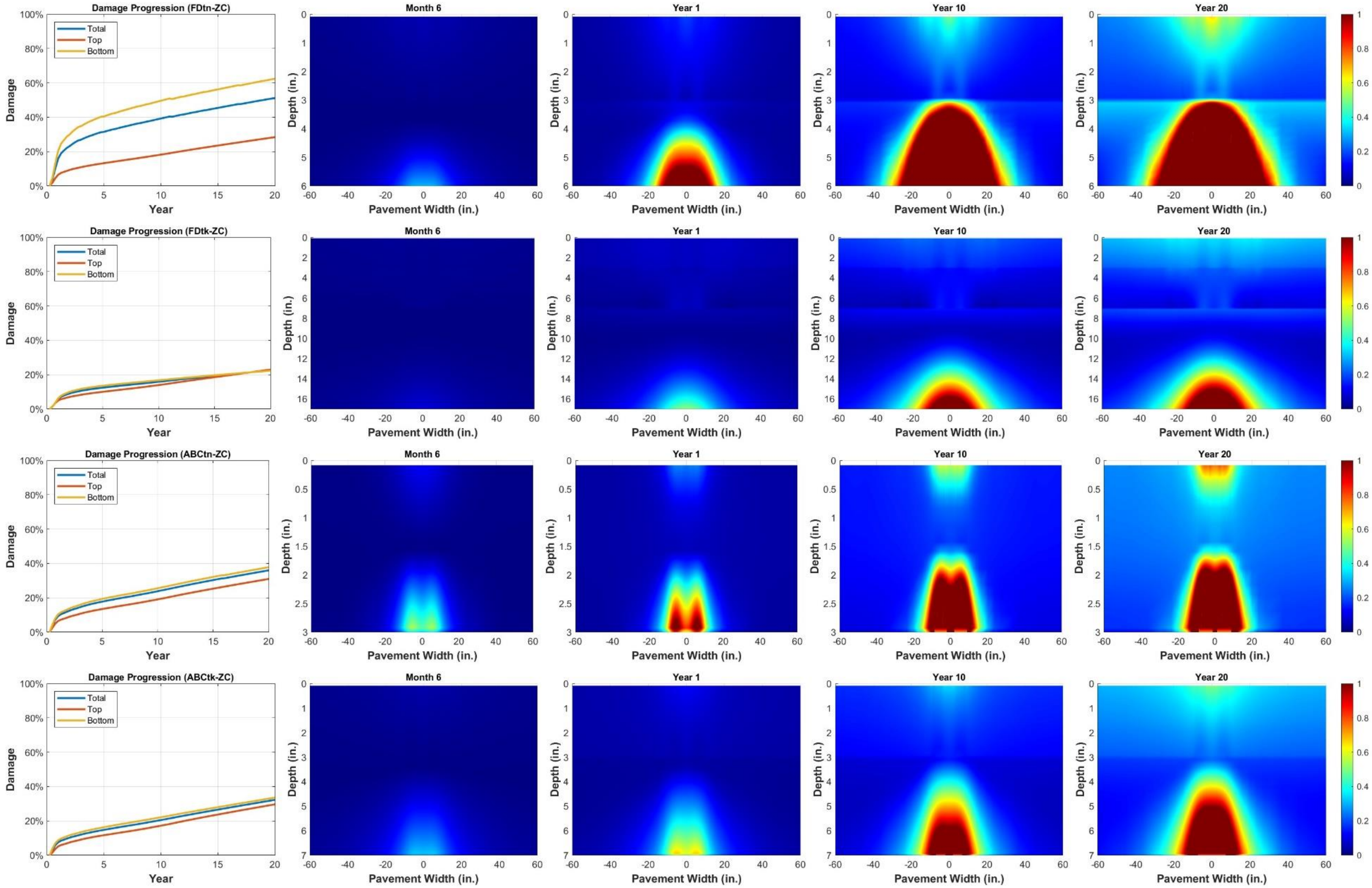


Figure E.7. Damage contours ZC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

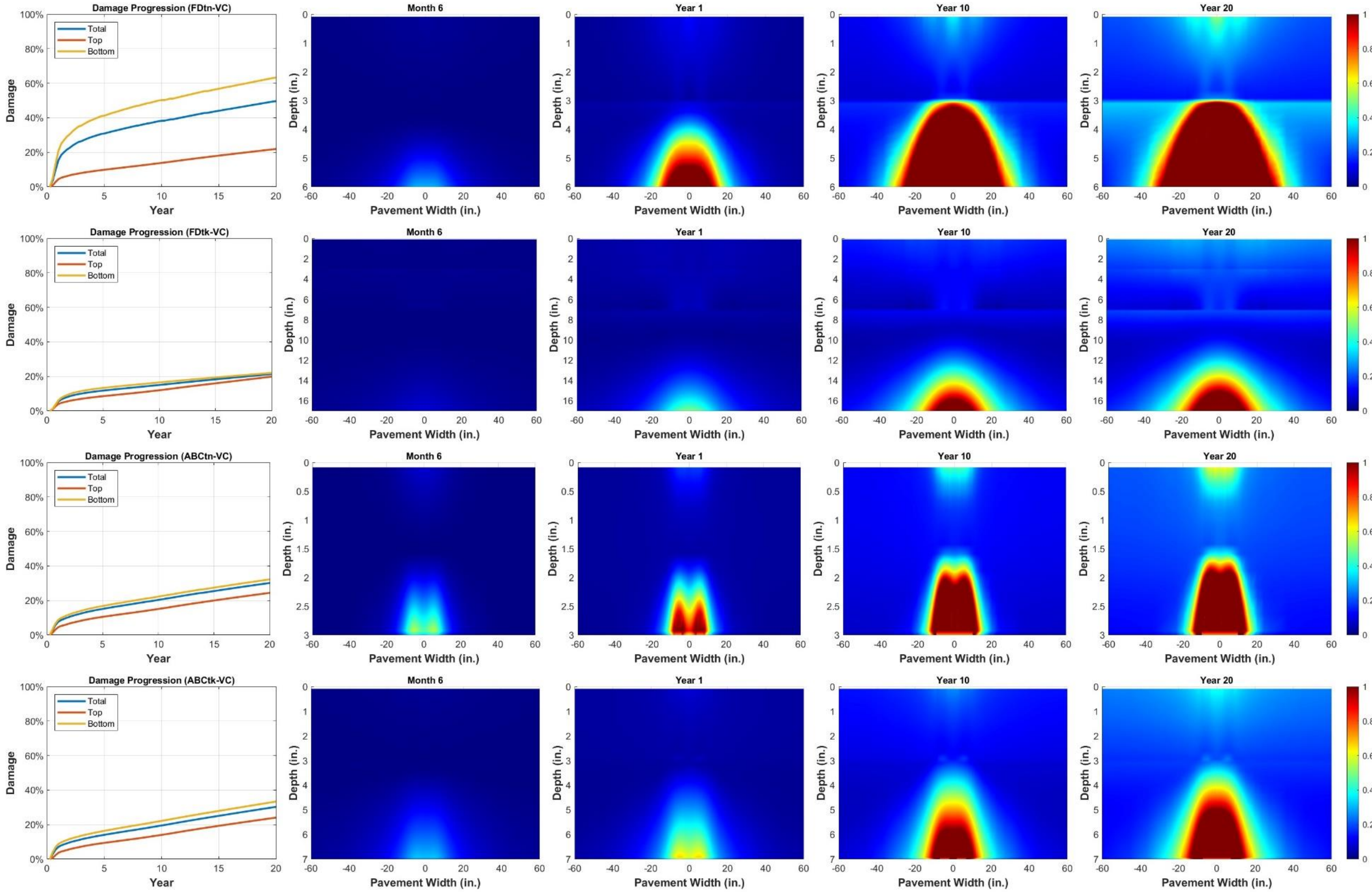


Figure E.8. Damage contours VC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

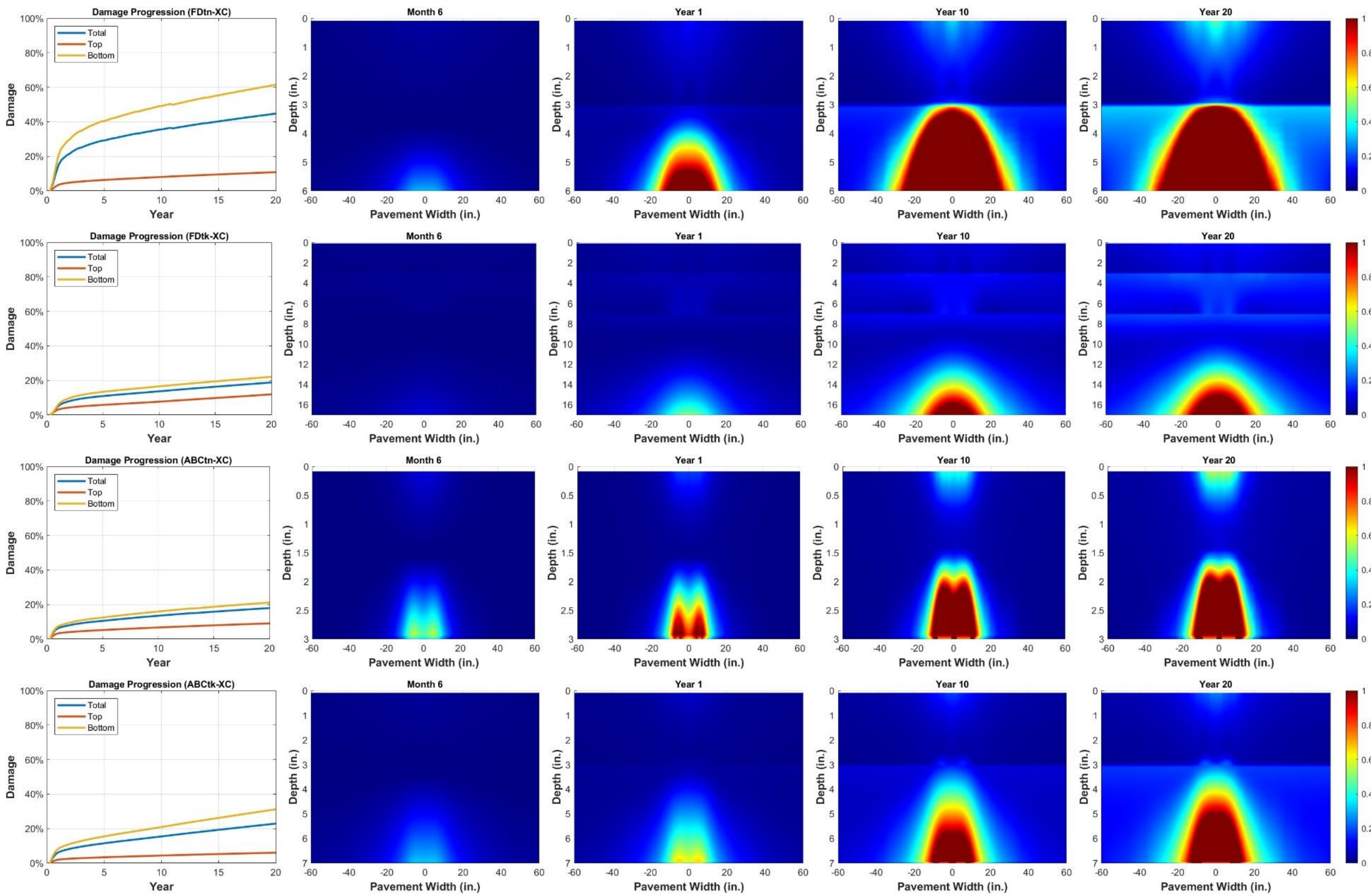


Figure E.9. Damage contours XC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

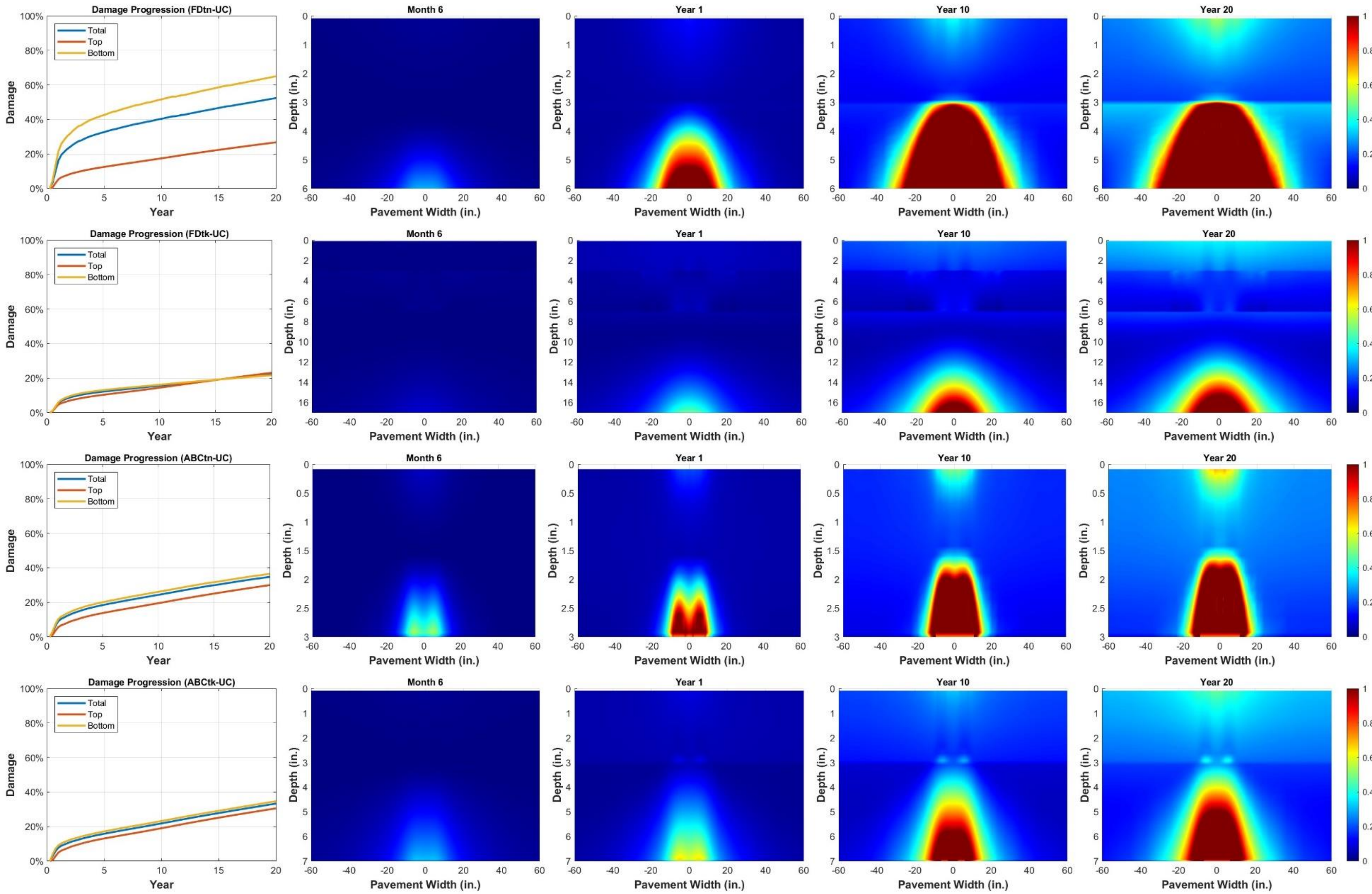


Figure E.10. Damage contours UC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

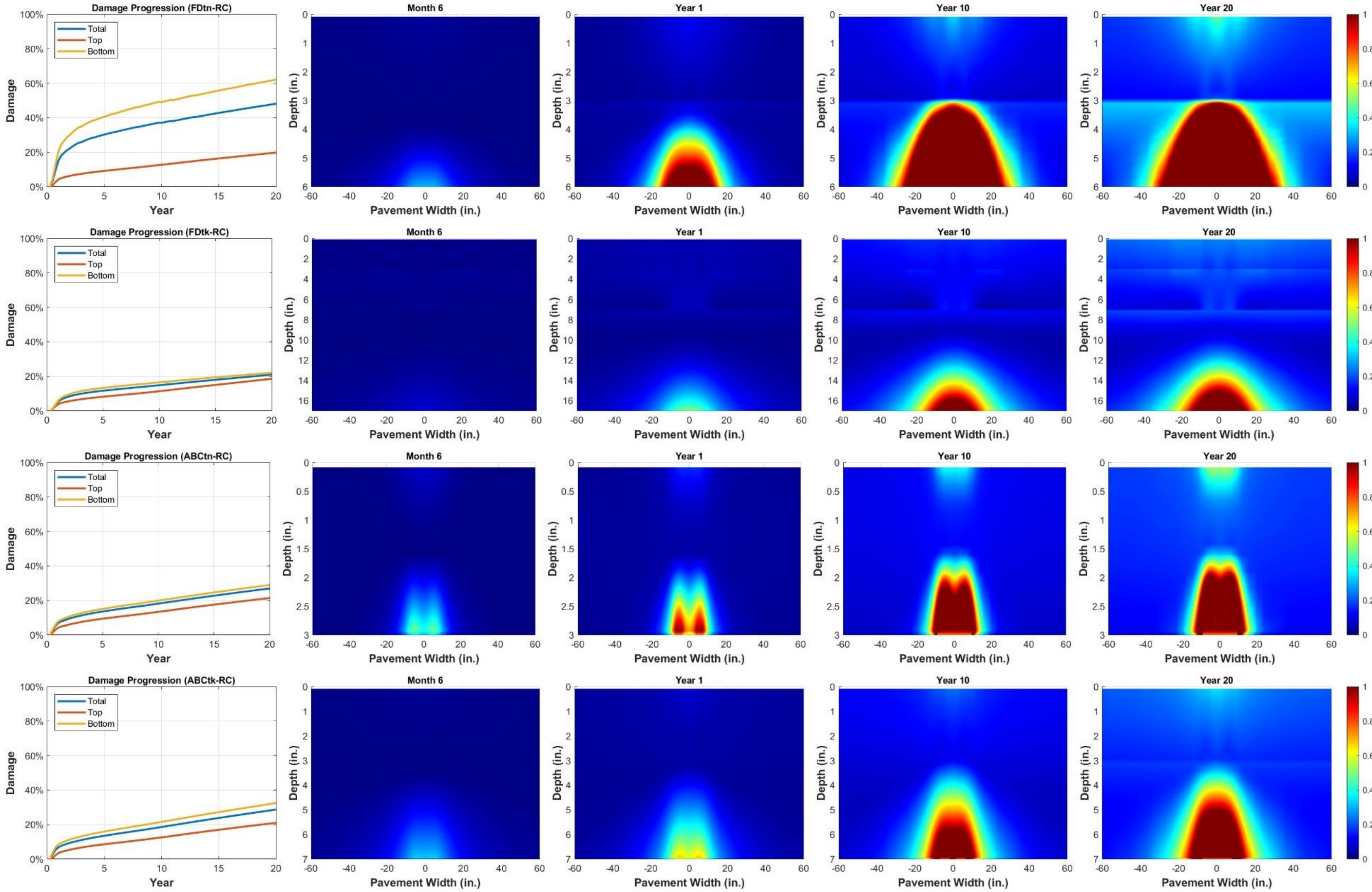


Figure E.11. Damage contours RC mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

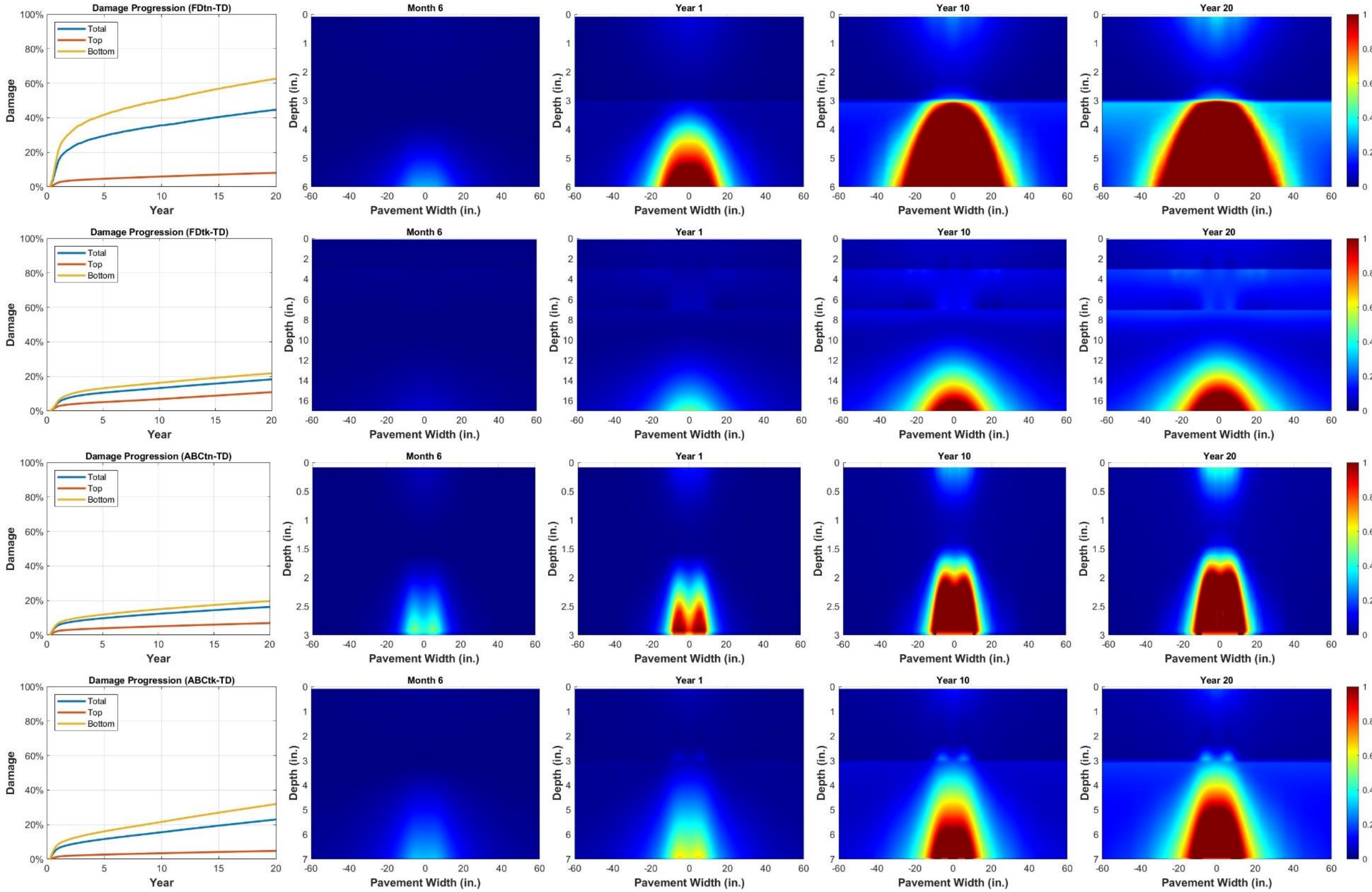


Figure E.12. Damage contours VD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

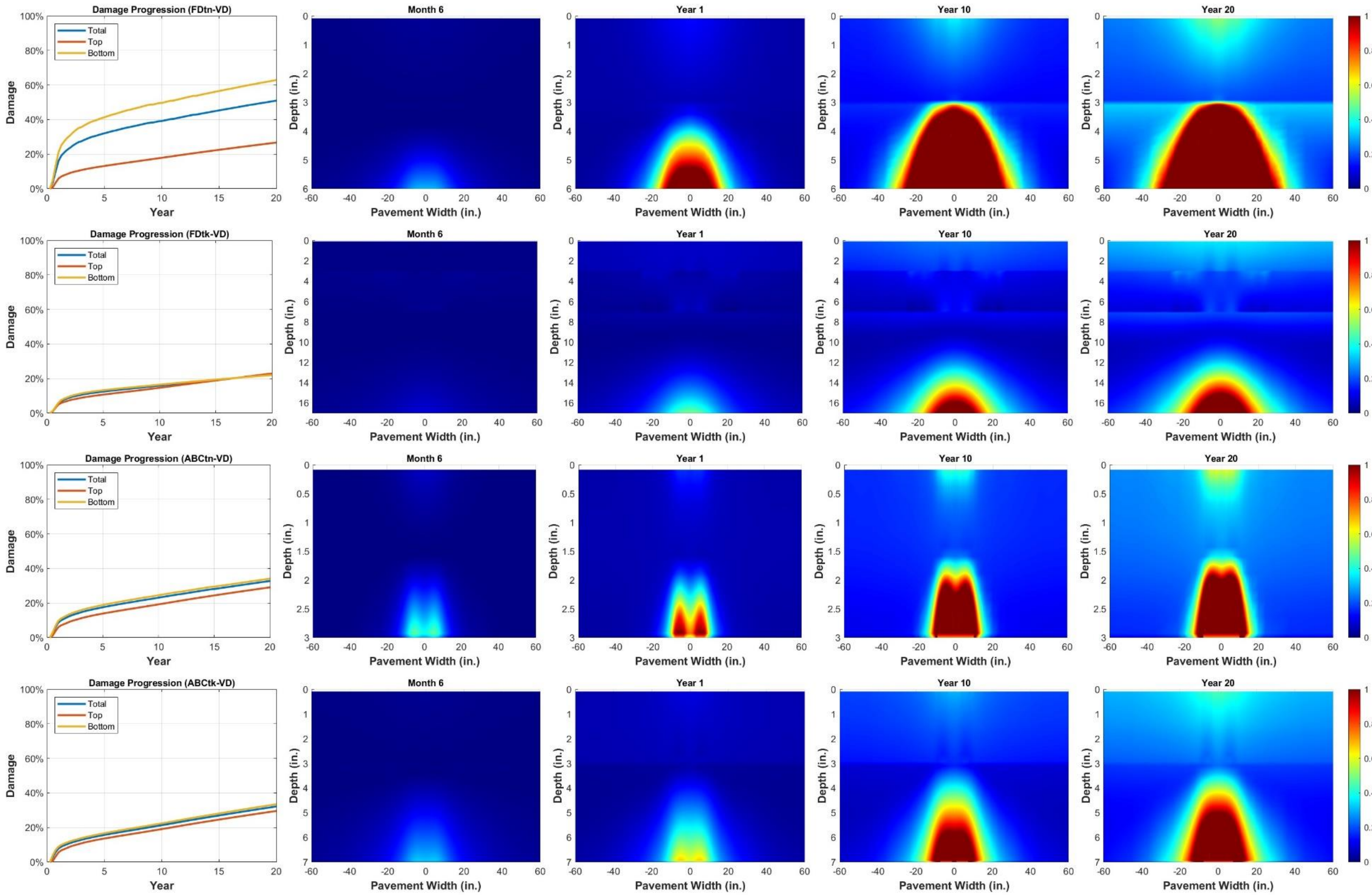


Figure E.13. Damage contours VD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

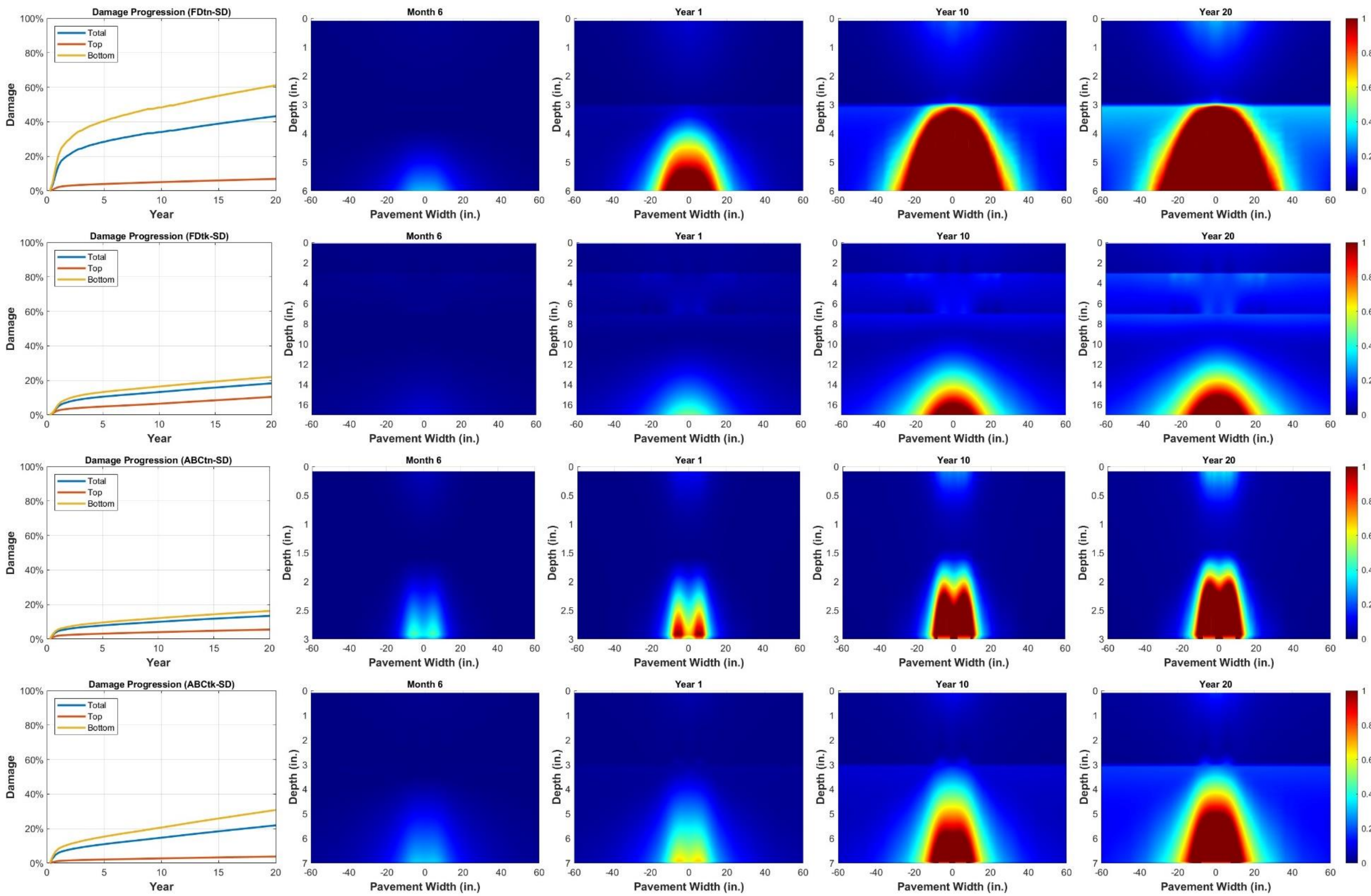


Figure E.14. Damage contours SD mixture structures: (first) FD thin, (second) FD thick, (third) ABC thin, (fourth) ABC thick.

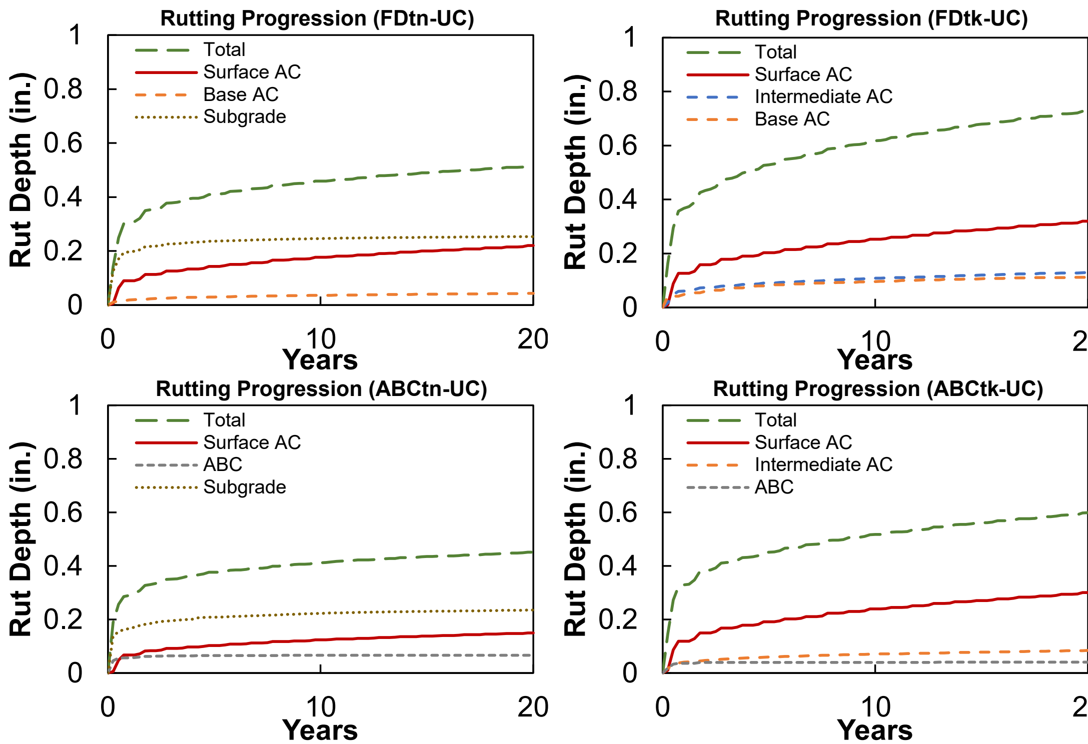


Figure E.15. Rutting prediction for UC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

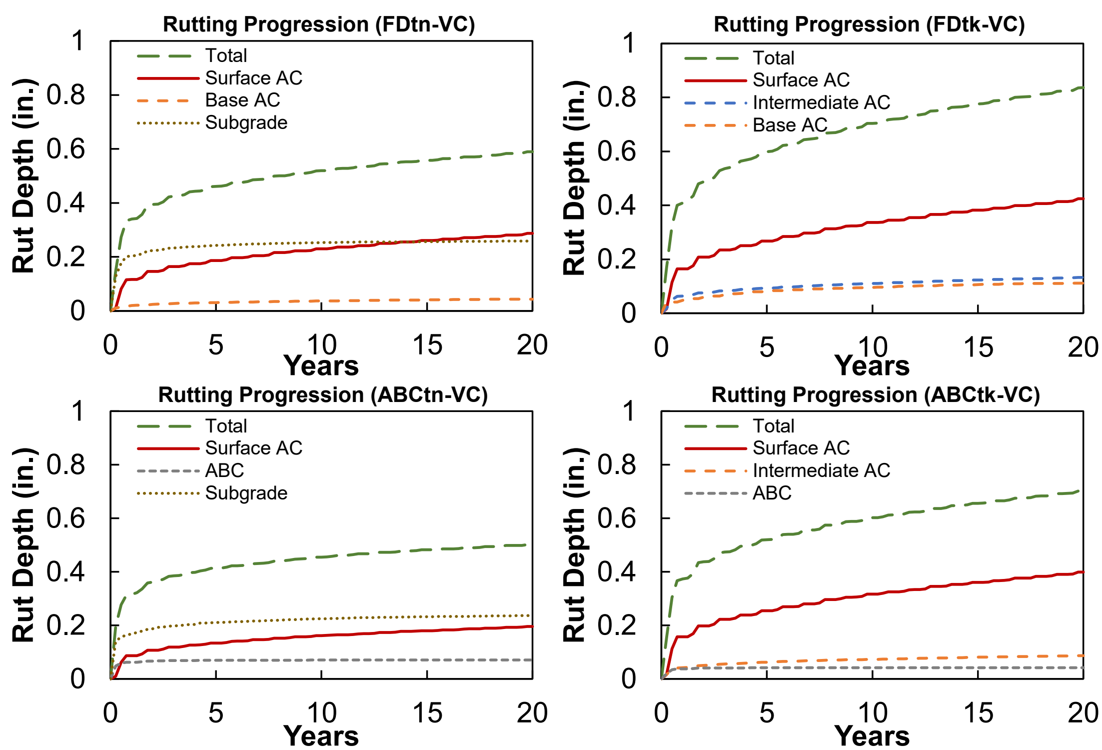


Figure E.16. Rutting prediction for VC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

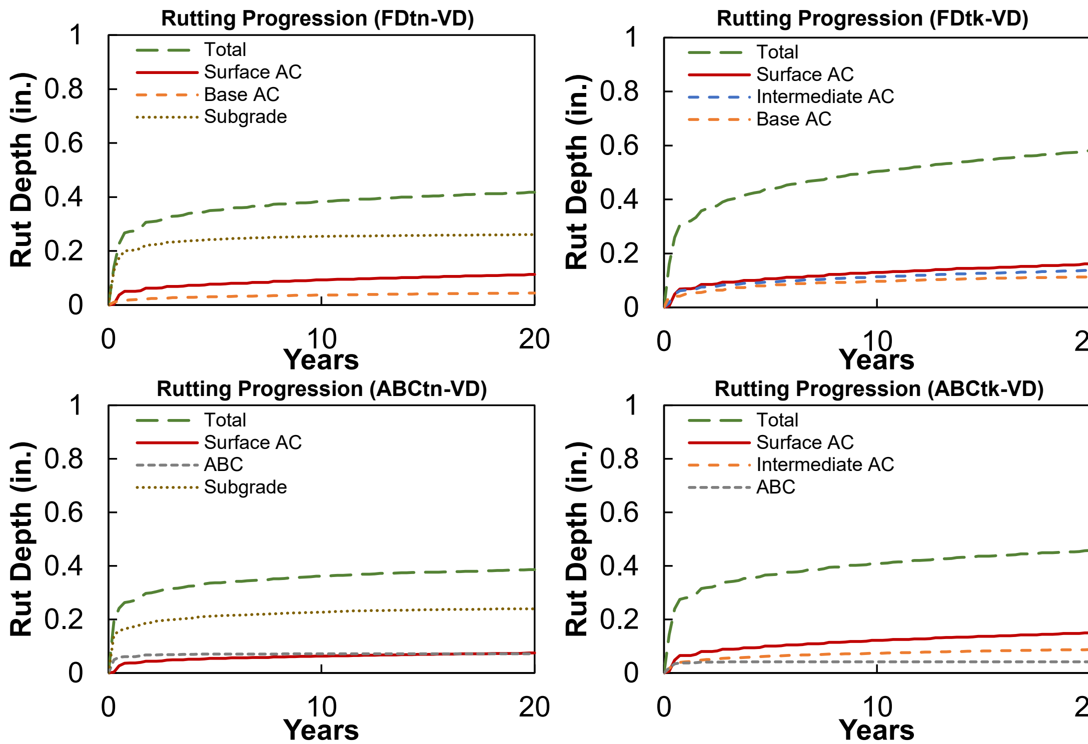


Figure E.17. Rutting prediction for VD structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

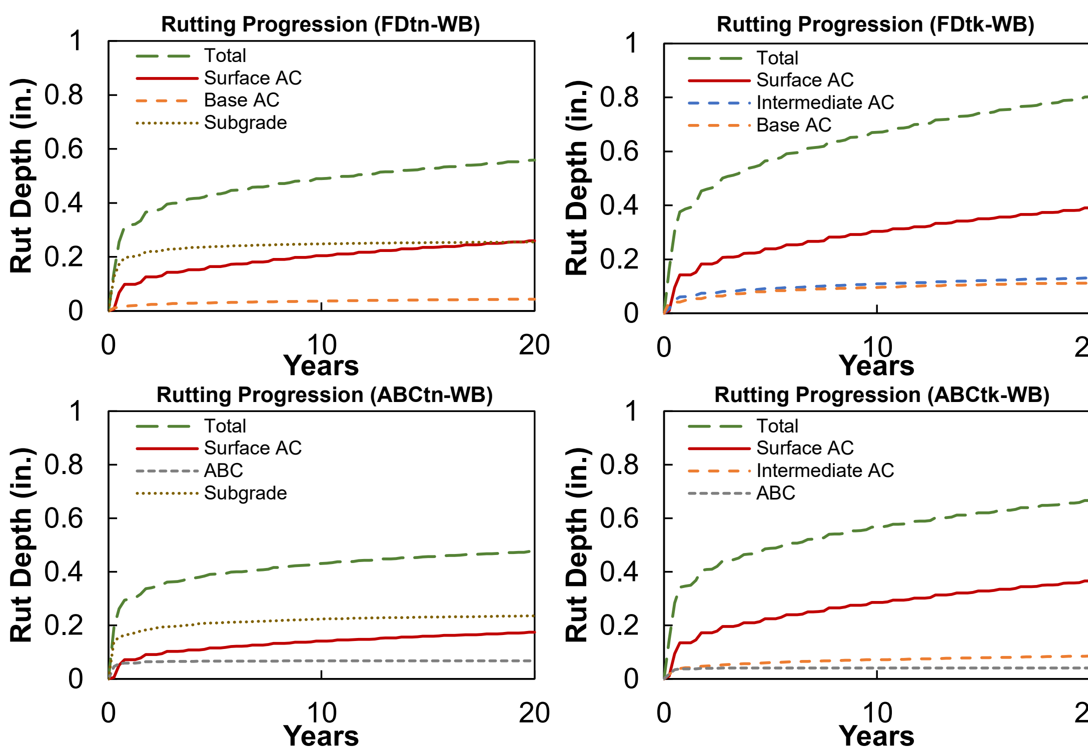


Figure E.18. Rutting prediction for WB structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

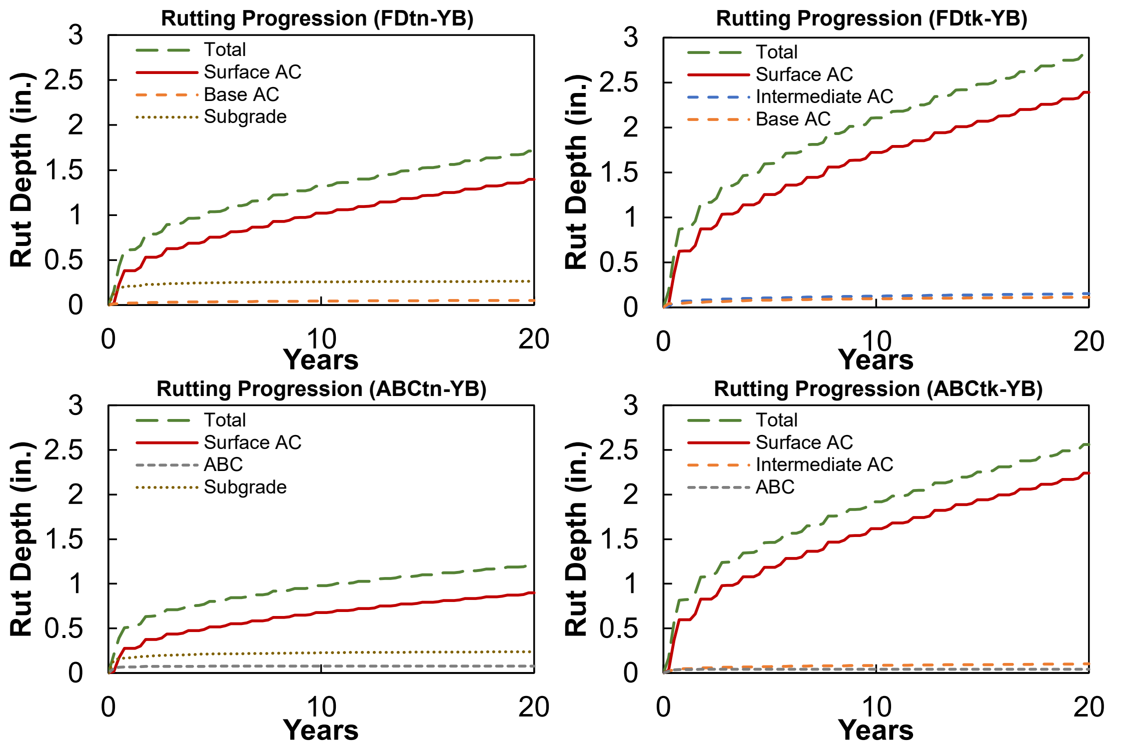


Figure E.19. Rutting prediction for YB structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

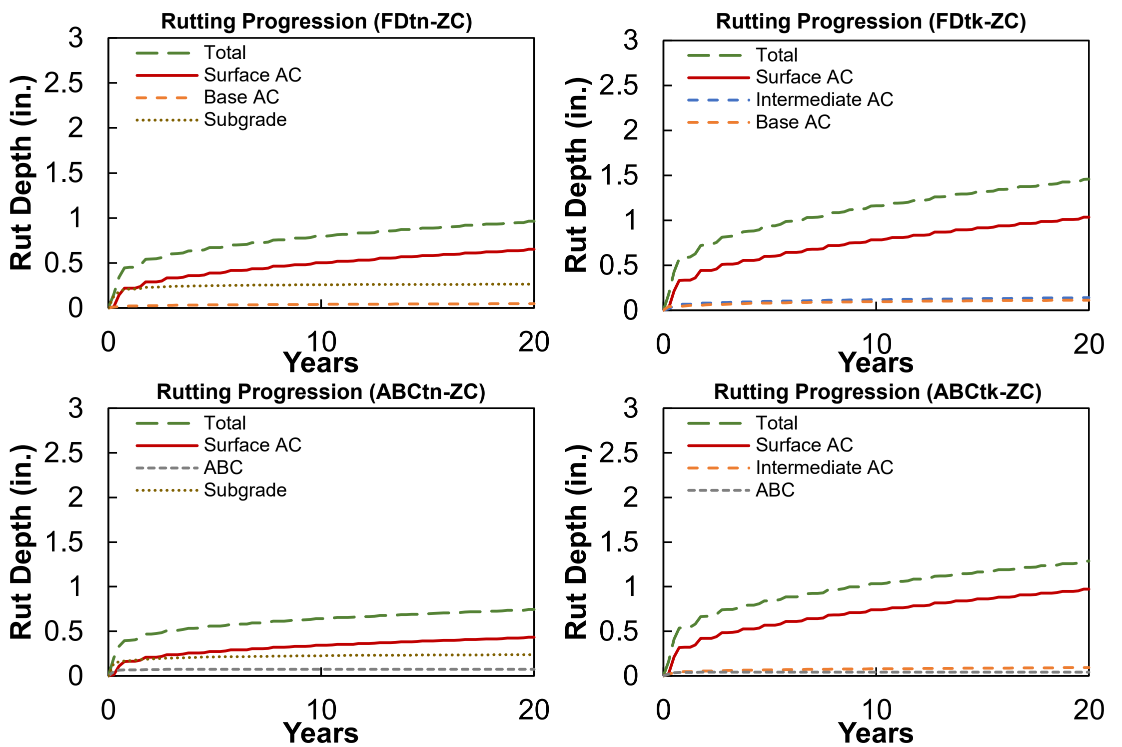


Figure E.20. Rutting prediction for ZC structures: (top-left) FD thin, (top-right) FD thick, (bottom-left) ABC thin, (bottom-right) ABC thick.

Table E.1. Asphalt Mixtures Fatigue and Shift Factor Inputs for FlexPAVE™.

Mixture	*a*1	*a*2	*a*3	*alpha*	C_{11}	C_{12}	D^R
PI_RI19.0C	5.56E-04	-1.47E-01	2.85E+00	3.43	4.10E-03	4.51E-01	0.58
MO_RB25C	5.97E-04	-1.54E-01	2.99E+00	3.09	1.25E-03	5.60E-01	0.39
FB	9.07E-04	-1.81E-01	3.42E+00	3.66	2.52E-03	4.58E-01	0.58
YB	7.15E-04	-1.64E-01	3.15E+00	3.38	2.32E-03	4.63E-01	0.48
WB	6.28E-04	-1.68E-01	3.27E+00	3.38	1.34E-03	4.95E-01	0.43
SB	7.07E-04	-1.73E-01	3.34E+00	3.71	1.91E-03	4.76E-01	0.54
UB	7.95E-04	-1.73E-01	3.29E+00	3.73	1.55E-03	4.90E-01	0.53
HC	7.20E-04	-1.70E-01	3.27E+00	3.72	2.16E-03	4.41E-01	0.57
ZC	8.01E-04	-0.17E-01	3.23E+00	3.36	2.97E-03	4.43E-01	0.48
VC	6.74E-04	-1.67E-01	3.23E+00	3.67	1.38E-03	5.00E-01	0.49
XC	8.60E-04	-1.73E-01	3.27E+00	3.55	1.79E-03	4.82E-01	0.48
UC	8.14E-04	-1.77E-01	3.38E+00	3.67	1.17E-03	5.10E-01	0.49
RC	9.52E-04	-1.80E-01	3.37E+00	3.73	1.84E-03	4.80E-01	0.56
TD	5.38E-04	-1.71E-01	3.37E+00	3.90	1.27E-03	5.00E-01	0.53
VD	6.06E-04	-1.76E-01	3.44E+00	3.92	1.66E-03	4.83E-01	0.55
SD	7.48E-04	-1.79E-01	3.44E+00	3.90	1.30E-03	5.01E-01	0.55

Table E.2. Asphalt Mixtures 2S2P1D Inputs for FlexPAVE™.

Mixture	E_{00} (MPa)	E_0 (MPa)	*k*	*h*	δ	τ_E (s)	β
PI_RI19.0C	18.2	40000	0.1273	0.4474	2.0113	1.59E-03	1E+12
MO_RB25C	0.5	40000	0.1643	0.4899	2.9933	5.32E-02	1E+12
FB	3.5	40000	0.1175	0.3974	2.8750	6.36E-03	1E+12
YB	4.9	40000	0.1285	0.4434	2.7355	8.65E-03	1E+12
WB	0.6	40000	0.1270	0.4306	2.7817	5.32E-02	1E+12
SB	2.4	40000	0.1115	0.3866	3.0444	8.09E-03	1E+12
UB	5.9	40000	0.1141	0.3899	2.4708	1.18E-02	1E+12
HC	6.7	40000	0.1125	0.3939	2.6451	7.14E-03	1E+12
ZC	5.7	40000	0.1332	0.4485	2.5903	8.91E-03	1E+12
VC	7.8	40000	0.1195	0.4018	2.3880	1.41E-02	1E+12
XC	6.4	40000	0.1227	0.4186	2.5022	1.05E-02	1E+12
UC	1.1	40000	0.1237	0.3877	2.4768	4.44E-02	1E+12
RC	4.1	40000	0.1131	0.3879	2.6796	8.31E-03	1E+12
TD	2.0	40000	0.1156	0.3600	2.3586	4.75E-02	1E+12
VD	0.4	40000	0.1090	0.3507	2.9279	2.24E-02	1E+12
SD	1.0	40000	0.1054	0.3562	2.8572	1.70E-02	1E+12

Table E.3. Asphalt Mixtures Rutting Inputs for FlexPAVE™.

Mixture	ε_0	N_1	β	*Tref* (°C)	P_1	P_2	D_1	D_2
PI_RI19.0C	2.86E-03	2.72	0.74	44.98	0.87	0.35	0.19	-2.71
MO_RB25C	1.45E-03	4.09	0.74	40.02	0.80	0.32	0.26	-4.23
YB	2.39E-03	0.01	0.54	51.02	0.76	0.30	0.10	-0.80
WB	2.01E-03	2.00	0.66	51.06	0.83	0.33	0.19	-4.09
ZC	2.67E-03	0.75	0.61	51.06	0.79	0.31	0.13	-1.84
VC	2.66E-03	1.57	0.68	51.06	0.80	0.32	0.18	-3.48
UC	2.49E-03	1.96	0.69	51.00	0.86	0.34	0.18	-3.62
VD	1.78E-03	2.05	0.72	50.94	0.89	0.36	0.25	-5.27

APPENDIX F. DRAFT SPECIFICATION ON MIX DESIGN

SECTION 610 (2024 Standards)
ASPHALT CONCRETE PLANT MIX PAVEMENTS – SUGGESTED CHANGES

610-1 DESCRIPTION

Perform the work covered by this section including, but not limited to, the construction of one or more courses of asphalt mixture placed on a prepared surface in accordance with these specifications and in reasonably close conformity with the lines, grades, thickness and typical sections shown on the plans. This work includes producing, weighing, transporting, placing and compacting the plant mix; furnishing aggregate, asphalt binder, anti-strip additive and all other materials for the plant mix; furnishing and applying tack coat as specified; furnishing scales; maintaining the course until final acceptance of the project; making any repairs or corrections to the course that may become necessary; providing and conducting QC as specified in Section 609; and surface testing of the completed pavement. The design requirements for the various mix types are given in Section 610 for dense-graded mix types, Section 650 for OGFC, Section 652 for PADC and Section 661 for UBWC.

Perform all activities in accordance with the Department's *Asphalt Quality Management System (QMS) Manual* in effect on the date of contract advertisement

Provide and conduct the QC and required testing for acceptance of the asphalt mixture in accordance with Section 609.

Define "warm mix asphalt (WMA)" as additives or processes that allow a reduction in the temperature at which asphalt mixtures are produced and placed. Use only NCDOT approved WMA additives listed on the NCDOT APL maintained by the Materials and Tests Unit.

Comments

- No modifications are expected in 610-1.

610-2 MATERIALS

Refer to Division 10.

Item	Section
Anti-Strip Additives	1020-8
Asphalt Binder, Performance Grade	1020-2
Coarse Aggregate	1012-1(B)
Fine Aggregate	1012-1(C)
Mineral Filler	1012-1(D)
Reclaimed Asphalt Pavement (RAP)	1012-1(F)
Reclaimed Asphalt Shingles (RAS)	1012-1(E)
Silicone	1020-9

Comments

- No modifications are expected in 610-2 nor the referenced Division 10 items.

610-3 COMPOSITION OF MIXTURES (MIX DESIGN AND JOB MIX FORMULA)

(A) Mix Design-General

Prepare the asphalt mix design using a mixture of coarse and fine aggregate, asphalt binder, mineral filler and other additives when required. Size, uniformly grade and combine the several aggregate fractions in such proportions that the resulting mixture meets the grading and physical requirements of the *Standard Specifications* for the specified mix type. Materials that will not produce a mixture within the design criteria required by the specifications will be rejected, unless otherwise approved by the Engineer.

At least 10 days excluding official state holidays before start of asphalt mix production, submit the mix design and proposed JMF targets for each required mix type and combination of aggregates to the Engineer for review and approval. Prepare the mix design using a Department certified mix design technician in an approved mix design laboratory and in accordance with the procedures outlined in Section 4.5 of the *Asphalt QMS Manual.*

For the final surface layer of the specified mix type, use a mix design with an aggregate blend gradation above the maximum density line on the 2.36 mm and larger sieves.

The Contractor has the option to use a recycled plant mix in lieu of virgin plant mix.

However, all provisions of the specifications for virgin mixes apply to recycled mixes. This means that the same design criteria tests, test frequencies, and quality control requirements will apply.

Reclaimed Asphalt Pavement (RAP) or Reclaimed Asphalt Shingles (RAS) may be incorporated into asphalt plant mixes in accordance with Article 1012-1 and the following applicable requirements. However, use of RAP materials is not allowed in Open-Graded Friction Course (OGFC) mixes or Ultra-Thin Bonded Wearing Course (UBWC) mixes. Use of RAS materials is not allowed in Ultra-Thin Bonded Wearing Course (UBWC) mixes.

RAS material may constitute up to 6% by weight of total mixture, except for Open Graded Friction Course (OGFC) mixes, which are limited to 5% RAS by weight of total mixture. Also, when the percentage of RAP is greater than 30% by weight of total mixture, use Fractionated RAP (FRAP) meeting the requirements of Subarticle 1012-1(F)(c).

When RAP, RAS, or a combination of both is used in asphalt mixtures, the recycled binder replacement percentage (RBR%) shall not exceed the amounts specified in Table 610-4 for the mix type. For recycled mixtures, the virgin binder Performance Grade (PG) grade to be used is specified in Table 610-5 for the mix type based on the recycled binder replacement percentage (RBR%).

If the Contractor wishes to submit mix designs containing recycled material amounts exceeding the specified maximums, additional testing will be required to verify the Performance Grade (PG) of the reclaimed binder. Also, the Contractor has the option to have additional testing performed to determine if the mix can be approved using a virgin binder grade different than specified in Table 610-5. The Engineer will determine if the binder grade is acceptable for use based on the test data submitted with the mix design If the mix design is acceptable, the Engineer will establish and approve the grade and percentage of virgin asphalt binder to be used.

If a change in the source of RAP or RAS be made, a new mix design and JMF may be required in accordance with Article 1012-1. Samples of the completed recycled mixture may be taken by the Department on a random basis and the recovered asphalt binder will be tested in accordance with Article 1020-2. If the grading is determined to be a value other than required for the specified mix type, the Engineer may require the Contractor to adjust any combination of the grade, the percentage of additional asphalt binder or the blend of reclaimed material to bring the grade to the specified value.

(B) Mix Design Criteria

Design and produce asphalt concrete mixtures that conform to the gradation requirements and design criteria in Table 610-2 and Table 610-3 for the mix type specified. The mix type designates the nominal maximum aggregate size and the design traffic level.

Surface mix designs will be tested by the Department for rutting susceptibility and cracking resistance. Rut depth requirements for each surface mix type and traffic level are specified in Table 610-3. Indirect Tension Cracking Test requirements for surface mixes are specified in Table 610-3. Mix designs that fail to meet these requirements will be unacceptable and shall be redesigned by the Contractor such that rut depths and cracking index are acceptable.

Table 610-2 provides gradation control points to be adhered to in the development of the design aggregate structure for each mix type. Aggregate gradations shall be equal to or pass between the control points. Table 610-3 provides the mix design criteria for the various mix types.

Use an anti-strip additive in all asphalt mixes. It may be hydrated lime or a chemical additive or a combination of both as needed to meet the retained strength requirements as specified in Table 610-3. When a chemical additive is used, add at a rate of not less than 0.25% by weight of binder in the mix, or as approved by the Engineer. When hydrated lime is used, add at a rate of not less than 1.0% by weight of the total dry aggregate.

(C) Job Mix Formula (JMF)

Establish the JMF gradation target values within the design criteria specified for the particular type of asphalt mixture to be produced. Establish the JMF asphalt binder content at the percentage that will produce voids in total mix (VTM) at the midpoint of the specification design range for VTM, unless otherwise approved by the Engineer. The formula for each mixture will establish the following: blend percentage of each aggregate fraction, the percentage of reclaimed aggregate, if applicable, a single percentage of combined aggregate passing each required sieve size, the total percentage and grade of asphalt binder required for the mixture (by weight of total mixture), the percentage and grade of asphalt binder to be added to the mixture (for recycled mixtures), the percentage of chemical anti-strip additive to be added to the asphalt binder or percentage of hydrated lime to be added to the aggregate, the temperature at that the mixture is to be discharged from the plant, the required field density and other volumetric properties.

When WMA is used, document the additive or process used and recommended rate on the JMF submittal. Verify the JMF based on plant produced mixture from the trial batch.

The mixing temperature at the asphalt plant will be established on the JMF. The JMF mix temperature shall be within the ranges shown in Table 610-1 unless otherwise approved by the Engineer.

TABLE 610-1 MIXING TEMPERATURE AT THE ASPHALT PLANT	
Binder Grade	**JMF Temperature**
PG 58-28; PG 64-22	250 - 290°F
PG 76-22	300 - 325°F

When RAS is used, the JMF mix temperature shall be established at 275°F or higher.

Have on hand at the asphalt plant the approved mix design and JMF issued by the Department, before beginning the work.

The JMF for each mixture will remain in effect until modified in writing, provided the results of QMS tests performed in accordance with Section 609 on material currently being produced conform with specification requirements. When a change in sources of aggregate materials is to be made, a new mix design and JMF will be required before the new mixture is produced. When a change in sources of RAP or RAS material is to be made, a new JMF is required and a new mix design may be required. When unsatisfactory results or other conditions make it necessary, the Engineer may revoke the existing JMF or establish a new JMF.

TABLE 610-2
AGGREGATE GRADATION CRITERIA
(Percent Passing Control Points)

Standard Sieves (mm)	Mix Type (Nominal Max. Aggregate Size)							
	4.75 mm		9.5 mm[A]		19.0 mm		25.0 mm	
	Min	*Max*	*Min*	*Max*	*Min*	*Max*	*Min*	*Max*
50.0	-	-	-	-	-	-	-	-
37.5	-	-	-	-	-	-	100	-
25.0	-	-	-	-	100	-	90.0	100
19.0	-	-	-	-	90.0	100	-	90.0
12.5	100	-	100	-	-	90.0	-	-
9.50	95.0	100	90.0	100	-	-	-	-
4.75	90.0	100	-	90.0	-	-	-	-
2.36	-	-	32.0[B]	67.0[B]	23.0	49.0	19.0	45.0
1.18	30.0	60.0	-	-	-	-	-	-
0.075	6.0	12.0	4.0	8.0	3.0	8.0	3.0	7.0

A. For the final surface layer of the specified mix type, use a mix design with an aggregate blend gradation above the maximum density line on the 2.36 mm and larger sieves.

B. For Type S9.5B, the percent passing the 2.36 mm sieve shall be a minimum of 60% and a maximum of 70%.

TABLE 610-3 MIX DESIGN CRITERIA										
Mix Type	Design ESALs millions[A]	Binder PG Grade	Compaction Levels		Max. Rut Depth (mm)	Min. CTindex[B]	Volumetric Properties[C]			
			G_{mm} @				VMA	VTM	VFA	%G_{mm}
			N_{ini}	N_{des}			% Min.	%	Min.-Max.	@ N_{ini}
S4.75A	< 1	64 - 22	6	50	11.5	-	16.0	4.0 - 6.0	65 - 80	≤ 91.5
S9.5B	0 - 3	64 - 22	6	50	9.5	14	16.0	3.0 - 5.0	70 - 80	≤ 91.5
S9.5C	3 - 30	64 - 22	7	65	6.5	14	15.5	3.0 - 5.0	65 - 78	≤ 90.5
S9.5D	> 30	76 - 22	8	100	4.5	-	15.5	3.0 - 5.0	65 - 78	≤ 90.0
I19.0C	ALL	64 - 22	7	65	-	-	13.5	3.0 - 5.0	65 - 78	≤ 90.5
B25.0C	ALL	64 - 22	7	65	-	-	12.5	3.0 - 5.0	65 - 78	≤ 90.5
	Design Parameter						Design Criteria			
All Mix Types	Dust to Binder Ratio ($P_{0.075}$ / P_{be})						0.6 - 1.4[D]			
	Tensile Strength Ratio (TSR)[E]						85% Min[F]			

A. Based on 20 year design traffic.

B. ASTM D8225-19 and short-term oven aging performed using procedure to use 4 hours at compaction temperature.

C. Volumetric Properties based on specimens compacted to N_{des} as modified by the Department.

D. Dust to Binder Ratio ($P_{0.075}$ / P_{be}) for Type S4.75A is 1.0 - 2.0.

E. NCDOT-T-283 (No Freeze-Thaw cycle required).

F. TSR for Type S4.75A & B25.0C mixes is 80% minimum.

TABLE 610-4 MAXIMUM RECYCLED BINDER REPLACEMENT PERCENTAGE (RBR%)			
Recycled Material	**Intermediate & Base Mixes**	**Surface Mixes**	**Mixes Using PG 76-22**
RAS	23%	20%	18%
RAP or RAP/RAS Combination	45%	40%	18%

TABLE 610-5 BINDER GRADE REQUIREMENTS (BASED ON RBR%)			
Mix Type	**%RBR ≤ 20%**	**21% ≤ %RBR ≤ 30%**	**%RBR > 30%**
S4.75A, S9.5B, S9.5C, I19.0C, B25.0C	PG 64-22	PG 64-22[A]	PG 58-28
S9.5D, OGFC	PG 76-22[B]	n/a	n/a

A. If the mix contains any amount of RAS, the virgin binder shall be PG 58-28.
B. Maximum Recycled Binder Replacement (%RBR) is 18% for mixes using PG 76-22 binder.

Comments

- Section B modified to include the indirect tension cracking test.
- Table 610-3 updated to include the CT_{index} threshold limit for S9.5B and S9.5C mixes. A footnote is added to the table to specify that the specimen should be short-term oven aged for 4 hours at the mixture compaction temperature.
- Potential additional changes could include specifically requiring contractors to provide component materials so that the NCDOT could control aging and ensure tested materials conform to NCDOT expectations.

610-4 Through 610-16

No changes are suggested in the remaining sections.