Assessment of Pavement Acceptance Criteria and Quantifying its As-Constructed Material and Structural Properties

Karim Chatti, Ph.D.
Anshu Manik, Ph.D.
Imen Zaabar, Ph.D.

Michigan State University

Department of Civil and Environmental Engineering
Pavement Research Center of Excellence

Final Report
Project RC-1540
Contract/Authorization # 2006-0411 Z4
March 2011

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. MDOT Project Manager
RC-1540		Dave Smiley and Robert Pena

4. Title and Subtitle	5. Report Date
Assessment of Pavement Acceptance Criteria and Quantifying its As-constructed Material and Structural Properties	March 2011
	6. Performing Organization Code
7. Author(s) Karim Chatti, Ph.D. Anshu Manik, Ph.D. Imen Zaabar. Ph.D.	**8. Performing Org. Report No.**
9. Performing Organization Name and Address Michigan State University Department of Civil and Environmental Engineering, 3546 Engineering Building, Michigan State University East Lansing, MI 48824-1226	**10. Work Unit No. (TRAIS)**
	11. Contract No. 2006-0411 Z4
	11(a). Authorization No.
12. Sponsoring Agency Name and Address Michigan Department of Transportation Construction and Technology Division P.O. Box 30049 Lansing, MI 48909	**13. Type of Report & Period Covered** Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

State Highway Agencies (SHAs) have long realized the importance of quality assurance (QA) to ensure longer pavement service life. Construction QA programs are intended to ensure that the quality of the materials and construction in highway projects is satisfactory. Failure to conform to either material or construction specifications can result in the premature failure of pavements. The intent of this research project is to assess existing pavement acceptance criteria, conduct a feasibility study to consider the need of using new criteria that relate to pavement performance including the use of non-destructive testing, and develop a process to quantify as-constructed material and structural pavement properties and relate them to pavement service life. This was achieved by reviewing MDOT construction data, construction and performance data from the national LTPP database, reviewing QA programs of other state highway agencies, conducting a literature review of previous research studies on the subject, and finally running simulations using the new MEPDG software to relate quality characteristics to pavement performance. Finally, Monte-Carlo based simulations were developed to understand how these quality characteristics should be used in the specifications to encourage better quality and reduce the risk of overpayment or underpayment. A plan of action was also developed for MDOT to implement the recommendations in its quality assurance and design process. An important part of this is a feedback process for design.

17. Key Words	18. Distribution Statement
Quality assurance, acceptance criteria, as-constructed material and structural properties, MEPDG, risk analysis, end-result specifications, feedback to design.	No restrictions. This document is available to the public through the Michigan Department of Transportation.

19. Security Classification - report	20. Security Classification - page	21. No. of Pages	22. Price
Unclassified	Unclassified	411	

DISCLAMER

This document is disseminated under the sponsorship of the Michigan Department of Transportation (MDOT) in the interest of information exchange. MDOT assumes no liability for its contents or use thereof.

The contents of this report reflect the views of the contracting organization, which is responsible for the accuracy of the information presented herein. The contents may not necessarily reflect the views of MDOT and do not constitute standards, specifications, or regulations.

Table of contents

Table of contents i
List of Figures vii
List of Tables xiii
Introduction 1
CHAPTER 1: Introduction and Background 1
1.1 INTRODUCTION 1
1.2 PROBLEM STATEMENT 2
1.3 OBJECTIVES 3
1.4 RESEARCH APPROACH 3
1.5 ORGANIZATION OF THE REPORT 6
Part I: Flexible Pavements 9
CHAPTER 2: MDOT Quality Assurance Data for Flexible Pavements 10
2.1 SELECTION OF SAMPLE PROJECTS 10
2.2 SAMPLE DATA GATHERING 11
CHAPTER 3: Review of MDOT QA Program and Design Process for HMA Pavements 13
3.1 Michigan DOT Quality Assurance Program 13
3.1.1 Percent Within Limits 14
3.1.2 Pay Factors for Quality Acceptance Items 20
3.1.2.1 Pay Factor for Air Voids – PF_{AV}: 20
3.1.2.2 Pay Factor for Binder Content (PF_{BINDER}) 20
3.1.2.3 Pay Factor for Voids in Mineral Aggregates – PF_{VMA}: 21
3.1.2.4 Pay Factor for In-Place Density – PF_{D}: 21
3.1.2.5 Overall Lot Pay Factor 21
3.1.3 Payment 22
3.2 Michigan DOT Pavement Design Process 22
Chapter 4: HMA Quality Assurance Programs in the United States 25
4.1 Regional Level Analysis of the QA Practices 26
4.1.1 North- East Region 27
4.1.2 Southern Region 27
4.1.3 Mid-West Region 27
4.1.4 Western Region 31

4.2 Conclusion 31
Chapter 5: Literature Review - Flexible Pavements 33
5.1 Asphalt Content 33
5.2 Air Void Content 34
5.3 Mixture Density 36
5.4 Aggregate Size and Gradation 37
5.5 Pavement Thickness 39
5.6 Binder Performance Grade 39
5.7 Review of different Non-Destructive tests 40
5.7.1 Thickness 40
5.7.1.1 Ground-penetrating radar (GPR) 40
5.7.1.1.4 Advantages and Limitations of GPR 43
5.7.1.1.5 Accuracy and Interpretation of GPR 44
5.7.1.2 Summary 44
5.7.1.3 Conclusion 44
5.7.2 Moisture and Density (Radioisotope Gauges) 45
5.7.3 Modulus of Pavement Layers 45
5.7.3.1 Humboldt Stiffness Gauge 46
5.7.3.2 Falling Weight Deflectometer (FWD) 46
5.7.3.3 Dynamic Cone Penetrometer 47
5.7.3.4 Seismic Pavement Analyzer 48
5.7.3.4.1 Advantages and Limitations of SPA 48
5.7.3.5 Conclusion for Modulus of Pavement Layers 49
5.7.4 HMA Temperature 49
Chapter 6: Empirical Data Analysis – Flexible Pavements 51
6.1 Preliminary Analysis 55
6.1.1 Rutting 55
6.1.2 Fatigue 57
6.1.3 Longitudinal Cracking 58
6.2 Analysis Using Percent-Within-Limits Concept 60
6.3 Conclusions 68
Chapter 7: Mechanistic-Empirical Analysis 70

7.1 Purpose 70
7.2 Analysis Approach 70
7.3 Summary of Findings 85
7.4 Simulation Using MEPDG 86
7.4.1 Effect of varying air void distribution on average project performance 89
7.4.2 Effect of varying both air void distribution and asphalt content on sublot project performance by sublot 92
7.5 Performance criteria 93
7.6 Effect of varying HMA thickness on project performance by sublot 108
Chapter 8: ERS Risk Analysis Using Simulation 122
8.1 INTRODUCTION AND BACKGROUND 122
8.2 END-RESULT SPECIFICATIONS 122
8.3 ESTIMATING RISK 123
8.4 MAJOR FACTORS AFFECTING RISK 125
8.5 COMPOSITE RISK INDEX 126
8.6 RISK ANALYSIS 127
8.7 ANALYSIS OF VARIANCE (ANOVA) 135
8.8 Conclusion 138
Chapter 9: Feedback Process to Design for HMA Pavements 140
9.1 MDOT Data Analysis 142
9.2 Assumption of Normality 143
9.3 Comparison of QA and QC Data 149
9.4 Effect of Sample Size on Feedback Process using Simulation 155
9.5 Conclusion 166
Chapter 10: Conclusions and Recommendations for HMA QA Program 167
10.1 Identification of Suitable Quality Characteristics for QA Program 167
10.1.1 Comparison of MDOT QA programs with others in the US 169
10.1.2 QA Parameters Indentified by Other Studies 169
10.1.3 Summary of Results from Empirical Analysis 169
10.1.4 Summary of Results from Mechanistic-Empirical Analysis 170
Effect of varying air void distribution on average project performance 171
Effect of varying air void distribution and asphalt content on project sublot performance 171

Effect of varying HMA thickness on project performance by sublot 173
10.2 Summary of Results from ERS Risk Analysis 173
10.3 Feedback Process for Design 174
10.4 Use of Non-destructive Tests in QA Program 175
10.5 Overall Conclusion and Recommendations 175
Part II : Rigid Pavements 178
CHAPTER 11: MDOT Quality Assurance Data for Rigid Pavements 179
11.1 Selection of Sample Projects 179
11.2 Sample Data Gathering 180
CHAPTER 12: Review of MDOT QA Program and Design Process for PCC Pavements 181
12.1 MDOT PCC Pavement Quality Assurance Program 181
12.2 MDOT PCC Pavement Design Process 187
CHAPTER 13: Survey of QA Programs for Rigid Pavements in the United States 189
13.1 Regional Level Analysis of the QA Practices 190
CHAPTER 14: Literature Review – Rigid Pavements 196
14.1 Slab Thickness 196
14.2 Water-Cement Ratio 196
14.3 Coefficient of Thermal Expansion (CTE) 197
14.4 Aggregate Gradation 198
14.5 Tensile Strength 199
14.6 Cement Content 199
14.7 Construction Temperature 200
14.8 Other Factors 200
14.9 Review of Different Non-Destructive Tests 201
14.9.1 Thickness 201
14.9.1.1 Ground-penetrating radar (GPR) 201
14.9.1.2 Mechanical wave methods for concrete thickness evaluation 202
14.9.2 Density 204
14.9.2.1 Thickness and Density (Radioisotope Gauges) 204
14.9.2.2 Moisture Gauges 206
14.9.2.3 Air Void Analyzer 207
14.9.3 Uniformity test 207

14.9.3.1 Pulse velocity test 207
14.9.3.2 Surface Hardness Test 209
14.9.4 Strength of Concrete 210
14.9.4.1 Penetration Resistance or Windsor Probe Test 210
14.9.4.2 Pullout Test 212
14.9.4.3 Break-Off Test 213
14.9.4.4 Maturity Test 213
14.9.5 Hidden Flaws 214
14.9.5.1 Infrared Thermography 214
14.9.5.2 Betatron PXB - 7.5 MeV (Force technology, 1999) 215
14.9.6 Modulus of Pavement Layers 216
14.9.6.1 Falling Weight Deflectometer (FWD) 216
14.9.7 Seismic Pavement Analyzer 216
CHAPTER 15: Empirical Data analysis – Rigid Pavements 217
15.1 Analysis Using Percent-Within-Limits Concept 217
15.2 Summary of Findings from LTPP Data Analysis 221
CHAPTER 16: Mechanistic-Empirical Analysis for Rigid Pavements 223
16.1 Analysis using MEPDG 223
16.2 Analysis using HIPERPAV II 233
16.3 SUMMARY AND CONCLUSIONS 237
CHAPTER 17: ERS Risk Analysis for Rigid Pavements Using Simulation 238
17.1 INTRODUCTION AND BACKGROUND 238
17.2 END-RESULT SPECIFICATIONS 238
17.3 ESTIMATING RISK 239
17.4 MAJOR FACTORS AFFECTING RISK 241
17.5 COMPOSITE RISK INDEX 242
17.6 RISK ANALYSIS 243
17.7 Analysis of Variance (ANOVA) 251
17.8 Summary of Results from ERS Risk Analysis 254
CHAPTER 18: Feedback Process to Design for PCC Pavements 255
18.1 Effect of Sample Size on Feedback Process using Simulation 257
18.2 Conclusion 269

CHAPTER 19: Conclusions and Recommendations for Rigid Pavement QA Program 270
19.1 Identification of Suitable Quality Characteristics for QA Program 270
19.1.1 Comparison of MDOT QA programs with others in the US 272
19.1.2 QA Parameters Indentified by Other Studies 272
19.1.3 Summary of Results from Empirical Analysis 273
19.1.4 Summary of Results from Mechanistic-empirical Analysis 274
19.2 Summary of Results from ERS Risk Analysis 275
19.3 Summary of Results from Feedback Process for Design 276
19.4 Use of Non-destructive Tests in QA Program 276
19.5 Overall Conclusion and Recommendations 276
References 279
Appendix A: MDOT Construction Data Details 288
Appendix B: Review of Different Non-Destructive Tests 313

List of Figures

Figure 3.1 Flowchart for Acceptance Testing Procedure for HMA Pavements 16
Figure 4.1 Attributes most often used for QA of HMA (40 responses) (adapted from NCHRP Synthesis 346)..... 26
Figure 6.1 Number of projects from different states which have fatigue data. 52
Figure 6.2 Number of projects from different states which have rutting data..... 53
Figure 6.3 Distribution of projects based on years of fatigue data available..... 54
Figure 6.4 Distribution of projects based on years of rutting data available 54
Figure 6.5 Rutting vs. percent air voids for all SPS-1 sections in LTPP database 55
Figure 6.6 Rutting vs. asphalt content for all SPS-1 sections in LTPP database..... 56
Figure 6.7 fatigue vs. in-situ air voids content for all SPS-1 sections in LTPP database. 57
Figure 6.8 Fatigue vs. asphalt content for all SPS-1 sections in LTPP database..... 58
Figure 6.9 Longitudinal cracking vs. in-situ air voids content for all SPS-1 sections in LTPP database 59
Figure 6.10 Longitudinal cracking vs. asphalt content for all SPS-1 sections in LTPP database..... 59
Figure 6.11 PWL (in-situ density) and fatigue cracking..... 62
Figure 6.12 PWL (in-situ density) and longitudinal cracking 62
Figure 6.13 PWL (in-situ density) and transverse crack Spacing..... 64
Figure 6.14 PWL (in-situ density) and rutting..... 65
Figure 6.15 PWL (plant air voids) and fatigue cracking 65
Figure 6.16 PWL (plant air voids) and longitudinal cracking 66
Figure 6.17 PWL (plant air voids) and transverse cracking 67
Figure 6.18 PWL (plant air voids) and rutting..... 67
Figure 7. 1 Aggregate gradation for the four mixes chosen for the first project 72
Figure 7. 3 Comparing actual fatigue cracking verses MEPDG prediction for Section ID 18890N..... 78
Figure 7. 4 Comparing actual fatigue cracking verses MEPDG prediction for SPS Site 117..... 79
Figure 7. 5 Comparing actual rutting verses MEPDG prediction for SPS-1 Site 117..... 79
Figure 7. 6 Comparing actual IRI verses MEPDG prediction for SPS-1 Site 117 80

Figure 7. 8 Fatigue cracking (%) at the end of 20 years for Section ID 18890N with varying asphalt content and in-situ air voids 81
Figure 7. 9 Fatigue cracking (%) at the end of 20 years for SPS-1 site 0117 with varying asphalt content and in-situ air voids..... 82
Figure 7. 10 Fatigue cracking for Section ID 29581N (Good, fair and poor) 83
Figure 7. 11 Fatigue cracking for Section ID 18890N (Good, fair and poor) 84
Figure 7. 12 Fatigue cracking for SPS-1 Site 117 (Good, fair and poor) 84
Figure 7. 13 Fatigue cracking Vs PWL (in-situ air voids) for mix 1..... 88
Figure 7. 14 Fatigue at 30 years Vs PWL (in-situ air voids) for mix 1 90
Figure 7. 15 Rutting Vs PWL (in-situ air voids) for mix 1..... 91
Figure 7. 16 IRI Vs PWL (in-situ air voids) for mix 1 92
Figure 7. 17 Effect of AV distribution on fatigue performance of sublots for AC=Opt-0.4. 94
Figure 7. 18 Effect of AV distribution on fatigue performance of sublots for AC=Opt-0.2. 95
Figure 7. 19 Effect of AV distribution on fatigue performance of sublots for AC=Opt. . 96
Figure 7. 20 Effect of AV distribution on fatigue performance of sublots for AC=Opt+0.2. 97
Figure 7. 21 Effect of AV distribution on fatigue performance of sublots for AC=Opt+0.4 98
Figure 7.22 Pavement structure used in simulation to study the effect of AC thickness variation on pavement performance..... 108
Figure 7.23 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt-0.5. 110
Figure 7. 24 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt..... 111
Figure 7. 25 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt+0.5..... 112
Figure 7. 26 Effect of AC thickness variation on rutting performance of sublots for Th=Opt-0.5. 113

Figure 7. 27 Effect of AC thickness variation on rutting performance of sublots for Th=Opt. 114
Figure 7. 28 Effect of AC thickness variation on rutting performance of sublots for Th=Opt+0.5 115
Figure 7. 29 Effect of AC thickness variation on IRI of sublots for Th=Opt-0.5 118
Figure 7. 30 Effect of AC thickness variation on IRI of sublots for Th=Opt. 119
Figure 7. 31 Effect of AC thickness variation on IRI of sublots for Th=Opt+0.5 120
Figure 8.1 Effect of production variability on risk for flexible pavements 131
Figure 8.2 Effect of measurement variability on risk for flexible pavements 132
Figure 8.3 Effect of sample size on risk for flexible pavements 133
Figure 8.4 Effect of measurement bias on risk for flexible pavements 134
Figure 8.5 CRI as a function of production and measurement variability 138
Figure 9.1 Flowchart showing feedback process for design 140
Figure 9.2 Sample MDOT QA data worksheet 144
Figure 9.3 Test of assumption of normality – Asphalt content QA data 145
Figure 9.4 Test of assumption of normality – Asphalt content QA data (Continued) .. 146
Figure 9.5 p and h – values from Chi-square goodness of fit for asphalt content QA data 147
Figure 9.6 p and h – values from Chi-square goodness of fit for density QA data 148
Figure 9.7 p and h – values from Chi-square goodness of fit for VMA QA data 148
Figure 9.8 p and h – values from Chi-square goodness of fit for Air voids QA data 149
Figure 9.9 Comparison of state and contractor measured mean VMA for different mixes 150
Figure 9.10 Comparison of state and contractor measured mean Air Voids for different mixes 150
Figure 9.11 Comparison of state and contractor measured mean Asphalt Content for different mixes 151
Figure 9.12 Comparison of state and contractor measured mean In-situ density for different mixes 151
Figure 9.13 Comparison of state and contractor measured Air Voids for different mixes – standard deviation, 5th & 95th percentile and sample size 153

Figure 9.14 Comparison of state and contractor measured Asphalt Content for different mixes – standard deviation, 5th & 95th percentile and sample size 154
Figure 9.15 Error in estimate of mean quality characteristic with 90% confidence interval from feedback process 157
Figure 9.16 Error in estimate of variability in quality characteristic with 90% confidence interval from feedback process 158
Figure 9.17 Error in estimate of payment risk with 90% confidence interval from feedback process 158
Figure 9.18 Width of 90% confidence interval in estimate of Q/C mean from feedback process..... 159
Figure 9.19 Width of 90% confidence interval in estimate of Q/C variability from feedback process 159
Figure 9.20 Width of 90% confidence interval in estimate of payment risk from feedback process..... 160
Figure 9.21 Average width of 90% confidence interval of error in estimate of mean . 164
Figure 9.22 Average width of 90% confidence interval of error in estimate of variability 165
Figure 9.23 Average width of 90% confidence interval of risk..... 165
Figure 12.1a Flowchart for Acceptance Testing Procedure for PCC Pavements..... 183
Figure 12.2b Flowchart for Acceptance Testing Procedure for PCC Pavements..... 184
Figure 13.1 Attributes most often used for QA of PCC (Hughes, 2005)..... 190
Figure 15. 1 PWL (compressive strength) Vs faulting – cement type 41..... 219
Figure 15. 2 PWL (compressive strength) Vs faulting – cement type 42..... 220
Figure 15. 3 PWL (compressive strength) Vs longitudinal cracking – cement type 41 . 220
Figure 15. 4 PWL (compressive strength) Vs transverse cracking – cement type 41 221
Figure 16. 1 Percent slab cracked for 50 sublots with varying compressive strength and slab thicknesses- Cases 1 through 6..... 227
Figure 16. 2 Percent slab cracked for 50 sublots with varying compressive strength and slab thicknesses- Cases 7 through 12..... 228
Figure 16. 3 IRI for 50 sublots with varying compressive strength and slab thicknesses- Cases 1 though 6 229

Figure 16. 4 IRI for 50 sublots with varying compressive strength and slab thicknesses-cases 7 through 12 230
Figure 16. 5 Faulting for 50 sublots with varying compressive strength and slab thicknesses- Cases 1 through 6 231
Figure 16. 6 Faulting for 50 sublots with varying compressive strength and slab thicknesses-cases 7 through 12 232
Figure 16. 7 Difference in critical stresses and strength gain in concrete pavement for construction at different times of the day 234
Figure 16. 8 Difference in critical stresses and strength gain in concrete pavement for construction at different times of the year 235
Figure 16. 9 Difference in critical stresses and strength gain in concrete pavement for construction at different times of the year 235
Figure 16. 10 Difference in critical stresses and strength gain in concrete pavement for different temperature of concrete at the time of construction. 236
Figure 17. 1 Effect of production variability on risk for rigid pavements 247
Figure 17. 2 Effect of measurement variability on risk for rigid pavements 248
Figure 17. 3 Effect of sample size on risk for rigid pavements 249
Figure 17. 4 Effect of measurement bias on risk for rigid pavements 250
Figure 18.1 Flowchart showing feedback process for design 255
Figure 18.2 Error in estimate of mean quality characteristic with 90% confidence interval from feedback process 259
Figure 18.3 Error in estimate of variability in quality characteristic with 90% confidence interval from feedback process 260
Figure 18.4 Error in estimate of pavement life (ESALs) with 90% confidence interval from feedback process 261
Figure 18.5 Width of 90% confidence interval in estimate of Q/C mean from feedback process 262
Figure 18.6 Width of 90% confidence interval in estimate of Q/C variability from feedback process 263
Figure 18.7 Width of 90% confidence interval in estimate of life (ESALs) from feedback process 264

Figure 18.8 Average width of 90% confidence interval of error in estimate of mean ... 268
Figure 18.10 Average width of 90% confidence interval of error in pavement life (in ESALs)..... 269

List of Tables

Table 2.1 Sample flexible pavement projects selected for data mining. 11
Table 2.2 Summary of test results collected from sampled HMA pavement projects. 12
Table 3.1 Estimated Percent Within Limits (Table 106-1 of Standard Specifications).... 18
Table 3.2 (Continued) Estimated Percent Within Limits (Table 106-1 of Standard Specifications)..... 19
Table 3.3 Quality Assurance Testing Tolerance..... 22
Table 4.1 Most commonly used quality characteristics for QC and QA on HMA pavements 25
Table 4.2 Less often used QC/QA quality characteristics for HMA pavements..... 26
Table 4.3 Quality Assurance Practices in North-East States 28
Table 4.4 Quality Assurance Practices in Southern States 29
Table 4.5 Quality Assurance practices in Mid-West States..... 30
Table 4.6 Quality Assurance Practices in Western States 32
Table 5. 1 NDT methods for measuring thickness 44
Table 5. 2 Pavement properties estimated by the Seismic Pavement Analyzer 48
Table 6. 1 Categories of data extracted from LTPP database..... 51
Table 7.1 Asphalt content measurements from actual projects 73
Table 7.2 In-situ air void measurements from actual projects..... 73
Table 7.3 Matrix for the first project selected for analysis..... 75
Table 7.4 Matrix for the second project selected for analysis 76
Table 7.5 Matrix for the third project selected for analysis..... 77
Table 7.6 Air voids and asphalt content levels for the three scenarios for analysis 82
Table 7.7 PWL and pay factor for density for the three scenarios. 85
Table 7.8 Asphalt content and air voids levels for the case scenarios..... 99
Table 7.9 Standard deviations for asphalt content and in-situ air voids for the three sets of scenarios 99
Table 7.10 PWL values and fatigue cracking for the 1st set of MEPDG runs for performance criteria..... 101
Table 7.11 PWL values and fatigue cracking for the 2nd set of MEPDG runs for performance criteria..... 103

Table 7.12 PWL values and fatigue cracking for the 3rd set of MEPDG runs for performance criteria 104
Table 7.13 PWL values and rutting for the 1st set of MEPDG runs for performance criteria 105
Table 7.14 PWL values and rutting for the 2nd set of MEPDG runs for performance criteria 106
Table 7.15 PWL values and rutting for the 3rd set of MEPDG runs for performance criteria 107
Table 7.17 Summary of results for effect of AC thickness variation on rutting performance 109
Table 7.18 Summary of results for effect of AC thickness variation on IRI 116
Table 7.19 Sample LTPP table with representative AC layer thickness variation 117
Table 8.1 Run Matrix for Risk Analysis of Flexible Pavements 128
Table 8.2 Price Adjustment for Concrete Thickness Deficiency..... 129
Table 8.3 Calculated CRI Values for the Scenarios Identified in the Run Matrix. 136
Table 8.4 ANOVA Table for CRI of all 81 Cases in the Run Matrix 137
Table 9.1 Width of 90% confidence interval of error in estimate of mean quality characteristic for different sample sizes..... 161
Table 9.2 Width of 90% confidence interval of error in estimate of variability in quality characteristic for different sample sizes..... 162
Table 9.3 Width of 90% confidence interval of error in estimate of risk in payment for the quality characteristic for different sample sizes..... 163
Table 11. 1 Sample rigid pavement projects selected for data mining. 180
Table 11. 2 Summary of test results collected from sampled concrete pavement projects. 180
Table 12.1 Price Adjustment for Concrete Thickness Deficiency (MDOT Table 602-2) 185
Table 12.2 Price Adjustment for Depth of Steel from Pavement Surface (MDOT Table 602-3A) 185
Table 12.3 Adjustment for Deviation of Depth of Steel from Design Range ((MDOT Table 602-3B) 186

Table 12.4 Sampling Rates, Re-Sampling Rates, & Rejection Limits for Concrete QA 186
Table 13.1 Most commonly used quality characteristics for QC and QA on PCC pavements (Hughes, 2005)..... 189
Table 13.2 Quality Assurance Practices in Southern States versus Michigan..... 192
Table 13.3 Quality Assurance Practices in Western States versus Michigan..... 193
Table 13.4 QA Practices in Some of the States for PCC Pavements..... 194
Table 14.1 Summary of Previous Impact Echo Concrete Pavement Thickness Studies 203
Table 14.2 Advantages and Limitations Various Gamma Radiometry Techniques..... 206
Table 15. 1 Data extracted from LTTP database. 218
Table 16. 1 Compressive strength and slab thicknesses for MEPDG runs for rigid pavements 224
Table 16. 2 Details of cracking simulation runs with summary cracking results..... 225
Table 17. 1 Run matrix for Risk Analysis of Rigid Pavements..... 244
Table 17. 2 Price Adjustment for Concrete Thickness Deficiency..... 244
Table 17. 3 Calculated CRI Values for the Scenarios Identified in the Run Matrix. 251
Table 17.4 ANOVA table for CRI for all 81 cases in the run matrix 253
Table 18.1 Width of 90% confidence interval of error in estimate of mean quality characteristic (strength, in psi) for different sample sizes 265
Table 18.2 Width of 90% confidence interval of error in estimate of variability in quality characteristic (strength, in psi) for different sample sizes 266
Table 18.3 Width of 90% confidence interval of error in estimate of pavement life (in 100,000 ESALs) for the quality characteristic for different sample sizes 267

Introduction

CHAPTER 1: Introduction and Background

1.1 INTRODUCTION

State Highway Agencies (SHAs) have long realized the importance of quality assurance (QA) to ensure longer pavement service life. Construction QA programs are intended to ensure that the quality of the materials and construction in highway projects is satisfactory. Failure to conform to either material or construction specifications can result in the premature failure of pavements.

Currently, the strategies and practices used by state highway agencies to ensure quality employ a wide variety of QA approaches to meet the FHWA's Quality Assurance Procedures for Construction regulation (*1,2*). In many SHA specifications, quality assurance procedures require contractors to perform quality control (QC) tests and the state to perform acceptance tests. Frequently, these tests measure the same engineering properties. However, it is known that these results vary, even when taken from the same population. The effect of this variability in terms of the difference between as-constructed and design values on pavement performance needs to be assessed.

In general, materials/construction-related distresses refer to pavement failures that are a direct result of the properties of the materials, construction quality and their interactions with the environment to which they are exposed. In this sense, these failures are differentiated from others that might be associated with inadequate design for the traffic and environmental loading or the use of improper practices during pavement construction (*3*).

For flexible pavements, test methods used for in-place quality control and acceptance of individual pavement layers and of new and rehabilitated pavement systems have not changed much in past decades. Such quality control and acceptance operations typically rely on nuclear density measurements or the results of testing conducted on pavement cores. Roughness measurements are often used to confirm that the newly constructed pavement has an adequate initial smoothness. More recently, nondestructive testing (NDT) methods, including lasers, ground-penetrating radar, falling weight

deflectometers, cone penetrometers, and infrared and seismic technologies, have been significantly improved and have shown potential for use in the quality control and acceptance of flexible pavement construction. Furthermore, the new Mechanistic-Empirical Pavement Design Guide (ME-PDG) will use pavement layer stiffness as a key material property. This will lead to increased measurement of layer moduli by owner agencies, an activity that is not at present a typical component in the acceptance of a completed project (*4*).

For rigid pavements, many methods exist for the nondestructive testing of concrete materials (*5-8*). Most of these testing techniques are highly refined and many have been standardized. However, these techniques are directed at strength determination rather than the identification of materials-related distress (MRD is normally a durability issue rather than a strength issue); the test methods that are directed at condition assessment as well as MRD will be considered.

This research project will assess pavement acceptance criteria, investigate the application of new non-destructive tests for measuring the quality of flexible and rigid pavements, and develop a process of quantifying as-constructed material and structural properties and relating their variability to pavement performance for different conditions including new, rehabilitation and preservation projects.

1.2 PROBLEM STATEMENT

The acceptance criteria for concrete (strength, air content, slump) and HMA (density, air voids, and asphalt content) have been used for decades and, therefore, need to be updated. The current acceptance criteria for pavements need to take into account other factors causing pavements to deteriorate. To be reliable, any acceptance criteria should relate to the performance experienced by the pavement structure (performance related specifications). Over time, numerous department investigations and research studies regarding distress initiation have found other factors that have contributed to the pavement's performance, both good and poor. These factors include adequate pavement support, varying material properties/characteristics, air-void distribution and sizes in concrete, excessive rates of concrete and HMA permeability, post-material (concrete) interaction chemistry, and asphalt film thickness.

Furthermore, when a pavement type and its cross section are determined during design, pre-established input values are used for material properties to counter the effects from truck loading. These same values are assumed to exist after construction occurs, if specifications are met. However, historical field sampling and testing for investigation projects after construction have found a wide variance in pavement material properties and the pavement's base/subbase support characteristics. Only a small portion of this variance can be attributed to natural aging and environmental effects. The remaining variance in pavement material properties occurs during construction that can cause pavement service life to vary within construction project limits.

The intent of this research project is to assess existing pavement acceptance criteria, conduct a feasibility study to consider the need of using new criteria that relate to pavement performance including the use of non-destructive testing, and develop a process to quantify as-constructed material and structural pavement properties and relate them to pavement service life.

1.3 OBJECTIVES

The objectives of this study are:

1. To conduct a feasibility study to consider the need of using new acceptance criteria/testing for pavements
2. To develop a feedback process for implementation during pavement design
3. To identify the pros/cons together with the impacts of using new criteria
4. To develop a plan of action for implementation

1.4 RESEARCH APPROACH

To achieve the objectives identified above, the following steps were taken.

(1) Collection and analysis of MDOT construction and performance data: Since the objective of this project is to assess the quality assurance program being used by the Michigan Department of Transportation, the most direct approach would be to gather construction data and specific details about projects, to assess if the QA program helped in ensuring that pavement construction quality was maintained up to the desired level and to verify if that did actually translate into a better performing pavement.

a. Determine if QA led to identification of projects on which quality control was good, fair or poor.

b. Assess if projects with poor quality control were corrected through suitable means and that the contractor was sufficiently penalized to recover the loss that MDOT would incur when improving the quality of the pavement.

c. Determine if those projects which were assessed to be of good quality have performed as well as or better than they were designed for.

This approach was meant to build on MDOT's on-going data gathering efforts to include collecting QC/QA data, design and construction data (including physical inventory data, material properties from in-situ and laboratory tests) from past projects as well as performance data (DI, RQI or IRI, rut depths and faulting etc.). This information was extracted from MDOT PMS database, records, and data files for pavement projects (both new construction and rehabilitation/preservation) that are longer than 1-mile. The information was extracted for the period between 1992 and 2001 (pavement performance data are not available for projects completed prior to 1992 and it is too early to analyze the pavement distress data for projects completed after the year 2001). The information was expected to include, if available, project identification (control section, project number, and BMP and EMP), project completion and so forth. It was decided that an effort will be made to collect data from a range of projects with varying levels of compliance in the respective specifications. However, an intensive search for MDOT's quality assurance and quality control data led to the finding that most of such data remain unaccounted for. Therefore, alternative approaches to relate the quality characteristics used in a QA program to actual pavement performance were indentified and are described in the following points.

(2) Analysis of LTPP data: It was decided that analysis of MDOT data as described in item (1) above would be supplemented with similar analysis on data extracted from the LTPP database. Chapter 6 and chapter 15 provide details of this effort for flexible and rigid pavements, respectively. Analysis performed on such data revealed that it is very difficult to reach firm conclusions using this data. This happens primarily because LTPP data was collected from real projects which have been constructed for a wide range of traffic using varied construction materials and in very different climates by different state agencies that differ in their design procedures and implementation details. This lead to too many factors varying among the projects and therefore it is almost impossible to isolate the effect of specific variables on pavement performance, although in some cases,

some general understandings can be derived regarding how different quality characteristics may affect pavement performance.

(3) Mechanistic-empirical analysis: Since empirical data analysis has shortcomings, mechanistic-empirical analysis is required to establish relationships between different candidate quality characteristics and pavement performance. The Mechanistic-Empirical Pavement Design Guide (MEPDG) is the most appropriate tool for such analysis. However, MEPDG does not include all the construction related inputs, particularly for rigid pavements (e.g., the time of the day when pouring of concrete is done). These variables are important for the prediction of early age cracking and/or built-in curling in concrete pavements. Therefore, an alternate software such as HIPERPAV can be used to study such construction related issues.

(4) Synthesis of empirical and mechanistic empirical analyses: This is the most important step in this project. Firstly, it requires preparing an exhaustive list of quality characteristics which should be considered for inclusion in the QA program. The different sources for preparing this list have been enumerated below.

a) Quality characteristics being used in other states' QA programs: Different states use varying combinations of quality characteristics. Some of these quality characteristics are used in determining payment to be made to the contractor for any project, while others are used merely to provide feedback for proper construction.
b) Quality characteristics being used in other states' QC program: Any of the quality characteristics that are used in a QC program, i.e., the variables that are monitored by the contractor and not used in the QA program should also be considered.
c) Quality characteristics used in the Mechanistic-Empirical Design Guide Software: The MEPDG software predicts pavement performance from the material, pavement structure, construction, traffic and environmental variables. The models used in the software are the result of studies carried out by many research teams after extensive testing and analysis to relate those variables to performance. Therefore, those variables or quality characteristics that are within the control of the contractor and can be tested at the time of construction should also be included in the list.
d) Quality characteristics studied in other research projects which have been shown to have impact on performance.

The second step is to shortlist those candidate QA variables that are known to affect pavement performance. These relationships can be established through steps described in items (1) through (3) above.

The third and last step is to identify those variables that should be included into the Michigan QA program. The criteria for including those variables are that:

a. They affect pavement performance either directly or in conjunction with other variables.
b. They need to be tested individually and cannot be estimated or calculated from other significant QA variables already being used in the QA program. For example there is a strong correlation between compressive strength and flexural strength of concrete.
c. It is feasible to test for them within a reasonable amount of time during the construction.
d. The testing for these candidate QA variables does not require very specialized or costly equipment.
e. It is possible for the contractor to control those variables through sound construction practices and tight quality control.

(5) Analyzing MDOT End-Result Specifications: Apart from how the individual quality characteristics affect pavement performance, it is important to understand how these quality characteristics should be used in the specifications to encourage better quality and reduce the risk of overpayment or underpayment. Monte-Carlo based simulations were developed to achieve this goal.

(6) Plan of Action: Using the conclusions and recommendations reached through steps (1) to (5) above, a plan of action was developed so that MDOT can implement the recommendations in its quality assurance and design process. An important part of this plan is a feedback process for design.

1.5 ORGANIZATION OF THE REPORT

This project addresses the quality assurance program for flexible and rigid pavements. Chapter 1 provides introduction, background, objectives and overall organization of the entire report. The rest of the report has been divided into two parts: Part I for flexible pavements and Part II for rigid pavements.

Part I - Flexible Pavements: This part begins with chapter 2.

Chapter 2 reports the effort that was made to collect actual quality assurance program data from Michigan pavement projects. An effort was made to collect data from MDOT projects with reasonably long performance history to be able to relate QA quality characteristics to pavement performance.

Chapter 3 contains the summary of MDOT's current QA program and the design process. Chapter 4 presents a broad picture of the types of specifications being used in QA programs across the United States with specific details from several states to be able to understand the current state of practice.

Chapter 5 documents the conclusions derived in other research studies which are of direct relevance to the objectives in this project. This knowledge base has been used in deriving conclusions regarding the importance of various quality characteristics from the point of view of pavement performance and in making recommendations.

Chapter 6 documents the details of analysis that was performed using data extracted from the Long Term Pavement Performance (LTPP) program. This analysis was intended to empirically explore relationships between different quality characteristics and pavement performance. Since it was found that firm conclusions are difficult to be drawn from empirical analysis, mechanistic-empirical analysis using MEPDG was conducted, and the details of this analysis have been presented in chapter 7.

Chapter 8 presents the details of a simulation which was developed to analyze the current MDOT QA program for flexible pavements. This analysis shows how to identify the strengths and weaknesses of a QA program and thereby improve it by suitable changes in the specifications. This was followed by MDOT construction data analysis and another simulation program to design an optimal feedback process for design presented in chapter 9. Chapter 10 presents the conclusions and recommendations for MDOT QA program for flexible pavements.

Part II presents the analysis performed and conclusions and recommendations for the rigid pavement QA program. The overall organization of this part is identical to that presented for flexible pavements above. Therefore, a similar outline is used for Part II.

Chapter 11: MDOT QA Data Collection Effort

Chapter 12: Review of MDOT's Current PCC QA Program and Design Process

Chapter 13: Survey of QA Programs in USA

Chapter 14: QA Variables and Rigid Pavement Performance-A Review

Chapter 15: Empirical Analysis

Chapter 16: Mechanistic-Empirical Analysis

Chapter 17: ERS Risk Analysis Using Simulation

Chapter 18: Feedback Process

Chapter 19: Conclusions and Recommendation for PCC QA Program

Part I: Flexible Pavements

CHAPTER 2: MDOT Quality Assurance Data for Flexible Pavements

Certain projects were identified to get a sample of construction and materials test data collected on highway construction projects. Although the objective of this project is to evaluate quality *assurance* procedures, data from quality *control* procedures were also collected as part of the data collection effort. Different quality assurance procedures were followed at different times in the history of highway construction in Michigan. However, if suitable data were collected under the quality control procedures, they could be used with the current quality assurance procedure (or a slightly modified version of it like a shadow specification) to assess how the construction would be rated according to those specifications. This could then be related to the actual performance of those pavements. In other words, any material test data, whether under QC or QA, can be evaluated for a possible relationship with pavement performance. If any of such test results show strong correlation to performance, then they can be considered for quality assurance.

2.1 SELECTION OF SAMPLE PROJECTS

Two sources of data were identified:

(1) A list of projects for which performance data had already been processed by MDOT was used to identify the first set of sample projects for data collection.

(2) Actual construction documents collected from the record center of Michigan Department of Transportation and microfilms stored in the MDOT Construction & Technology (C&T) office.

Table 2.1 lists the projects which were selected as sample projects. All the data available at MDOT related to these projects were collected. This was done to understand the types of data that will be available for analysis at the time of the project.

Table 2.1 Sample flexible pavement projects selected for data mining.

REGION	ROUTE	CS	JN	BMP	EMP	LET	LOCATION
University	M-99 NB	33011	00434	4.260	5.178	12/21/1976	Victor Avenue to Moores River Drive
University	M-99 SB	33011	00434	4.310	5.210	12/21/1976	Victor Avenue to Moores River Drive
Grand	M-44 NB	41051	25745/2574	4.287	5.155	6/6/1990	N of I-96 to Windcrest Court
Grand	M-44 SB	41051	25745/2574	4.241	5.155	6/6/1990	N of I-96 to Windcrest Court
University	I-96 EB	33083	29581	0.000	2.348	2/4/1994	Ingham/Eaton Co. Line to Richard Road
University	I-96 WB	33083	29581	0.041	2.371	2/4/1994	Ingham/Eaton Co. Line to Richard Road
University	M-99 NB	23092	10729	2.085	9.229	1/18/1978	N of Petrieville Highway to N of Holt Road
University	US-27 NB	19033	20046	8.526	12.775	2/10/1993	Price Road to S of Wildcat Road
Superior	US-2	36022	37563	0.203	0.717	3/6/1998	River Avenue to N of Cayuga Street
Bay	M-53	44031	36021	1.592	2.835	3/3/1999	S of Water Street to N of Kingsbrook Dr.

Key: CS: Control Section
JN: Job Number
BMP: Beginning Mile Point
EMP: Ending Mile Point

2.2 SAMPLE DATA GATHERING

All the boxes and microfilms (if available) were searched for data related to the sample projects listed in Table 2.1. Table 2.2 summarizes the test data available for asphalt pavements. Appendix A gives a more detailed list of the types of data available.

For asphalt pavements, we have observed that very few tests were performed during construction for older projects; i.e., prior to 1990. Two of the sampled projects constructed in 1993 and 1994 have limited amount of core density and VMA data in the case of HMA pavements. Other quality characteristics used in contemporary QA programs are scarce. From all the projects selected as sample projects no quality characteristic data was obtained.

Considering the fact that there is scarcity of data required for determining the influence of various quality characteristics on pavement performance, it was deemed necessary to consider other alternatives rather than just relying on Michigan data. The alternatives are: (1) to collect relevant data from the Long Term Pavement Performance (LTPP) database and (2) to conduct mechanistic-empirical analysis using MEPDG and other performance models.

Table 2.2 Summary of test results collected from sampled HMA pavement projects.

Material Property	Sample			Material		Source					
	Truck	Mat	Virgin Material	Aggregate	Binder Also Mod. Polymer	Plant Mix	Top Course	Leveling Course	Base Course	Shoulder	Sub-Base
Actual Depth Measurement									2		2
Air Voids		1		2		2					
Asphalt Content						1					
Chert			1	1							
Crushed Material	1	3	4	4		1			2		
Density (Marshal)						2	1	1	1		
Density (Theoretical Maximum)						2	1	1	1		
Density (Average Core)		1				1					
Ductility		1	4		5	5					
Ductility of Residue (TFO)			4		5	5					
Elastic Recovery					1	1					
Fineness Modulus			1	1							
Flash Point		1	2		5	4					
Gradation	5	5	5	5	4	5	5	4	5	1	
Hard Absorbent Particles			1	1							
Incrusted Particles Less than 1/3 Area	1	1	2	2							
Incrusted Particles More than 1/3 Area	1	1	2	2	2	2					
Loss on Heating (TFO)		1	3		2	2					
Penetration (Original & Recovered)	4	1	4	1	5	5	4	5	3	1	
Penetration of Residue (TFO)		1	4		5	5					
Softened Particle	1	1	2	2							
Softening Point		1	1		3	2					
Solubility in Trichloroethylene		1	4		5	5					
Specific Gravity		1	2		3	3					
Spot Test			4		5	5					
Temperature of Mix	2		1	1		3	4	3	1	1	
Thin or Elongated Piece	1	1	1	1							
V. M. A.		1		2		2					
Viscosity (Also TFO)			4		5	5					
Wear Index			1	2		2	2				

CHAPTER 3: Review of MDOT QA Program and Design Process for HMA Pavements

3.1 Michigan DOT Quality Assurance Program

The Michigan Department of Transportation currently uses the following quality characteristics in their Quality Assurance (QA) program for Hot Mix Asphalt mixtures when calculating the payment to be made to the Contractor:

1. Air Voids (AV)
2. Voids in Mineral Aggregates (VMA)
3. Asphalt Binder Content (AC)
4. In-Place Density

The testing for these parameters, evaluated by the Engineer under the department's Quality Assurance Acceptance program, for Hot Mix Asphalt materials is performed on a Lot-by-Lot basis. Prior to beginning the work, individual Lot size is agreed upon by the Engineer and each Lot is divided into Sublots of approximately equal size and smaller than 1,000 tons.

If the total tonnage of a specific mixture does not exceed 5,000 tons, the total quantity of that mixture will be considered as a Lot and will be divided into a minimum of three and up to a maximum of seven approximately equal Sublots for testing and acceptance.

Quality Assurance & dispute resolution samples are taken within each Sublot through a random process managed by the Engineer. Each sample, weighing approximately 20,000 grams, is assigned an identifier by the Engineer and delivered to the testing facility, as specified in the HMA quality assurance plan, where one is tested and one is retained for possible appeal testing.

Within four calendar days after sampling, the following tests are conducted by the Engineer:

- Maximum Specific Gravity, G_{mm} , (MTM 314)
- Bulk Compacted Density, N_{Max} , (AASHTO TP 4-97)
- Air Voids, N_{ini} , N_{des} , N_{max} , (AASHTO PP28-97) (*For information only*)
- Voids in Mineral Aggregate, VMA , (AASHTO PP28-97)
- Voids Filled with Asphalt, VFA , (AASHTO PP28-97) (*For information only*)
- Ratio of Fines to Effective Asphalt Binder, (Passing #200 / P_{be})

- Composition of the Mixture:
  - **Method 1**: Asphalt binder content based on calculated value using sublot maximum specific gravity (Gmm) and current job mix formula (JMF) effective specific gravity (Gse); Gradation (ASTM C 136, and ASTM C 117) and crushed particle content (MTM 117) from extracted (AASHTO T 164) or incinerated aggregate (MTM 319).
  - **Method 2**: Asphalt binder content based on vacuum extraction (MTM 325 and the checklist for HMA mixture analysis vacuum extraction of the HMA Production Manual); Gradation (ASTM C 136, C117) and crushed particle content (MTM 117) based on extracted aggregate (AASHTO T 164).

After completion and final rolling of each Sublot four core samples are taken by the Contractor from locations specified and marked by the Engineer. It is cautioned that these random cores should be taken at a time which is independent of paving operations. Each core, approximately 6 inches in diameter, needs to be measured for thickness at the time of extraction. Any disqualified core based on minimum thickness criteria is discarded and the Engineer selects a new core location.

A Pavement In-Place Density acceptance test is completed by the Engineer in accordance with MTM 315, (Michigan Test Method for Bulk Specific Gravity and Density of Compacted HMA Mixtures using Saturated Surface-Dry Specimens) within four calendar days of core extraction. If more than fifty percent of cores in a Lot are disqualified, production has to stop and will not continue until the Engineer approves the conformation of contract application and paving operation.

The Engineer's test results for the compacted HMA will be used as a basis of acceptance and payment. HMA pay items will be paid for according to contract prices for completed items on a Lot-by-Lot basis.

The Engineer will calculate percent within limits, pay factor, and payment in accordance with the procedures described in the following sub-sections.

3.1.1 Percent Within Limits

The percentage of each Lot within the specification limits established for each Quality Assurance parameter is determined as follows:

a) ***Arithmetic mean*** and ***standard deviation*** of the test results are computed:

$$\overline{X} = \frac{\sum X_i}{n}$$

$$S = \sqrt{\frac{\sum (X_i - \overline{X}^2}{n-1}}$$

where:

$\overline{X}$ = Arithmetic mean of test results

X_i = Test results

n = Number of test results

S = Standard deviation of the test results

b) Follow the flowchart below

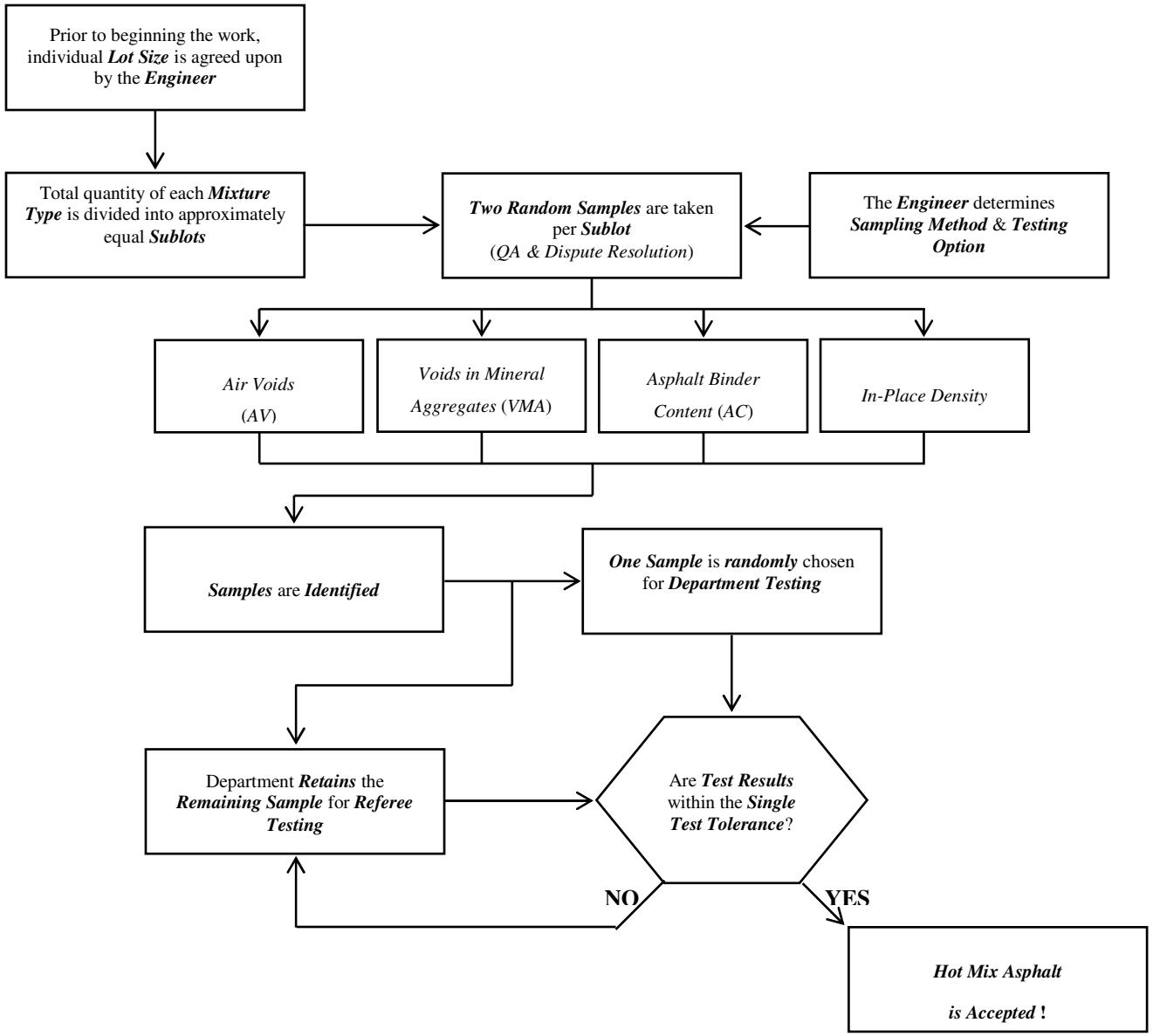


Figure 3.1 Flowchart for Acceptance Testing Procedure for HMA Pavements

c) ***Upper Quality Index***, QU, and ***Lower Quality Index***, QL, are determined using Upper Specifications Limit and Lower Specifications Limit. The results should be rounded to the nearest 0.05 according to ASTM E 29, Section 6.6. If any of Upper Specifications Limit or Lower Specifications Limit is not specified, the Upper Percent within Limits or Lower Specifications Limits will be 100:

$$Q_U = \frac{USL - \overline{X}}{S}$$

$$Q_L = \frac{\overline{X} - LSL}{S}$$

where:

$\overline{X}$ = Arithmetic mean of test results
USL = Upper Specifications Limit
LSL = Lower Specifications Limit
S = Standard deviation of the test results

d) ***Percentage of materials*** within the Upper Specification Limit, PU, and Lower Specification Limits, PL, is estimated according to Tables 3.1 and 3.2 (Table 106-1 of the Standard Specifications). For this purpose one should enter the table with QU or OL and follow the column which is appropriate to the total number of tests, n.

e) ***Quality Level*** stated as Percent within Limits, PWL, is calculated using the following formula. Note that all values of Percent within Limits, PWL, are Percents.

$$PWL = (P_U + P_L) - 100$$

If only an Upper Specification Limit or Lower Specification Limit applies, then PU or PL, respectively, is the Percent within Limits, PWL.

f) ***Percent Defective***, PD, if required to calculate pay adjustments, can be determined using

$$PD = 100 - PWL$$

Table 3.1 Estimated Percent Within Limits (Table 106-1 of Standard Specifications)

Q_U or Q_L	Percent Within Limits for Given Sample Sizes (a)									
	n=1 (b)	n=2 (b)	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10
0.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00
0.05	55.10	51.68	51.38	51.67	51.78	51.84	51.87	51.89	51.91	51.92
0.10	60.20	53.36	52.76	53.33	53.56	53.67	53.74	53.78	53.82	53.84
0.15	65.31	55.03	54.15	55.00	55.33	55.50	55.60	55.67	55.71	55.75
0.20	70.41	56.71	56.54	56.67	57.10	57.32	57.46	57.54	57.60	57.65
0.25	75.51	58.39	56.95	58.33	58.87	59.14	59.30	59.41	59.48	59.53
0.30	80.61	60.07	58.37	60.00	60.63	60.94	61.13	61.25	61.34	61.40
0.35	85.71	61.74	59.80	61.67	62.38	62.73	62.94	63.08	63.18	63.25
0.40	90.82	63.42	61.26	63.33	64.12	64.51	64.74	64.89	65.00	65.07
0.45	95.92	65.10	62.74	65.00	65.84	66.27	66.51	66.67	66.79	66.87
0.50	100.00	66.78	64.25	66.67	67.56	68.00	68.26	68.43	68.55	68.63
0.55	100.00	68.46	65.80	68.33	69.26	69.72	69.99	70.16	70.28	70.36
0.60	100.00	70.13	67.39	70.00	70.95	71.41	71.68	71.85	71.97	72.06
0.65	100.00	71.81	69.03	71.67	72.61	73.08	73.34	73.51	73.63	73.72
0.70	100.00	72.49	70.73	73.33	74.26	74.71	74.97	75.14	75.25	75.33
0.75	100.00	75.17	72.50	75.00	75.89	76.32	76.56	76.72	76.83	76.90
0.80	100.00	76.85	74.36	76.67	77.49	77.89	78.12	78.26	78.36	78.43
0.85	100.00	78.52	76.33	78.33	79.07	79.43	79.63	79.76	79.84	79.90
0.90	100.00	80.20	78.45	80.00	80.62	80.93	81.10	81.21	81.28	81.33
0.95	100.00	81.88	80.75	81.67	82.14	82.39	82.52	82.61	82.67	82.71
1.00	100.00	83.56	83.33	83.33	83.64	83.80	83.90	83.96	84.00	84.03
1.05	100.00	85.23	86.37	85.00	85.09	85.18	85.23	85.26	85.28	85.29
1.10	100.00	86.91	90.16	86.67	86.52	86.50	86.51	86.51	86.5	86.50
1.15	100.00	88.59	97.13	88.33	87.90	87.78	87.73	87.70	87.68	87.66
1.20	100.00	90.27	100.00	90.00	89.24	89.01	88.90	88.83	88.79	88.76

Table 3.2 (Continued) Estimated Percent Within Limits (Table 106-1 of Standard Specifications)

Q_U or Q_L	Percent Within Limits for Given Sample Sizes (AASHTO R 9 Appendix C) (a)									
	n=1 (b)	n=2 (b)	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10
1.25	100.00	91.95	100.00	91.67	90.54	90.19	90.02	89.91	89.85	89.79
1.30	100.00	93.62	100.00	93.33	91.79	91.31	91.07	90.94	90.84	90.78
1.35	100.00	95.30	100.00	95.00	92.98	92.37	92.08	91.90	91.78	91.70
1.40	100.00	96.98	100.00	96.67	94.12	93.37	93.02	92.81	92.67	92.56
1.45	100.00	98.66	100.00	98.33	95.19	94.32	93.90	93.65	93.49	93.37
1.50	100.00	100.00	100.00	100.00	96.20	95.19	94.72	94.44	94.26	94.13
1.55	100.00	100.00	100.00	100.00	97.13	96.00	95.48	95.17	94.97	94.82
1.60	100.00	100.00	100.00	100.00	97.97	96.75	96.17	95.84	95.62	95.46
1.65	100.00	100.00	100.00	100.00	98.72	97.42	96.81	96.45	96.22	96.05
1.70	100.00	100.00	100.00	100.00	99.34	98.02	97.38	97.01	96.76	96.59
1.75	100.00	100.00	100.00	100.00	99.81	98.55	97.89	97.51	97.25	97.07
1.80	100.00	100.00	100.00	100.00	100.00	98.99	98.35	97.96	97.70	97.51
1.85	100.00	100.00	100.00	100.00	100.00	99.36	98.74	98.35	98.09	97.91
1.90	100.00	100.00	100.00	100.00	100.00	99.65	99.07	98.69	98.44	98.25
1.95	100.00	100.00	100.00	100.00	100.00	99.85	99.35	98.99	98.74	98.56
2.00	100.00	100.00	100.00	100.00	100.00	99.97	99.57	99.24	99.00	98.83
2.05	100.00	100.00	100.00	100.00	100.00	100.00	99.74	99.45	99.23	99.06
2.10	100.00	100.00	100.00	100.00	100.00	100.00	99.86	99.61	99.41	99.26
2.15	100.00	100.00	100.00	100.00	100.00	100.00	99.94	9.74	99.57	99.42
2.20	100.00	100.00	100.00	100.00	100.00	100.00	99.99	99.84	99.69	99.56
2.25	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.91	99.79	99.68
2.30	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.96	99.86	99.77
2.35	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.98	99.92	99.84
2.40	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.95	99.89
2.45	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.98	99.93
2.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.99	99.96
2.55	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.98
2.60	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.99
2.65	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

3.1.2 Pay Factors for Quality Acceptance Items

3.1.2.1 Pay Factor for Air Voids – PF_{AV}:

A. If PWL for Air Voids (PWL_{AV}) is between 100 and 70, the following formula is used to determine PF_{AV}. The value of PF_{AV} is rounded off to two decimal places.

$$PF_{AV} = 55 + (0.5 \times PWL)$$

B. If PWL for Air Voids is between 70 and 50 inclusive, the following formula is used to determine PF_{AV}. The value of PF_{AV} is rounded off to two decimal places.

$$PF_{AV} = 37.5 + (0.75 \times PWL)$$

C. If PWL for Air Voids is less than 50, the Engineer may elect to do one of the following:

(1) Require removal and replacement of the entire Lot with new QA sampling and testing and repeat the evaluation procedure.
(2) Allow the Lot to remain in place and apply an Overall Lot Pay Factor of 50.00.
(3) Allow submittal of a corrective action plan for the Engineer's approval. The corrective action plan may include removal and replacement of one or more sublots. If one or more sublots are replaced, the Sublot(s) will be retested and the Overall Lot Pay Factor will be recalculated according to this special provision. If the Engineer does not approve the plan for corrective action, subsections (1) or (2) above will be applied.

3.1.2.2 Pay Factor for Binder Content (PF_{BINDER})

A. If PWL for Binder Content (PWL_{BINDER}) is between 100 and 70, the following formula is used to determine PF_{BINDER}. The value of PF_{BINDER} is rounded off to two decimal places.

$$PF_{BINDER} = 55 + (0.5 \times PWL)$$

B. If PWL for Binder Content is between 70 and 50 inclusive, the following formula is used to determine PF_{BINDER}. The value of PF_{BINDER} is rounded off to two decimal places.

$$PF_{BINDER} = 37.5 + (0.75 \times PWL)$$

C. If PWL for Binder Content is less than 50, the Engineer may elect to take one of the actions specified in description for PF_{BINDER} above.

3.1.2.3 Pay Factor for Voids in Mineral Aggregates – PF_{VMA}:

A. If PWL for VMA (PWL_{VMA}) is between 100 and 70, the following formula is used to determine PF_{VMA}. The value of PF_{VMA} is rounded off to two decimal places.

$$PF_{VMA} = 55 + (0.5 \times PWL)$$

B. If PWL for VMA is between 70 and 50 inclusive, the following formula is used to determine PF_{VMA}. The value of PF_{VMA} is rounded off to two decimal places.

$$PF_{VMA} = 37.5 + (0.75 \times PWL)$$

C. If PWL for VMA is less than 50, the Engineer may elect to take one of the actions specified in description for PF_{AV} above.

3.1.2.4 Pay Factor for In-Place Density – PF_D:

A. If PWL for In-Place Density (PWL_D) is between 100 and 70, the following formula is used to determine PF_D. The value of PF_D is rounded off to two decimal places.

$$PF_D = 55 + (0.5 \times PWL)$$

B. If PWL for In-Place Density is between 70 and 50 inclusive, the following formula is used to determine PF_D. The value of PF_D is rounded off to two decimal places.

$$PF_D = 37.5 + (0.75 \times PWL)$$

C. If PWL for In-Place Density is less than 50; the Engineer may elect to take one of the actions specified in description for PF_D above.

3.1.2.5 Overall Lot Pay Factor

$$OLPF = (0.20 \times PF_{BINDER}) + (0.20 \times PF_{VMA}) + (0.60 \times PF_D)$$

3.1.3 Payment

Payment for HMA Pay Items are based on the Contract Prices for the completed items of work as adjusted according to Special Provisions for Hot Mix Asphalt Percent within Limits (PWL). Adjusted Payment for HMA Type is calculated on a Lot-By-Lot basis. The Overall Lot Pay Factor, OLPF, is used to determine the Lot Pay Adjustment as follows:

$$\text{Lot Payment Adjustment} = \frac{(OLPF - 100) \times (\text{Contract Unit Price}) \times (\text{Lot Quantity})}{100}$$

Table 3.3 Quality Assurance Testing Tolerance

Quality Characteristic	Initial Production Lot Single Test Tolerance
G_{ms}	± 0.019
G_{mb}	± 0.020
Air Voids	± 1.00 %
VMA	± 1.20 %

3.2 Michigan DOT Pavement Design Process

An effective pavement design is highly dependent upon performing an adequate investigation of the existing pavement structure. Therefore, prior to construction/reconstruction of a pavement some investigations such as:

- Reviewing As-Built Plans
- Reviewing and Analyzing Existing Pavement Distress Condition
- Determining Causes of Pavement Surface Distresses
- Evaluating Pavement Ride Quality
- Reviewing Pavement Remaining Service Life
- Evaluating Drainage System
- Evaluating Subgrade

should be conducted. With no doubt, a comprehensive investigation of the pavement structure not only ensures the Engineer is employing the proper reconstruction or rehabilitation strategies, but also aids the Designer in selection of appropriate input values for pavement design.

M-DOT uses the pavement design methodology recommended by the American Association of State Highway and Transportation Officials, AASHTO:

- 1993 AASHTO Guide for Design of Pavement Structures
- AASHTO Pavement Design Software "DARWin"

The followings summarize "Typical Values" recommended for "Design Inputs":

- New / Reconstruction Life Design Life …… 20
- Accumulated ESAL's Years ……. 20
- Initial Serviceability …….. 4.5
- Terminal Serviceability ……. 2.5
- Reliability Level ……... 95%
- Overall Standard Deviation …….. 0.49

- Structural Coefficient:
  - HMA Top & Leveling Course ……..... 0.42
  - HMA Base Course …….. 0.36
  - Rubblized Concrete ……. 0.18
  - Crush & Shaped HMA …… 0.20
  - Aggregate Base ……... 0.14
  - Sand Subbase …….. 0.10
  - ASCRL & Stabilized Base …….. 0.30

- Elastic Modulus:
  - HMA Top & Leveling Course ……... 390,000 – 410,000 psi
  - HMA Base Course ……. 275,000 – 320 000 psi
  - Rubblized Concrete …….. 45,000 – 55,000 psi
  - Crush & Shaped HMA …….. 100,000 – 150,000 psi
  - Aggregate Base …….. 30,000 psi
  - Sand Subbase ……..... 13,500 psi
  - ASCRL & HMA Stabilized Base …….. 160,000 psi

- Roadbed Soil Resilient Modulus: Use "Falling Weight Deflectometer" (FWD) data when possible, otherwise a value is chosen based on the predominant subgrade soil type. A correlation can be made between "Soil Type" and "Resilient Modulus".

- Drainage Coefficient: (Refer to Table 2.4, Page II-25, AASHTO Guide for Design of Pavement Structures)

  - HMA Top & Leveling Course …… 1
  - HMA Base Course …….. 1
  - Rubblized Concrete ……. 1
  - Crush & Shaped HMA …… 1
  - Aggregate Base ……... 1
  - Sand Subbase …….. 1

- Stage Construction …….....1

Chapter 4: HMA Quality Assurance Programs in the United States

A detailed study of the current quality assurance practices of state and federal departments of transportation with regard to highway materials and construction was carried out. The motivation behind this exercise was to assess how the MDOT QA program compares to those being used by other states and agencies and to identify the items that can be considered for possible inclusion into the MDOT QA program. The motivation was not necessarily to exhaustively gather information on all the QA programs.

Different agencies view QA differently. In a nutshell, it can be stated that different states follow different combinations of end-result and materials and methods requirements with varying emphasis on different requirements. More recently there has been a trend to orient QA programs towards ensuring performance rather than checking if the construction methods were followed properly. Therefore, quality characteristics, like volumetric properties for HMA, which affect performance, are tested. It was also observed that most of the states use the same quality characteristics and test methods for quality assurance as for quality control. Only 10 states in the US specify test methods only for QA.

The most common quality characteristics used in QA programs in the US along with the number of agencies using them (NCHRP Synthesis 346) are shown below.

Table 4.1 Most commonly used quality characteristics for QC and QA on HMA pavements

Quality Characteristic	No. of Agencies	
	QC	QA
Asphalt content	40	40
Gradation	43	33
Compaction	28	44
Ride quality	16	39
Voids in total mix	20	26
Voids in mineral aggregate	26	23
Aggregate fractured faces	25	23
Thickness	13	22
Voids filled with asphalt	19	13

Table 4.1 and figure 4.1 show that the attributes most often used for the acceptance of HMA pavements are asphalt content, used by 40 agencies; compaction by 44; and ride quality by 39.

Thirty three agencies accept HMA pavements based on gradation and 26 accept voids in total mix. The lesser-used acceptance attributes are aggregate fractured faces, thickness and voids filled with asphalt.

Table 4.2 presents other quality characteristics which are used by some of the agencies although far less often than those mentioned in table 4.1 above.

Table 4.2 Less often used QC/QA quality characteristics for HMA pavements

Other Quality Characteristics	
Retained tensile strength	Joint density
Fine aggregate angularity	Aggregate moisture
Stability	Percent lime
Water sensitivity	Sand equivalence
Wheel tracking test	Flat and elongated particles
Plant mix temperature	Indirect tensile strength
Asphalt temperature	Percent lightweight particles

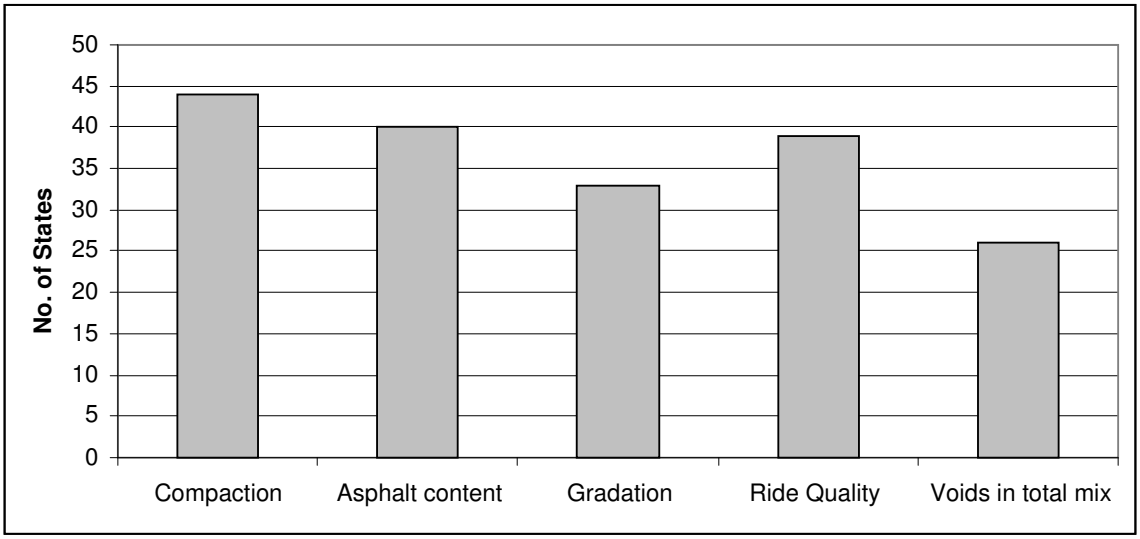


Figure 4.1 Attributes most often used for QA of HMA (40 responses) (adapted from NCHRP Synthesis 346)

4.1 Regional Level Analysis of the QA Practices

Another approach to analyze the similarities and differences in agencies' QA norms is to categorize them on the regional level as has been attempted below.

4.1.1 North- East Region

New Hampshire and Maine use similar testing and sampling methods for gradation- AASHTO T30, while New York employs AASTO T27. For testing the asphalt binder content, numerous sampling and testing procedures are in use. They are AASHTO T 164, NHDOT B-2, NHDOT B-6 (New Hampshire) & NYSDOT MM 5.0 (New York). Michigan uses AASHTO T 164 for the sampling and testing purpose.

The tolerance level for the air void at N_d is 2.5 to 5.5% in Maine while it is 3 to 5% in Michigan. 3 to 5% is more commonly used all across United States for air voids.

We have summarized the QA variables for the north-east states and compared them with Michigan in table 4.3.

4.1.2 Southern Region

The flexible pavement's QA testing and sampling method for aggregate gradation in the state of Arkansas is AASHTO T 30 & T 308, while AASHTO T 168 is used in Alabama and GDT-38 in Georgia. In Arkansas, the frequency at which samples are tested for gradation purpose is 1 per 750 metric tons. In Georgia, two samples are collected after 5 days of production. Table 4.4 presents information on the QA requirements for the southern states.

4.1.3 Mid-West Region

In Wisconsin, the sample for the Asphalt Material Content (AMC) is not collected if its quantity is less than 501 tons, while one sample is collected for quantities ranging between 510 tons - 30,000 tons. In South Dakota, SD 202 procedure is used in flexible pavements in order to carry out the sampling and testing of the mineral aggregate gradation, whereas in Michigan, the department of transportation employs AASHTO T 164. We have provided more information in table 4.5.

The overall lot pay factor for Michigan is = $(0.4 \times PF_D) + (0.3 \times PF_{AV}) + (0.15 \times PF_{BINDER}) + (0.15 \times PF_{VMA})$

Table 4.3 Quality Assurance Practices in North-East States

	Maine	**New Hampshire**	**New York**	**Michigan**
Sample & Test				
Aggregate Gradation	AASHTO T 30	AASHTO T 30	AASHTO T 27	AASHTO T2
Asphalt Binder Content(ABC)	AASHTO T 308	AASTO T 164 NHDOT B-2	NYSDOT MM 5	AASHTO T 164
Air Void	AASHTO T 312	Per QC Plan	AASHTO T 166 AASHTO T 209	AASHTO PP28-97
Sample Location				
Aggregate Gradation	Paver Hooper	Haul Unit		
Asphalt Binder Content	Paver Hooper	Haul Unit		
Air Void	Paver Hooper	Haul Unit		
Frequency				
Aggregate Gradation	1 per 500 Mg (500 ton)			
Air Void	1 per 500 Mg (500 ton)		1 per sub lot	1 per day
Asphalt Binder Content	1 per 500 Mg (500 ton)		4 per day per mix	1 per day
Tolerance				
Asphalt Binder Content	Target ± 0.4%			JMF ± 0.40%
Air Voids at N_d	4.0% ± 1.5%			3% - 5%
In Place Air Voids in total mix		± 2%		0%-8%

Table 4.4 Quality Assurance Practices in Southern States

	Alabama	Arkansas	Georgia	South Carolina	Michigan
Sample & Test					
ABC	AASHTO T 168	AASHTO T308	GDT-125	SC-T-64	AASHTO T 164
Air Void	AASHTO T 168	AASHTO T 269		SC-T-66	AASHTO PP28-97
Material Density	ALDOT 210	AASHTO T 209		SC-T-101	MTM 314
Voids in Mineral Aggregatge (VMA)	ALDOT 210	AHTD 464		SC-T-101	AASHTO PP28-97
Sample Location					
ABC	Loaded Truck		Truck/Road way	Plant	
Air Void	Loaded Truck			Plant	
Material density	Roadway			Road (in place)	
VMA	Loaded Truck			Plant	
Frequency					
ABC	1 per day per Lot		2 samples per 5 days production	1 per Sub lot	1 per day
Air Void	1 per day per Lot		2 samples per 5 days production	1 per Sub lot	1 per day
Material density	1 per 3,000 lane feet/lift			1 per1,500 ft. Sub lot	1 per day
VMA	as needed			1 per Sub lot	1 per day
Tolerance					
AC		± 0.3% for mix design value	± 0.4%	± 0.4%	JMF ± 0.40%
Air Void		3% - 5%			3% - 5%
Material density		Bases, Binder, Surfaces : 92% - 96%			Bases, Binder, Surfaces : 92% - 100%
VMA		ACHM Base, Binder, Surface Course : 11.0%-17%			±1%
Gradation		± 7.0%	± 4.0%		

Table 4.5 Quality Assurance practices in Mid-West States

	South Dakota	Wisconsin	Michigan
Sample & Test			AASHTO T2
Aggregate Gradation	SD 202		AASHTO T 164
ABC	SD 314		
Maximum Specific Gravity of Asphalt Concrete	SD 312	AASHTO T 209	MTM 314
Bulk Specific Gravity of Asphalt Concrete	SD 313	AASHTO T 166	AASHTO T 168
Air Voids		AASHTO T 269	AASHTO PP28-97
Frequency			
Aggregate Gradation	1 per1000 tons	No test : 0 to 501 tons 1 sample : 501 to 30,000 tons More than 30,000 tons: Add One Test for Each Additional 30,000-ton Increment	
ABC	1 per day		1 per day
Maximum Specific Gravity of Asphalt Concrete	1per 1000 tons		
Bulk Specific Gravity of Asphalt Concrete	1 per 1000 tons		
Air Voids			1 per day
Tolerance			
Aggregate Gradation	± 5.0%	37.5 mm ± 6.0	
ABC		± 0.4%	JMF ± 0.40%
Maximum Specific Gravity of Asphalt Concrete	± 0.020 %		
Bulk Specific Gravity of Asphalt Concrete	± 0.020 %		
Air Voids	± 1.2 %	± 1.3%	3% - 5%

4.1.4 Western Region

In accordance with procedure CP 41, samples are collected for testing the asphalt content in Colorado. One sample is collected for every 500 tons of asphalt. Procedure for asphalt content sampling and testing in Montana is MT-302. Again, 1 sample is collected for 500 tons of asphalt. Colorado and Montana have developed their own procedures for carrying out the QA exercise for the flexible pavements. We have summarized sampling and testing procedures, sample location, tolerance limits for western region states in table 4.6.

4.2 Conclusion

The MDOT HMA QA program is similar to that being used by the states who have adopted end result specifications for their QA program. This similarity exists in (1) the tests that are used for verifying the quality of the constructed pavement, (2) the specifications limits that are used and (3) the quality characteristics, the statistical method and the pay formula used for calculating the payment to be made to the contractor. In other words there is nothing alarmingly different in the QA program being used by MDOT as compared to other "ERS" states.

Table 4.6 Quality Assurance Practices in Western States

	Colorado	Montana	Utah	Michigan
Sample & Test				
Gradation	CP 30 CP 31A CP 31B	MT-202	AASHTO T 30	AASHTO T2
ABC	CP 41 CP 55	MT-302	AASHTO T 308	
VMA	CP 41		AASHTO T 269	AASHTO PP28-97
Maximum Specific Gravity	CP 41 CP 55	MT-212	AASHTO T 166	MTM 314
Sample location				
Gradation	Aggregate from the cold feed, pug mill discharge or extraction	Before Bitumen Is Added to Mix	Grade Behind the Paver	
ABC	Plant discharge, at/or behind paver. For Central Lab Correction Factor, sample from belt and binder from Contractors tank	At the Plant	Grade Behind the Paver	
VMA	Plant Discharge, windrow, at/or behind paver			
Maximum Specific Gravity	Plant Discharge, windrow, at/or behind paver	Randomly After Rolling and Before Opening to Traffic	Grade After Compaction Prior to Traffic	
Frequency				
Gradation	1 per 2000 tons	1 per 600 tons	4 per Lot	
ABC	1 per1000 tons	1 per 500 tons	4 per Lot	1 per day
Maximum Specific Gravity		1 per 600 tons	10:2 in each of 5 equal Sub lots	
Tolerance				
Gradation	3/8": ± 6%		3/4": ± 6% 1/2": ± 6%	
ABC	± 0.3%	5% below minimum requirement	± 0.35%	JMF ± 0.40%
Maximum Specific Gravity			Lower Limit: -2.0% Upper Limit: + 3.0%	

Chapter 5: Literature Review - Flexible Pavements

Several research projects were reviewed to verify the significance of candidate QA variables identified in this project for pavement performance. This section presents a brief mention of some of these findings.

5.1 Asphalt Content

Asphalt content primarily controls the voids contents of a mixture. It has been predicted that each 0.4 percent of increase in asphalt content results in about 1.0 to 1.5 percent decrease in in-place air void (Mauplin et al., 2006), which in fact improves durability of the mixture.

The effect of asphalt content is found to be more important during the initial stages of compaction compared to the final stages. Also, it has been found that an increase in asphalt content results in a significant decrease in work or force load required for compaction to 8 % air voids content, (typical for the construction stage in the field). The increase in asphalt content also resulted in a significant decrease, on the order of 2% to 6%, in air voids at the initial number of gyrations, $N_{Initial}$ (Stakston and Bahia, 2003). This confirms that higher asphalt content will result in a mixture less resistance to compaction. Besides, it seems that higher asphalt content will lead to lower resilient modulus and dynamic modulus to all frequencies (Flintsch et al., 2007). This was experimentally proved by 33 samples of 3 different mixtures in Virginia. It was also observed in 25 samples of four different Performance-Grade mixtures that for a 0.5-percent increase in asphalt content, flexural stiffness of mixture peaked (Mauplin et al., 2006). Such an increase did also make an upward trend for fatigue life of the sample, however, an additional 1.0-percent of asphalt was needed for a major beneficial effect (Mauplin et al., 2006).

"Development of Simplified Asphalt Concrete Stiffness-Fatigue Testing Device" (Tran and Hall, 2004) study suggests similar conclusion. That is, mixtures with asphalt content of 0.5 percent below optimum are stiffer than those of 0.5 percent above the optimum.

A study of 84 asphalt mix samples of two aggregate sources in Nevada also experimentally proved that mixes which violated the binder content on the low side performed better than the optimum mix especially in the fatigue or thermal cracking resistance (Sebaaly and Bazi 2004).

The results of laboratory-controlled strain flexural beam testing, that is fatigue life and flexural stiffness, for one aggregate and asphalt cement combination, five asphalt contents, and three air void contents clearly indicate that increased asphalt content increases fatigue life and reduce stiffness (Harvey and Tsal, 1996). It has been also experimented that for relatively thicker pavements, fatigue life increases approximately 10 percent for each 0.5 percent increase in asphalt content. For the thin structure, on the other hand, thinner structures experience approximately 20 percent increase in fatigue life for each 0.5 percent increase in asphalt content (Harvey and Tsal, 1996).

5.2 Air Void Content

Air Voids Content (AVC), which is the amount of voids in a compacted HMA pavement, can have a detrimental and significant effect on performance of the pavement (Harrigan, 2002). However, it has been suggested that this effect is not as significant as conditioning method is on pavement fatigue life (Vivar and Haddock, 2006).

Many studies have shown that the initial in-place voids should be no more than approximately 8% and that the in-place voids should never fall below approximately 3% during the life of the pavement. There is evidence that when in-place air void content drops below 3.090 to 3.5%, the probability of rutting increases drastically (Brown and Cross, 1992). That is, in-place air void content above 3.0% is needed to decrease the probability of premature rutting throughout the life of the pavement (Brown et al., 1991).

Mixes meeting $N_{Initial}$ and N_{Max}, initial and maximum number of gyrations, criteria that have been specified in order to avoid tender mixes and mixes prone to rutting, respectively, do not necessarily show less rutting potential than mixes which do not meet these criteria (Kandhal and Mallick, 1999). An experimental study of mixture of different aggregates, gradation, Nominal Maximum Aggregate Size (NMAS), and asphalt content in Alabama proved that no correlation could be established between APA rut depths and the gyratory compaction slopes (between $N_{Initial}$ and N_{Design}) of all mixes (Kandhal and Mallick, 1999).

Since permeability is directly related to the amount of interconnected voids within the pavement, high air void content leads to permeability to water and air, resulting in water damage, oxidation, raveling, and cracking where as low air void content results in rutting and shoving of the HMA (Brown et al, 2004). In addition, due to consolidation under wheel loading, high AVC can also contribute to the development of rutting in the wheel paths and low AVC, however, increases the likelihood of bleeding, shear flow, and permanent deformation (i.e., rutting) in the wheel paths (Harrigan, 2002).

In general, it has been investigated and the findings suggest that for each percent drop in air void content, there is a corresponding 10 percent loss of pavement life (Linden et al., 1988). This statement was also proven by three other separate sources: existing literature on the effect of air void content on pavement performance, a questionnaire survey of 48 state highway agencies on compaction practice, and performance data from the Washington State Pavement Management System emerge that, as a rule-of-thumb, for each 1 percent increase in air voids (over a base air void level of 7 percent), there is about a 10 percent (approximately one year) loss in pavement life (Linden et al., 1988). As it appears, more air voids will lead to lower resilient modulus and lower dynamic modulus (Flintsch et al., 2007). Monitoring 33 samples of three different mixtures in Virginia proves the foregoing statement.

It has been observed from a study of 60 samples of 5 asphalt pavement sites for rutting, that low voids (in re-compacted samples and/or field samples) are the cause of most rutting in the pavements (Brown et al., 1989). It is possible the relationship between rutting and air voids is affected by an increase in air voids once the pavement begins to rut and shove. Also, there is a good possibility that the void level decreases under compaction to some point at which rutting begins to occur and at which time the void level begins to increase due to shoving of the mixture (Brown et al., 1989).

The in-place rutting study of asphalt pavements clearly shows that very little rutting occurs when the re-compacted air voids are 3.0 percent or higher for compactive efforts of 75 blow Marshall and Gyratory with 120 psi, 1 degree and 300 revolutions. Three of the four mixes with more than 3 percent air voids have no rutting while the other mix has only 10 percent rutting. Significant rutting occurs in those layers having less than 3.0 percent re-compacted air voids. In this case three of the eight pavement layers have more than 20 percent rutting which is significant (Brown et al., 1989).

Mallick et. al. (2003) determined that there is a significant effect of air void content (as measured by voids in total mix) of dense graded HMA on in-place permeability of pavements and of NMAS on permeability of coarse-graded Superpave designed mixes. They recommended that State DOTs consider designing mixes as less permeable than coarse graded mixes at similar void levels i.e. placed 100 mm below the pavement surface on the fine side of the maximum density line, and thereby less susceptible to allowing moisture or moisture vapor to propagate upward through the pavement structure. This in turn would reduce the possibility of moisture damage within such structures.

A Wisconsin DOT study (WisDOT, 2004) estimates and measures permeability during mixture design and defines a relationship between lift thickness and aggregate gradations that minimizes the densification problem and addresses the permeability concerns. They state that it is well recognized that field density is significantly affected by maximum aggregate size of aggregates, gradation, and lift thickness and that permeability of asphalt mixtures is a function of aggregate gradation, density achieved, and distribution of air voids. With a shift in mixture designs to Superpave methods, gradations, which are unique in their densification characteristics and claimed to be more permeable, on the coarse side of the maximum density line are being widely used as well as recommended. This trend, according to them, might be due to changes in the air voids distribution, the lower densities being achieved, or both. Permeability is also a directional property such that orientation of the aggregates, which is affected by lift thickness and level of compaction, has a significant effect on total permeability.

5.3 Mixture Density

One of the most important parameters in construction of asphalt mixtures is density. It must be closely controlled to insure that the voids stay within an acceptable range. However, effect of mixture density on performance of mixture is dependent upon the mixture gradation and aggregate size, and NMAS (Vivar and Haddock, 2006).

The Asphalt Institute recommends that the mix design density should closely approach the maximum density obtained in the pavement under traffic. A review of several state DOT specifications has shown that in-place density, measured as a percent of maximum theoretical density, ranges between 91% and 98% (with many falling between 92% and 97%), supporting

the previous statement (Brown et al., 2004). HMA mixtures of lower density tend to have higher permeability, lower dynamic modulus, lower flexural stiffness, and shorter fatigue life (Vivar and Haddock, 2006).

Much of the loss in pavement life is a direct result of low density. Study of 36 mixture samples in lab and field surveys have both concluded that inadequate density is a significant problem on a high percentage of paving projects (Brown et al., 1991).

A mixture that is properly designed and compacted will contain enough air voids to prevent rutting due to plastic flow but low enough air voids to prevent permeability of air and water. Since density of an asphalt mixture varies throughout its life, the voids must be low enough initially to prevent permeability of air and water and high enough after a few years of traffic to prevent plastic flow.

In general, it has been shown that the rut potential is significantly reduced with increases in HMA density, which is the reduction in air voids content (Vivar and Haddock, 2006). Insufficient compaction during mix design and mixture testing results in a higher required asphalt content to obtain the specified void level. For that matter, compaction during construction and under traffic loads results in a density which is higher than laboratory density and consequently lower voids than measured during mix design. Thus, compaction during mix design and field quality control has to produce a density in the laboratory equal to what will be obtained in the field after a few years of traffic (Brown et al., 1989).

5.4 Aggregate Size and Gradation

It is well known that the density which could be achieved in the field and its effects on HMA mixture performance (Vivar and Haddock, 2006) are significantly affected by aggregate size and gradation. Also well recognized is that permeability of asphalt mixture is a function of aggregate gradation. It has been shown that for any given in-place air void content, permeability of HMA mixture increases by one order of magnitude as the maximum aggregate size increases (Mallick et al., 2003).

Aggregate size, shape, and gradation influence the size of voids within a pavement structure. Coarse graded mixes require more asphalt content (Haddock et al., 1991). Such mixtures contain a relatively higher percentage of coarse aggregate than fine graded mixes. This higher percentage of coarse aggregate leads to larger voids within the mix matrix. The

combination of larger voids and fewer fine aggregates to fill the voids likely result in more interconnected voids which make the mix more permeable. In addition, coarse graded mixtures tend to have lower dynamic modulus and flexural stiffness in comparison with fine graded mixtures but higher permeability and longer fatigue life (Vivar and Haddock, 2006).

Increasing the size of the largest aggregate in a gradation will increase the mix quality with respect to creep performance, resilient modulus, and also tensile strength. However, there will not be any significant effect on Marshall Stability (Brown and Bassett, 1990). For instance, monitoring 84 asphalt mixture samples of two aggregate sources in Nevada proved that mixtures which were low on the # 4 sieve and high on the # 200 sieve never achieved a performance that is equivalent to or better than the optimum mixture. In fact, mixtures of high percent passing # 200 were always worse than optimum mixture unless higher binder content was introduced (Sebaaly and Bazi, 2004). It can be said, thus, that mixes with larger maximum aggregate size are stiffer and will reduce stresses in the underlying layers (Brown and Bassett, 1990). This conclusion has been experimentally proven in Arkansas by 36 mixture samples of two aggregate size and four different asphalt content that large aggregate size of 25 mm has higher dynamic modulus than small aggregate size of 12.5 mm (Tran and Hall, 2004).

In addition, a study of 4 HMA mixtures with different air void content, aggregate size and gradation showed that mixtures with a 19.0-mm NMAS tend to have higher dynamic modulus and flexural stiffness, higher permeability and moisture damage, and lower fatigue life than mixtures with a 9.5-mm NMAS (Vivar and Haddock, 2006).

It is also true in the case of coarse-graded against fine-graded mixes. Coarse-graded mixtures tend to have lower dynamic modulus and flexural stiffness, but higher permeability and fatigue life compared to fine-graded mixes (Vivar and Haddock, 2006). In the same study, at constant 96 percent density, a coarse-graded mixture with 9.5-mm NMAS showed the best fatigue life where fine-graded mixtures of higher dynamic modulus with 19.0-mm NMAS manifested less of rut performance (Vivar and Haddock, 2006).

Studies on effects of aggregate angularity (Stakston and Bahia, 2003) show that the influence of aggregate gradation in predicting mixture performance, such as control of voids, is highly specific to the angularity of aggregate and its source.

“Evaluation of Asphalt Pavement Analyzer for HMA Mix Design” (Kandhal and Mallick, 1999) states that the effect of gradation on granite and limestone wearing and binder courses

with PG 64-22 asphalt is significant, with below restricted zone gradation showing higher rutting compared to above and through restricted zone. This experiment studied mixture samples of different aggregates type, gradation, nominal maximum size aggregates and binder in Alabama and proved that effect of gradation is similar and significant for granite PG 58-22 wearing courses but not significant for granite binder course. However, such an effect it is not significant for rutting of gravel wearing and binder course mixes with PG 64-22. The above and through restricted zone mixes showed slightly higher rutting compared to below zone mixes.

5.5 Pavement Thickness

In general, the thicker the asphalt layer, the stiffer it is. This fact has been studied in Arkansas where 25-mm thick HMA layer samples presented higher dynamic modulus than samples of 12.5-mm HMA thickness (Tran and Hall, 2004). However, roles of aggregate size and binder content on this effect should not be overlooked.

Investigations of 18 national projects along with Specific Pavement Studies in LTPP database (Von Quintus and Simpson, 2003) have clearly concluded that greater amounts of fatigue cracking occur on pavement structures of thinner HMA layers. Chatti et al. (2005) confirmed that pavements with “thin” (102 mm = 4 in.) HMA surface layer have shown more fatigue cracking, slightly more rutting, and higher changes in IRI than those with “thick” (178 mm = 7 in.) HMA surface layer.

Selezneva et. al. (2002) studied the quality and completeness of pavement layering information and layer thickness data for LTPP sites. They evaluated the consistency of material type and thickness data between different data sources and also layer thickness variability indicators, within-section material type consistency, and material type and thickness reasonableness. They also analyzed experiments to determine characteristics of within-section layer thickness variation, including layers with different material and functional types and computed descriptive statistics such as mean, standard deviation, skewness, and kurtosis for each section.

5.6 Binder Performance Grade

The asphalt binder affects various performance aspects of the asphalt mixtures such as permanent deformation, fatigue cracking, and low temperature cracking. The Performance Grade,

PG, binder specification is intended to select the binder to optimize its effect on the performance of the pavement.

The increase in the performance grade resulted in a marginal increase in air voids, on the order of 2% to 6%, at the initial number of gyrations, Ninit, (Stakston and Bahia, 2003). This confirms that higher performance grade asphalt could result in a mixture which is more resistant to compaction (Stakston and Bahia, 2003). Also, in Alabama it was observed from asphalt mixtures with different aggregates, gradation, nominal maximum aggregate size, and binder, that rut depths of mixes with PG 58-22 asphalt binder (tested at 58 °C) were higher than those of mixes with PG 64-22 asphalt binder (tested at 64 °C).

5.7 Review of different Non-Destructive tests

A detailed review of non-destructive tests is provided in Appendix B. The following sections give a brief overview of some of the relevant tests for flexible pavements.

5.7.1 Thickness

Pavement layer thickness is an important factor in determining the quality of newly constructed pavements and overlays because deficiencies in thickness reduce the life of the pavement. In order to use pavement thickness as a measure of quality assurance, it is necessary to have an accurate and reliable method for making the thickness measurement. Cores are accurate, but they are time consuming, they damage the pavement, and they represent a very small sample of the actual pavement. Therefore, it is desirable to have a thickness measuring method which is quick, non-destructive, and which can generate an accurate and representative population of pavement thickness data points. Some of the non destructive test methods available for thickness measurements have been listed here with their key features from the point of view of their suitability of use in a quality assurance (QA) program.

5.7.1.1 Ground-penetrating radar (GPR)

Ground-penetrating radar (GPR) is a high resolution geophysical technique that utilizes electromagnetic radar waves to scan shallow subsurface objects. It can provide information on pavement layer thickness or locate targets (Daniels,1990; Hasted, 1973; Ulriksen, 1982, Harris, 1998). Frequency of the GPR antenna affects depth of penetration. Lower frequency antennas

penetrate further, but higher frequency antennas yield higher resolution. To successfully provide pavement thickness information or scan an interface, the following conditions have to be present:

- Physical properties of the pavement layers must allow for penetration of the radar wave.
- Interface between pavement layers must reflect the radar wave with sufficient energy to be recorded.
- Difference in physical properties between layers separated by interfaces must be significant.

Physical (electrical) properties of pavement layers, thickness of pavement layers, and magnitude of difference between electrical properties of successive pavement layers impact the ability to detect thickness information using GPR. Conductive losses occur when electromagnetic energy is transformed into thermal energy to provide for transport of charge carriers through a specific medium. Presence of moisture or clay content in a pavement layer will cause significant conductive losses and hence will increase the dielectric permittivity and decrease the depth of penetration.

For asphalt pavements, GPR is by far the most established technology for measuring pavement thickness. Evaluation studies have been carried out by over ten state highway agencies, by SHRP, MnROAD, and by the FHWA, all of which have documented the accuracy of GPR asphalt thickness vs. core samples (Maser, 1999; Wenzlick and Maser, 1999). The studies have generally compared the GPR results to cores, and have shown differences that range from 2 to 10%. The lower differences (2-5%) are generally associated with newly constructed pavements, while the bigger differences are generally associated with older pavements (Infrasense, 2003). In general, where there are large deviations between GPR and core values, the GPR gave the larger values, and the difference appeared to be due to portions of the core that remained in the hole (Infrasense, 2006). Studies have also shown that with proper equipment and data processing, GPR can accurately determine thickness for overlays as thin as 25 mm (1 inch) (Maser and Scullion, 1992).

5.7.1.1.1 Horn Antenna GPR

The air-launched antenna is routinely used at highway speeds and is not physically affected by rough road conditions. Most importantly, it is not necessary to obtain cores to calibrate the air-launched horn antenna system. Another important advantage of the horn antenna is the ability to measure thin pavement layers.

Advantages:

- Only highway-speed subsurface pavement testing tool
- Excellent for flexible pavement rehabilitation projects
- Can be merged with surface video and other NDT data

Limitations (Wimsatt et al., 2008):

- Depth limited to top 20 – 24 inches
- Attenuation problems with concrete layers
- Initially limited software available for processing data

Barriers to Implementation:

- FCC restrictions on manufacturers in the US
- Oversold - initial results disappointing
- No certification of equipment and vendors

5.7.1.1.2 Ground-coupled GPR

As the name suggests, a ground-coupled antenna needs to remain in contact with the ground (or suspended very slightly above the ground) to properly couple the electromagnetic energy to and from the antenna. To calibrate the ground-coupled system it is necessary to obtain cores from the pavement and physically measure the actual pavement thickness.

Advantages:

- Fairly inexpensive
- Robust equipment – technology and software widely available
- Deep investigations possible with low frequency equipment

Limitations (Wimsatt et al., 2008):

- Speed typically less than 10 mph
- Limited near surface information
- Penetration limited in clay material
- Qualitative information requiring an expert for interpretation

Barriers to implementation:

- Technology not well understood by DOT's
- No other significant barriers

5.7.1.1.3 GPR applications

Air coupled GPR:

- Thickness of pavement layers
- Moisture or density related defects in HMA and base layers

- Density of new HMA layers
- Delaminations in bridge decks (with HMA surfaces)
- Section uniformity

Ground coupled GPR:

- Detecting buried objects
- Voids under thick concrete slabs
- Detecting steel presence and depth
- Locations where deep investigations are required

It is capable of detecting a number of parameters in reinforced concrete structures:

- location of reinforcement
- depth of cover
- location of voids
- location of cracks
- in situ density
- moisture content variations

User expertise

Users must have good knowledge of wave propagation behavior in materials in order to meaningfully collect and interpret results. Training and experience are required.

5.7.1.1.4 Advantages and Limitations of GPR

- It can be used to survey large areas rapidly for locating reinforcement, voids and cracks. GPR can be collected continuously at various speeds, and thus allowing for the availability of a large number of thickness data points to be collected economically.
- Results must be correlated to test results on samples obtained. Any features screened by steel reinforcement will not be recorded.
- GPR has also been effectively used to determine variations in asphalt density (Saarenketo and Roimela, 1998).
- With increasing depth, low level signals from small targets are harder to detect due to signal attenuation.
- It is expensive to use and uneconomical for surveying small areas.
- GPR technology lacks the ability to differentiate between the AC layers and layers of asphalt-treated materials in thickness estimation (Hanna, 2002).
- Most of these GPR layer thickness studies have been carried out with air-coupled horn antennas since these can be implemented at driving speed without lane closures. However, for the purposes of quality assurance, lower data collection speeds permit consideration of ground-coupled antennas as well.

5.7.1.1.5 Accuracy and Interpretation of GPR

There are a number of factors to be taken into account when interpreting radar data and signals:

- Hyperbolic shapes typically represent a point reflector
- The diameter of cylindrical objects ranging from rebars to metallic oil drums cannot be determined from radargrams
- Radar wave velocity reduces when travelling through wet concrete
- Radar waves are more rapidly attenuated when travelling through wet concrete
- Radar waves cannot penetrate conductors such as metals, clays, salt water
- Radar antennas cannot identify objects in the near field which are closer to the surface than $\lambda/3$, where λ is the wave length (IAEA, 2002)

5.7.1.2 Summary

The methods described in this section are summarized in Table 5.1 below.

Table 5. 1 NDT methods for measuring thickness

Method	Technology	Application	Measurement Type	Measurement Rate	Prior Experience
Horn antenna	Non-Contact GPR (electromagnetic)	asphalt	continuous	up to 9 m/sec (30 feet/sec)	extensive
Calibrated Single Antenna	Ground-Coupled GPR (electromagnetic)	asphalt or concrete	continuous	up to 1.5 m/sec (5 feet/sec)	none documented
Dual Antenna CMP	Ground-Coupled GPR (electromagnetic)	asphalt or concrete	Point	estimated 2 min./point	limited for pavement

The summary table distinguishes between continuous and “point” methods. The continuous methods can collect data while the equipment is moved continuously along the pavement. The "point" methods must be set up to make a measurement at a particular point. An estimated rate of data collection has been indicated. Note that some of the methods are well established, while others are relatively new for this application.

5.7.1.3 Conclusion

The GPR system is capable of estimating the layer thicknesses accurately, especially for HMA layers. It should be emphasized that the accuracy of the GPR system may be significantly affected when noise is present in the data due to external interferences. According to Holzschuhe

et al., (2007), the repeatability of the GPR system is excellent for speeds ranging from less than 15 mph up to 70 mph. They also reported that the thickness predictions are very reliable at highway speeds. However, it is strongly recommended that when the data is collected at highway speeds, more markers be inserted in the GPR data in order to minimize the offset errors. These markers should be linked to physical objects with known mileposts. It is strongly recommended that the GPR system be used as a tool for assisting in pavement thickness determination. More accurate thickness information can be obtained when the core thicknesses are used as feedback into the GPR analysis for calibration of radar velocities (Holzschuhe et al., 2007).

5.7.2 Moisture and Density (Radioisotope Gauges)

Moisture gauges consist of a source of neutron radiation, which irradiates the material under test. As a result of radiation, gamma rays are created and detected. The result is a series of counts, which are a measure of the composition of the material. It can be used to measure moisture content of concrete, soil and bituminous materials and to map moisture migration patterns in masonry walls.

5.7.2.1 Advantages:

- Instrument is portable
- Moisture measurements can be made rapidly

5.7.2.2 Limitations:

- A minimum thickness of surface layer is required for backscatter to be measured,
- It measures only the moisture content of surface layer (50 mm),
- It emits radiation,
- Results are inaccurate because hydrogen atoms of building materials are measured in addition to those of water,
- Its use in concrete is limited and requires calibration in order to calculate density or moisture content

5.7.3 Modulus of Pavement Layers

Making accurate assessments of the structural condition of roads during construction helps tremendously in locating weak areas prone to localized failure and correcting them prior to completion of the pavement.

5.7.3.1 Humboldt Stiffness Gauge

The Humboldt Stiffness Gauge (HSG) provides a simple, quick and accurate means of directly measuring stiffness of the upper lift of unbound material. The HSG measures impedance at the soil surface by generating vibrations at 100 and 200 Hz that impart a very small change in the applied load (Humboldt, 1999). The stiffness of the pavement material in resisting this load is determined at each frequency and the average is displayed on the Stiffness Gauge display window. The entire process takes about one minute. The HGS weighs about 10 kg, is 28 cm in diameter, 25.4 cm tall and rests on the soil surface via a ring-shaped foot. The advantage of the HSG is that it is lightweight and can be used at many locations to assess variability in individual compacted lifts of unbound material. The disadvantage of the HSG is that its depth of penetration is limited to about six inches.

5.7.3.2 Falling Weight Deflectometer (FWD)

The Falling Weight Deflectometer (FWD) is a nondestructive testing device widely used for assessing the structural condition of a pavement. When complete deflection basins are available, deflection testing can provide key properties for the existing pavement structure through backcalculation of the measured pavement responses, Specifically, for hot-mix asphalt (HMA) pavements, the elastic modulus (E) of the individual paving layers can be determined, along with the resilient modulus (MR) of the subgrade. Deflection testing has also seen some limited use as a means of monitoring the quality of a pavement during construction.

Portable light weight FWD has been developed and used in Europe and has gained the interest of many DOTs. Its applications are:

- Rapid stiffness testing of bases and subgrades but discrete measurement of bearing capacity of granular layers.
- Alternative to Nuclear density gauges

5.7.3.2.1 Advantages and Disadvantages of FWD

Advantages:

Pavement deflection testing provides some distinct advantages over destructive testing, including the following (Hudson et al. 1987):

- More rapid testing operation.
- Relative ease of operation.
- Lower operating cost.

- Less manpower requirements.
- Less intrusive procedure.
- Increased number of test points.

Specific advantages of the FWD are (NHI, 1994):

- Realistic simulation of actual wheel loading.
- High productivity.
- Ability to measure deflection basin.
- Ability to measure joint/crack load transfer.

Disadvantages:

- High initial cost.
- The need for traffic control.
- Relatively complex electro-mechanical system.

Specific issues in backcalculation of layer moduli include:

- Depth of influence unknown
- Number of layers for backcalculation is typically limited to three
- Layer thicknesses must be known
- Difficult to distinguish between layers of similar stiffness

5.7.3.3 Dynamic Cone Penetrometer

The Dynamic Cone Penetrometer (DCP) is a quick, simple, automated field test method for evaluating the in-situ stiffness of existing highway pavements. The greatest advantage associated with the DCP is its ability to penetrate into underlying layers and accurately locate zones of weakness within the pavement structure. It measures the strength and stiffness of unstabilized base and subgrade layers. The unit has software for storing DCP data. The DCP drives the penetrometer rod into the ground using constant energy for each blow, and the penetration index determined with the DCP is calculated as a running depth of penetration per blow.

5.7.3.3.1 Applications

- Quality assurance testing of subgrade and embankment materials
- Alternative to Nuclear density gauges

5.7.3.3.2 Advantages and limitations

Advantages (Wimsatt, 2008):

- Cheap/portable/simple
- Related to CBR and stiffness

Limitations:

- Slow, labor intensive
- Point specific
- Problems with granular materials
- Rod friction should be accounted for in clays

Barriers to implementations:

- No specifications
- Influence of layer moisture content

5.7.3.4 Seismic Pavement Analyzer

The Seismic Pavement Analyzer (SPA) lowers transducers and sources to the pavement and digitally records surface deformations induced by a large pneumatic hammer which generates low-frequency vibrations, and a small pneumatic hammer which generates high-frequency vibrations. The SPA differs from the FWD in that more and higher frequency transducers are used, and more sophisticated interpretation techniques are applied. All measurements are spot measurements; that is, the device has to be towed and situated at a specific point before measurements can be made. A complete testing cycle at one point takes less than one minute. A summary of pavement properties estimated by the SPA is provided in table 5.2.

Table 5. 2 Pavement properties estimated by the Seismic Pavement Analyzer

Pavement component	Parameter measured				
	Young's Modulus	Shear Modulus	Thickness	Damping	Other
Paving layer	yes	yes	yes	no	temperature
Base	yes	yes	yes*	no	
Subgrade	no	yes	no	yes	

*Thickness estimate of base depends on shear modulus contrast with subgrade.

5.7.3.4.1 Advantages and Limitations of SPA

Advantages (Wimsatt et al., 2008):

- Reduces number of destructive tests required for determining pavement layer properties
- Results can be obtained within two minutes, since the data is analyzed on site

Limitations (Wimsatt et al., 2008):

- The testing is discrete by nature (i.e. the testing measures properties at a single point per test, and it takes two minutes per test)
- Not suitable for rapid 100% coverage testing
- Unsuitable for testing composite pavements (Hanna, 2002)
- Unproven equipment reliability
- Need for high skills relevant to data reduction and analysis

5.7.3.5 Conclusion for Modulus of Pavement Layers

It is difficult to directly compare results of the various deflection testing equipment because they measure to different depths, they utilize different technologies to induce load and measure in-situ response, and different equations are used to convert surface deformation to layer modulus, particularly on two layered pavement structures. Data obtained in a study indicate strongly that the devices do give similar magnitudes of stiffness and modulus, and similar trends in the data with regard to relative stiffness of the in-situ layers (Hanna, 2002).

The Humboldt Stiffness Gauge is an effective tool for monitoring the integrity of individual material lifts as they are constructed, since the measurements are limited to that lift. Conversely, the FWD is effective in measuring the total composite stiffness of in-situ pavement structures. The FWD has a definite advantage over the Plate Load Test in being faster, less labor intensive and able to provide much better coverage within a given period of time. If specific areas of the pavement are identified with the FWD as having unusually low stiffness, the Dynamic Cone Penetrometer can be used to identify the cause(s) of low stiffness and locate specific layers within the structure which will likely cause premature distress. Engineers can then assess the cost and benefits of correcting the problem early to extend the service life of the pavement, and avoid higher maintenance costs and public inconvenience later.

5.7.4 HMA Temperature

Temperature measurement of the HMA mat during construction using infra-red cameras is very useful to investigate temperature uniformity of new HMA layers, detect thermal segregation, create a permanent log of paving operations, and locate and establish duration of paver stops.

Advantages

- Segregation of hot mix a continuing problem
- Newer lower cost camera systems widely available

- Automated system with 100% coverage
- Cameras and guns available

Limitations

- Equipment not widely available

Barriers to Implementation

- Unknown targets given the variability of PG gradations and mix types
- Not currently included in specifications

Chapter 6: Empirical Data Analysis – Flexible Pavements

As stated earlier, data from Michigan projects were rather scarce. Therefore, alternative sources of data needed to be explored to determine how quality characteristics used in QA programs affect pavement performance. A preliminary analysis was first performed to study the relationship of acceptance parameters like air voids (or density) and asphalt content to performance parameters like rutting and fatigue using data from Long Term Pavement Performance (LTPP) projects. In this analysis, data from several states was used. These states geographically lie in different climatic zones. The LTPP database contains performance data (rut depth, fatigue cracking, longitudinal cracking, transverse cracking, IRI etc.) and design and construction data (including physical inventory data, material properties from in-situ and laboratory tests). For the preliminary analysis all the data were derived from the Specific Pavement Studies – 1 (SPS -1) experiment. This analysis was followed by alternative analysis with data from General Pavement Studies (GPS) experiments.

Table 6.1 lists categories of data that were extracted from the LTPP database. The data comes from multiple states. There were very few data points available for the state of Michigan.

Table 6. 1 Categories of data extracted from LTPP database.

Construction number	Asphalt Content
Traffic opening date	Mean
Maximum specific gravity	Minimum
Voids in mineral aggregate	Maximum
Effective asphalt content	Standard deviation
	Number of samples
Bulk specific gravity	% Air voids (in-situ)
Mean	Mean
Minimum	Minimum
Maximum	Maximum
Standard deviation	Standard deviation
Number of samples	Number of samples

Different regions in the United States are broadly divided into four climatic zones: (a) Wet freeze (b) Wet non-freeze (c) Dry freeze and (d) Dry non-freeze. Michigan falls in the wet freeze zone. Therefore, data corresponding to all sections in the wet freeze zone were separated for analysis. Other states falling in the wet freeze zone, for which data was available and was extracted from the LTPP database, are Delaware, Iowa and Virginia. However, it was found that these four states have relatively little data available. Therefore, we decided to include states from other climatic zones in the analysis. It is important to mention that different projects from these states have partial results. Figure 6.1 and 6.2 show the number of projects for each of the states which have fatigue and rutting data respectively. The state code for Michigan is 26.

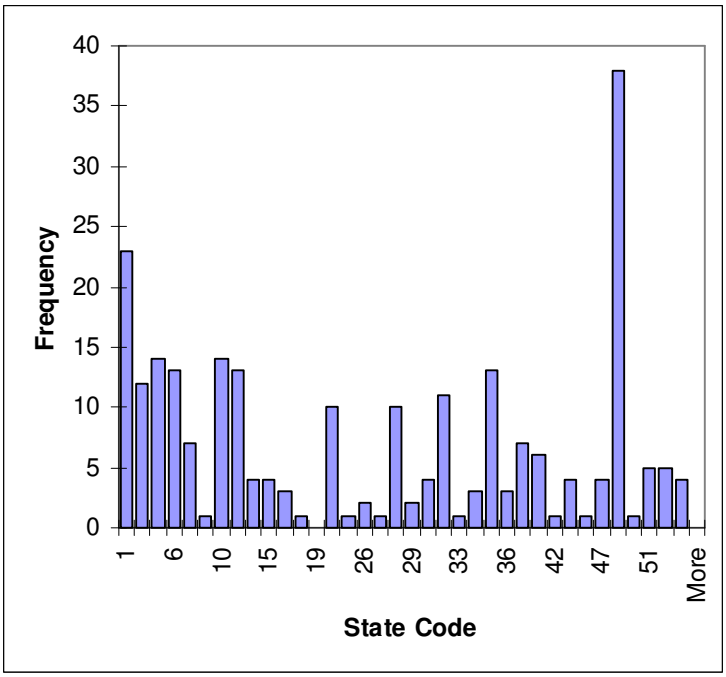


Figure 6.1 Number of projects from different states which have fatigue data.

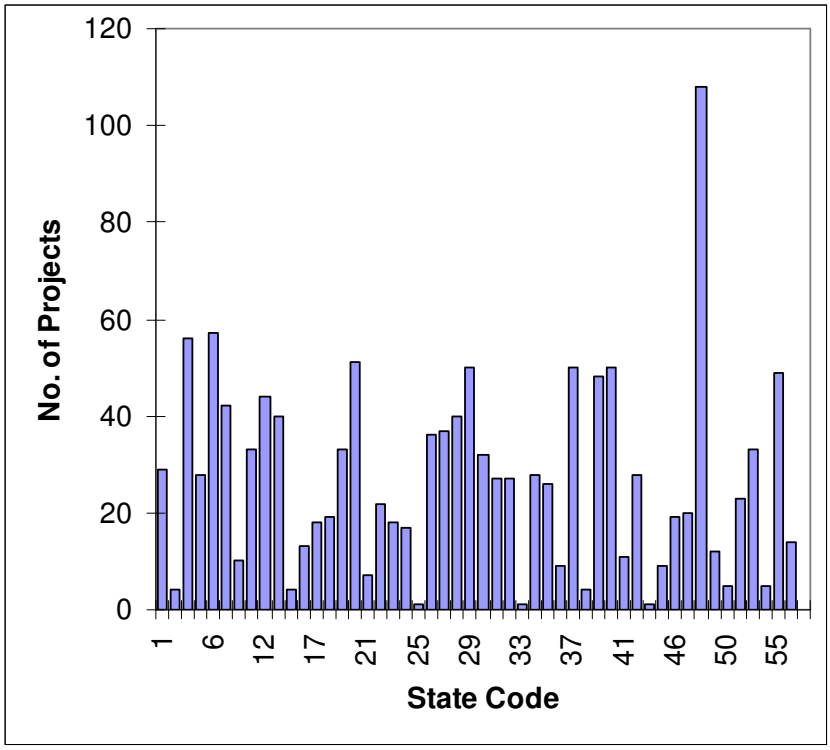


Figure 6.2 Number of projects from different states which have rutting data

Since different projects have varying age, the length of period for which performance data is available varies from project to project. Figure 6.3 and 6.4 show the distribution of all the projects with respect to number of years of performance data available. To be able to objectively compare their performances, the area under the performance curve was normalized by the length of period for which performance data is available.

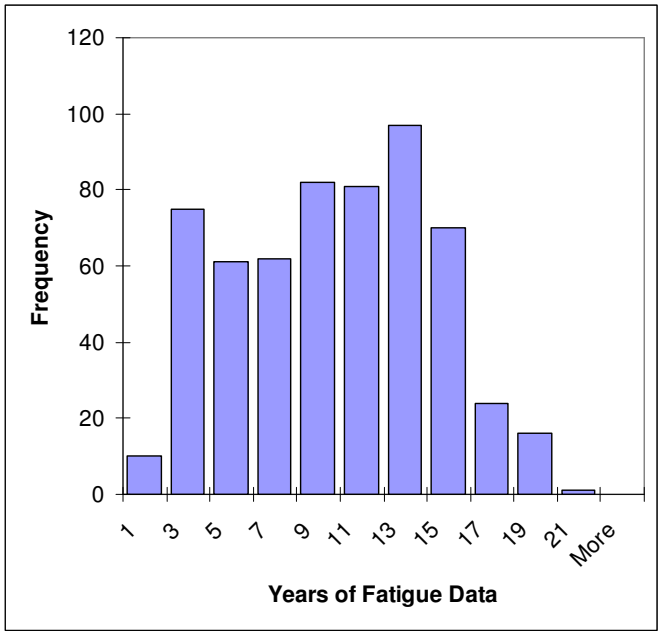


Figure 6.3 Distribution of projects based on years of fatigue data available

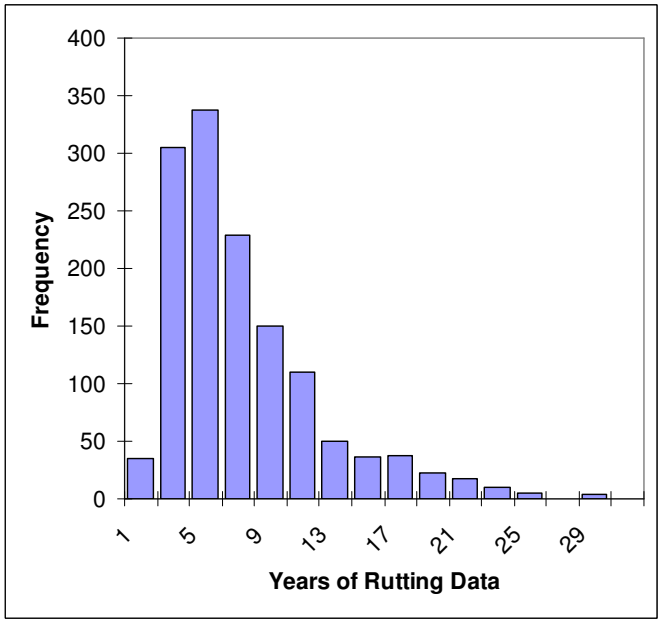


Figure 6.4 Distribution of projects based on years of rutting data available

6.1 Preliminary Analysis

6.1.1 Rutting

Figure 6.5 shows the plot of pavement rutting verses percent in-situ air voids immediately after construction. The purpose of this study was to see if certain levels of air voids affect rutting performance. It should be noted that a trend cannot be derived from such a plot because many other parameters were different between projects from which data has been used. However, it can be clearly seen that pavements with low (< 4%) or high (>8%) in-situ air voids have higher probability of rutting than those in which air voids fall in the range of 4 to 8%. This relationship is significant because it is a clear indication that in-situ air voids (or density) immediately after construction is an important parameter that needs to be controlled to ensure good rutting performance.

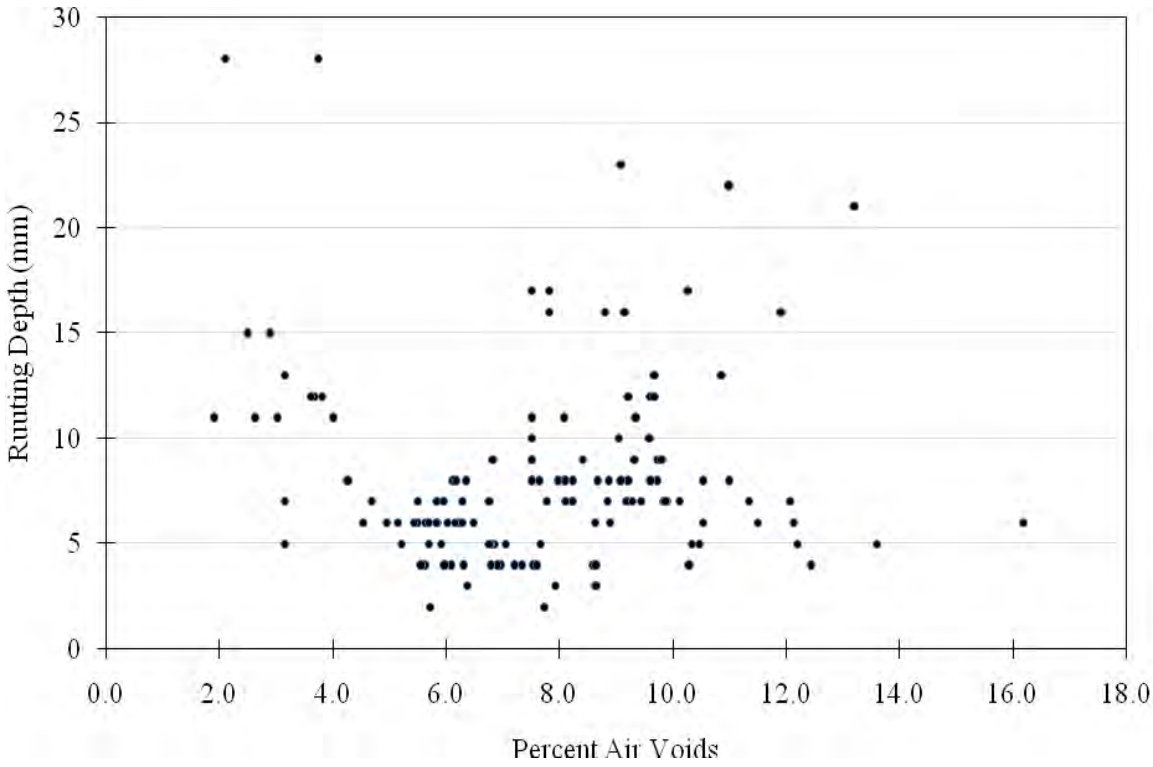


Figure 6.5 Rutting vs. percent air voids for all SPS-1 sections in LTPP database

Figure 6.6 shows a plot of rutting vs percent asphalt content. This scatter plot does not seem to show any particular relationship between them. However, intuitively it would be

expected that very high values of asphalt content should lead to a richer mix which would rut more. A possible explanation for this apparent anomaly is presented below.

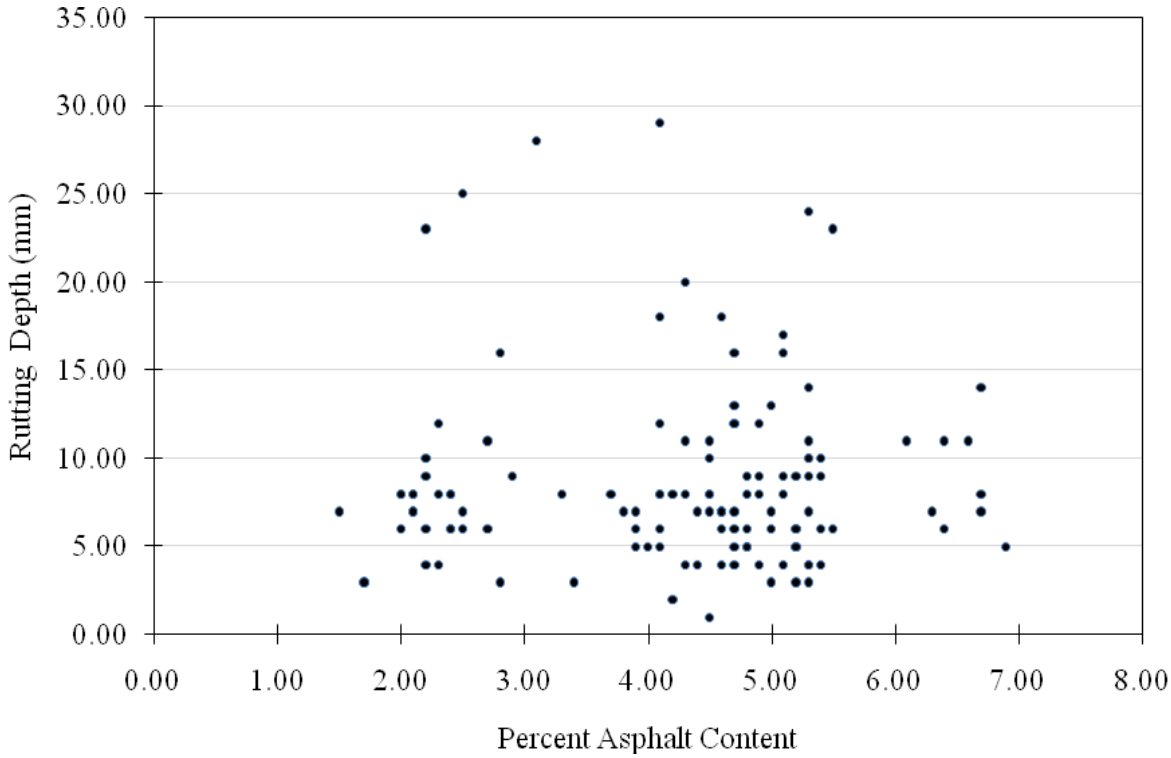


Figure 6.6 Rutting vs. asphalt content for all SPS-1 sections in LTPP database

From field experience, it is known that it is much easier to precisely control asphalt content in a mix than certain other parameters like air voids. SPS-1 was an experiment performed under strict quality controls. Therefore, most of the mixes have optimum quantity of asphalt as estimated during mix design. One may conclude that from this set of data asphalt content does not seem to be affecting rutting. However, a better conculsion would be that since there are very few projects with exceptionally high asphalt content, a relationship between asphalt content and rutting can not be established from such data.

To better understand the relationship between asphalt content and rutting projects with similar aggregate gradation and pavement structure but with varying asphalt content would be required.

6.1.2 Fatigue

It would be expected that generally higher air voids would lead to greater amount of fatigue on HMA pavements. Figure 6.7 shows this trend to some extent. Historically, it is also known that lean HMA mixes (i.e., HMA with low asphalt content) fatigue prematurely. This trend is not visible in Figure 6.8. However, there is some indication that higher asphalt content may contribute to poor fatigue performance. This needs to be verified with data from a more controlled experiment. The data plotted in either of the cases shown here come from the SPS-1 experiment of LTPP in which several other factors were also varied.

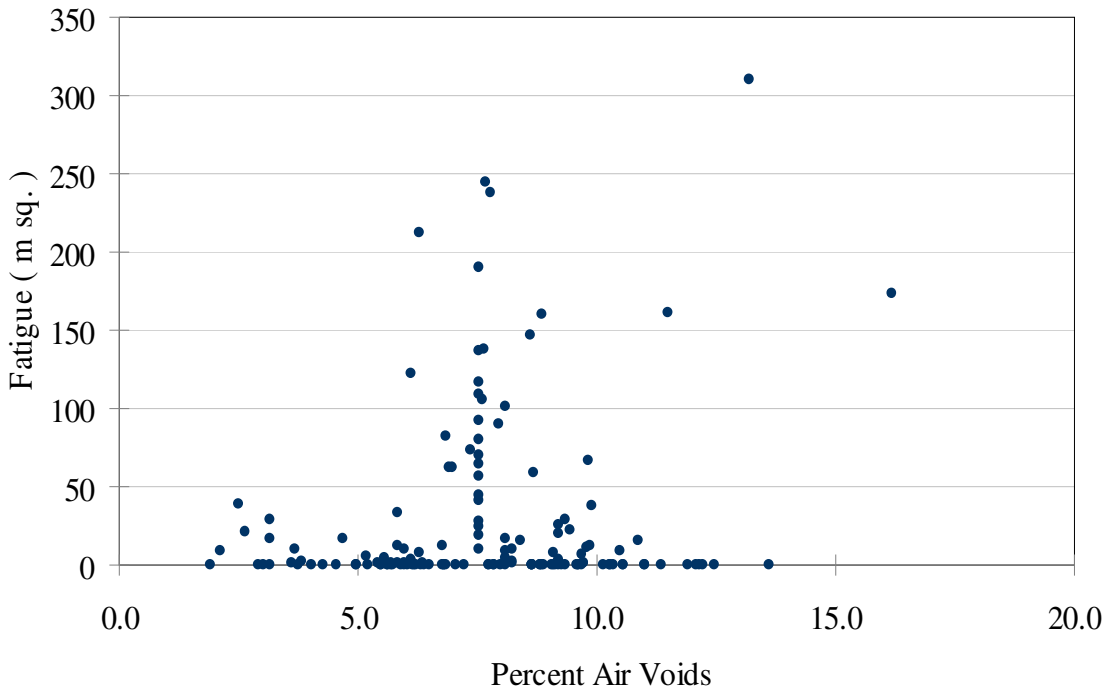


Figure 6.7 fatigue vs. in-situ air voids content for all SPS-1 sections in LTPP database

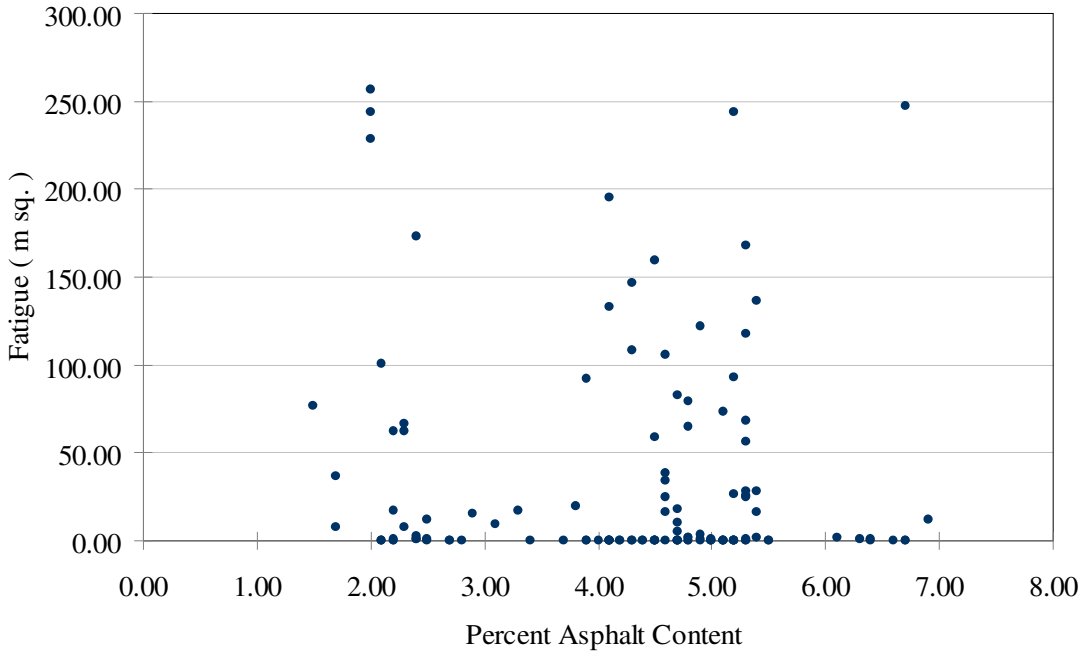


Figure 6.8 Fatigue vs. asphalt content for all SPS-1 sections in LTPP database

6.1.3 Longitudinal Cracking

Hot Mix Asphalt properties such as air voids content and amount of asphalt in the mixture, can have an effect on longitudinal cracking. However, Figure 6.9 does not show a particular trend. The cluster of high level cracking corresponding to low air voids (2 to 4%) is counterintuitive. However, at this point, no definite conclusions can be derived from this data. Further study of other possible factors would need to be undertaken to explain such poor performance.

These results highlight the need for more detailed analyses, with preferably the inclusion of forensic studies. LTPP database has data from several states. There were different approaches used in this analysis to be able to decipher credible relationships between QA tests and pavement performance.

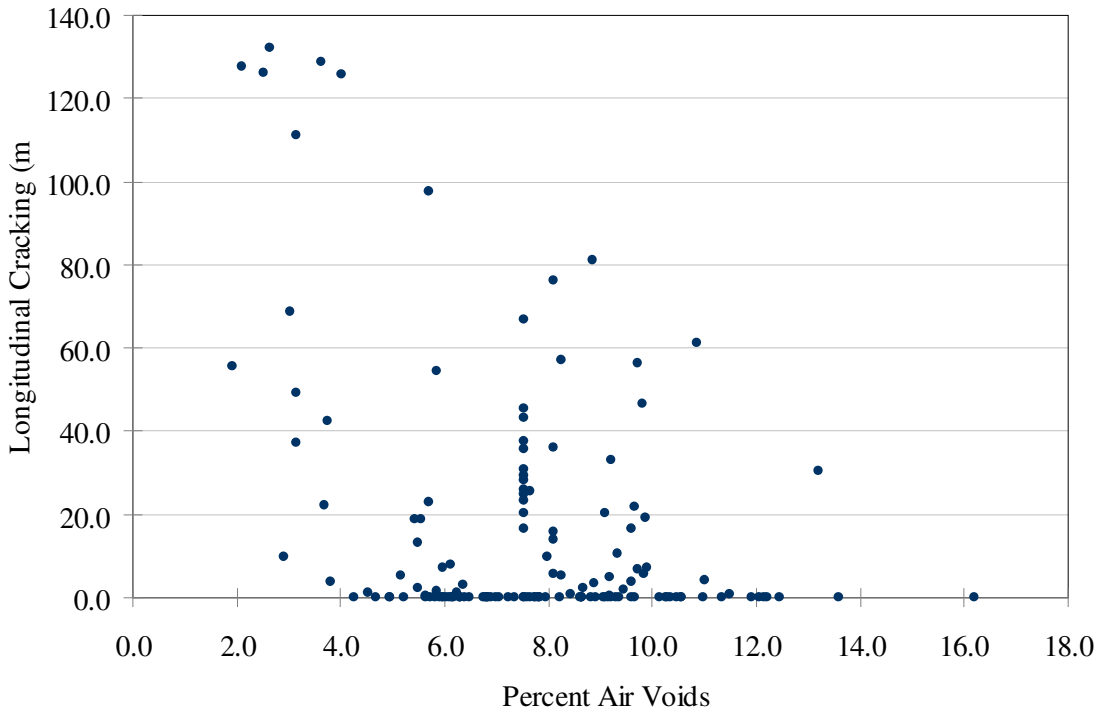


Figure 6.9 Longitudinal cracking vs. in-situ air voids content for all SPS-1 sections in LTPP database

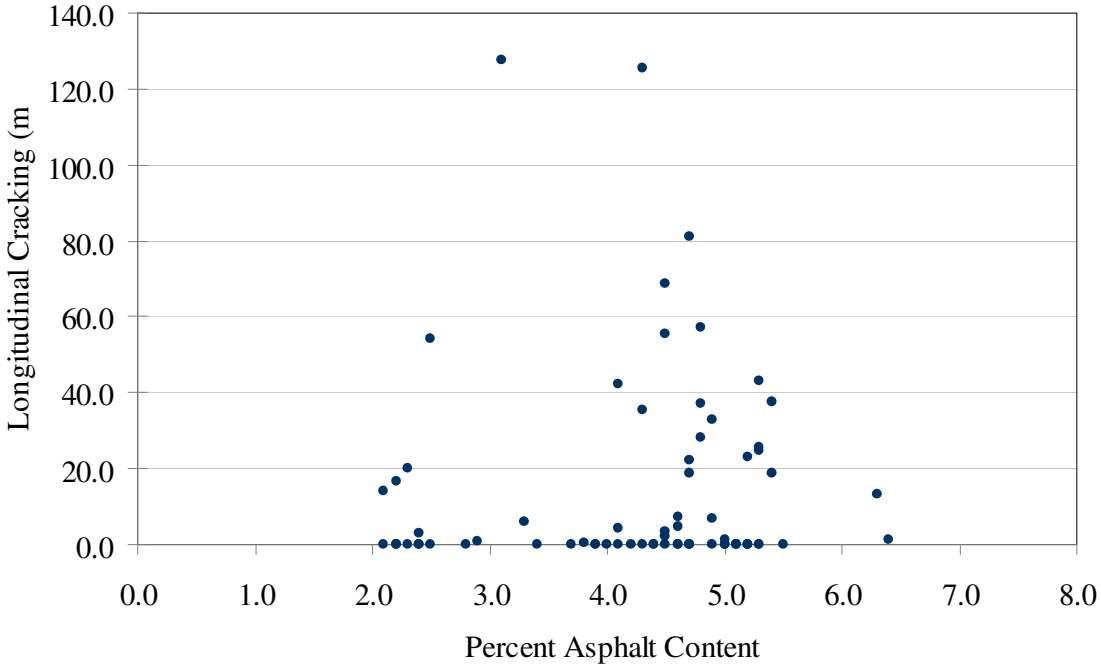


Figure 6.10 Longitudinal cracking vs. asphalt content for all SPS-1 sections in LTPP database

6.2 Analysis Using Percent-Within-Limits Concept

In a QA program, like the one that Michigan uses, a portion of payment completely depends on the percent within limits (PWL) achieved for one quality characteristic. For example, 40% of payment in Michigan flexible pavement construction depends on the in-situ density achieved immediately after final rolling. Therefore, the PWL for in-situ density is expected to be directly related to performance. Although PWL is calculated from quality achieved (in-situ density in this example), it takes into account mean as well as deviation from the mean. PWL of any project should be indicative of the expected performance irrespective of the site (state) and the HMA mix used. Therefore, relationships between PWL for various quality characteristics and performance would be of interest in this project.

A study of end-result specifications from many states showed that the majority of states follow statistical specifications using PWL. It was also observed that the specification limits used by the states are very similar to the ones being used in the state of Michigan. For example, most of the states use 91.5% or 92% as their lower specification limits for in-situ density, target ±0.4% or ±0.5% as limits for asphalt content and 4±1% or 4±1.2% for air voids at N_{design} for plant HMA samples. Also, the procedures followed by different states for calculating PWL are almost always identical. Therefore, it is possible to apply a common procedure to calculate PWL to data from other states, and then relate them to observed performance.

The inventory database in LTPP has mean bulk specific gravity for the as-placed mixture along with the maximum specific gravity. Therefore, mean in-situ density can be calculated for these projects. The same database also provides standard deviations of bulk specific gravity and number of samples used for calculating the standard deviations. However, standard deviations have not been provided for all the projects. Calculation of percent within limits requires mean and standard deviation for the projects. Therefore, those projects which had both types of data were filtered and extracted from the database. This filtering reduced the available data points from 2027 to 306.

Cracking performance for the projects is reported in the LTPP “monitoring” data table and is reported in the form of low, medium and high severity fatigue, longitudinal and transverse cracking. LTPP recommends that if distinction between severity levels does not have to be made in an analysis they can be added together. This was done to obtain the three categories of distress

cracking. Then, the projects which had enough details for calculation of percent-within-limits in the inventory database were matched to those which had cracking performance reported.

Different projects have different ages, and distress surveys are collected at different times. To be able to compare cracking performance, it was first decided that a normalized cracking be calculated. Therefore, time histories of each of the projects reported were separately plotted for each project and the area under the fatigue cracking curve was calculated. Thereafter, the area was divided by the maximum age at the time of final survey. PWL values for those projects were then plotted against the normalized fatigue cracking performance. However, the plots did not show observable trends in terms of whether a change in PWL, which means change in quality of construction, reflects in cracking performance. Therefore, to avoid any distorting effect that the normalization process may have, three dimensional plots were generated with each data point having its own time in the third dimension.

It was also found that standard plotting options available in Microsoft Excel or Matlab do not represent such data very well. Therefore, a new plotting system was developed in-house using Matlab. This allowed plotting of each point as a column in the three dimensional space. Each column or point on the X-Y plane is plotted in color which is proportional to the magnitude of the distress being plotted. The color coding of the magnitude is shown in the color bar next to the plots. It should be noted that color coding is based on the range of distress values plotted in each plot, and therefore, the same color in two different plots may represent different magnitudes.

Figure 6.11 shows a three-dimensional plot of fatigue cracking at different times during the service life of projects versus percent-within-limits for in-situ density. This plot has 306 data points. However, any one project can have more than one point corresponding to different times at which the distress was measured. According to Figure 6.11, projects having PWL close to or equal to 100% show more distress than those with lower PWL values. However, this is not necessarily true. In QA programs generally a window around the target for each of the quality characteristic is allowed. As long as the quality characteristics remain within that window PWL values would be 100%. Therefore, a very large percentage of projects would have 100% PWL even if the mean and standard deviations of the quality characteristics may be varying. Also, many of such points with 100% or near-100% PWL plot right on top of each other and the points which have more fatigue cracking become more visible while hiding other points.

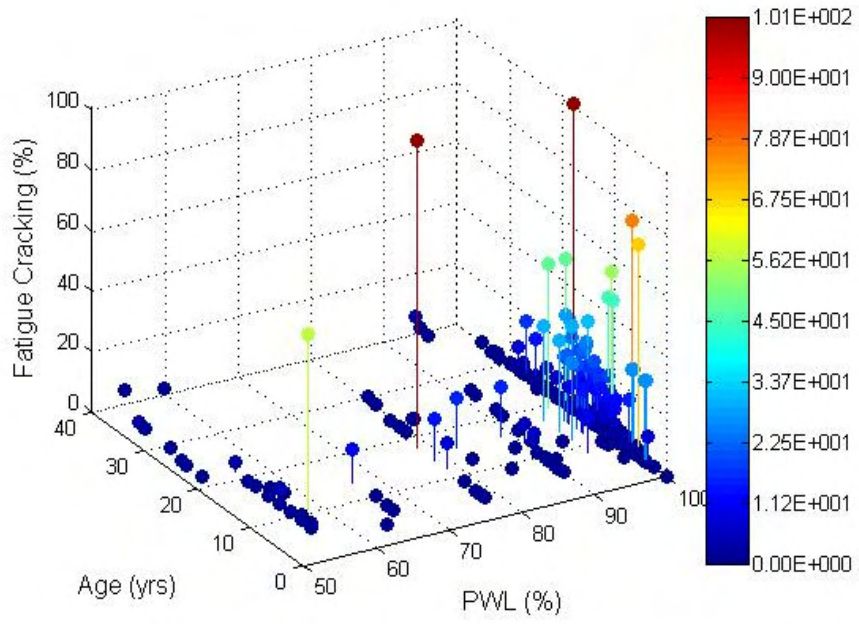


Figure 6.11 PWL (in-situ density) and fatigue cracking

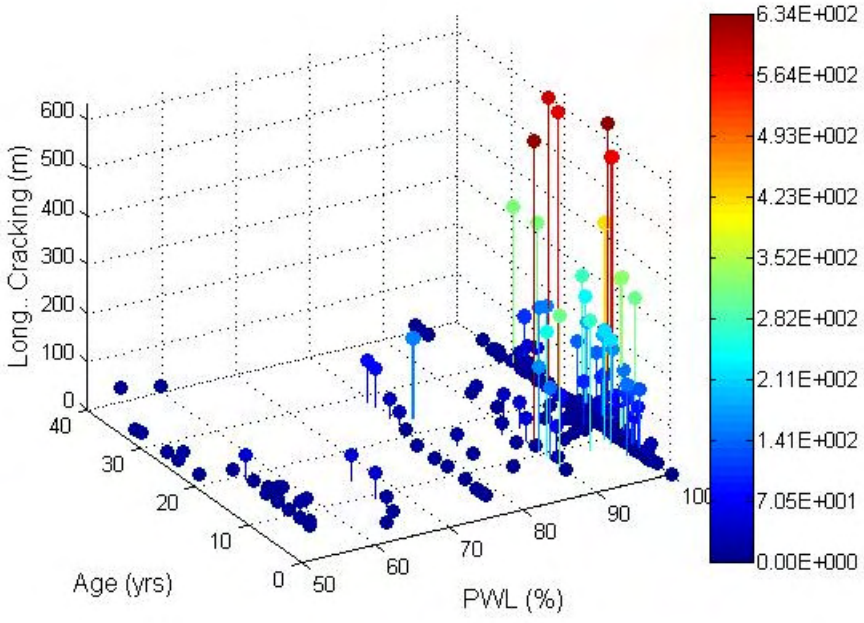


Figure 6.12 PWL (in-situ density) and longitudinal cracking

Figure 6.11 does show several projects for which PWL values are much lower than 100%, in some cases as low as 50%, which is the trigger point for "remove and replace" in QA programs, and yet they have no or almost no fatigue cracking. This is a very important

observation from a QA point of view. Figures 6.12 and 6.13 also show many similar projects. This corroborates with some of the results that we have seen from MEPDG runs as well. Fatigue cracking occurs because of repeated loading and picks up towards the later phases of the pavement. Therefore, it is possible that two different projects may have been constructed to different quality levels, and still neither of them shows any cracking at the end of the 10th year. It is believed that if the pavements are monitored for 20 years or more, the differences would start becoming more visible in terms of magnitude of cracking at that age and the rate of deterioration after a certain level of initial deterioration.

Presence of data points with lower PWL with good cracking performance does raise issues about PWL being the criteria for payment. While PWL certainly represents quality achieved during construction, the design itself may be too good for the design life of the pavement. Our design methods are largely empirical in nature. The pavement constructed with such a design may perform well for much longer or shorter than the theoretical design life associated with the design. This issue has been further studied using simulation and discussed in greater detail in the following chapters of this report.

Figure 6.13 shows the results for transverse cracking. The dark brown colored columns represent cases where either there was no transverse cracking or the crack spacing would have been greater than 500 feet which is greater than the length of each LTPP section.

Figure 6.14 shows rutting performance for several projects against in-situ density. This plot also does not show any clear trend between PWL (in-situ density) and rutting. In this plot, there are a total of 433 points out of which 323 points have PWL higher than 95% and 281 of them have PWL higher than 99%.

The LTPP database also has plant air voids documented for several projects. Some of these projects have standard deviations also reported for air voids. All such project data points, numbering 452, were filtered and extracted from the database. These projects were then matched with their cracking and rutting performance from the monitoring data table and their age from the inventory data table. Figure 6.15 presents the three-dimensional plot of fatigue performance versus PWL (plant air voids) and age. This plot shows a greater number of projects which have poor PWL (plant air voids) having more fatigue cracking compared to those with PWL values closer to 100%. In most of the states, including Michigan, any projects having PWL between 90%

and 100% are not penalized; i.e., they are paid in full. Therefore, in this case, most of the projects being paid in full did not see much fatigue cracking. This tends to show that PWL (plant air voids) may be directly related to cracking performance. There are several projects, though, which have lower PWL, which means that they would have been penalized even though they do not show fatigue cracking.

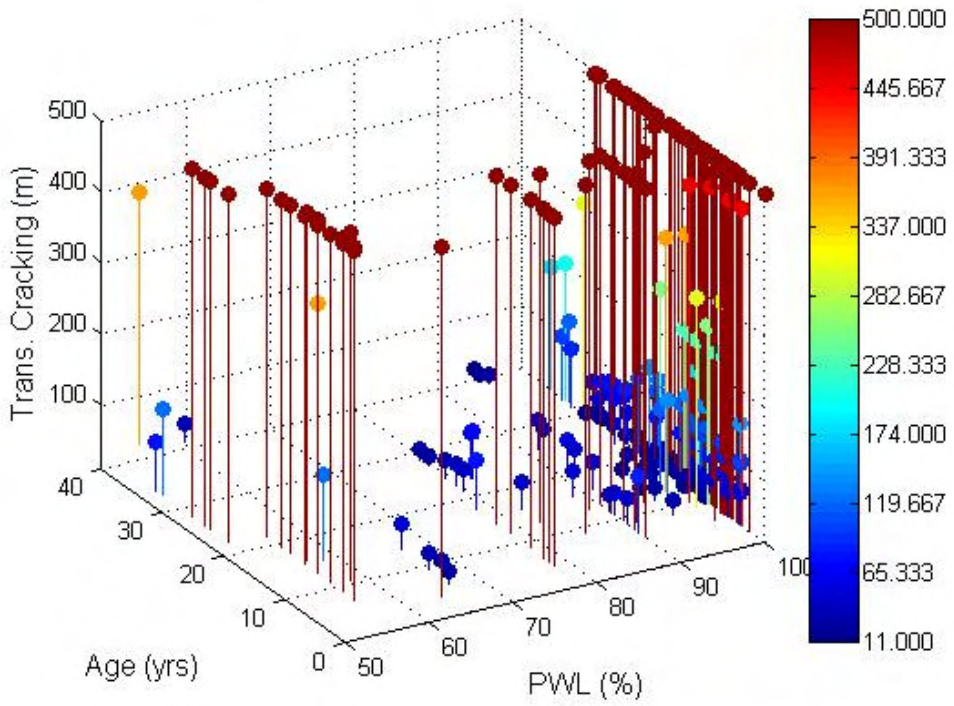


Figure 6.13 PWL (in-situ density) and transverse crack Spacing

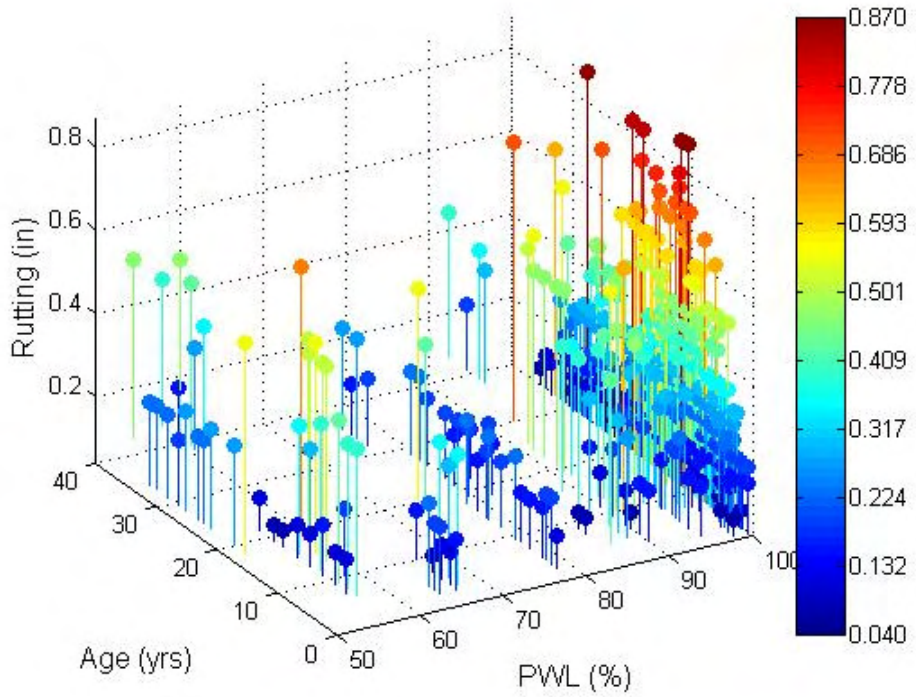


Figure 6.14 PWL (in-situ density) and rutting

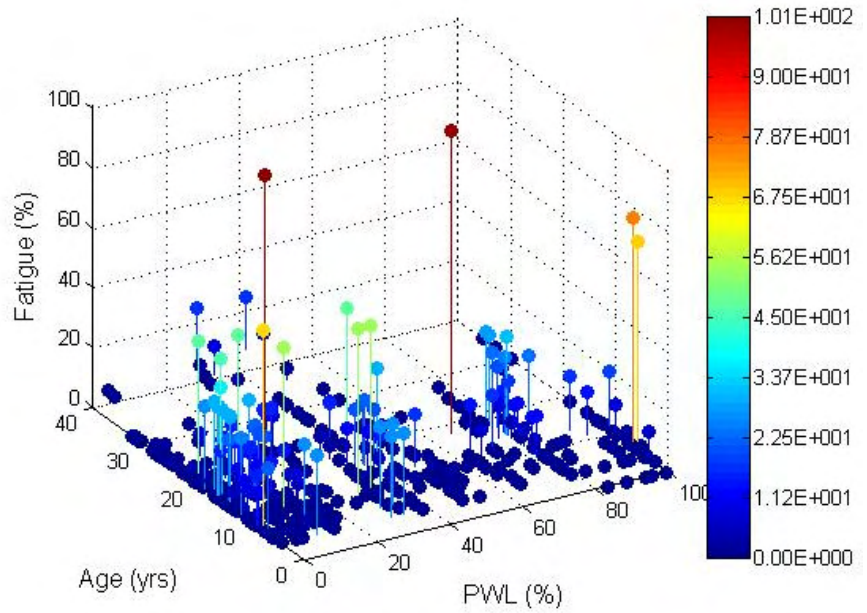


Figure 6.15 PWL (plant air voids) and fatigue cracking

Figures 6.16 and 6.17 present plots of longitudinal cracking and transverse cracking against PWL (plant air voids). Both plots show that as PWL (plant air voids) become higher, i.e. air voids are within the specification limits to a greater extent, there is better longitudinal and transverse cracking performance. Each plot has 453 points represented. Figure 6.18 shows rutting performance, which shows an increase in rutting with time; however, it does not show any observable relationship between PWL (plant air voids) and rutting.

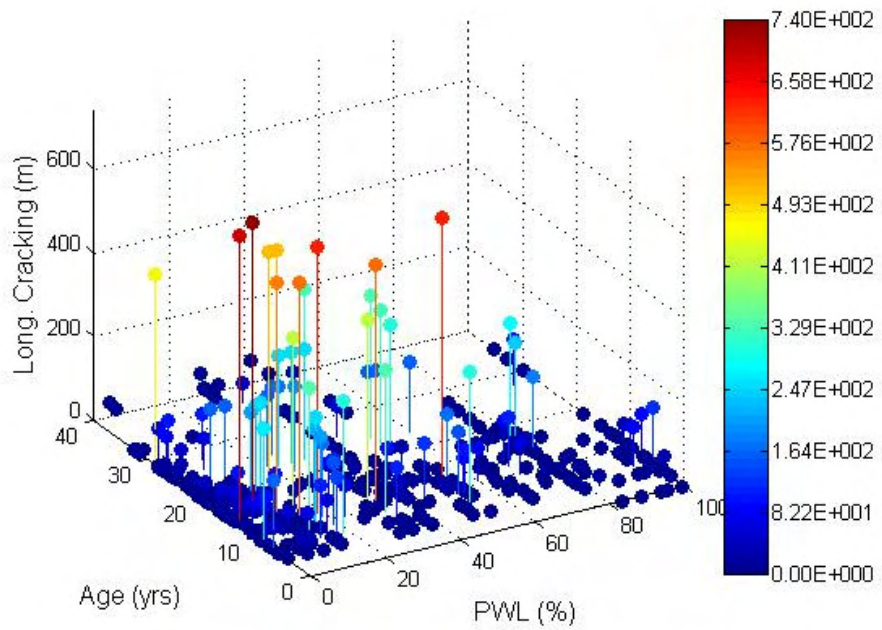


Figure 6.16 PWL (plant air voids) and longitudinal cracking

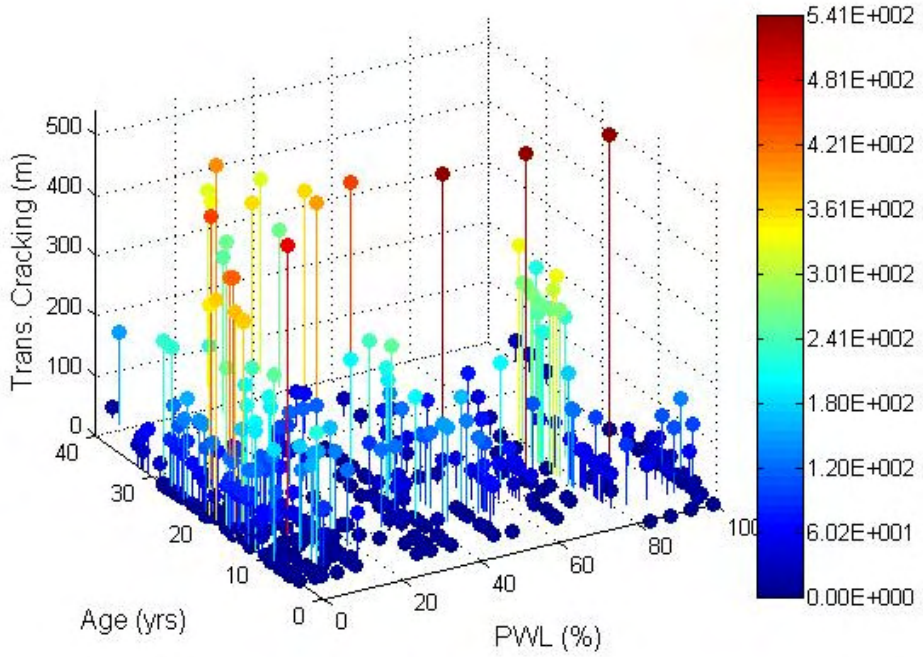


Figure 6.17 PWL (plant air voids) and transverse cracking

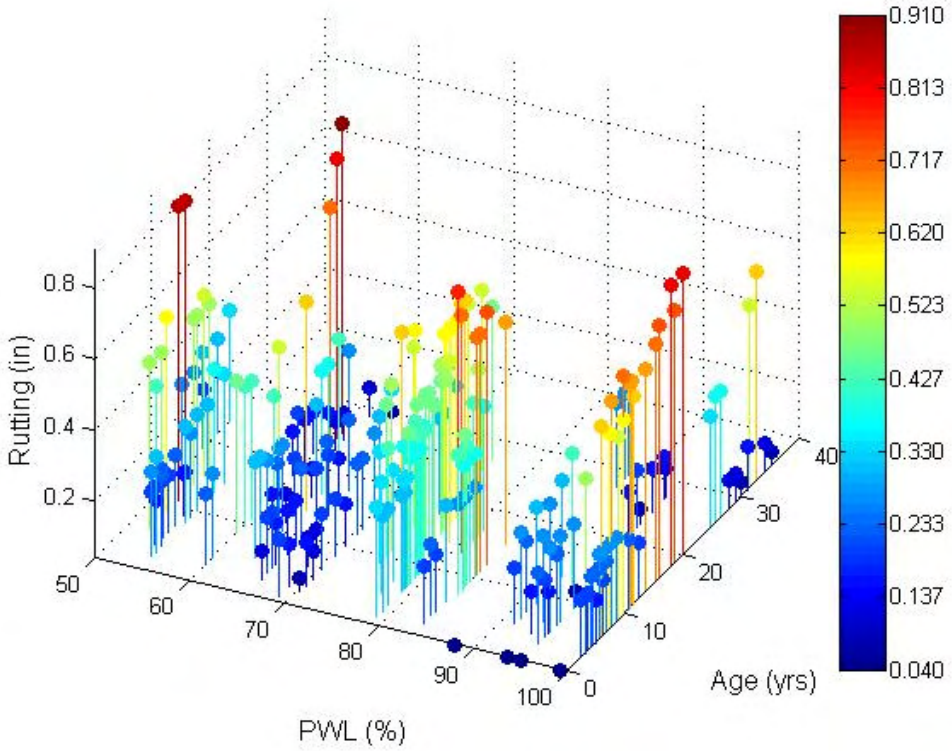


Figure 6.18 PWL (plant air voids) and rutting

6.3 Conclusions

This chapter presented details of empirical data analysis that was performed to determine if the quality characteristics most commonly used in quality assurance programs, such as those used in Michigan, affect pavement performance and can therefore be used to estimate the expected performance. If yes, then a quality assurance program can use suitable engineering and logistics model to rationally determine payment being made to the contractor and minimize losses because of poor quality construction by the contractor.

In this chapter, it was shown that analysis using empirical data generally cannot be used to develop a performance model. Such analysis can, at best, show some trends as to how a particular quality characteristic can affect certain aspects of pavement performance.

The analysis of data extracted from the LTPP database showed that pavements with low (< 4%) or high (>8%) in-situ air voids have higher probability of rutting than those in which air voids fall in the range of 4 to 8%. This indicates that in-situ air voids immediately after construction is an important test for a quality assurance program. Contrary to intuition, the asphalt content data did not show any clear trend in its effect on rutting. This is most likely because of not having data points corresponding to very low or very high asphalt content compared to the corresponding design values. Also, it is not possible to determine the contribution of the HMA layer to rutting directly from the LTPP; therefore some of the observed rutting may be caused by the unbound layers. The trends observed in the case of fatigue cracking and longitudinal cracking were not very clear. This is most likely due to the fact that several other factors are also affecting the fatigue performance, and these cannot be controlled in the analysis.

An argument was presented in this chapter as to how percent-within-limits for a particular quality characteristic should be related to pavement performance to rationally justify its use in pay factor equations. The analysis did not show clear relationships between PWL for in-situ density and fatigue cracking, longitudinal cracking and transverse cracking. Based on the plots showing PWL (plant air voids) versus cracking performance, it can be stated that while lower PWL does not necessarily mean poor performance, the probability of better performance certainly goes up with higher PWL. The analysis also showed that as PWL (plant air voids) become higher, i.e. air voids are within the specification limits to a greater extent, there is better

longitudinal and transverse cracking performance. No observable relationship between PWL (plant air voids) and rutting was found in the analysis.

In conclusion, the results derived from the empirical data analysis presented in this chapter are mixed and unclear. Therefore, a mechanistic-empirical analysis needs to be performed to derive firm conclusions required for assessing quality assurance programs such as the one being used by the state of Michigan.

Chapter 7: Mechanistic-Empirical Analysis

7.1 Purpose

A major issue with this research effort was the lack of construction data from actual projects. As reported earlier in the report, asphalt pavement projects that are older than 10 years were not built with Superpave mixes. The current QA program as adopted by MDOT is based on Superpave. Therefore, even if some test data is available for these older projects they cannot be easily used to assess the current QA program. For example, in the case of older projects, no data would be available for air voids and VMA at N_{design}. Both these factors account for forty five percent of the total pay factor. When the projects are less than 10 years old, it is generally difficult to assess as to how they are performing because they have not gone through their design service life unless they are already performing poorly. In addition to these factors, even in the case of very well controlled experiments like SPS-1, it is not very easy to draw clear relationships between cause and effect because of the number of variables involved, several of which may not be controllable.

An alternative approach would be to assess the QA program by relating the QA parameters to expected performance if reliable predictions can be made for performance. The ME-PDG provides an opportunity to be able to predict pavement performance when sufficiently accurate inputs are provided. While ME-PDG predictions may not have a high level of accuracy, several studies have shown that M-E PDG predictions can be reasonable provided that proper (local) calibration is done. Also, it is important to mention that ME-PDG uses the best knowledge available today to the pavement engineering community. Given the lack of real construction data and recognizing that many other factors (other than QA factors) may be affecting in-service pavement performance over time, it is useful to attempt this exercise using the tools provided in ME-PDG software.

7.2 Analysis Approach

The Superpave mix design process requires specimen preparation and testing of the mixture being designed. This is done in order to determine the optimal asphalt content and achieve the desired 4 percent air voids. These tests lead to valuable information not just in the

form of optimal asphalt content but also in terms of differences in mix characteristics that would result from the use of actual aggregate blends (not just theoretical blending) and varying asphalt content in the mix. These provide realistic inputs for predicting the change in performance as a result of such changes in the mix, thus simulating what actually happens during the construction process.

The following paragraphs present the steps that were taken to come up with inputs based on actual testing. These inputs represent the variation that would occur because of material variability, construction process itself, operator error etc. For instance, the bold line in Figure 7.1 shows gradation curve as target in a project. However, the actual gradation of the aggregates in the mix could be slightly different from the target and could look like any of the other gradation curves in the same figure. In reality, the other gradation curves were taken from other projects in this case. The QA program is meant to identify these deviations and also limit them so that the overall pavement quality does not diminish appreciably. In this strategy, we introduce such deviations in the mixture and use ME-PDG to quantify the change in performance. Assessing the type and magnitude of changes in performance provides knowledge about the level of variabilities that would be acceptable under a QA program. In all the analyses that follow, only MDOT mixes have been used. An example of this strategy is presented below.

Step 1: Identify the project, or a typical example of the type of project that needs to be analyzed with respect to the QA program. We chose the following three projects with varying performance.

(i) Section ID 29581W located in Lansing, MI and constructed in 1995
(ii) Section ID 18890N located in Ludington, MI and constructed in 1989
(iii) SPS-1 site 0117 located near Lansing, MI and constructed in 1994

Step 2: Match the mix used in the identified projects as closely as possible to a Superpave mix currently being used by MDOT. The criteria used for this was

(i) Traffic level
(ii) Climatic conditions
(iii) Asphalt grade
(iv) Aggregate gradation of the mix
(v) Asphalt content of the mix

Using these criteria, the following mixes were identified from the inventory of recently used Michigan mixes available to us.

(i) Section ID 29581W: Mix Design# 07MD161, 07MD090, 07MD0170 and 07MD0234
(ii) Section ID 18890N: Mix Design# 07MD046 and 07MD071
(iii) SPS-1 site 0117: Mix Design# 07MD042 and 07MD290

The list above gives project identification number/name followed by some mix design numbers. The first mix design number corresponds to the MDOT mix that was found to be closest to the original mix used in the project based on the criteria listed earlier in this section. The remaining mixes are similar in aggregate gradation, asphalt grade etc. but have some differences that would be representative of the difference that would exist in a real project because of variability in aggregate gradation etc. In summary, the first listed mix is expected to represent the original mix that was supposed to be used in the construction process; i.e., the target mix, referred to as "target mix" henceforth. The remaining mixes are the ones that could result because of material and construction variability.

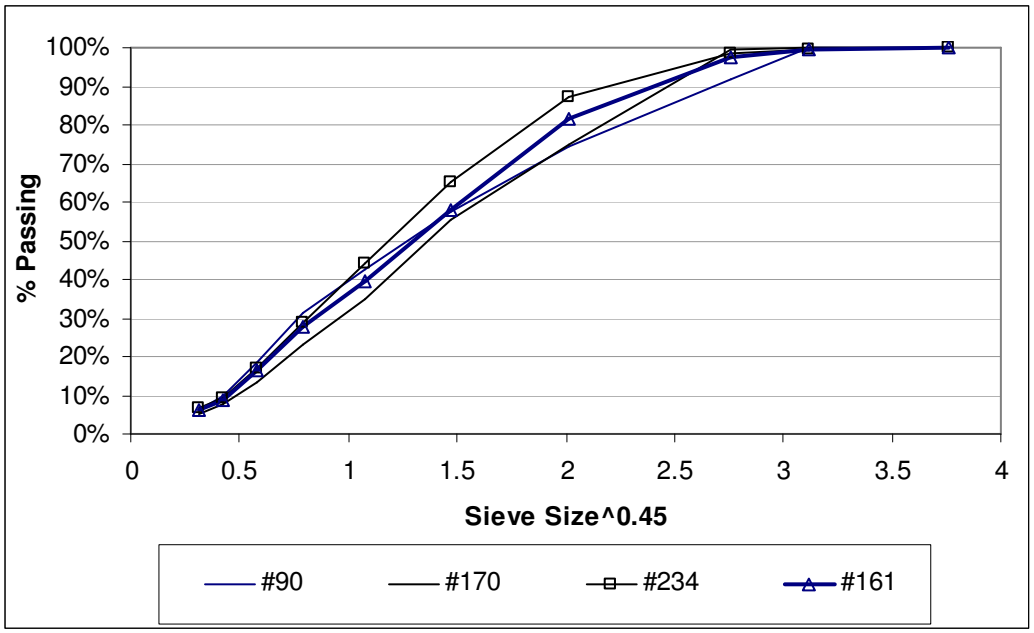


Figure 7. 1 Aggregate gradation for the four mixes chosen for the first project

Step 3: Use the target mix to determine inputs for ME-PDG software. The inputs available from the mix design would be for the mix as produced in the plant like asphalt content and aggregate gradation. Also, choose other levels of asphalt content that can result because of variability in the production process. Table 7.1 presents the asphalt content measurements from actual projects

constructed in 2002. Superpave mix designs were used in all these projects. In this analysis, five levels of asphalt content were chosen:

(i) Optimum AC – 0.5%
(ii) Optimum AC – 0.25%
(iii) Optimum AC
(iv) Optimum AC +0.25%
(v) Optimum AC +0.5%

Table 7.1 Asphalt content measurements from actual projects

	Measured Asphalt Content (%)		Target (%)	Deviation from Target (%)	
	Min	Max			
Project 1	5.35	5.74	5.5	-0.15	0.24
Project 2	5.3	6.1	5.6	-0.3	0.5
Project 3	4.5	4.9	4.6	-0.1	0.3
Project 4	5.2	5.7	5.4	-0.2	0.3
Project 5	4.5	4.8	4.5	0	0.3
Project 6	5.3	5.7	5.4	-0.1	0.3
Project 7	5	5.4	5.2	-0.2	0.2
Project 8	5.2	5.4	5.3	-0.1	0.1
Project 9	5.2	5.6	5.4	-0.2	0.2
Project 10	4.4	4.8	4.7	-0.3	0.1

Step 4: Choose the range of air voids which can result because of the lay down of the mix, i.e. in-situ air voids. Table 7.2 presents the in-situ air voids from the same projects as those mentioned in Table 7.1. Based on these actual observations the following levels of in-situ air voids were chosen in this analysis.

(i) 12.0%
(ii) 9.5%
(iii) 7.0%
(iv) 4.5%
(v) 2.0%

Table 7.2 In-situ air void measurements from actual projects

	Measured In-situ Air Voids (%)		Standard Deviation (%)
	Min	Max	
Project 1	2.7	10.2	1.45
Project 2	2.9	9.1	1.80
Project 3	2.6	9.0	1.25
Project 4	4.0	8.6	0.98
Project 5	2.8	9.6	1.75
Project 6	2.0	10.0	1.84
Project 7	2.5	6.1	1.18
Project 8	4.4	11.9	1.15
Project 9	4.2	10.5	1.39
Project 10	3.0	10.1	1.23

Step 5: Make a matrix with all possible combinations of the levels chosen for asphalt content and air voids. This should be followed by calculating the other parameters with reference to the test results from the mix design procedure and using the volumetric relationships of the mixes. For example, for cases in the matrix with asphalt content less than the optimal value, one can refer to the volumetric properties corresponding to that from the table obtained during the Superpave mix design procedure. The said table lists air voids, VMA, VFA at N_{design} gyrations for the same mix but with varying amounts of asphalt content. This step helps ensure that the inputs selected for ME-PDG are characteristics of the original mix even though the mix components may not have been added in the right proportion according to the design. This is the same as what actually happens in real mix production and construction of the pavements.

Step 6: Step 5 should be repeated for each mix identified for each of the projects. However, when determining the inputs for the mix other than the target mix, optimal asphalt content for the target mix should be used. Tables 7.3, 7.4 and 7.5 present the matrices for the three cases that we analyzed.

Step 7: Run ME-PDG for each case within the matrix prepared in steps 5 and 6.

Step 8: Analyze a project by simulating quality characteristic test results from QA point of view and then relate them to performance.

It is important to determine the accuracy of performance predicted by MEPDG for this strategy of analysis to be successful. Figures 7.2 through 7.6 compare the performance predicted by MEPDG software with the actual performance observed. For the sake of brevity, only some cases have been presented here. Also, in some cases, the data for actual performance seem to be erroneous. Therefore, they could not be used for this verification purpose. These comparisons show that there is reasonable agreement between actual and predicted performance. It is also possible to calibrate the models used in MEPDG using local data. That would certainly further improve the accuracy of predictions.

Table 7.3 Matrix for the first project selected for analysis

Run No.	Design#	AC	Density (in-situ) (% Gmm)	% Retained on 3/4 inch	% Retained on 3/8 inch	% Retained on #4	P#200	Asphalt Grade	Pbe (vol.) (%)	In-Situ Air Voids (%)	Tot. Ut. Wt. (lb/ft3)
1	07MD161	JMF-0.5	88	0	2.4	18.4	6	PG 64-22	10.270	12.0	136.545
2	07MD161	JMF-0.5	90.5	0	2.4	18.4	6	PG 64-22	10.562	9.5	140.425
3	07MD161	JMF-0.5	93	0	2.4	18.4	6	PG 64-22	10.853	7.0	144.304
4	07MD161	JMF-0.5	95.5	0	2.4	18.4	6	PG 64-22	11.145	4.5	148.183
5	07MD161	JMF-0.5	98	0	2.4	18.4	6	PG 64-22	11.437	2.0	152.062
6	07MD161	JMF-0.25	88	0	2.4	18.4	6	PG 64-22	10.764	12.0	136.015
7	07MD161	JMF-0.25	90.5	0	2.4	18.4	6	PG 64-22	11.070	9.5	139.879
8	07MD161	JMF-0.25	93	0	2.4	18.4	6	PG 64-22	11.376	7.0	143.743
9	07MD161	JMF-0.25	95.5	0	2.4	18.4	6	PG 64-22	11.681	4.5	147.607
10	07MD161	JMF-0.25	98	0	2.4	18.4	6	PG 64-22	11.987	2.0	151.472
11	07MD161	JMF	88	0	2.4	18.4	6	PG 64-22	11.257	12.0	135.508
12	07MD161	JMF	90.5	0	2.4	18.4	6	PG 64-22	11.576	9.5	139.357
13	07MD161	JMF	93	0	2.4	18.4	6	PG 64-22	11.896	7.0	143.200
14	07MD161	JMF	95.5	0	2.4	18.4	6	PG 64-22	12.216	4.5	147.056
15	07MD161	JMF	98	0	2.4	18.4	6	PG 64-22	12.536	2.0	150.905
16	07MD161	JMF+0.25	88	0	2.4	18.4	6	PG 64-22	11.736	12.0	135.005
17	07MD161	JMF+0.25	90.5	0	2.4	18.4	6	PG 64-22	12.069	9.5	138.840
18	07MD161	JMF+0.25	93	0	2.4	18.4	6	PG 64-22	12.403	7.0	142.675
19	07MD161	JMF+0.25	95.5	0	2.4	18.4	6	PG 64-22	12.736	4.5	146.511
20	07MD161	JMF+0.25	98	0	2.4	18.4	6	PG 64-22	13.070	2.0	150.346
21	07MD161	JMF+0.5	88	0	2.4	18.4	6	PG 64-22	12.224	12.0	134.522
22	07MD161	JMF+0.5	90.5	0	2.4	18.4	6	PG 64-22	12.571	9.5	138.344
23	07MD161	JMF+0.5	93	0	2.4	18.4	6	PG 64-22	12.918	7.0	142.165
24	07MD161	JMF+0.5	95.5	0	2.4	18.4	6	PG 64-22	13.266	4.5	145.987
25	07MD161	JMF+0.5	98	0	2.4	18.4	6	PG 64-22	13.613	2.0	149.809
26	07MD090	JMF-0.5	88	0	8.1	25.7	6.1	PG 64-22	10.486	12.0	136.625
27	07MD090	JMF-0.5	90.5	0	8.1	25.7	6.1	PG 64-22	10.784	9.5	140.507
28	07MD090	JMF-0.5	93	0	8.1	25.7	6.1	PG 64-22	11.082	7.0	144.388
29	07MD090	JMF-0.5	95.5	0	8.1	25.7	6.1	PG 64-22	11.380	4.5	148.269
30	07MD090	JMF-0.5	98	0	8.1	25.7	6.1	PG 64-22	11.678	2.0	152.151
31	07MD090	JMF-0.25	88	0	8.1	25.7	6.1	PG 64-22	10.982	12.0	136.099
32	07MD090	JMF-0.25	90.5	0	8.1	25.7	6.1	PG 64-22	11.294	9.5	139.966
33	07MD090	JMF-0.25	93	0	8.1	25.7	6.1	PG 64-22	11.606	7.0	143.832
34	07MD090	JMF-0.25	95.5	0	8.1	25.7	6.1	PG 64-22	11.918	4.5	147.699
35	07MD090	JMF-0.25	98	0	8.1	25.7	6.1	PG 64-22	12.230	2.0	151.565
36	07MD090	JMF	88	0	8.1	25.7	6.1	PG 64-22	11.479	12.0	135.612
37	07MD090	JMF	90.5	0	8.1	25.7	6.1	PG 64-22	11.805	9.5	139.465
38	07MD090	JMF	93	0	8.1	25.7	6.1	PG 64-22	12.131	7.0	143.318
39	07MD090	JMF	95.5	0	8.1	25.7	6.1	PG 64-22	12.457	4.5	147.170
40	07MD090	JMF	98	0	8.1	25.7	6.1	PG 64-22	12.783	2.0	151.023
41	07MD090	JMF+0.25	88	0	8.1	25.7	6.1	PG 64-22	11.957	12.0	135.092
42	07MD090	JMF+0.25	90.5	0	8.1	25.7	6.1	PG 64-22	12.296	9.5	138.929
43	07MD090	JMF+0.25	93	0	8.1	25.7	6.1	PG 64-22	12.636	7.0	142.767
44	07MD090	JMF+0.25	95.5	0	8.1	25.7	6.1	PG 64-22	12.976	4.5	146.605
45	07MD090	JMF+0.25	98	0	8.1	25.7	6.1	PG 64-22	13.316	2.0	150.443
46	07MD090	JMF+0.5	88	0	8.1	25.7	6.1	PG 64-22	12.448	12.0	134.653
47	07MD090	JMF+0.5	90.5	0	8.1	25.7	6.1	PG 64-22	12.802	9.5	138.479
48	07MD090	JMF+0.5	93	0	8.1	25.7	6.1	PG 64-22	13.155	7.0	142.304
49	07MD090	JMF+0.5	95.5	0	8.1	25.7	6.1	PG 64-22	13.509	4.5	146.129
50	07MD090	JMF+0.5	98	0	8.1	25.7	6.1	PG 64-22	13.863	2.0	149.955

Table 7.4 Matrix for the second project selected for analysis

Run No.	Design#	AC	Density (in-situ) (% Gmm)	% Retained on 3/4 inch	% Retained on 3/8 inch	% Retained on #4	P#200	Asphalt Grade	Pbe (vol.) (%)	In-Situ Air Voids (%)	Tot. Ut. Wt. (lb/ft3)
1	07MD046	JMF-0.5	88	0	15.3	31.1	5.4	PG64-28	9.203	12.0	137.287
2	07MD046	JMF-0.5	90.5	0	15.3	31.1	5.4	PG64-28	9.464	9.5	141.187
3	07MD046	JMF-0.5	93	0	15.3	31.1	5.4	PG64-28	9.726	7.0	145.088
4	07MD046	JMF-0.5	95.5	0	15.3	31.1	5.4	PG64-28	9.987	4.5	148.988
5	07MD046	JMF-0.5	98	0	15.3	31.1	5.4	PG64-28	10.248	2.0	152.888
6	07MD046	JMF-0.25	88	0	15.3	31.1	5.4	PG64-28	9.706	12.0	136.805
7	07MD046	JMF-0.25	90.5	0	15.3	31.1	5.4	PG64-28	9.982	9.5	140.691
8	07MD046	JMF-0.25	93	0	15.3	31.1	5.4	PG64-28	10.258	7.0	144.578
9	07MD046	JMF-0.25	95.5	0	15.3	31.1	5.4	PG64-28	10.533	4.5	148.464
10	07MD046	JMF-0.25	98	0	15.3	31.1	5.4	PG64-28	10.809	2.0	152.351
11	07MD046	JMF	88	0	15.3	31.1	5.4	PG64-28	10.202	12.0	136.243
12	07MD046	JMF	90.5	0	15.3	31.1	5.4	PG64-28	10.492	9.5	140.114
13	07MD046	JMF	93	0	15.3	31.1	5.4	PG64-28	10.782	7.0	143.984
14	07MD046	JMF	95.5	0	15.3	31.1	5.4	PG64-28	11.072	4.5	147.855
15	07MD046	JMF	98	0	15.3	31.1	5.4	PG64-28	11.362	2.0	151.725
16	07MD046	JMF+0.25	88	0	15.3	31.1	5.4	PG64-28	10.696	12.0	135.732
17	07MD046	JMF+0.25	90.5	0	15.3	31.1	5.4	PG64-28	11.000	9.5	139.588
18	07MD046	JMF+0.25	93	0	15.3	31.1	5.4	PG64-28	11.304	7.0	143.444
19	07MD046	JMF+0.25	95.5	0	15.3	31.1	5.4	PG64-28	11.608	4.5	147.300
20	07MD046	JMF+0.25	98	0	15.3	31.1	5.4	PG64-28	11.912	2.0	151.156
21	07MD046	JMF+0.5	88	0	15.3	31.1	5.4	PG64-28	11.185	12.0	135.241
22	07MD046	JMF+0.5	90.5	0	15.3	31.1	5.4	PG64-28	11.503	9.5	139.083
23	07MD046	JMF+0.5	93	0	15.3	31.1	5.4	PG64-28	11.821	7.0	142.925
24	07MD046	JMF+0.5	95.5	0	15.3	31.1	5.4	PG64-28	12.139	4.5	146.767
25	07MD046	JMF+0.5	98	0	15.3	31.1	5.4	PG64-28	12.456	2.0	150.609
26	07MD071	JMF-0.5	88	0	10.1	27.8	5.3	PG64-28	9.488	12.0	135.612
27	07MD071	JMF-0.5	90.5	0	10.1	27.8	5.3	PG64-28	9.757	9.5	139.464
28	07MD071	JMF-0.5	93	0	10.1	27.8	5.3	PG64-28	10.027	7.0	143.317
29	07MD071	JMF-0.5	95.5	0	10.1	27.8	5.3	PG64-28	10.296	4.5	147.170
30	07MD071	JMF-0.5	98	0	10.1	27.8	5.3	PG64-28	10.566	2.0	151.022
31	07MD071	JMF-0.25	88	0	10.1	27.8	5.3	PG64-28	9.988	12.0	135.080
32	07MD071	JMF-0.25	90.5	0	10.1	27.8	5.3	PG64-28	10.272	9.5	138.918
33	07MD071	JMF-0.25	93	0	10.1	27.8	5.3	PG64-28	10.556	7.0	142.755
34	07MD071	JMF-0.25	95.5	0	10.1	27.8	5.3	PG64-28	10.839	4.5	146.593
35	07MD071	JMF-0.25	98	0	10.1	27.8	5.3	PG64-28	11.123	2.0	150.430
36	07MD071	JMF	88	0	10.1	27.8	5.3	PG64-28	10.480	12.0	134.597
37	07MD071	JMF	90.5	0	10.1	27.8	5.3	PG64-28	10.778	9.5	138.420
38	07MD071	JMF	93	0	10.1	27.8	5.3	PG64-28	11.076	7.0	142.244
39	07MD071	JMF	95.5	0	10.1	27.8	5.3	PG64-28	11.373	4.5	146.068
40	07MD071	JMF	98	0	10.1	27.8	5.3	PG64-28	11.671	2.0	149.892
41	07MD071	JMF+0.25	88	0	10.1	27.8	5.3	PG64-28	10.961	12.0	134.118
42	07MD071	JMF+0.25	90.5	0	10.1	27.8	5.3	PG64-28	11.273	9.5	137.928
43	07MD071	JMF+0.25	93	0	10.1	27.8	5.3	PG64-28	11.584	7.0	141.738
44	07MD071	JMF+0.25	95.5	0	10.1	27.8	5.3	PG64-28	11.896	4.5	145.548
45	07MD071	JMF+0.25	98	0	10.1	27.8	5.3	PG64-28	12.207	2.0	149.358
46	07MD071	JMF+0.5	88	0	10.1	27.8	5.3	PG64-28	11.454	12.0	133.628
47	07MD071	JMF+0.5	90.5	0	10.1	27.8	5.3	PG64-28	11.779	9.5	137.424
48	07MD071	JMF+0.5	93	0	10.1	27.8	5.3	PG64-28	12.104	7.0	141.221
49	07MD071	JMF+0.5	95.5	0	10.1	27.8	5.3	PG64-28	12.430	4.5	145.017
50	07MD071	JMF+0.5	98	0	10.1	27.8	5.3	PG64-28	12.755	2.0	148.813

Table 7.5 Matrix for the third project selected for analysis

Run No.	Design#	AC	Density (in-situ) (% Gmm)	% Retained on 3/4 inch	% Retained on 3/8 inch	% Retained on #4	P#200	Asphalt Grade	Pbe (vol.) (%)	In-Situ Air Voids (%)	Tot. Ut. Wt. (lb/ft3)
1	07MD042	JMF-0.5	88	0.1	24.4	58.6	4.2	PG58-22	8.516	12.0	138.243
2	07MD042	JMF-0.5	90.5	0.1	24.4	58.6	4.2	PG58-22	8.758	9.5	142.170
3	07MD042	JMF-0.5	93	0.1	24.4	58.6	4.2	PG58-22	9.000	7.0	146.098
4	07MD042	JMF-0.5	95.5	0.1	24.4	58.6	4.2	PG58-22	9.242	4.5	150.025
5	07MD042	JMF-0.5	98	0.1	24.4	58.6	4.2	PG58-22	9.484	2.0	153.953
6	07MD042	JMF-0.25	88	0.1	24.4	58.6	4.2	PG58-22	9.028	12.0	137.745
7	07MD042	JMF-0.25	90.5	0.1	24.4	58.6	4.2	PG58-22	9.284	9.5	141.658
8	07MD042	JMF-0.25	93	0.1	24.4	58.6	4.2	PG58-22	9.541	7.0	145.571
9	07MD042	JMF-0.25	95.5	0.1	24.4	58.6	4.2	PG58-22	9.797	4.5	149.485
10	07MD042	JMF-0.25	98	0.1	24.4	58.6	4.2	PG58-22	10.054	2.0	153.398
11	07MD042	JMF	88	0.1	24.4	58.6	4.2	PG58-22	9.536	12.0	137.200
12	07MD042	JMF	90.5	0.1	24.4	58.6	4.2	PG58-22	9.807	9.5	141.098
13	07MD042	JMF	93	0.1	24.4	58.6	4.2	PG58-22	10.077	7.0	144.995
14	07MD042	JMF	95.5	0.1	24.4	58.6	4.2	PG58-22	10.348	4.5	148.893
15	07MD042	JMF	98	0.1	24.4	58.6	4.2	PG58-22	10.619	2.0	152.791
16	07MD042	JMF+0.25	88	0.1	24.4	58.6	4.2	PG58-22	10.037	12.0	136.676
17	07MD042	JMF+0.25	90.5	0.1	24.4	58.6	4.2	PG58-22	10.322	9.5	140.559
18	07MD042	JMF+0.25	93	0.1	24.4	58.6	4.2	PG58-22	10.607	7.0	144.442
19	07MD042	JMF+0.25	95.5	0.1	24.4	58.6	4.2	PG58-22	10.892	4.5	148.325
20	07MD042	JMF+0.25	98	0.1	24.4	58.6	4.2	PG58-22	11.177	2.0	152.207
21	07MD042	JMF+0.5	88	0.1	24.4	58.6	4.2	PG58-22	10.529	12.0	136.186
22	07MD042	JMF+0.5	90.5	0.1	24.4	58.6	4.2	PG58-22	10.828	9.5	140.055
23	07MD042	JMF+0.5	93	0.1	24.4	58.6	4.2	PG58-22	11.127	7.0	143.924
24	07MD042	JMF+0.5	95.5	0.1	24.4	58.6	4.2	PG58-22	11.426	4.5	147.793
25	07MD042	JMF+0.5	98	0.1	24.4	58.6	4.2	PG58-22	11.726	2.0	151.662
26	07MD290	JMF-0.5	88	0	26	53.6	5	PG58-22	9.113	12.0	137.832
27	07MD290	JMF-0.5	90.5	0	26	53.6	5	PG58-22	9.371	9.5	141.748
28	07MD290	JMF-0.5	93	0	26	53.6	5	PG58-22	9.630	7.0	145.663
29	07MD290	JMF-0.5	95.5	0	26	53.6	5	PG58-22	9.889	4.5	149.579
30	07MD290	JMF-0.5	98	0	26	53.6	5	PG58-22	10.148	2.0	153.495
31	07MD290	JMF-0.25	88	0	26	53.6	5	PG58-22	9.617	12.0	137.321
32	07MD290	JMF-0.25	90.5	0	26	53.6	5	PG58-22	9.890	9.5	141.222
33	07MD290	JMF-0.25	93	0	26	53.6	5	PG58-22	10.164	7.0	145.124
34	07MD290	JMF-0.25	95.5	0	26	53.6	5	PG58-22	10.437	4.5	149.025
35	07MD290	JMF-0.25	98	0	26	53.6	5	PG58-22	10.710	2.0	152.926
36	07MD290	JMF	88	0	26	53.6	5	PG58-22	10.118	12.0	136.811
37	07MD290	JMF	90.5	0	26	53.6	5	PG58-22	10.406	9.5	140.698
38	07MD290	JMF	93	0	26	53.6	5	PG58-22	10.693	7.0	144.585
39	07MD290	JMF	95.5	0	26	53.6	5	PG58-22	10.981	4.5	148.471
40	07MD290	JMF	98	0	26	53.6	5	PG58-22	11.268	2.0	152.358
41	07MD290	JMF+0.25	88	0	26	53.6	5	PG58-22	10.609	12.0	136.295
42	07MD290	JMF+0.25	90.5	0	26	53.6	5	PG58-22	10.910	9.5	140.167
43	07MD290	JMF+0.25	93	0	26	53.6	5	PG58-22	11.212	7.0	144.039
44	07MD290	JMF+0.25	95.5	0	26	53.6	5	PG58-22	11.513	4.5	147.911
45	07MD290	JMF+0.25	98	0	26	53.6	5	PG58-22	11.814	2.0	151.783
46	07MD290	JMF+0.5	88	0	26	53.6	5	PG58-22	11.094	12.0	135.785
47	07MD290	JMF+0.5	90.5	0	26	53.6	5	PG58-22	11.409	9.5	139.643
48	07MD290	JMF+0.5	93	0	26	53.6	5	PG58-22	11.724	7.0	143.500
49	07MD290	JMF+0.5	95.5	0	26	53.6	5	PG58-22	12.039	4.5	147.358
50	07MD290	JMF+0.5	98	0	26	53.6	5	PG58-22	12.354	2.0	151.215

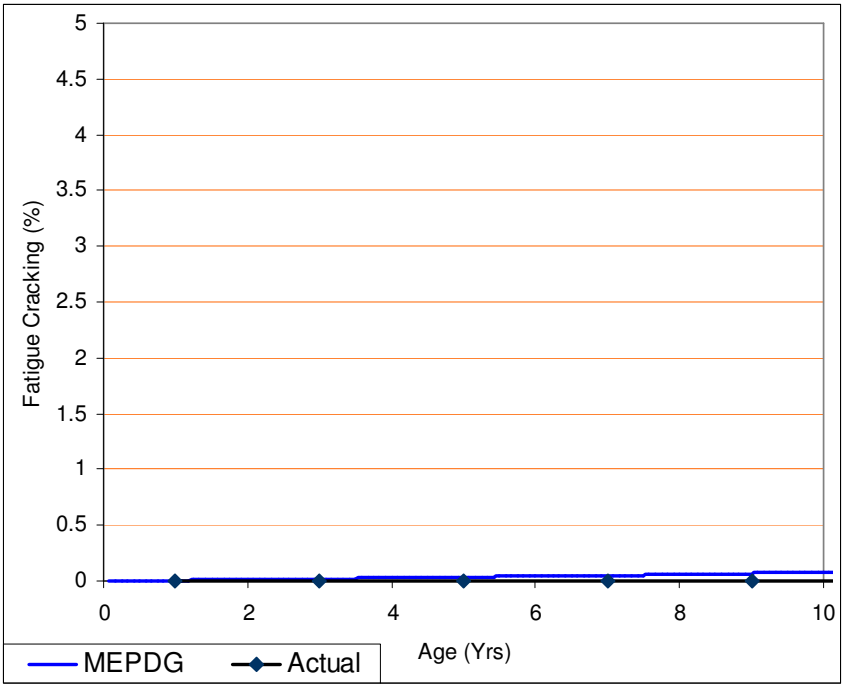


Figure 7. 2 Comparing actual fatigue cracking verses MEPDG prediction for Section ID 29581N

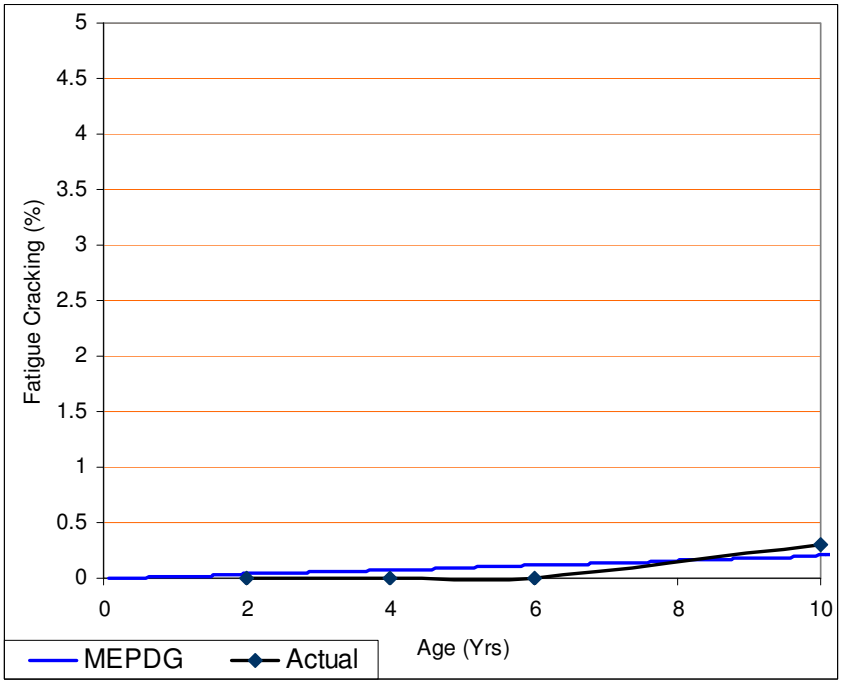


Figure 7. 3 Comparing actual fatigue cracking verses MEPDG prediction for Section ID 18890N

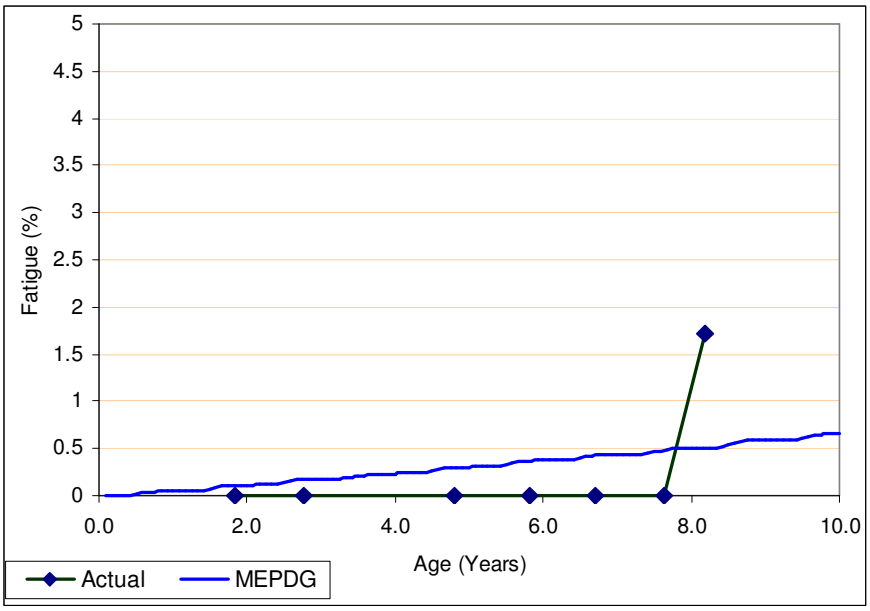


Figure 7. 4 Comparing actual fatigue cracking verses MEPDG prediction for SPS Site 117

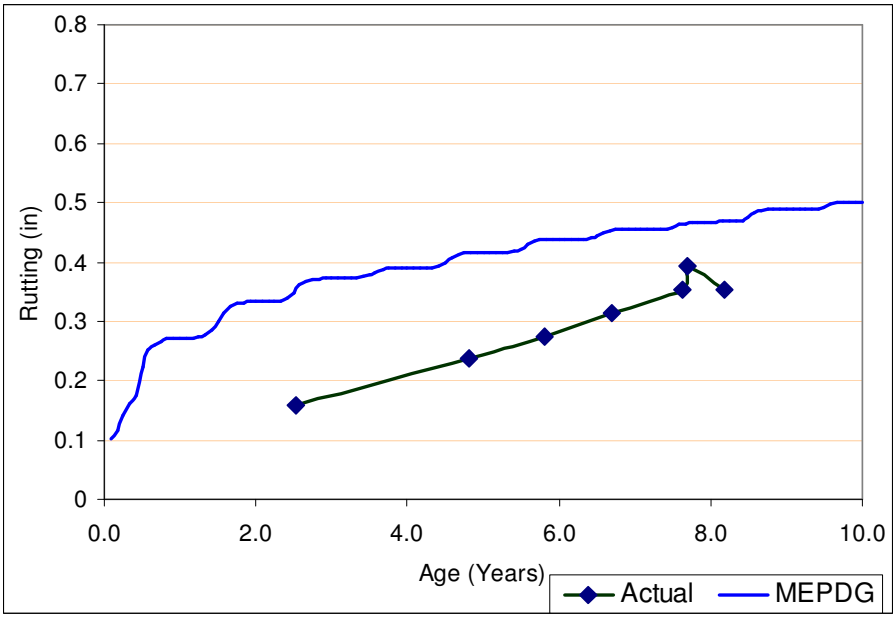


Figure 7. 5 Comparing actual rutting verses MEPDG prediction for SPS-1 Site 117

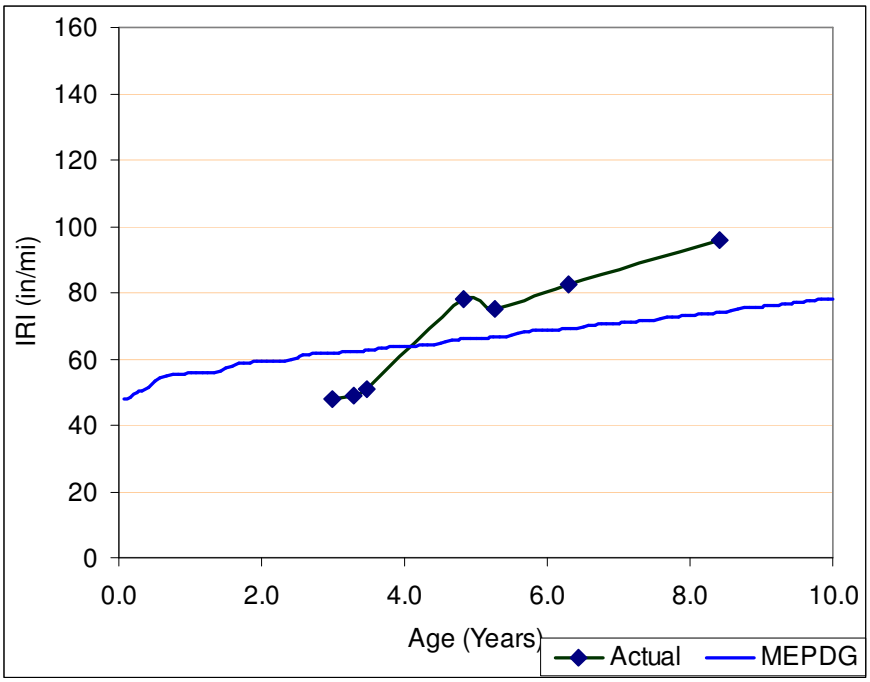


Figure 7. 6 Comparing actual IRI verses MEPDG prediction for SPS-1 Site 117

ME-PDG was run for all of the 150 cases corresponding to the three examples, and performance data was compiled. The real value of such analysis is to be able to study different scenarios and see how the current QA program is performing with respect to ensuring a quality product. Figures 7.7, 7.8 and 7.9 show predicted fatigue performance for all three cases with varying asphalt content and air voids.

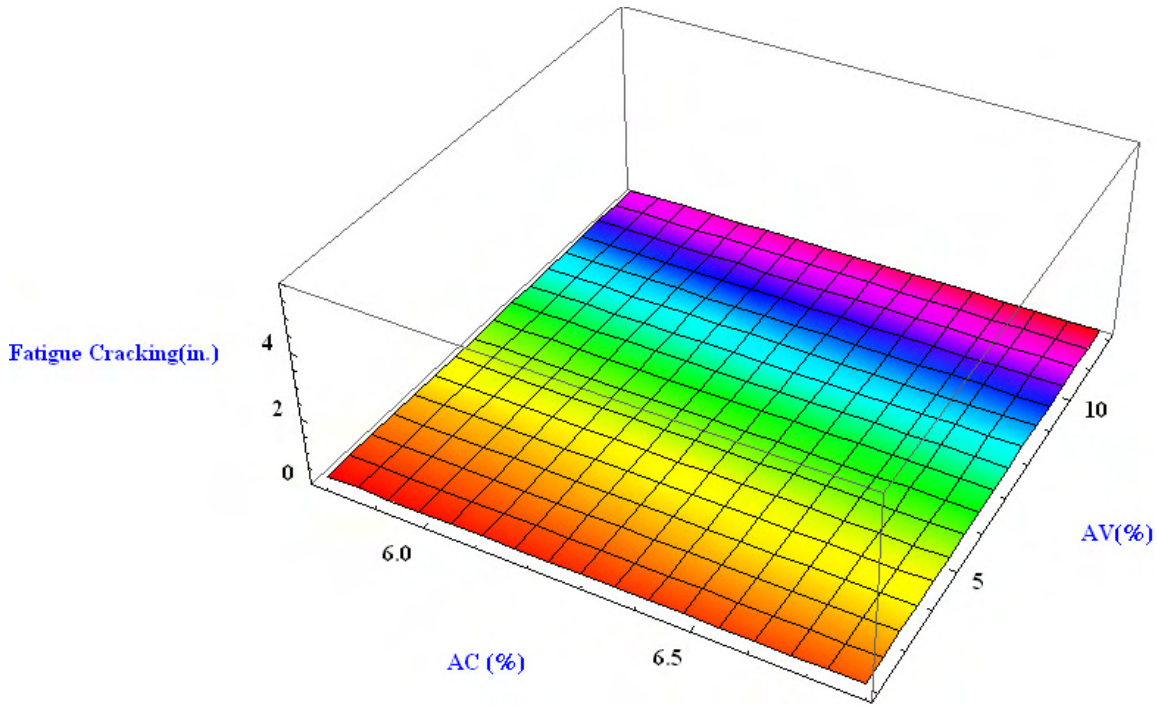


Figure 7. 7 Fatigue cracking (%) at the end of 20 years for Section ID 29581W with varying asphalt content and in-situ air voids

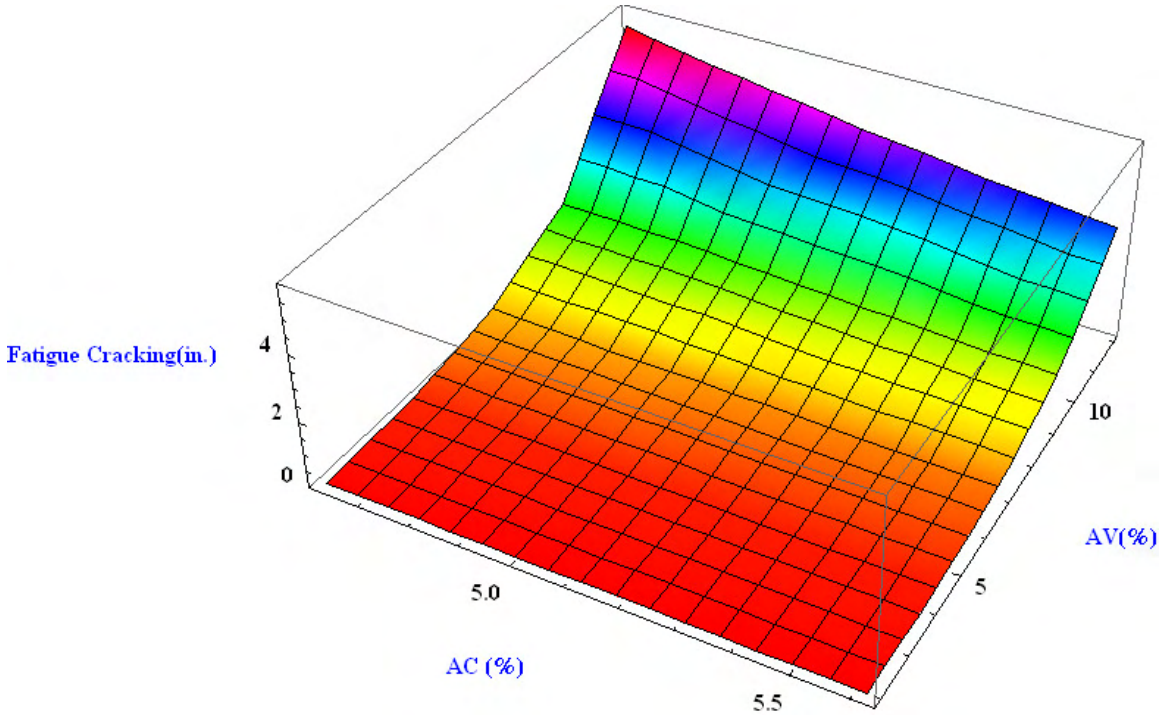


Figure 7. 8 Fatigue cracking (%) at the end of 20 years for Section ID 18890N with varying asphalt content and in-situ air voids

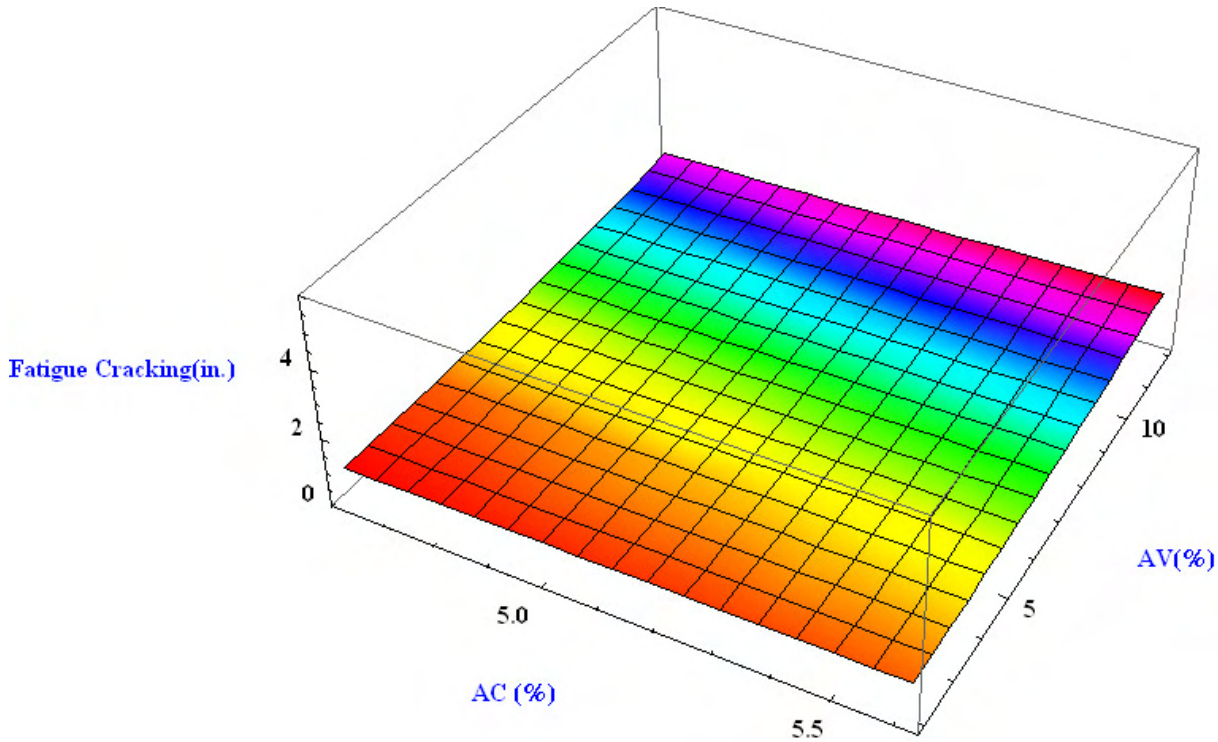


Figure 7. 9 Fatigue cracking (%) at the end of 20 years for SPS-1 site 0117 with varying asphalt content and in-situ air voids

To be able to assess how the current QA program is performing in terms of ensuring good quality, we need to present some contrasting case scenarios. For the three mixes that are being analyzed, three scenarios with respect to the quality of construction were considered. Table 7.6 presents the details of these scenarios. Note that the quality of construction is exhibited through the standard deviation; for example, while the "poor" case scenario has an acceptable air voids mean of 7% , taking into account two standard deviations from the mean, the in-situ air void will reach 10.7%, which is on the high side. It was assumed that the project would have 50 sublots.

Table 7.6 Air voids and asphalt content levels for the three scenarios for analysis

Scenario	In-situ Air Voids (%)		Asphalt Content (%)	
	Mean	Standard Deviation	Mean	Standard Deviation
Good	6	0.98	Target	0.05
Fair	6.5	1.45	Target	0.1
Poor	7	1.84	Target - 0.1	0.15

Performance curves obtained from all the MEPDG runs can be used to estimate performance for all these scenarios using interpolation methods. In this case, piecewise bi-cubic interpolation was used to accurately determine all the performance curves as required for the analysis. Figures 7.10 through 7.12 show fatigue performance for each of the sublots for all three scenarios and the three different mixes being analyzed. The mix corresponding to Section ID 29581N shows good performance even when construction quality is poor, while the same is not true for the other mixes. Also, while all the three mixes seem to perform well when the construction quality is good, the second and the third mixes start showing poor performance, or loss of service life, as construction quality becomes worse.

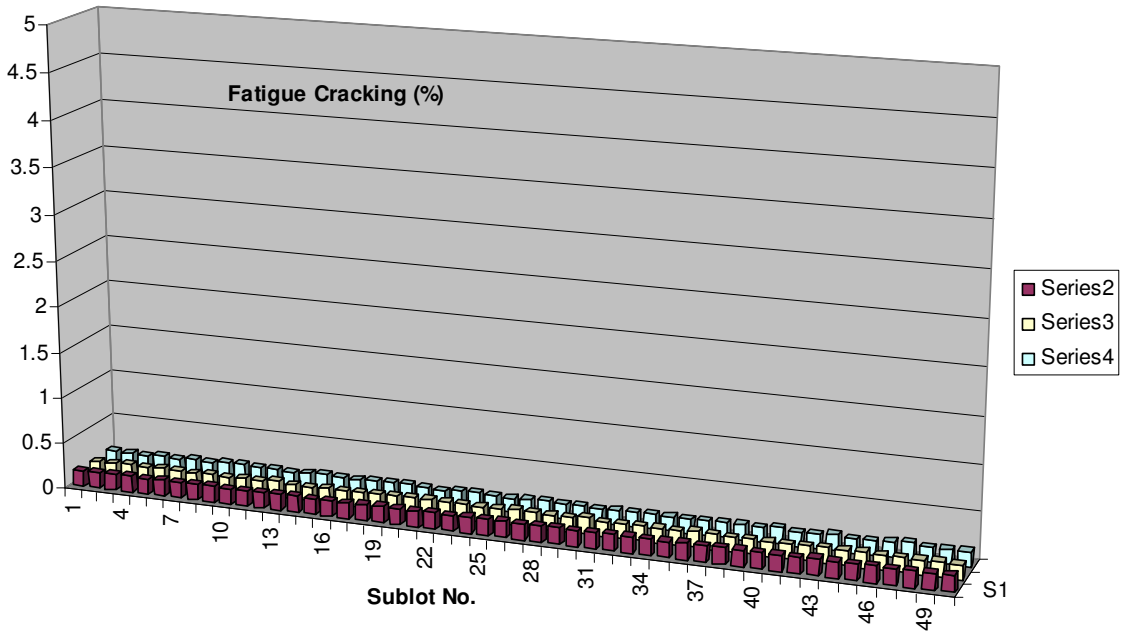


Figure 7. 10 Fatigue cracking for Section ID 29581N (Good, fair and poor)

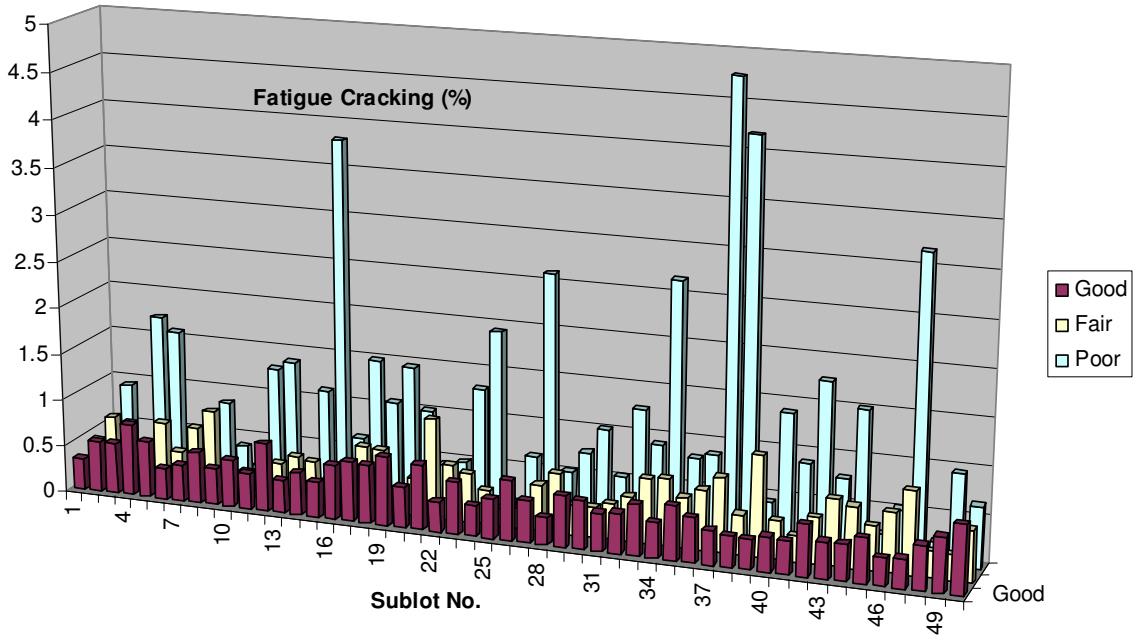


Figure 7. 11 Fatigue cracking for Section ID 18890N (Good, fair and poor)

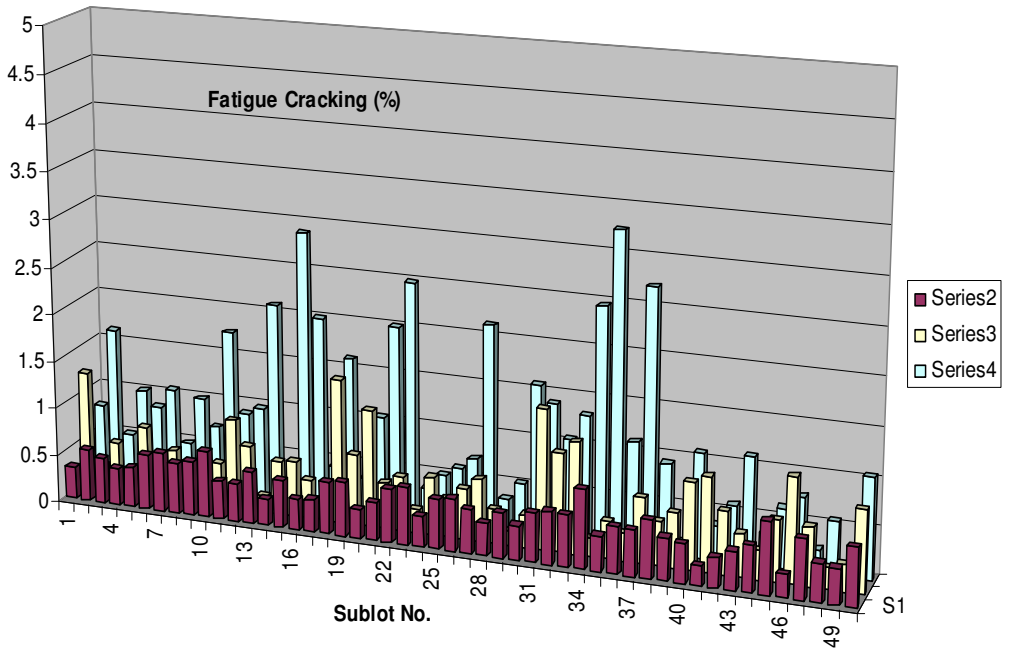


Figure 7. 12 Fatigue cracking for SPS-1 Site 117 (Good, fair and poor)

In a good QA program, properties that lead to poor long term results should translate into a lower pay factor. In the current QA program, pay factor is determined based on percent within limits achieved for plant air voids, plant VMA and in-situ density/air-voids. Table 7.7 shows the percent within limits and pay factor for density. It shows that, according to the current QA program, good construction quality in the example presented here would result in a pay bonus of 1.6%. Also, the pay factor for the fair construction is very close to 100% with a penalty of 0.3%. It also penalizes the poor construction much more severely with 4% penalty in pay.

Table 7.7 PWL and pay factor for density for the three scenarios.

Scenario	Density PWL	Density PF
Good	98	101.6
Fair	85	99.7
Poor	70	96.0

If the highway agencies were to have a reliable estimate of the performance of the pavements, they can also determine a rational QA program. Such a program would not only allow for rational payment to the contractor, it would also encourage them to achieve quality in construction that would lead to good performance. The analysis presented here is intended to show the possibility of doing this. The next step would be to perform exhaustive analysis considering all possible aspects of construction and determine their effect on performance. The following step would be to determine the change in service life as a result of construction quality.

7.3 Summary of Findings

There are three very important aspects of a QA program:

(1) The tests that are conducted within the QA program should relate to pavement performance, and sufficient and appropriate number of tests should be done to ensure quality in different aspects of mix preparation and construction process.

(2) The targets for the quality characteristics such as air voids, VMA and density should be set such that they ensure good performance. The window of variability allowed around the target should be achievable and at the same time tight enough to ensure consistency in pavement quality.

(3) The pay factor should be rational. In other words the penalty or bonus should be proportional to the gain or loss in life of the pavement as a result of good or poor construction quality.

Based on the limited number of verifications, MEPDG seems to give reasonable prediction trends for different pavement distresses. Therefore, MEPDG can be used to explore various aspects of HMA mix and construction process that influence pavement performance. However, pending to verification through other methods, other aspects of mix production or lay down that are not being accounted for in the current QA program can be identified through the strategy presented in this report. The proposed strategy can also be directly used to determine the targets and allowable windows for each of the quality characteristics being used in the QA program. MEPDG can be used to simulate the variability encountered in construction and the resulting pavement performance. A preliminary example of that has been presented in this report. When extended with possible life cycle cost analysis, such simulation process can lead to a rational pay factor formula/procedure. In the next section, several case scenarios are simulated using the MEPDG and the results are discussed.

7.4 Simulation Using MEPDG

The MEPDG includes many different models for predicting pavement properties and performance. This makes it possible to study the effect of different quality characteristics on pavement performance. Also, the effect of different input variables or quality characteristics can be studied together rather than individually because there is an interaction between their effects on performance. The models used in the MEPDG, like any other mechanistic-empirical model, have limitations and inaccuracies associated with them. This is more or less expected as we are trying to model natural materials which vary by their location and with changing environment. However, it should also be accepted that these models are the best that the pavement community can put together at this time.

There is a substantial amount of data in the LTPP database for different categories of input variables as well as performance from survey on real life pavements. However, any construction project is unique because of a variety of factors affecting it. For example, different construction projects would have different construction crews, different climatic conditions during the days of construction, different brands of equipment used, including the paver, different material transportation and discharge practices and conditions, different oversight managers etc. Also, during the life of the pavement one may have some unique factors influencing it like a traffic pattern which is not common to other pavements. The end result is

that when distress survey data is collected from such project, the data is a result of *many* factors that it is almost impossible to separate their individual effects. Therefore, it is difficult to account for all the factors even in very carefully controlled road tests like the AASHO road test or Westrack.

Although many factors affect pavement performance, not all of them can be controlled at the time of construction, even though they do affect pavement performance. This can be easily known in those cases where quality control is very poor and the pavement develops premature distresses. However, in the majority of the pavements constructed, the quality control is fair to excellent. It is important to study how relatively minor changes in quality affect service life or performance of the pavement. MEPDG lends itself well to such analysis even though such simulation results are indicative rather than predictive. This section presents analysis performed using MEPDG to study the influence of QA variables, including plant air voids, in-situ density and asphalt content, on HMA pavement performance. Similar analysis on PCC pavements is also conducted and discussed later on.

In the previous sections, some analysis was performed where the effects of individual input variables on performance were studied. In this section, several case scenarios were simulated using MEPDG, and statistical analysis on the corresponding results are presented.

As stated earlier, the LTPP database has mean and standard deviation values for input variables like in-situ density and asphalt content along with the number of samples used to determine standard deviation, although individual data used for the calculation are not documented. Statistical methods were used to generate simulated values of in-situ air voids and asphalt content which would have the same means and standard deviations as those reported in the LTPP database. This was done by writing a computer program in Matlab. It is assumed that since these values simulate the LTPP means and standard deviations, they would also be representative of actual projects under a QA program.

The next step was to input each scenario in MEPDG and determine expected pavement performance in terms of fatigue cracking, longitudinal cracking, rutting etc. To get an appreciation for what this analysis entails, let us take the example of data plotted in Figure 7.13. Each data point in the plot represents the mean fatigue cracking value for one project. Each project has a number of samples tested for determining mean and standard deviation of in-situ density and asphalt content. Combining the number of samples from all the projects plotted in

the figure for density and asphalt content, the total number of MEPDG runs required would be 7,884. Each run using MEPDG to predict distresses for 30 years would take an average of 50 minutes. This would mean that the total time required for carrying out all the runs for a single plot would be 273 days, assuming that there are no errors or crashes during the execution of the runs. For all practical purposes, it is almost impossible to execute this task. We present below an alternative, which can make this task much more efficient.

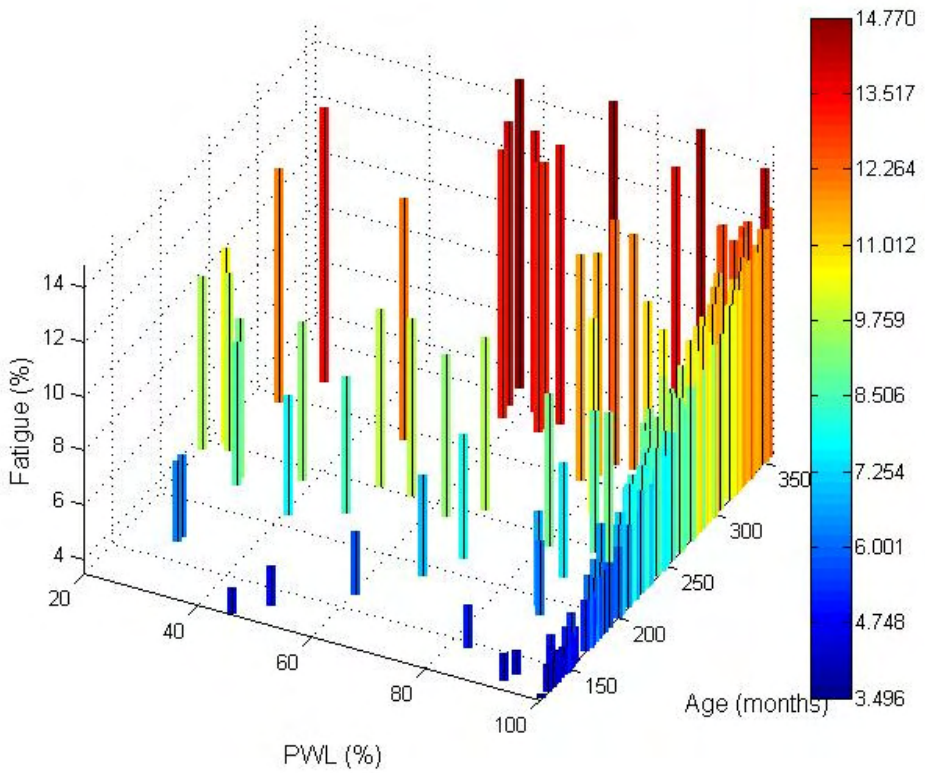


Figure 7. 13 Fatigue cracking Vs PWL (in-situ air voids) for mix 1

The alternative, in principle, is to run MEPDG to develop response surfaces and use these surfaces to determine performance for all 7,884 case scenarios required for one plot. For developing the response surface, the entire range of in-situ density and asphalt content should be identified. Then some points in between the maximum and minimum values should be identified. In this analysis, we identified 3 more values between the extremes, making for a total of 5 values each for density and asphalt content. Then a full factorial matrix is defined using all possible combinations of the values of asphalt content and density. This would make for 5*5= 25 runs. In order to keep the number of simulation runs manageable, we identified two different mixtures

and input all of their characteristics in MEPDG. Then all 25 runs were executed for each of the mixtures. The results from the 50 runs were compiled based on performance category (fatigue, rutting, etc.). All the 25 cases defined here were real combinations of input variables derived using MDOT designs. One can then imagine a response surface where the four dimensions represent density, asphalt content, age and fatigue (or rutting etc.). Such response surfaces were created in Matlab.

Each of the required 7,884 runs of MEPDG correspond to one combination of density and asphalt content, and the response is obtained for varying ages in each case. The response surface created in the analysis actually contains each of these points as long as density and asphalt content fall within the range identified for running the 25 cases with MEPDG. Therefore, using the piecewise cubic spline interpolation technique in four dimensions the various performances (fatigue, rutting, etc.) were generated for all of the 7,884 cases. This was done using the MatLab computer program. Then, the performance (e.g., fatigue) corresponding to each project was calculated by averaging all the values corresponding to different samples.

Piecewise cubic spline interpolation technique was chosen in this case because it fits an n-dimensional cubic surface locally to each portion of the surface while maintaining continuity in magnitude and slope of the surface with the neighboring areas within the surface in all dimensions. This makes this technique extremely versatile to be used on diverse types of surfaces while maintaining excellent accuracy all across the surface. Regression or model fitting can lead to appreciable errors in certain ranges of input variables especially when it has to be fit in more than two dimensions.

7.4.1 Effect of varying air void distribution on average project performance

First, we present the results corresponding to the cases of varying in-situ air voids to achieve various PWL ranges, while fixing the asphalt content (at -0.4% of optimum). Figure 7.13 shows mean fatigue cracking estimated using the response surface for different case scenarios corresponding to data available in LTPP database as well as using additional input data to cover lower PWL ranges more uniformly. These cases were added to get a better picture of what happens when PWL is lower than 90%. The plot shows the expected trend of higher fatigue cracking with increasing age of the pavement. It is also clear that as PWL (in-situ air voids/density) increases the projects have lower expected fatigue cracking for the same age. To get a better appreciation of the magnitude of difference in cracking we estimated fatigue cracking

at the end of 30 years for all the cases. The results are shown in Figure 7.14. The maximum difference in fatigue is only about 2% between the worst PWL and the acceptable PWL (greater than 90%); a negligible difference.

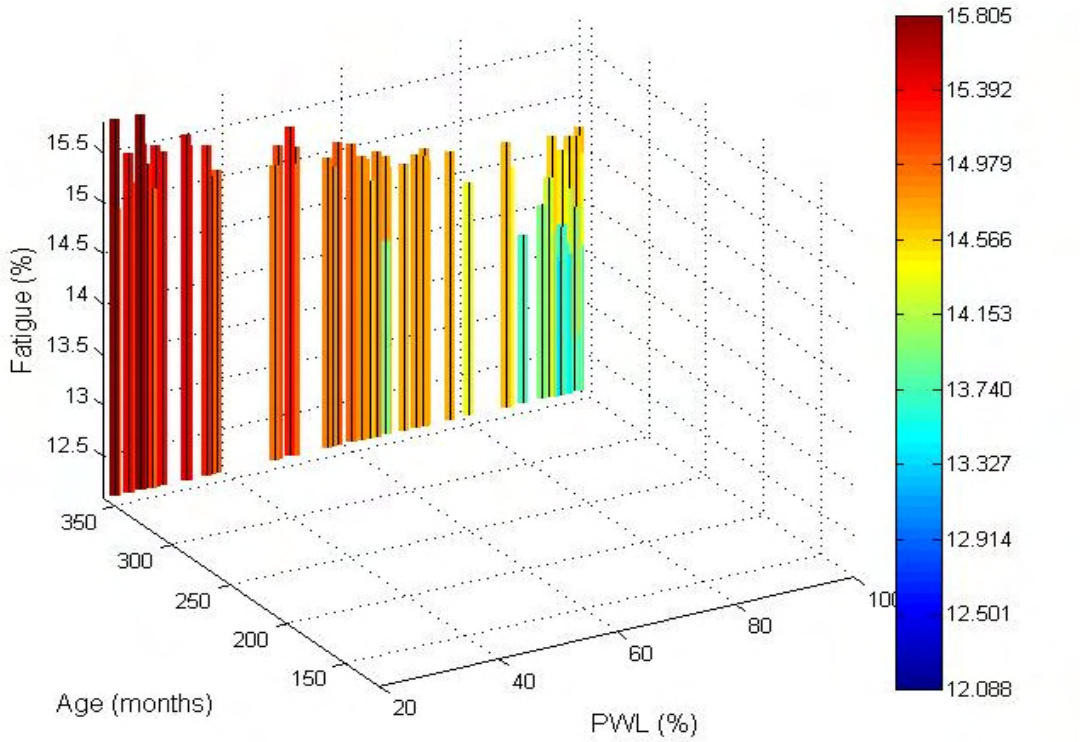


Figure 7. 14 Fatigue at 30 years Vs PWL (in-situ air voids) for mix 1

It is also noticeable that for very similar PWL values, different projects show different amounts of fatigue cracking. This is because fatigue cracking in each project is estimated by sampling. As long as the range, mean and standard deviation of two samples are the same they would have the same PWL. But the average fatigue cracking can vary depending where each of the samples falls within the specifications. For example, if a project has all densities equal to 93%, PWL would be 100%. If the density of all the samples were 98%, still PWL would be 100%. However, fatigue performance for the 98% density project would be much better than that for the project with mean density of 93%.

Rutting performance estimated through the simulation is shown in Figure 7.15. This figure shows that rutting performance improves as PWL approaches 100%. However, this

improvement is very small. In other words, rutting performance is not so sensitive to PWL (in-situ air voids) according to MEPDG results. Figure 7.16 shows the same result with IRI. The plot shows that IRI also does not seem to be getting much affected by PWL (in-situ air voids).

It must be mentioned that all these results are for one type of mix. It is important to study different types of mixes to assess how PWL values affect performance for each one of them. It is quite possible that this effect may be much more pronounced in other types of mixes.

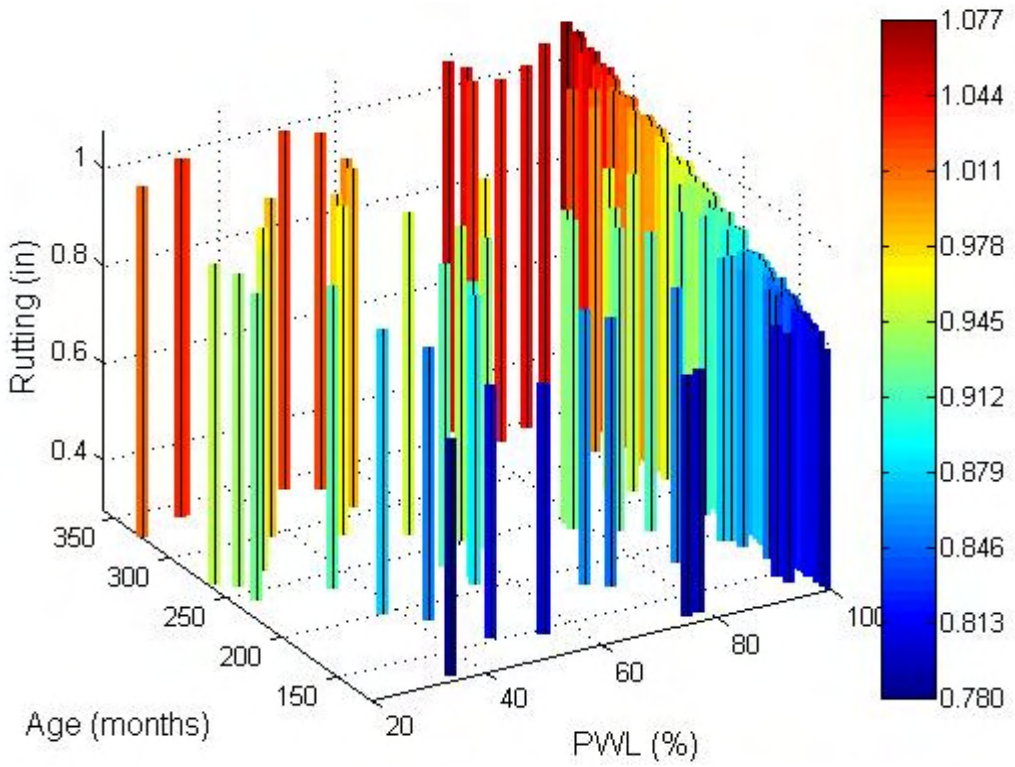


Figure 7. 15 Rutting Vs PWL (in-situ air voids) for mix 1

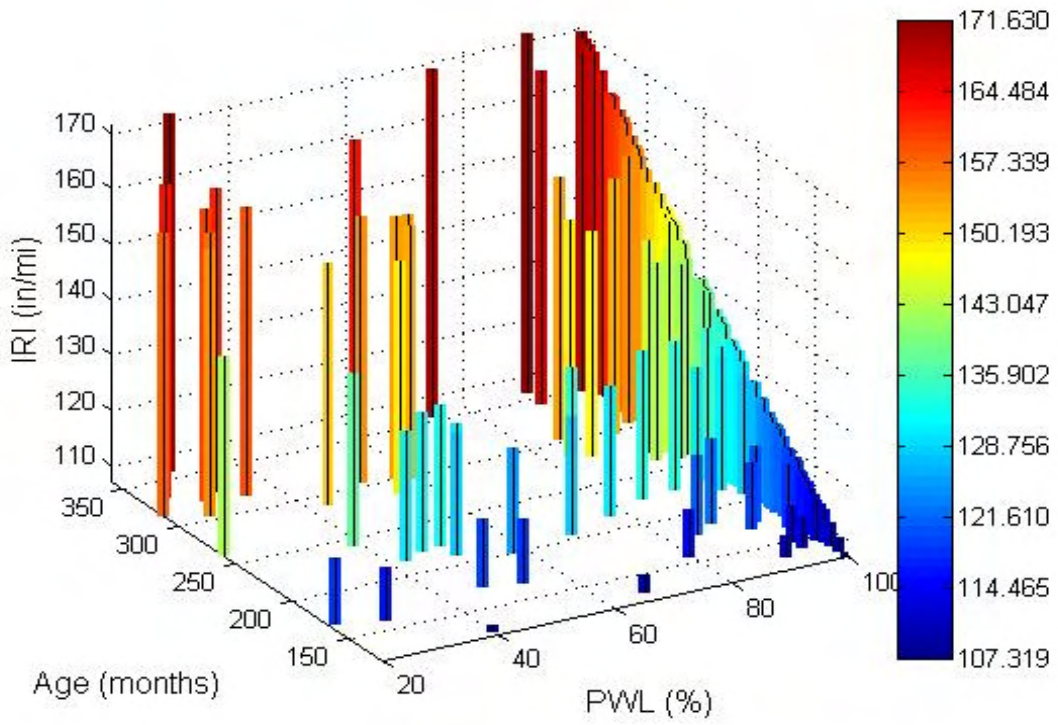


Figure 7. 16 IRI Vs PWL (in-situ air voids) for mix 1

7.4.2 Effect of varying both air void distribution and asphalt content on sublot project performance by sublot

The next exercise is to explore the effects of varying air voids and asphalt content together. This means that different sublots in the same project have varying air void and asphalt content. The previous analysis showed average performance for a given project. The observed effects of PWL on performance were not as pronounced as one would expect intuitively. The main reason for this, as will be shown in this section, is the attenuation of distress levels because of the averaging of performance across sublots. Therefore, the results presented in this separate analysis are shown in terms of performance by individual sublots. Figures 7.17 through 7.21 show the fatigue results for various scenarios (three air void and five asphalt content distributions). The figures show that:

1. performance of different sublots within a project can vary significantly;
2. within a given asphalt content range, there are more underperforming sublots when the air void distribution moves towards the upper specification limit ;

3. as the asphalt content range moves dry-of-optimum, more sublots are underperforming in fatigue, even at relatively high combined PWL [for example, compare Figures 7.17(a) through 7.21(a)].
4. as the asphalt content range moves wet-of-optimum, more sublots perform better in fatigue, even at relatively low combined PWL [for example, compare Figures 7.17(c) through 7.21(c)]. This may however lead to more rutting within the HMA layer.

7.5 Performance criteria

In the previous section, results from multi-factor analysis were presented for flexible pavements. The results showed how certain factors affect pavement performance. The results were presented in the form of comparative plots. However, Task 7 of this project would require firm decisions to be made whether a candidate QA variable should be included in the list of important QA variables. This section presents an objective strategy that can be used to make such decisions

Many different criteria were considered which could possibly be used to compare different project scenarios and make a decision whether a certain candidate QA variable is important for a QA program. Sets of possible case scenarios were developed to test these criteria. The results from these case scenarios would also give deeper insight into the workings of percent-within-limits statistical approach used in QA programs by the Michigan Department of Transportation and several other states. Each scenario represents one flexible pavement project with 50 sublots. Each of the sublots would have varying levels of in-situ air voids and asphalt content. Table 7.8 shows the asphalt content and in-situ air void levels for the 25 scenarios. This forms one set of scenarios. Three such sets were planned for this study which would have different levels of variability associated with asphalt content and air voids. Table 7.9 presents the variability for the three sets.

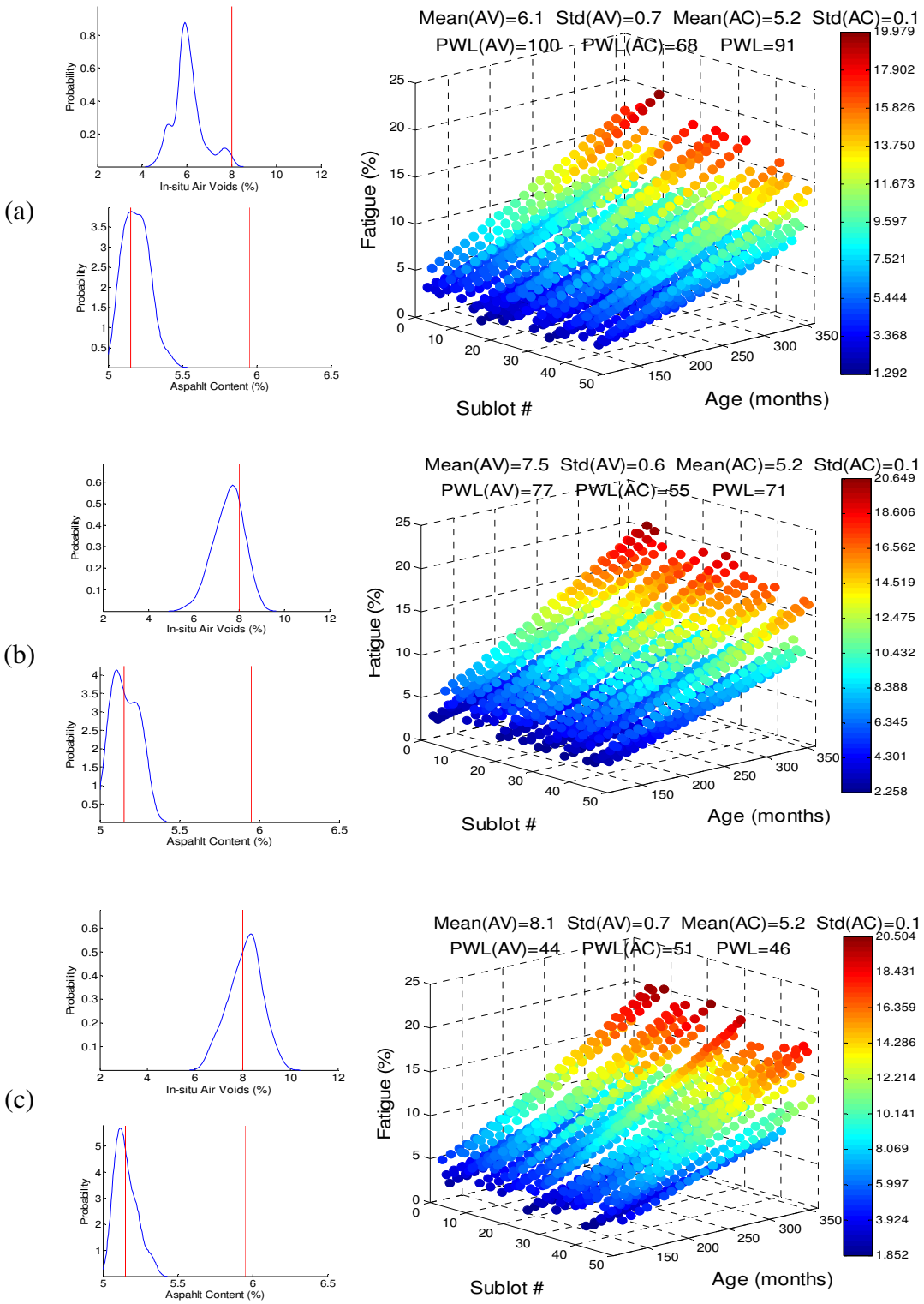


Figure 7. 17 Effect of AV distribution on fatigue performance of sublots for AC=Opt-0.4.

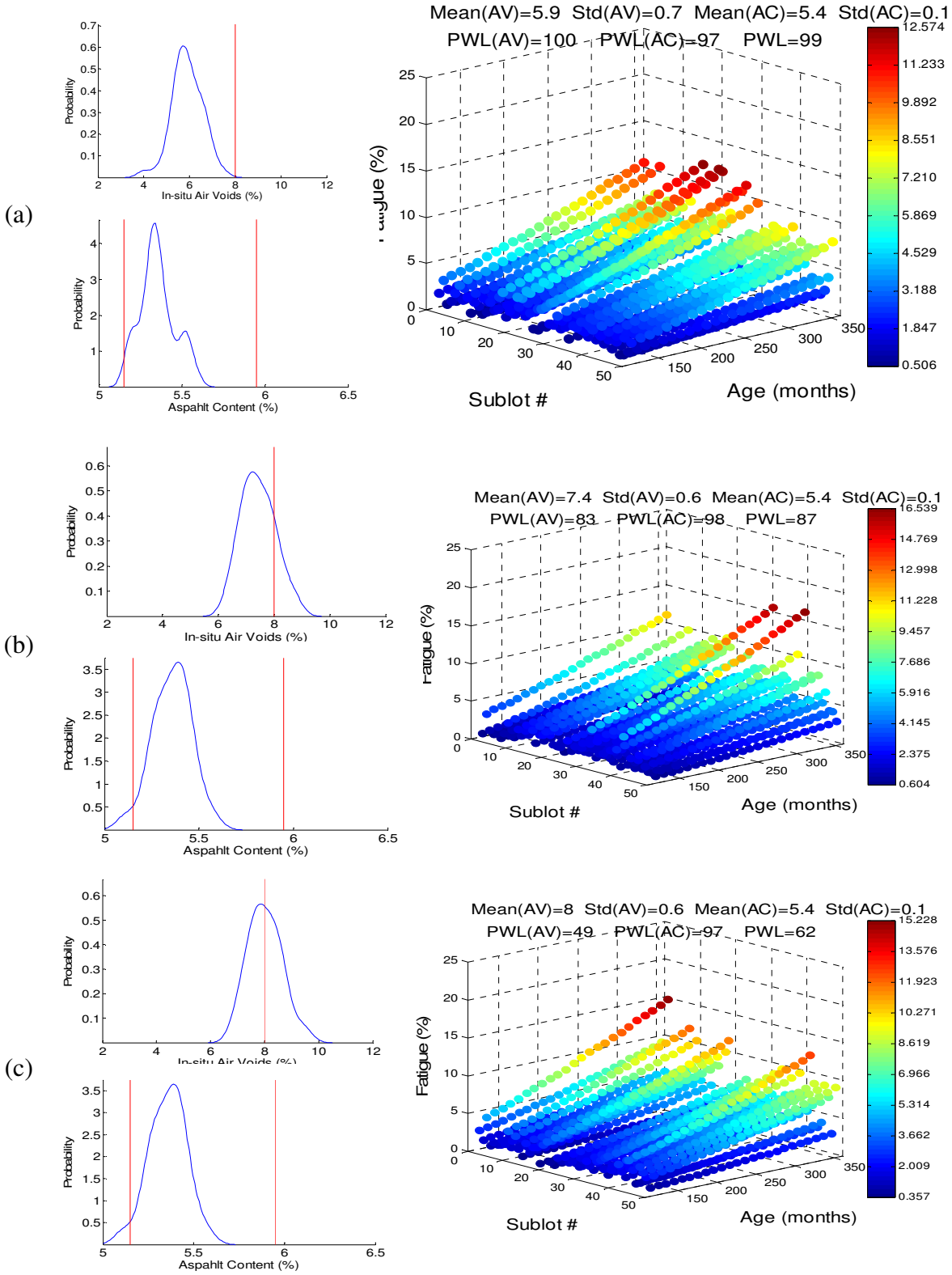


Figure 7. 18 Effect of AV distribution on fatigue performance of sublots for AC=Opt-0.2.

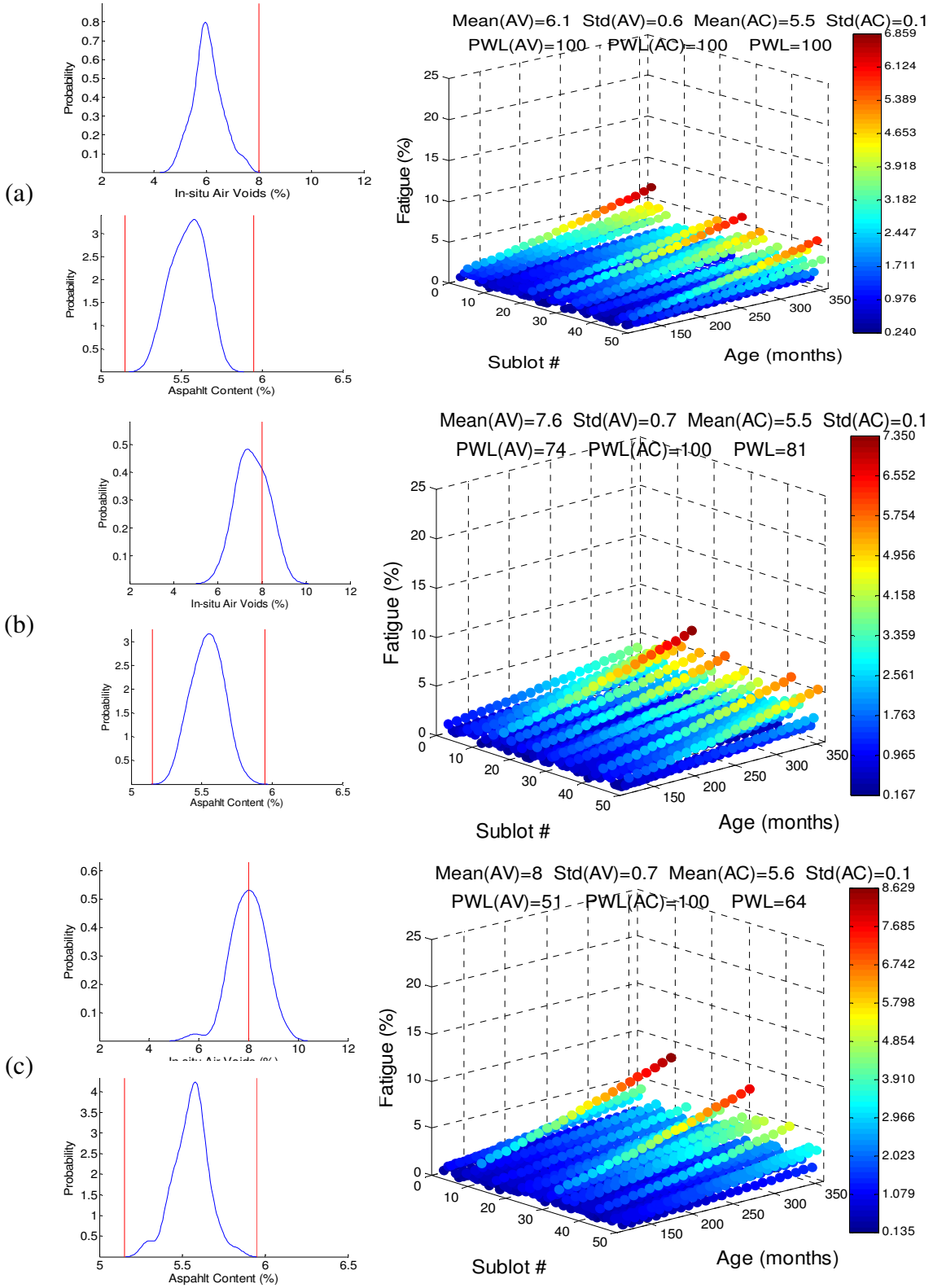


Figure 7. 19 Effect of AV distribution on fatigue performance of sublots for AC=Opt.

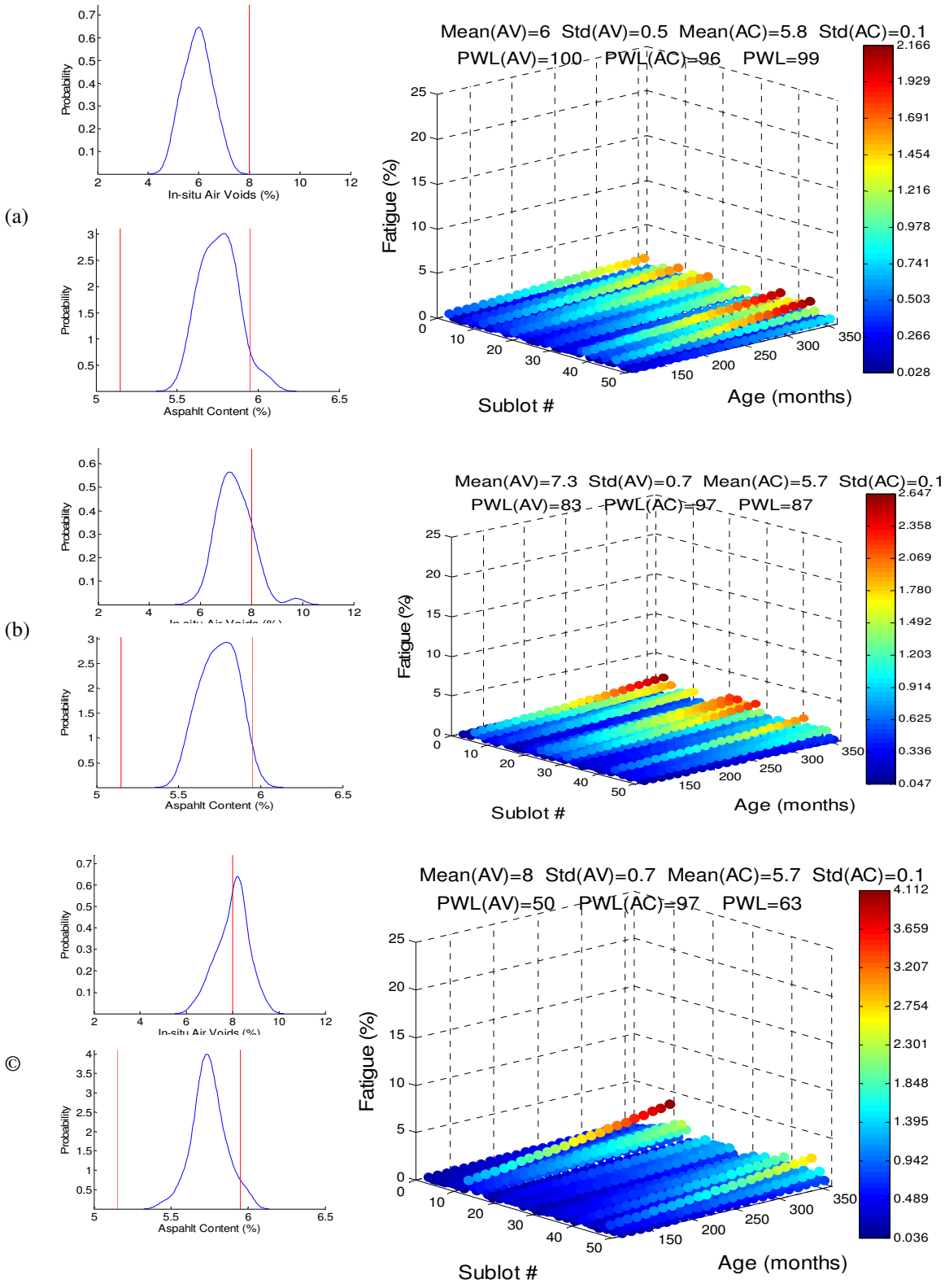


Figure 7. 20 Effect of AV distribution on fatigue performance of sublots for AC=Opt+0.2.

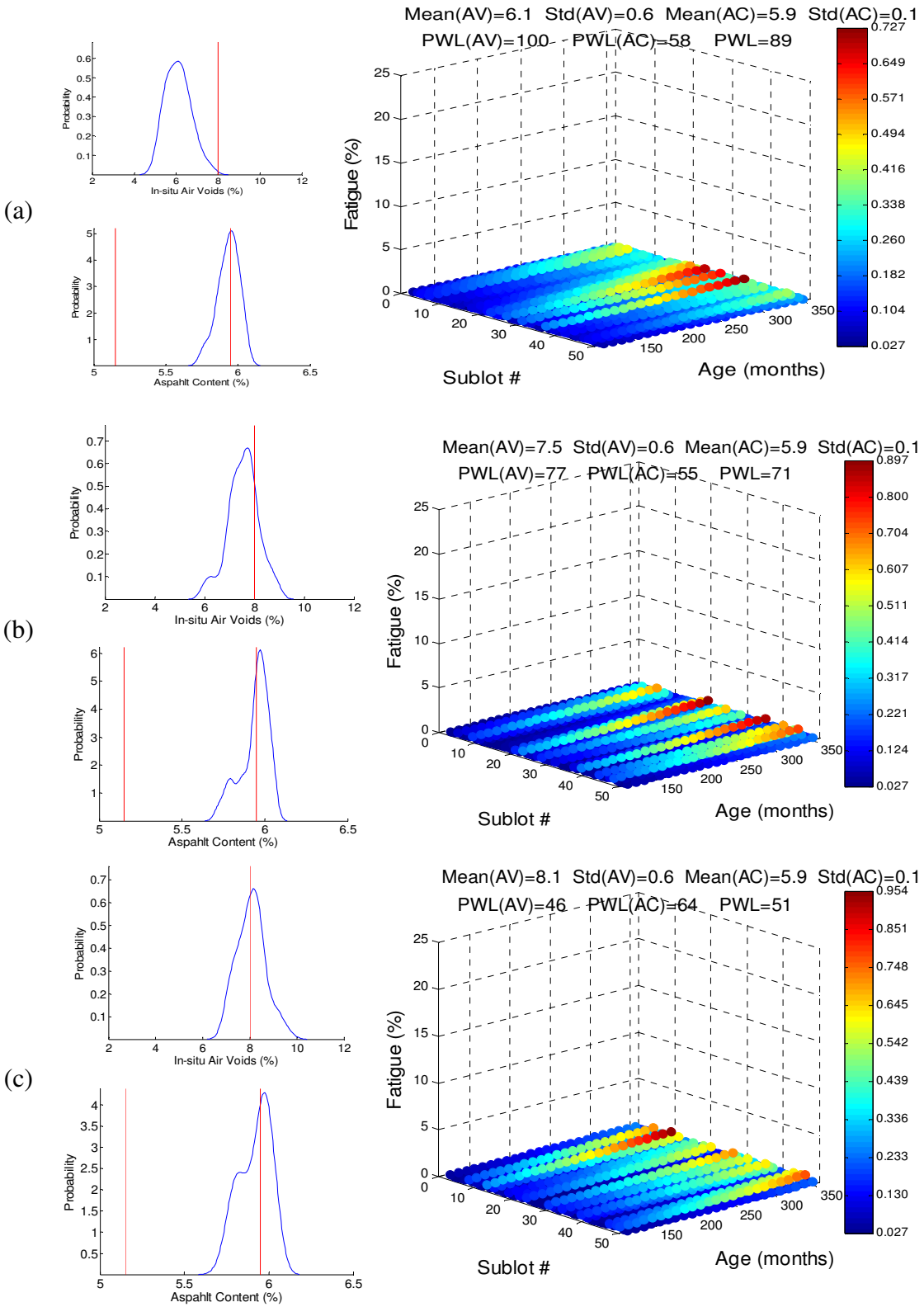


Figure 7. 21 Effect of AV distribution on fatigue performance of sublots for AC=Opt+0.4

Therefore, a total of 75 case scenarios were studied, each having 50 simulated sublots. This equates to running MEPDG software 3750 times. However, instead of running MEPDG, the strategy involving response surfaces was used. The project simulated in this study was a flexible pavement constructed in Ludington, MI in 1988. This pavement has 7.5 inch asphalt concrete layer over 4 inches thick base and 18 inches thick subbase.

Table 7.8 Asphalt content and air voids levels for the case scenarios

Run Number	Asphalt Content (%)	In-situ Air Voids (%)
1	Optimum-0.4	3.5
2		6.0
3		7.5
4		8.0
5		9.5
6	Optimum-0.2	3.5
7		6.0
8		7.5
9		8.0
10		9.5
11	Optimum	3.5
12		6.0
13		7.5
14		8.0
15		9.5
16	Optimum+0.2	3.5
17		6.0
18		7.5
19		8.0
20		9.5
21	Optimum+0.4	3.5
22		6.0
23		7.5
24		8.0
25		9.5

Table 7.9 Standard deviations for asphalt content and in-situ air voids for the three sets of scenarios

Set Number	Standard Deviation (AC) (%)	Standard Deviation (AV) (%)
1	0.10	0.75
2	0.15	1.00
3	0.15	1.20

Different criteria considered are enumerated below. Each project represents a set of 50 sublots with fixed means and standard deviations for air voids and asphalt content.

(i) Average distress (for example, fatigue cracking) at the end of 30 years for the 50 sublots

(ii) Average distress for the sublots considering distresses for each of the months during 30 year service life

(iii) Area beneath the curve showing distress at the end of 30 years versus sublot number

(iv) Total area beneath all the distress curves for each project

(v) 50th, 75th, 90th and 95th percentile of distresses at the end of 30 years for each project.

(vi) Area beneath the distress curve when distress is raised to a certain power. This was done to especially identify those scenarios which lead to unacceptably high distresses in even a few sublots within a project.

Tables 7.10 through 7.12 present the average fatigue and 75th and 90th percentile for different projects which represent different realistic scenarios. The rest of the criteria as mentioned in the list above were studied in a similar way for all of the scenarios, although the details are not being included in the report for the sake of brevity. Comparing all the above criteria it was found that average fatigue cracking after 30 years combined with 75th or 90th percentile are good to compare the projects (or scenarios).

Some of the rows in Table 7.10 have been highlighted. These cases have very similar combined percent-within-limits (PWL) values. Combined PWL is generally used directly to calculate payment for the contractor, which ideally is in accordance with the quality of the pavement constructed by him. Although the PWL values are very similar in these cases the average fatigue cracking at 30 years varies from 0.69% to 6.76%. This much of variation is very significant because they would be categorized as being in very good and nearly poor condition respectively. The 75th percentile value for the same cases varies from 0.83% to 8.68% and the 95th percentile varies from 1.1% to 9.86%. This clearly shows that the candidate QA variable causing this variation, asphalt content in this case, has significant influence on fatigue performance.

Variations in fatigue performance as a result of change in in-situ air voids in these cases is also appreciable, although not as large as that because of asphalt content. When the air voids increase from 3.52% to 7.97% between run number 1 and 4, the 75th percentile for fatigue

cracking goes up from 13.9% to 17.87%. An advantage of percentile distresses, as used here, is that it can possibly identify those cases where a few sublots perform very poorly while rest of the sublots may have acceptable performance. Such a scenario should not be acceptable because the distressed sublots may force early repair work for the entire project.

Table 7.10 PWL values and fatigue cracking for the 1st set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL(AV)	PWL(AC)	PWL Combined	Fatigue Avg. (%)	Fatigue 75th Percentile	Fatigue 90th Percentile
3.52	5.17	1	100.0	59.6	89.0	11.81	13.90	15.54
5.85	5.17	2	100.0	58.6	88.7	13.17	15.86	16.94
7.48	5.16	3	79.4	55.0	72.7	14.11	16.05	17.67
7.97	5.16	4	52.2	56.3	53.3	14.25	17.87	19.65
9.50	5.14	5	1.1	42.7	12.4	16.82	19.00	20.68
3.62	5.32	6	100.0	95.3	98.7	6.76	8.68	9.86
6.07	5.35	7	100.0	98.9	99.7	6.64	8.39	9.67
7.46	5.35	8	80.7	97.3	85.2	7.09	9.06	11.66
8.09	5.36	9	43.9	98.9	58.9	6.74	7.83	11.21
9.38	5.34	10	0.7	97.8	27.2	8.05	9.44	13.30
3.41	5.57	11	100.0	99.8	100.0	2.33	2.95	4.24
5.95	5.54	12	99.8	100.0	99.8	2.83	3.48	4.87
7.46	5.54	13	83.4	100.0	87.9	2.96	3.53	5.02
8.06	5.57	14	46.4	99.9	61.0	2.66	3.29	4.91
9.57	5.57	15	0.5	100.0	27.6	2.76	3.54	4.81
3.60	5.77	16	100.0	98.6	99.6	0.69	0.83	1.10
6.20	5.76	17	99.6	98.7	99.3	0.83	1.00	1.38
7.40	5.75	18	83.7	96.5	87.2	0.99	1.31	1.70
7.97	5.74	19	51.7	98.7	64.5	1.05	1.27	1.68
9.63	5.75	20	0.5	96.8	26.8	1.06	1.21	1.59
3.64	5.93	21	100.0	57.8	88.5	0.27	0.30	0.50
6.00	5.94	22	99.8	56.9	88.1	0.30	0.37	0.44
7.48	5.94	23	80.8	55.6	74.0	0.28	0.33	0.54
8.23	5.92	24	37.0	63.0	44.1	0.33	0.42	0.66
9.48	5.94	25	0.4	53.8	15.0	0.33	0.36	0.61

Tables 7.11 and 7.12 present the fatigue cracking results obtained for the 2nd and 3rd sets of scenarios, and show similar trends. Tables 7.13 through 7.15 present the rutting results for different scenarios. The range of rutting values corresponding to the average and 75th and 90th percentiles for all the cases is rather small. It should be noted that in these runs the asphalt content was allowed to vary only up to 0.5% more than the optimal. If higher asphalt contents were allowed, the increase in rutting would probably have been more significant. However, the

selected criteria, namely the average distress at 30 years, rutting in this case, and 75th and 90th percentiles are still good indicators of the effect of QA variables on performance.

The results also show that even though different projects may have very similar combined PWL values they may perform differently in reality. However, this does not necessarily mean that PWL methodology does not work. Different factors may have relatively different influences on performance and they may affect each others' effect as well. It is the goal for a good QA program to include this interaction when calculating the combined PWL for the entire project.

Table 7.11 PWL values and fatigue cracking for the 2nd set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL av	PWL ac	PWL comb	Avg. Fatigue (%)	75th Percentile	90th Percentile
3.45	5.21	1	100.0	73.3	92.7	10.24	13.42	15.32
6.26	5.17	2	95.8	56.5	85.1	13.59	17.22	18.69
7.40	5.17	3	71.9	59.6	68.6	13.66	16.77	18.63
7.91	5.18	4	53.3	59.6	55.0	13.88	17.60	19.04
9.59	5.16	5	5.9	53.8	19.0	16.67	19.78	21.01
3.55	5.35	6	100.0	91.1	97.6	6.50	8.62	12.66
5.97	5.34	7	99.5	88.5	96.5	7.65	10.56	14.03
7.52	5.38	8	71.4	93.5	77.5	6.70	7.62	12.85
8.03	5.36	9	48.7	94.0	61.1	7.21	9.63	12.57
9.71	5.36	10	6.8	91.4	29.9	9.27	12.81	15.63
3.33	5.54	11	100.0	99.4	99.8	2.73	3.89	5.24
6.14	5.57	12	98.7	99.1	98.8	2.71	3.17	5.25
7.46	5.53	13	73.5	99.2	80.5	3.53	4.73	5.95
8.07	5.55	14	47.5	99.9	61.8	3.08	3.82	5.57
9.57	5.58	15	6.6	98.2	31.6	3.29	4.16	6.06
3.52	5.74	16	100.0	92.1	97.8	1.02	1.33	1.95
6.04	5.72	17	97.8	92.5	96.4	1.33	1.53	3.29
7.40	5.76	18	71.0	88.5	75.7	1.09	1.52	2.23
8.05	5.74	19	48.2	90.9	59.9	1.29	1.50	2.77
9.61	5.72	20	6.0	94.7	30.2	1.43	1.70	2.17
3.70	5.91	21	100.0	63.6	90.1	0.34	0.40	0.69
5.87	5.93	22	99.0	58.1	87.8	0.36	0.44	0.80
7.47	5.92	23	73.1	60.0	69.5	0.38	0.38	1.09
8.17	5.91	24	42.9	61.8	48.1	0.50	0.41	0.87
9.56	5.90	25	5.3	65.2	21.7	0.60	0.65	1.09

Table 7.12 PWL values and fatigue cracking for the 3rd set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL av	PWL ac	PWL comb	Avg. Fatigue (%)	75th Percentile	90th Percentile
3.19	5.19	1	99.9	65.6	90.5	10.86	13.75	15.51
5.85	5.21	2	97.1	70.8	89.9	11.57	14.67	16.67
7.32	5.17	3	68.7	56.8	65.4	13.98	17.12	18.83
7.99	5.18	4	50.4	63.2	53.9	13.71	16.69	19.75
9.09	5.22	5	19.1	71.5	33.4	13.59	18.89	21.17
3.25	5.34	6	99.9	88.5	96.8	6.64	8.26	11.77
6.17	5.34	7	94.0	92.7	93.7	7.31	9.00	12.17
7.12	5.35	8	73.7	94.1	79.3	7.33	9.44	13.11
8.37	5.37	9	38.6	94.4	53.8	7.25	8.77	11.82
9.54	5.33	10	11.8	90.0	33.1	10.25	14.84	17.70
3.17	5.53	11	100.0	99.4	99.8	2.85	4.03	5.02
6.22	5.59	12	91.6	99.8	93.8	2.39	2.73	5.07
7.58	5.55	13	64.5	97.8	73.6	3.33	4.65	7.23
7.72	5.56	14	60.8	99.9	71.4	2.89	3.60	5.46
9.44	5.59	15	11.5	98.5	35.3	2.98	3.16	5.85
3.27	5.75	16	99.9	90.6	97.3	0.97	1.12	2.15
5.62	5.76	17	97.6	90.5	95.7	1.05	1.33	2.08
7.54	5.75	18	68.3	91.5	74.6	1.10	1.46	1.96
7.91	5.72	19	52.8	93.8	64.0	1.38	2.02	2.83
9.37	5.74	20	15.7	91.3	36.3	1.37	1.53	2.62
3.29	5.91	21	100.0	61.1	89.4	0.36	0.35	0.70
5.83	5.93	22	95.0	59.9	85.4	0.33	0.36	0.52
7.21	5.91	23	72.6	64.5	70.4	0.40	0.41	0.82
8.09	5.92	24	47.1	59.9	50.6	0.36	0.40	0.75
9.41	5.89	25	11.1	69.4	27.0	0.55	0.68	1.15

Table 7.13 PWL values and rutting for the 1st set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL av	PWL ac	PWL comb	Avg. Rut (in)	h ntile	90th Percentile
3.49	5.17	1	100.0	60.8	89.3	1.06	1.08	1.10
6.15	5.17	2	99.9	61.8	89.5	1.04	1.07	1.09
7.43	5.17	3	77.4	60.7	72.9	1.04	1.07	1.08
7.95	5.17	4	53.3	61.7	55.6	1.03	1.06	1.08
9.54	5.18	5	0.3	64.7	17.9	1.02	1.05	1.07
3.43	5.34	6	100.0	97.5	99.3	0.98	1.01	1.03
6.07	5.36	7	99.9	97.1	99.1	0.96	0.98	1.03
7.47	5.32	8	83.0	95.4	86.4	0.97	1.00	1.02
7.94	5.36	9	54.1	98.7	66.2	0.95	0.98	0.99
9.46	5.36	10	0.4	97.4	26.8	0.95	0.97	0.99
3.54	5.54	11	100.0	100.0	100.0	0.91	0.94	0.95
6.22	5.59	12	99.9	100.0	99.9	0.89	0.90	0.92
7.37	5.54	13	81.9	100.0	86.8	0.90	0.91	0.93
7.90	5.55	14	55.7	100.0	67.8	0.89	0.91	0.92
9.57	5.53	15	0.1	100.0	27.3	0.89	0.91	0.92
3.64	5.76	16	100.0	98.2	99.5	0.85	0.86	0.87
6.15	5.75	17	99.9	98.4	99.5	0.85	0.86	0.86
7.58	5.73	18	73.1	99.3	80.3	0.85	0.86	0.87
8.06	5.74	19	46.3	98.4	60.5	0.84	0.85	0.87
9.49	5.74	20	0.4	98.1	27.0	0.84	0.85	0.87
3.47	5.95	21	100.0	52.6	87.1	0.82	0.83	0.83
5.83	5.94	22	100.0	55.5	87.9	0.82	0.82	0.83
7.28	5.93	23	92.2	60.2	83.5	0.82	0.82	0.83
8.07	5.93	24	46.1	60.5	50.0	0.82	0.82	0.83
9.63	5.93	25	0.7	62.0	17.4	0.81	0.81	0.83

Table 7.14 PWL values and rutting for the 2nd set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL av	PWL ac	PWL comb	Avg. Rut (in)	75th Percentile	90th Percentile
3.56	5.20	1	100.0	69.3	91.6	1.04	1.08	1.10
6.02	5.17	2	98.9	58.6	87.9	1.04	1.08	1.10
7.42	5.18	3	69.4	60.4	66.9	1.04	1.08	1.09
8.03	5.20	4	48.6	67.8	53.9	1.02	1.07	1.08
9.57	5.18	5	7.0	62.0	22.0	1.02	1.06	1.07
3.60	5.37	6	100.0	92.1	97.9	0.97	1.01	1.06
5.99	5.35	7	95.4	91.6	94.4	0.97	1.00	1.07
7.48	5.39	8	67.8	91.0	74.1	0.95	0.99	1.05
7.90	5.35	9	54.4	92.0	64.7	0.96	0.99	1.03
9.72	5.35	10	4.9	91.1	28.4	0.95	0.97	1.01
3.37	5.55	11	100.0	98.5	99.6	0.91	0.93	0.97
5.96	5.54	12	97.2	99.6	97.8	0.90	0.93	0.95
7.66	5.58	13	61.2	97.7	71.2	0.89	0.91	0.96
8.11	5.55	14	46.0	99.6	60.6	0.89	0.93	0.94
9.58	5.57	15	4.3	99.7	30.3	0.88	0.90	0.91
3.65	5.74	16	100.0	94.4	98.5	0.86	0.87	0.89
5.87	5.75	17	99.5	94.4	98.1	0.85	0.87	0.88
7.47	5.76	18	70.3	91.1	75.9	0.84	0.86	0.88
8.04	5.72	19	48.4	95.9	61.4	0.85	0.87	0.88
9.43	5.75	20	7.5	89.6	29.9	0.84	0.86	0.87
3.60	5.91	21	100.0	64.6	90.3	0.83	0.84	0.85
5.80	5.91	22	99.7	67.6	91.0	0.82	0.83	0.84
7.40	5.91	23	72.8	63.5	70.3	0.82	0.82	0.84
7.89	5.91	24	54.4	65.4	57.4	0.82	0.82	0.84
9.60	5.93	25	2.4	58.8	17.7	0.81	0.82	0.83

Table 7.15 PWL values and rutting for the 3rd set of MEPDG runs for performance criteria

Mean AV	Mean AC	Run No.	PWL av	PWL ac	PWL comb	Avg. Rut (in)	75th Percentile	90th Percentile
3.62	5.18	1	99.8	60.8	89.2	1.05	1.09	1.11
6.12	5.20	2	96.0	66.2	87.9	1.03	1.07	1.09
7.71	5.19	3	59.6	64.5	60.9	1.03	1.07	1.08
8.12	5.19	4	45.6	64.6	50.8	1.03	1.07	1.09
9.37	5.18	5	12.7	60.0	25.6	1.02	1.06	1.08
3.53	5.35	6	99.9	90.1	97.3	0.98	1.03	1.05
6.04	5.35	7	94.2	91.2	93.4	0.97	1.00	1.04
7.74	5.33	8	60.1	92.0	68.8	0.97	1.01	1.04
7.95	5.36	9	51.5	91.6	62.5	0.96	1.00	1.03
9.37	5.38	10	11.5	93.5	33.9	0.94	0.97	1.01
3.56	5.53	11	100.0	99.6	99.9	0.91	0.95	0.97
5.97	5.53	12	93.2	98.4	94.7	0.91	0.94	0.97
7.40	5.55	13	67.2	98.9	75.8	0.90	0.91	0.96
8.19	5.56	14	43.7	99.8	59.0	0.89	0.91	0.93
9.44	5.58	15	6.7	98.4	31.7	0.88	0.90	0.94
3.53	5.76	16	100.0	91.8	97.7	0.86	0.87	0.90
5.80	5.76	17	96.1	91.5	94.9	0.85	0.86	0.88
7.48	5.78	18	69.7	91.0	75.5	0.84	0.86	0.88
7.87	5.75	19	54.1	93.1	64.7	0.84	0.85	0.89
9.49	5.77	20	6.9	85.4	28.3	0.84	0.87	0.89
3.64	5.91	21	100.0	65.4	90.5	0.83	0.84	0.85
6.12	5.93	22	89.1	57.3	80.4	0.82	0.83	0.84
7.49	5.92	23	68.0	61.0	66.1	0.82	0.83	0.84
7.66	5.90	24	60.3	67.9	62.4	0.82	0.83	0.84
9.48	5.92	25	12.3	63.0	26.2	0.81	0.82	0.83

7.6 Effect of varying HMA thickness on project performance by sublot

In this section, we present the results corresponding to the case of varying HMA surface layer thickness to achieve various PWL ranges and show its effect on rutting and fatigue performance. The pavement structure that was simulated to analyze the effect of variability in thickness on pavement performance is shown in Figure 7.22. This is from an SPS 1 site in Michigan. This section was chosen because it had shown poor performance. This would highlight the effect of variability. If the pavement was performing very well, the problems that will result because of thickness deficiency would get compensated by other strengths of the pavement. The results presented subsequently correspond to fatigue and rutting performance as well as IRI for a design life of 20 years.

Layer No.	Layer Description
5	Original Surface Layer (Layer Type:AC)1.8 Inch
4	AC Layer Below Surface (Binder Course) (Layer Type:AC)1.8 Inch
3	Base Layer (Layer Type:PATB)4 Inch
2	Base Layer (Layer Type:GB)8 Inch
1	Subgrade (Layer Type:SS)

Section 120

Layer No.	Layer Description
5	Original Surface Layer (Layer Type:AC)1.9 Inch
4	AC Layer Below Surface (Binder Course) (Layer Type:AC)2 Inch
3	Base Layer (Layer Type:PATB)4 Inch
2	Base Layer (Layer Type:GB)8 Inch
1	Subgrade (Layer Type:SS)

Section 121

Figure 7.22 Pavement structure used in simulation to study the effect of AC thickness variation on pavement performance

To generate and study different realistic case scenarios, the AC thicknesses were varied over a mean thickness of 7.6 inches as shown in Table 7.16. The variations correspond to mean thickness ± 3 standard deviations. Table 7.16 also summarizes the results for effect of AC thickness variation on fatigue performance. Since MDOT does not have thickness specification limits for HMA pavements, the lower thickness specification limit was set as the target thickness minus 0.75 inch for the purpose of this case study. Figures 7.23 through 7.25 capture the effect of AC thickness variation on fatigue performance using the response surface for the nine different case scenarios listed in Table 7.16.

Similarly, the effect of AC thickness variation on rutting performance is displayed in Figures 7.26 through 7.28 for the nine different case scenarios listed in Table 7.17. Figures 7.29 through

7.31 capture the effect of AC thickness variation on IRI of sublots for nine different case scenarios listed in Table 7.18.

Table 7.16 Summary of results for effect of AC thickness variation on fatigue performance

Mean Thickness (in)	Std_dev of Thickness (in)	PWL	Mean Fatigue (%)	Fatigue 75th %tile (%)	Fatigue 90th %tile (%)
7.1	0.10	99.2	40.6	42.3	43.5
7.1	0.31	78.0	41.1	46.1	50.0
7.1	0.51	67.2	41.7	50.2	56.2
7.6	0.10	100.0	32.1	33.4	34.2
7.6	0.31	99.2	32.3	36.0	39.3
7.6	0.51	92.3	32.9	39.4	45.3
8.1	0.10	100.0	24.5	25.6	26.4
8.1	0.31	100.0	24.9	28.3	31.2
8.1	0.51	99.2	25.3	31.3	35.4

Table 7.17 Summary of results for effect of AC thickness variation on rutting performance

Mean Thickness (in)	Std of Thickness (in)	PWL	Rutting (in)	Rutting 75th %tile (in)	Rutting 90th %tile (in)
7.1	0.09	99.8	0.86	0.86	0.87
7.1	0.27	83.7	0.86	0.87	0.89
7.1	0.45	73.0	0.87	0.87	0.93
7.6	0.09	100.0	0.85	0.86	0.86
7.6	0.27	99.8	0.85	0.85	0.86
7.6	0.45	95.7	0.85	0.85	0.87
8.1	0.09	100.0	0.83	0.83	0.83
8.1	0.27	100.0	0.83	0.83	0.85
8.1	0.45	99.8	0.82	0.84	0.85

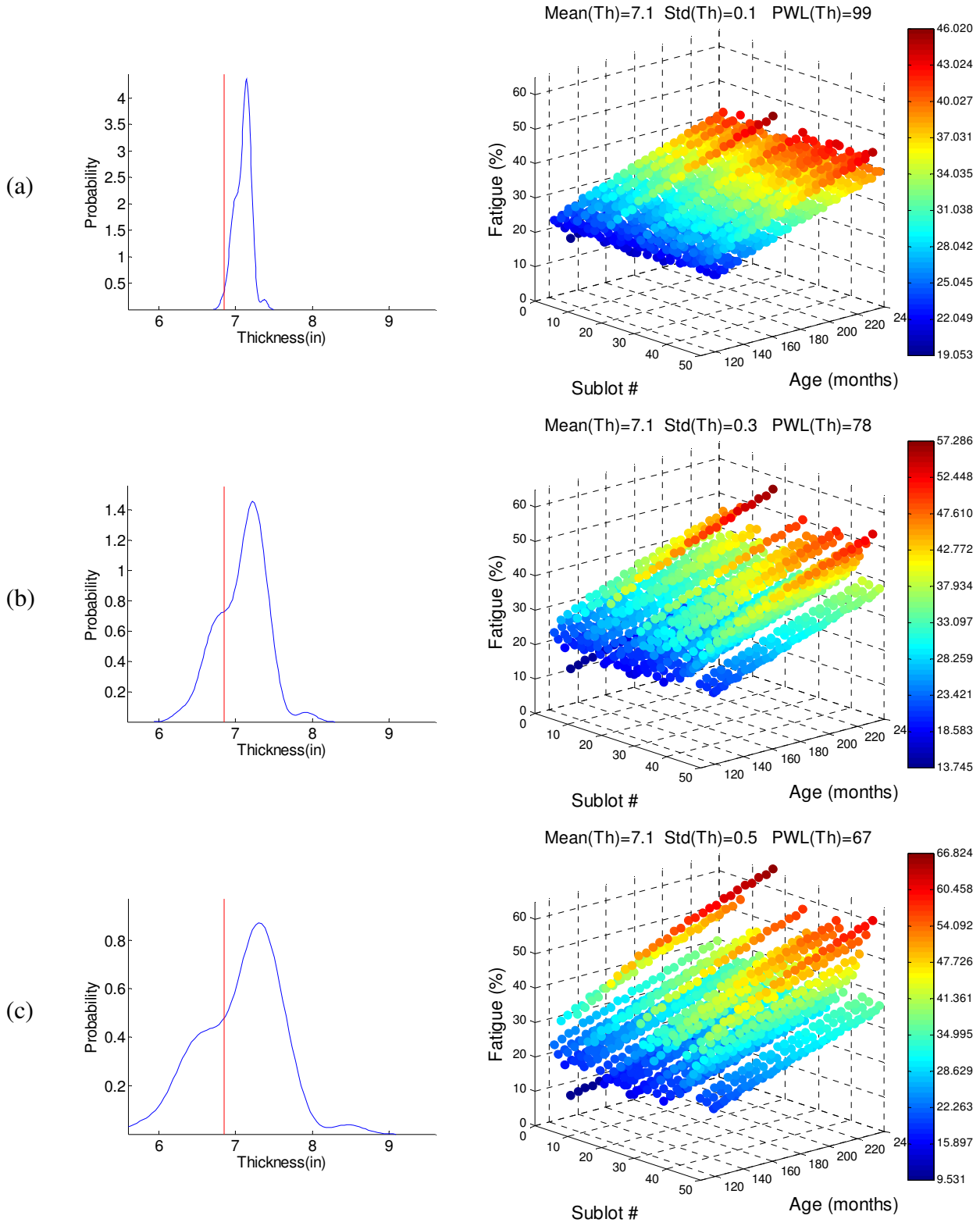


Figure 7.23 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt-0.5.

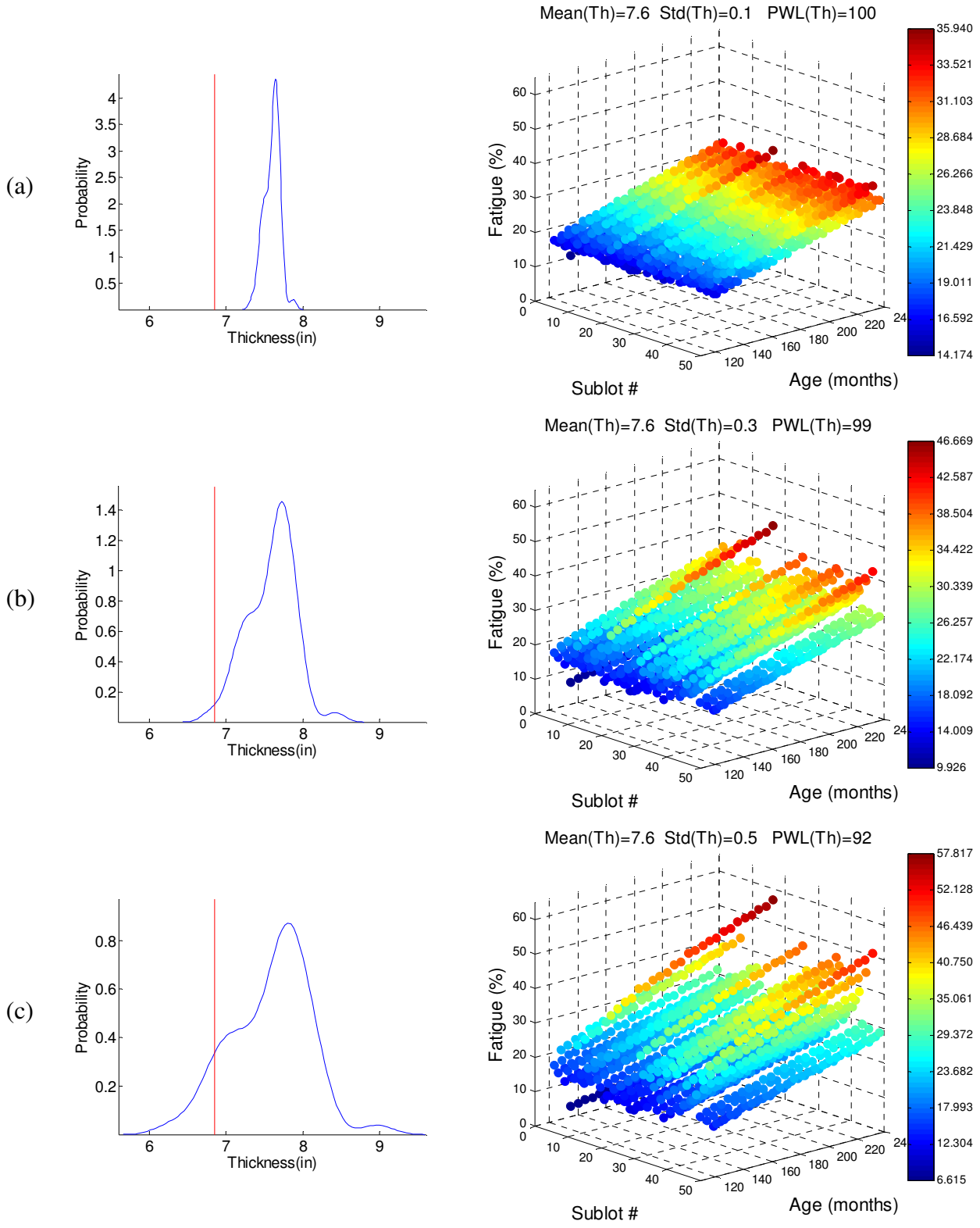


Figure 7. 24 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt.

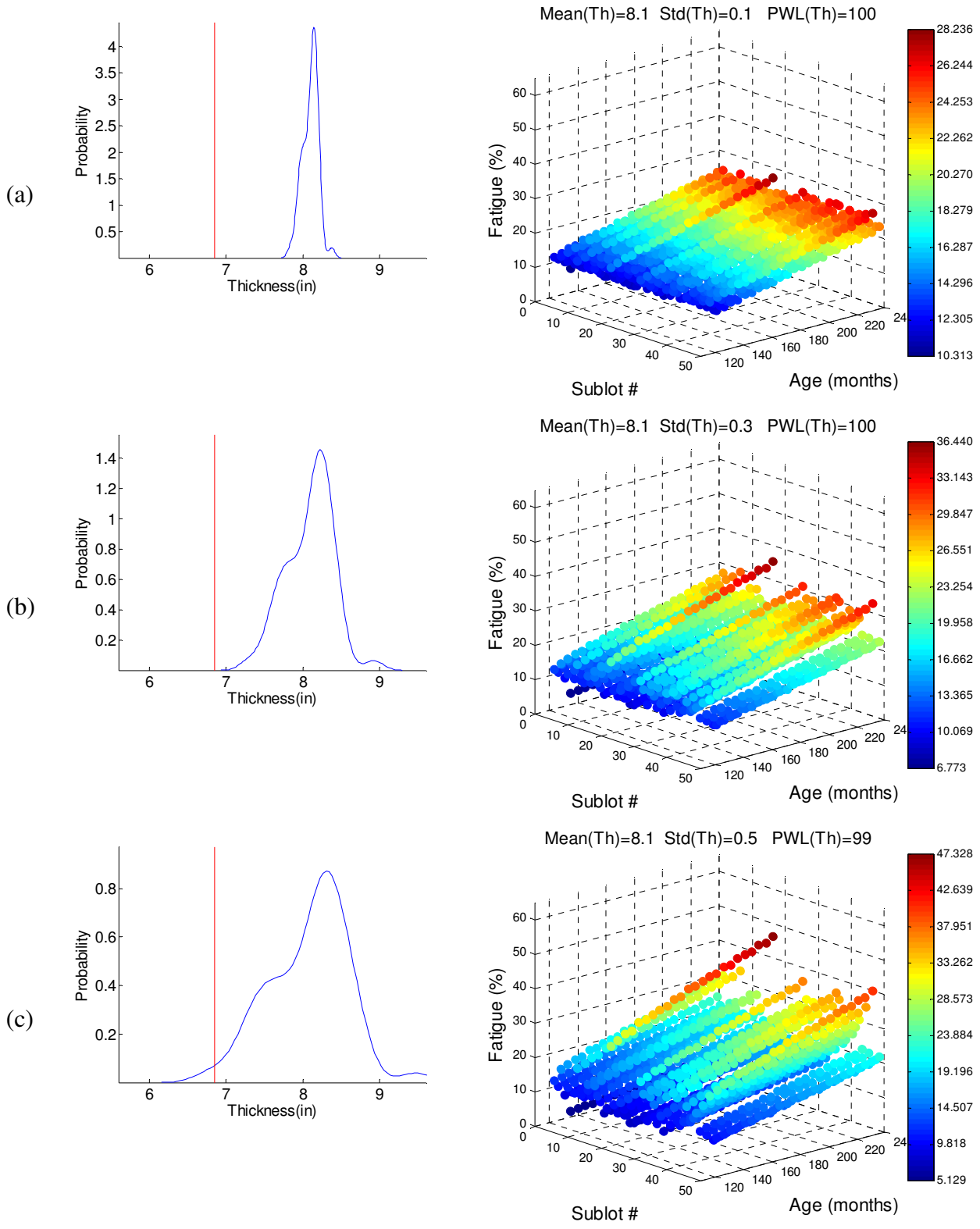


Figure 7. 25 Effect of AC thickness variation on fatigue performance of sublots for Th=Opt+0.5.

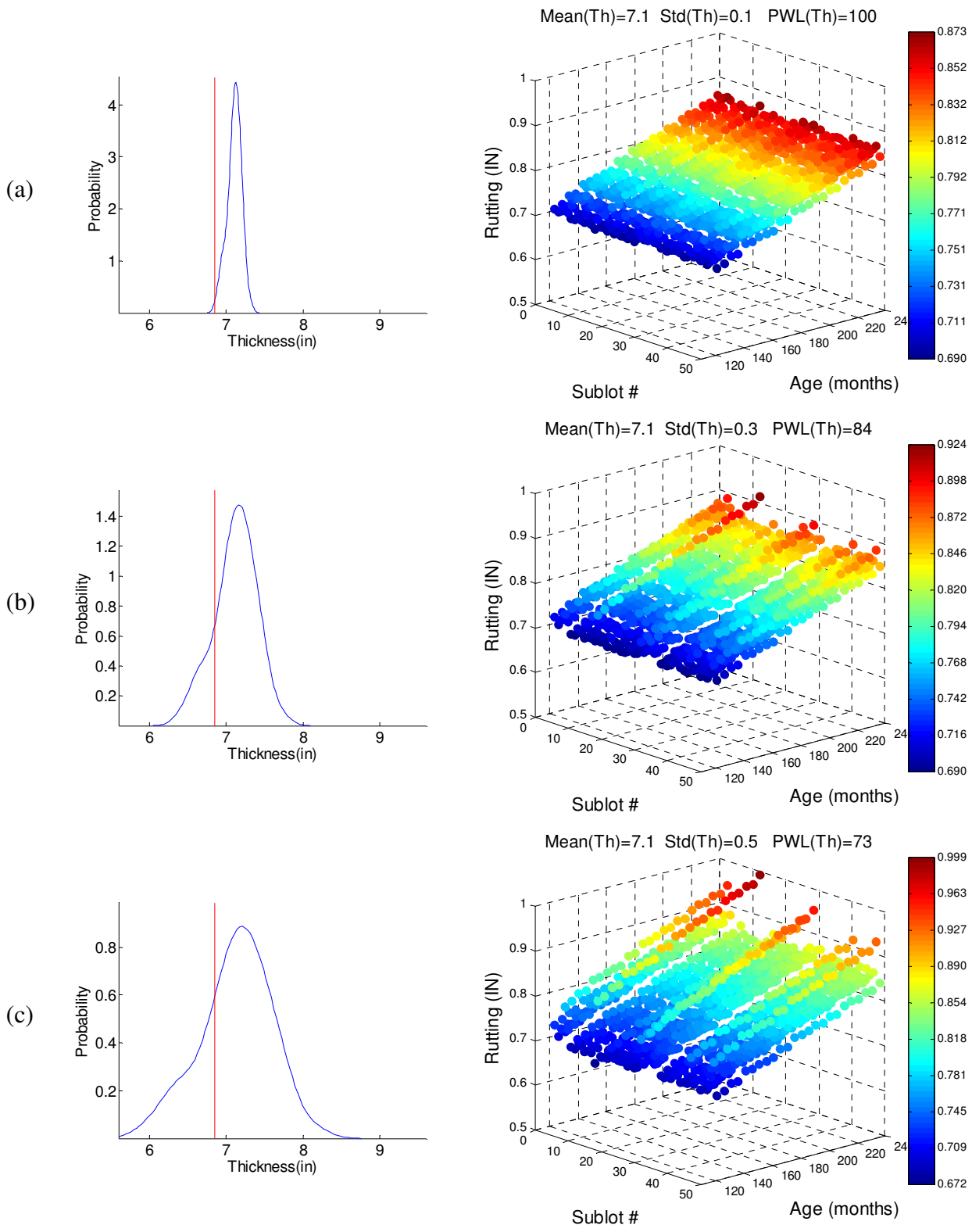


Figure 7. 26 Effect of AC thickness variation on rutting performance of sublots for Th=Opt-0.5.

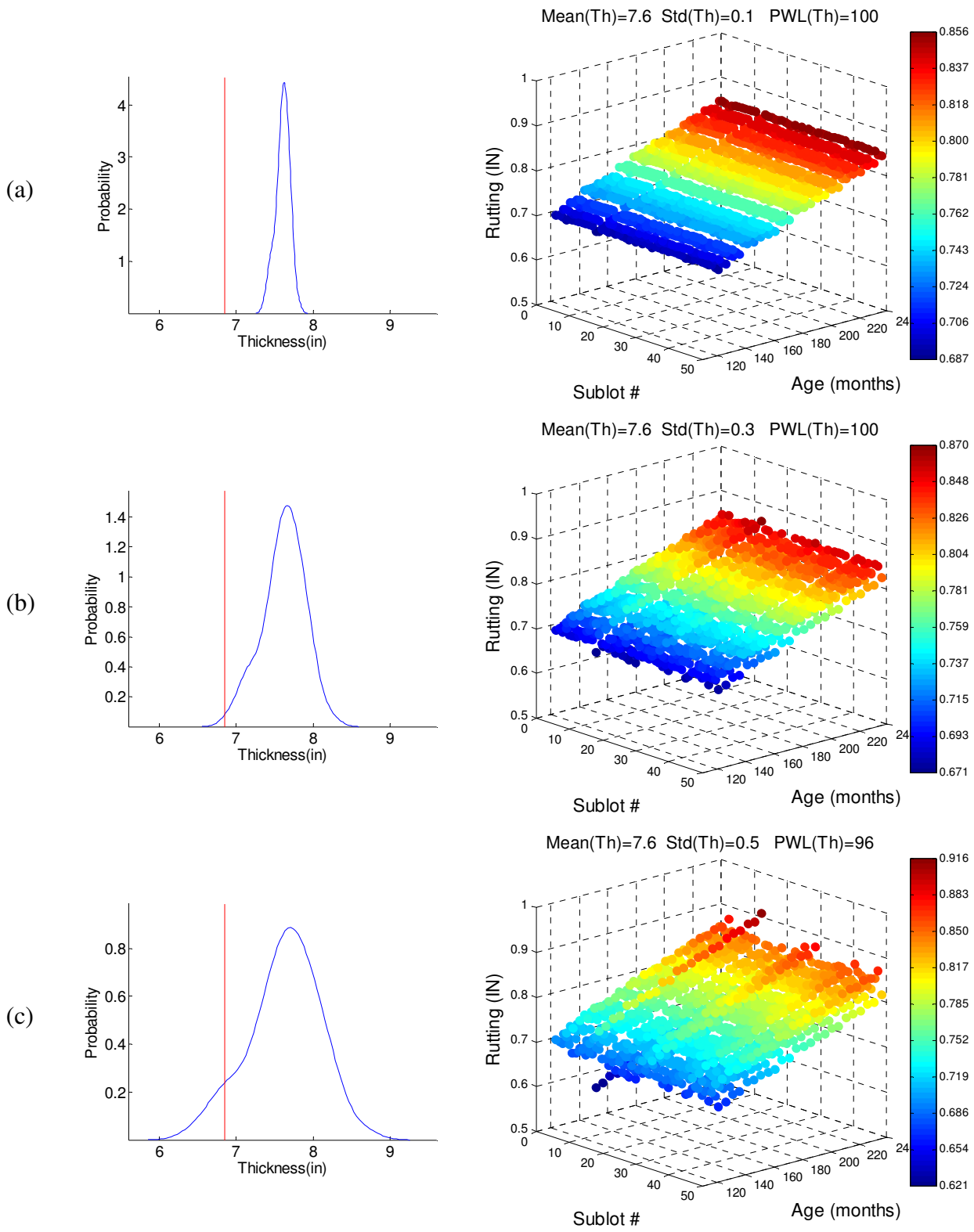


Figure 7. 27 Effect of AC thickness variation on rutting performance of sublots for Th=Opt.

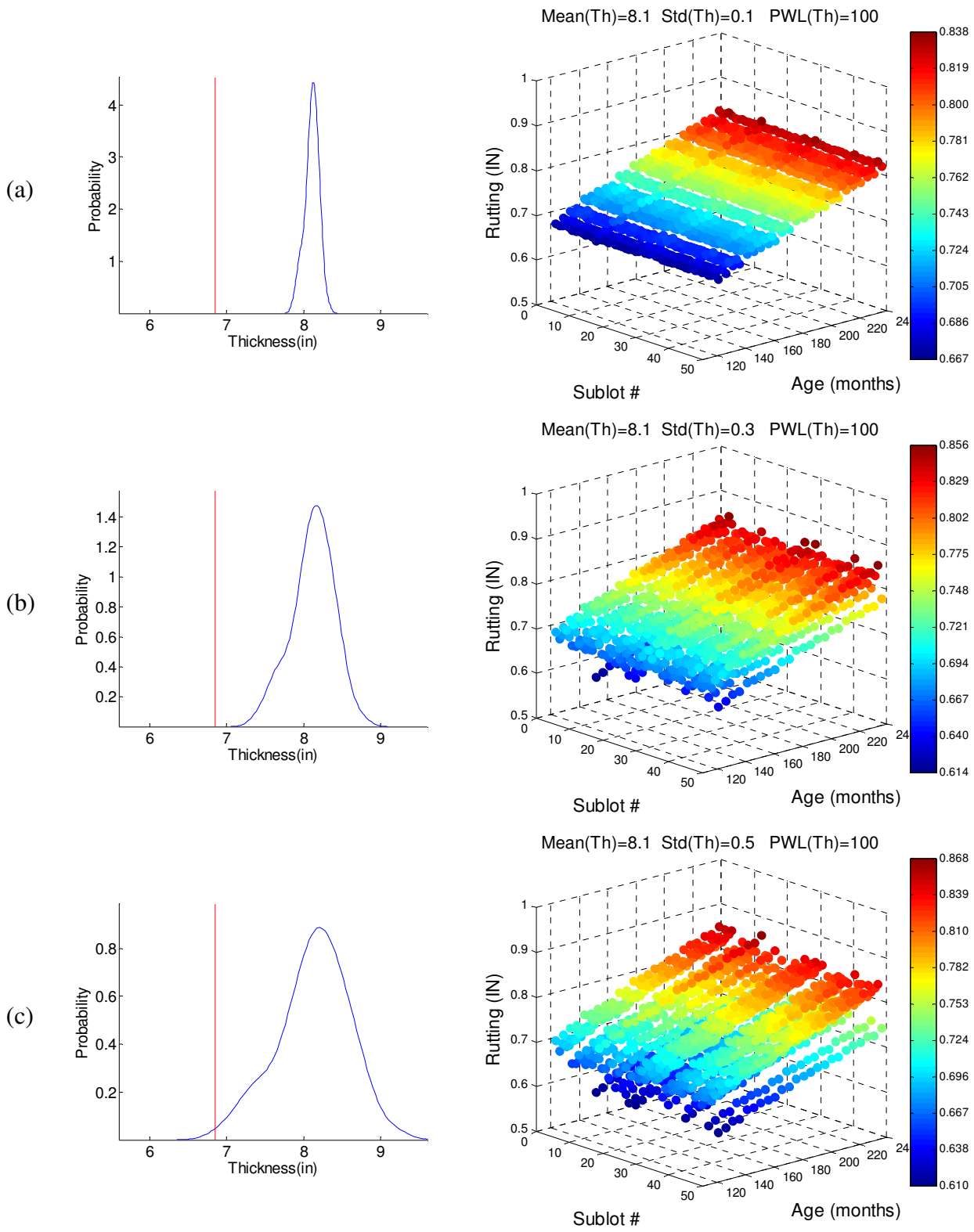


Figure 7. 28 Effect of AC thickness variation on rutting performance of sublots for Th=Opt+0.5.

Table 7.18 Summary of results for effect of AC thickness variation on IRI

Mean Thickness (in)	Std of Thickness (in)	PWL	IRI (in/mi)	IRI 75th %tile (in/mi)	IRI 90th %tile (in/mi)
7.1	0.09	99.5	158	159	161
7.0	0.27	76.0	159	164	170
7.0	0.45	63.2	162	170	182
7.6	0.09	100.0	149	150	150
7.5	0.27	99.5	150	152	156
7.5	0.45	92.6	151	155	164
8.1	0.09	100.0	142	143	143
8.0	0.27	100.0	143	145	148
8.0	0.45	99.5	143	148	152

Table 7.19 displays a sample table from the LTPP database with representative AC layer thickness variation, which also confirms that ranges used for the current study are reasonable and realistic.

It must be mentioned that all these results are for one type of mix and pavement structure. It is important to study different types of mixes and pavement structures to assess how PWL values affect performance for each one of them. It is quite possible that this effect may be much more pronounced in other types of mixes.

Table 7.19 Sample LTPP table with representative AC layer thickness variation

STATE_CODE	SHRP_ID	LAYER_NO	MEAN_THICKNESS	MIN_THICKNESS	MAX_THICKNESS	STD_DEV_THICKNESS
12	4136	4	1.4	1.3	1.5	0.1
12	4137	4	2.8	2.7	2.9	0.1
16	3017	4	4	3.4	4.9	0.1
19	1044	5	1.9	1.6	2.2	0.1
19	1044	6	0.9	0.7	1.2	0.1
24	0500	7	1	0.8	1.1	0.1
88	1647	5	1.6	1.4	1.8	0.1
12	4105	4	2.2	1.8	2.7	0.2
12	4135	4	1.6	1.5	1.8	0.2
19	6049	8	1.6	1.4	2.5	0.2
24	0500	5	2	1.8	2.3	0.2
13	4112	3	3.9	3.3	4.4	0.3
13	4113	3	3.9	3.3	4.4	0.3
19	6049	6	1.6	1.3	2.4	0.3
19	6049	7	1.7	1.1	3	0.3
24	0500	6	1.7	1.2	2	0.3
30	8129	5	3	3	4	0.3
40	1015	2	8.8	8	9	0.3
48	5328	3	4.3	4	4.7	0.3
88	1645	4	1.6	1	2	0.3
12	0900	4	2.3	1.2	3	0.4
13	4112	2	12.2	11.5	13	0.4
13	4113	2	12.2	11.5	13	0.4
19	5042	2	4	3.5	5	0.4
19	6049	3	1.9	1.1	3.5	0.5
19	9116	2	4	3.5	5.4	0.5
19	1044	4	12.6	12	14	0.6
88	1647	4	5.2	4	7	0.6
34	0500	5	2.4	1	3	0.8
34	0500	4	5.7	3	6	0.9
29	7054	5	1.8	-	-	1

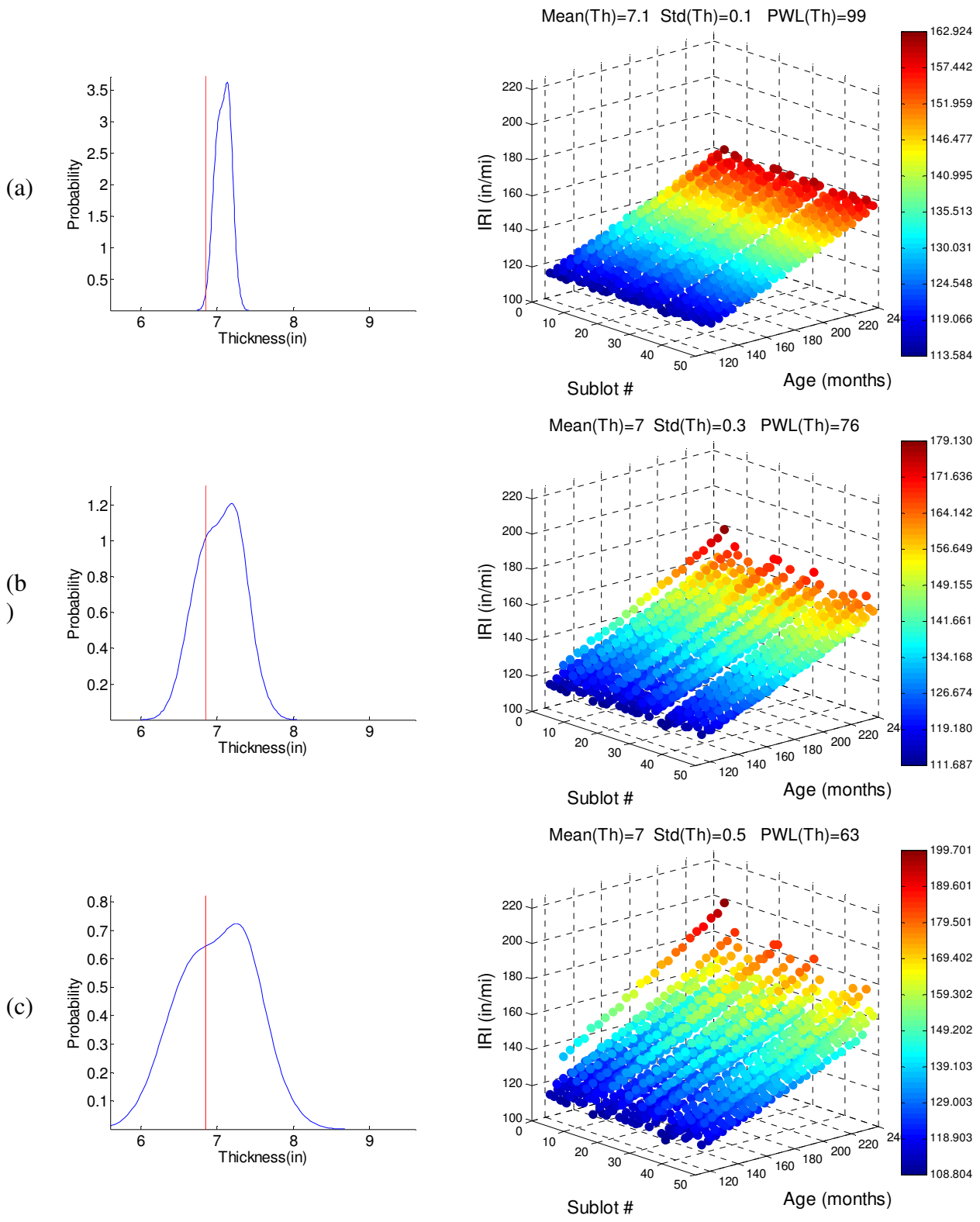


Figure 7. 29 Effect of AC thickness variation on IRI of sublots for Th=Opt-0.5.

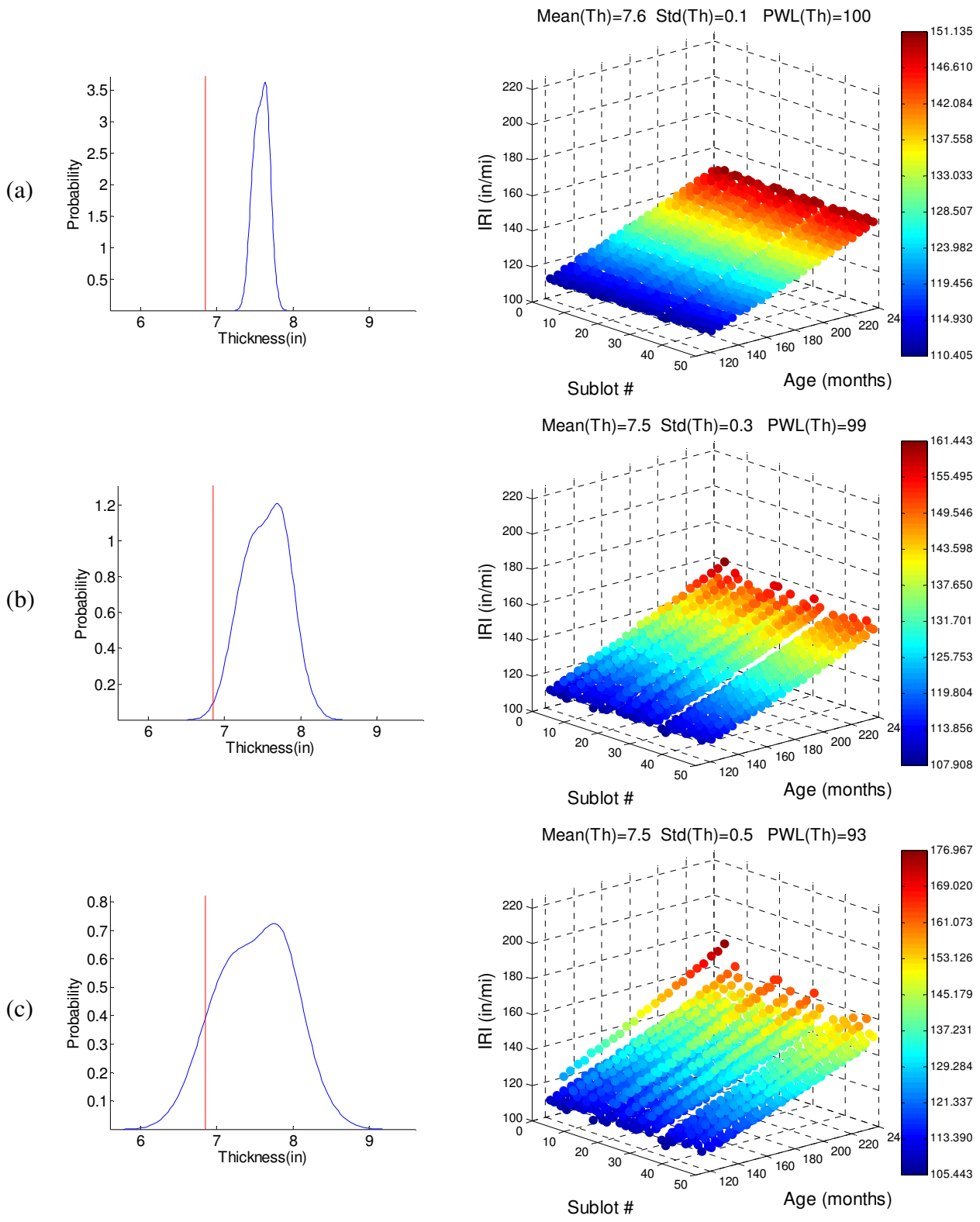


Figure 7. 30 Effect of AC thickness variation on IRI of sublots for Th=Opt.

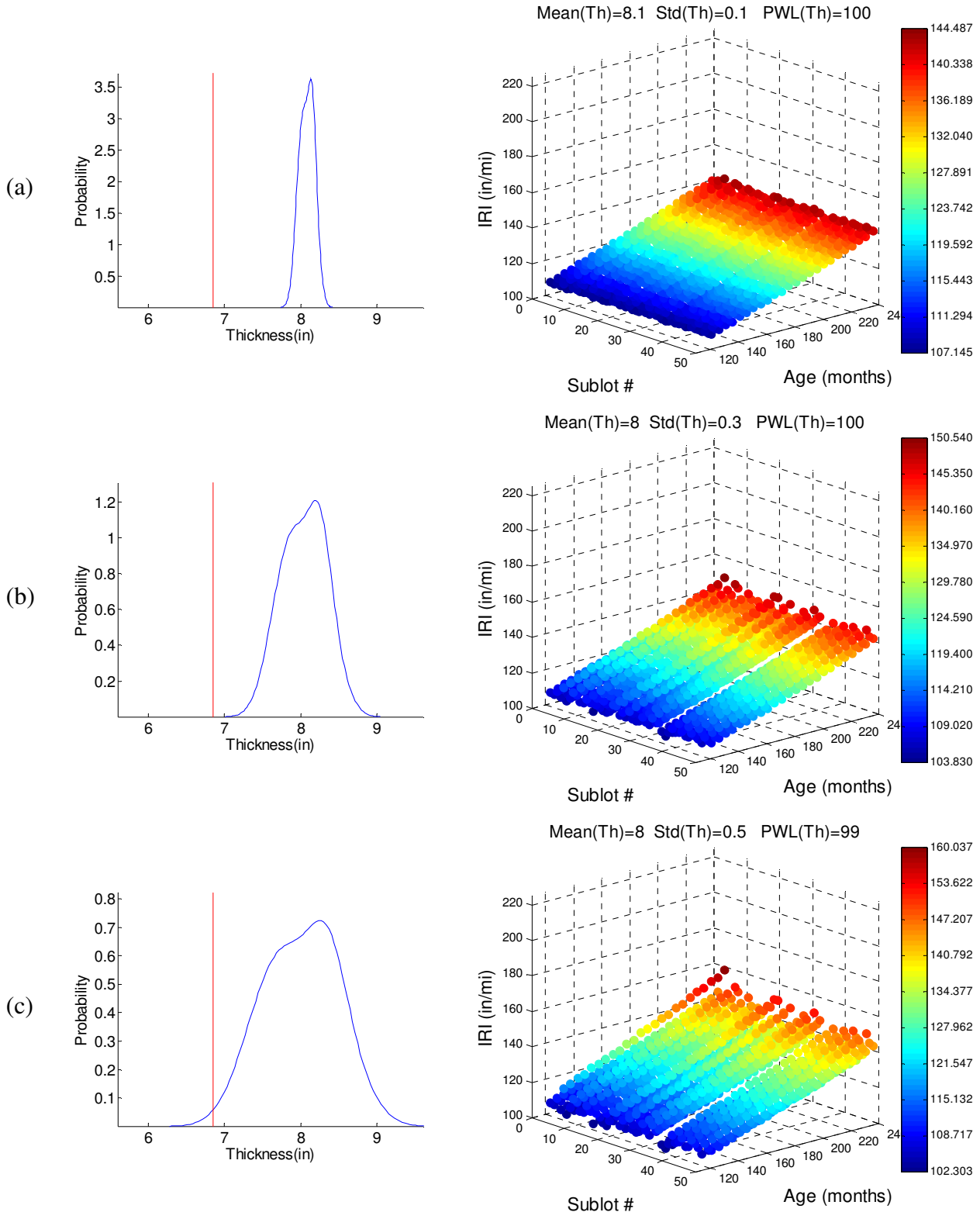


Figure 7. 31 Effect of AC thickness variation on IRI of sublots for Th=Opt+0.5.

Based on the results from the analysis of the effect of AC thickness variation on project performance, the following conclusions are drawn:

1. The performance of different sublots within a project can vary significantly.
2. AC mat thickness variation has a significant influence on fatigue performance. The mean fatigue ranges from 25.3% to 40.6% for AC mat thickness ranging from 7.1 in. to 8.1 in., respectively. Thus, the loss in fatigue performance can be significant with as little as 0.5-in. reduction in AC mat thickness.
3. There seems to be almost 10% increase in fatigue distress for every 0.5-in. decrease in AC mat thickness from 8.1 in. to 7.1 in.
4. The mean IRI ranges from 143 in./mile to 158 in./mile for thicknesses ranging from 8.0 in. to 7.1 in., respectively. It should be noted that in MEPDG, the IRI values are influenced by rutting and fatigue distresses.
5. These simulations and plots provide valuable information in helping us understand how performance varies within a project for such narrow ranges and variations in inputs, which is not possible to gain any other way.

Chapter 8: ERS Risk Analysis Using Simulation

8.1 INTRODUCTION AND BACKGROUND

Over the years many highway agencies in North America have made a valued commitment to End Results Specifications (ERS). As a direct result, it is believed that the quality of our roadways has improved (Smith, 1998, Benson, 1999). A quality Assurance (QA) program, which involves material testing, plays an important role in measuring this quality and is an integral ERS component. The results from the material testing are used to determine payment to be made to the contractor.

Aurilio et al. (2002) considered the effect of differences between laboratory test results on payment to the contractor. Further analysis of actual ERS project data indicated that in addition to test bias several other factors, like measurement variability, production variability etc. can have significant effect on payment. It has also been demonstrated that the concept of simulation program can be used to take into account all the parameters that could be identified to be affecting payment calculation (Aurilio et al. 2002, Manik and Buttlar, 2006). This chapter reports on the development of a Monte-Carlo based simulation program for assessing Michigan Department of Transportation's QA program and estimate the errors or risk involved with payment made to the contractor according to the provisions in the QA program. The chapter presents the details of the program, analysis and conclusions that can be derived from those. Such study would provide valuable insight into how QA programs can be formulated to reduce risk to the contractor as well as the agency and also balance risk.

8.2 END-RESULT SPECIFICATIONS

End-result specification places full responsibility of producing a pavement of a certain specified quality on the contractor. The contractor has full freedom to choose methodologies for construction process and take strategic decisions. He conducts quality control tests at a specified frequency to monitor the quality of the pavement being constructed. The responsibility of the state highway authority (SHA) is to check from their own side that the quality is acceptable,

through quality assurance tests (AASHTO, 1996). The SHA can decide, based on criterion laid out in the specification, whether the quality is acceptable or rejectable, or that the pavement be accepted but with penalty to the contractor in terms of reduced pay. Adjustment in pay is one of the most significant aspects of ERS in present day practices. Rather than setting pass and fail criteria, a percent of the material produced is judged to be within acceptable limits and payment is determined accordingly. This calls for use of statistical methods (Box and Wilson, 1951)

The quality characteristics (defined as that characteristic of a unit or product that is actually measured to determine conformance with a given requirement) that are being used to determine "quality" of the pavement are generally air voids, binder content, voids in mineral aggregates, density etc. for flexible pavements. These quality characteristics are believed to be related to performance but the exact relationships are not yet firmly quantitatively established for all of them. Therefore, the pay adjustments are based on the values of the quality characteristics themselves and not on expected performance of the constructed pavement (Smith, 1998).

8.3 ESTIMATING RISK

In the past, researchers have attempted to develop statistical or simulation tools to help understand and balance risks in asphalt construction specifications. A computer simulation program called OCPLOT, developed in FHWA Demonstration Project 89 by Weed (Weed, 1996) is available for generating Operating Characteristics (OC) curves. OCPLOT was found to be user-friendly and very useful for initial assessment of relative risks, allowing the user to vary the following factors: sample size, pay factor equation, specification limits, and retest provisions. The program allows the user to assess the probability of acceptable material being rejected (defined as contractor risk) and the probability of rejectable quality material being accepted (defined as agency risk) over the long run (e.g., when considering the characteristics of the specification over a long period of time). However, a number of the factors that appear to be related to risk, including measurement variability and testing bias are not considered in OCPLOT. In addition, it can be argued that the most tangible measurement of risk should be linked to the financial impact on the project, i.e., how *risk* affects *what is actually paid versus what should have been paid*.

One of the necessary steps in the assessment of payment risk is to clearly define the risk metric. A very straight-forward and yet very effective way of defining risk could be as shown in equation 1, where baseline pay represents the ideal or correct payment. .

$$\text{Payment Risk} = \text{Payment made to the contractor} - \text{Base Line pay} \quad (1)$$

Ideally, tests performed by different parties on the same material should give very similar results. However, in practice even split samples will show different results when the tests are carried out by two different agencies or in two different labs. Because of these uncertainties there is a risk of accepting rejectable quality and vice-versa. In the ERS approach, a percentage of acceptable quality (Percent Within Limits-PWL) is determined rather than pass/fail criteria used in typical QC/QA. Then, payment is made based on this percent within limits value (Patel, 1996). Because of the uncertainties involved with the test results, the payment made also may be more or less than what it would be if the actual quality of the construction would have been exactly determined (Weed, 1996; Willenbrock, 1976; Bowery and Hudson, 1976; Barros et al. 1983; Puangchit et al., 1983; Afferton et al., 1992; AASHTO, 1995). Overpayment of the contractor is often referred to as 'agency risk' while underpayment is often termed as 'contractor risk'. Throughout this report, *positive values of risk refer to the instance where the agency paid more than required (agency risk) and negative values of risk indicate that the agency paid less than what the contractor deserved (contractor risk).*

Buttlar and his coworkers at the University of Illinois at Urbana-Champaign have developed a series of risk simulation models that provide the user a virtual environment to quickly generate and analyze thousands of realistic ERS data sets. The first simulation model developed was ILLISIM (Buttlar and Hausman, 2000). This was followed by PaySim and BiasSim (Aurillio et al. 2002) which used different models and catered to different aspects of risk analysis and simulation. The latest model developed for the Illinois Department of Transportation is called Simulate Risk Analysis, or SRA, which combines the capabilities of all earlier programs into a single program, with added features to simplify the process of conducting sensitivity analyses (Buttlar and Manik, 2007). Using the same principles, a new simulation model called AMSim has been developed to analyze the MDOT QA program by the authors.

This chapter presents the details of this simulation model along with the analysis performed and conclusions derived from the analysis.

8.4 MAJOR FACTORS AFFECTING RISK

Analysis of data obtained from actual construction projects corresponding to various quality characteristics like density, binder content, air voids etc., have shown that such data are generally normally distributed *(Hall and Williams, 2002)*. Generally, the target values for these quality characteristics are fixed. This indicates that, as it would be expected in the real world, there are certain factors involved in the construction and quality characteristic measurement procedures which are not completely controllable, or even predictable. They tend to induce variability (Benson, 1995) in the quality around the targeted quality level.

Variability observed in the field, however, has at least two components, namely production variability and measurement variability. Production variability includes all variability introduced due to workability of concrete, variability in the quality and physical characteristics of source materials, changes in the relative proportions of ingredients in the mix, changes in plant operational characteristics, changes in equipment operators, changes in ambient temperature etc. Measurement variability is the variability which is introduced by the measuring devices, test procedures, and operator techniques and human error. In addition to variability around the actual value, a measurement bias may be introduced as well. Bias refers to a consistent shift in data and can be introduced by device calibration errors, human error, or by the intentional biasing of measurements and/or recorded data (Aurilio et al. 2002).

Every choice made in the development of an ERS comes with an associated risk. Risks are undertaken by both the contractor and the agency. The introduction or manipulation of certain specification attributes can shift the risk from the contractor to the agency and vice-versa. Other specification attributes can widen or narrow the range of risk. In summary, the key contributors to risk in ERS are:

- Contractor data versus agency data
- Frequency of testing and/or number of samples
- Variability and/or bias of test device and/or test procedure
- Specification parameters, including:
    - Specification limits
    - Pay factor equation
    - Pay “caps”

- Acceptance test logic and frequency and acceptance tolerance
- Third party testing provisions

In the procedure used for determining the pay factor in ERS, a sample of data with finite measurements is used to estimate the quality of a population which this sample belongs to. Therefore, mean and standard deviation of any quality characteristic of the sample is considered as equal to the mean and standard deviation of the entire material in the lot or pavement produced in that project. However, the finite sample being used may not have exactly the same mean and standard deviation as it could have been if a much larger number of samples were collected. Theoretically, actual quality of the material can be determined only if the sample collected is infinite. Such an infinite size sample, or rather population, would give payment called as "ideal payment". Therefore, finite size samples would lead to a deviation from the ideal pay. In addition, the use of imperfect measuring devices would also lead to error in measurements. The error in turn would lead to deviation from the ideal pay. Therefore, to be able to determine the ideal pay, thousands of data with similar characteristics would need to be simulated, with each simulation representing an actual individual project. Pay calculated for each individual project coupled with the ideal or base-line pay for the entire population would provide distribution of risk on a project with those characteristics.

In order to simulate variability in an asphalt pavements material properties, one must be able to sequentially simulate, in this order: 1) production or construction variability; 2) results of random samples taken from that variable material; 3) the effects of measurement variability on the estimated properties; and finally; 4) the effects of bias on the final reported test measurement values. In order to estimate risk in terms of effects on pay, the software must also simulate the formulas and decision tree logic contained in the construction specification.

8.5 COMPOSITE RISK INDEX

A simulation tool like AMSim or SRA helps estimate and analyze risk in payment that can be expected in different scenarios using a certain set of end-result specifications. The main advantage of such a tool is that it can provide invaluable information in what-if scenarios without the need of a demonstration project or shadow specification. This can greatly help in determining the effect of different aspects or values in the specification used in end-result projects.

The main format in which AMSim would provide information would be risk plots. A risk plot presents the expected mean risk and associated confidence interval for the entire range of quality characteristic possible on a project. This means that a risk plot can give a very good understanding of how "well" a set of specifications would do for that quality characteristic.

A wealth of information can be gained from the risk plots generated by SRA. However, the interpretation of the risk plot could be subjective. This may make it difficult to compare risk scenarios arising because of two different specifications or any combination of other parameters affecting risk. In addition to this, if an algorithm needs to be developed for comparing risk plots for the purpose of comparing specifications etc. various quantitative characteristics of the plot would have to be used. Manik (2006) developed a composite risk index (CRI) to quantitatively characterize the risk plots. The concept of CRI was tested on a wide range of risk plots and was found to be very objective and promising in its purpose. The analysis presented in this chapter also uses CRI.

8.6 RISK ANALYSIS

One of the earlier sections in this chapter identified several factors associated with a QA program that affect the risk involved in payment made to the contractor through that program. It is very important to assess how exactly these factors affect payment risk. In addition to that, it is also important to determine how their contribution to other factors influence payment risk. This section presents risk analysis performed for the MDOT QA program with the following four factors in focus.

(1) Production Variability
(2) Measurement Variability
(3) Sample size and
(4) Bias

Three levels were identified for each of these four quality characteristics and a full factorial run matrix was constructed as shown in Table 8.1.

Table 8.1 Run Matrix for Risk Analysis of Flexible Pavements

Sample Size	Bias	Prod Var** = 0.45			Prod Var = 1.4			Prod Var = 2.3		
		Meas var*=0.1	Meas var=0.5	Meas var=1	Meas var=0.1	Meas var=0.5	Meas var=1.0	Meas var=0.1	Meas var=0.5	Meas var=1.0
10	0.1	1	10	19	28	37	46	55	64	73
	0.3	2	11	20	29	38	47	56	65	74
	0.5	3	12	21	30	39	48	57	66	75
40	0.1	4	13	22	31	40	49	58	67	76
	0.3	5	14	23	32	41	50	59	68	77
	0.5	6	15	24	33	42	51	60	69	78
70	0.1	7	16	25	34	43	52	61	70	79
	0.3	8	17	26	35	44	53	62	71	80
	0.5	9	18	27	36	45	54	63	72	81

* Measurement Variability in %

** Production Variability in %

AMSim simulates the entire MDOT QA program including the specification limits, sampling scheme and decision logic ending with pay factor calculations. The MDOT QA program has similar specifications for flexible pavement in-situ density, air voids, binder content and VMA. The sampling scheme for in-situ density is different because samples are collected from the mat immediately after compaction. In this case, randomization is used not only in the longitudinal direction but in transverse direction also. For the other factors, quality characteristics samples are collected in the form of mixture and, therefore, randomization is applied to the entire mixture in each sublot. The analysis presented here first corresponds to in-situ density followed by plant air voids. For all the four quality characteristics, pay formula is used instead of pay-schedule. In the past, pay schedules were very commonly used. However, with the increasing use of statistical methods in ERS, pay schedules have been replaced by percent-within-limits concept and pay formula (Buttlar and Harrell, 1998). Table 8.2 shows an example of what a pay schedule looks like. This is the pay schedule used by MDOT for thickness QA program for rigid pavements. Equations 1 and 2 show the pay formulae used for density QA program for flexible pavements. For PWL less than 50 corrective action is required as described below (MDOT specifications).

Table 8.2 Price Adjustment for Concrete Thickness Deficiency

Initial Core Type	Deficiency in Thickness (Inch)	Price Adjustment (Percent)
A	0.20 or Less	0
B	0.30	-5.0
B	0.40	-15.0
B	0.50	-25.0
B	0.60 To 1.0	-50.0
C	1.10 and Over	-100 [a]

A. If PWL for In-Place Density (PWL_D) is between 100 and 70, use the following formula to determine PF_D. Round the value of PF_D two decimal places.

$$PF_D = 55+(0.5xPWL) \quad (1)$$

B. If PWL for In-Place Density is between 70 and 50 inclusive, use the following equation to determine PF_D. Round the value of PF_D two decimal places.

$$PF_D=37.5+(0.75xPWL) \quad (2)$$

C. If PWL for In-Place Density is less than 50; the Engineer may elect to do one of the following:

(1) Require removal and replacement of the entire lot with new QA sampling and testing and repeat the evaluation procedure.

(2) Allow the lot to remain in place and apply an Overall Lot Pay Factor of 50.00.

(3) Allow submittal of a corrective action plan for the Engineer's approval. The corrective action plan may include removal and replacement of one or more sublots. If one or more sublots are replaced, the sublot(s) will be retested and the Overall Lot Pay Factor will be recalculated according to this special provision. If the Engineer does not approve the plan for corrective action, subsections (1) or (2) above will be applied.

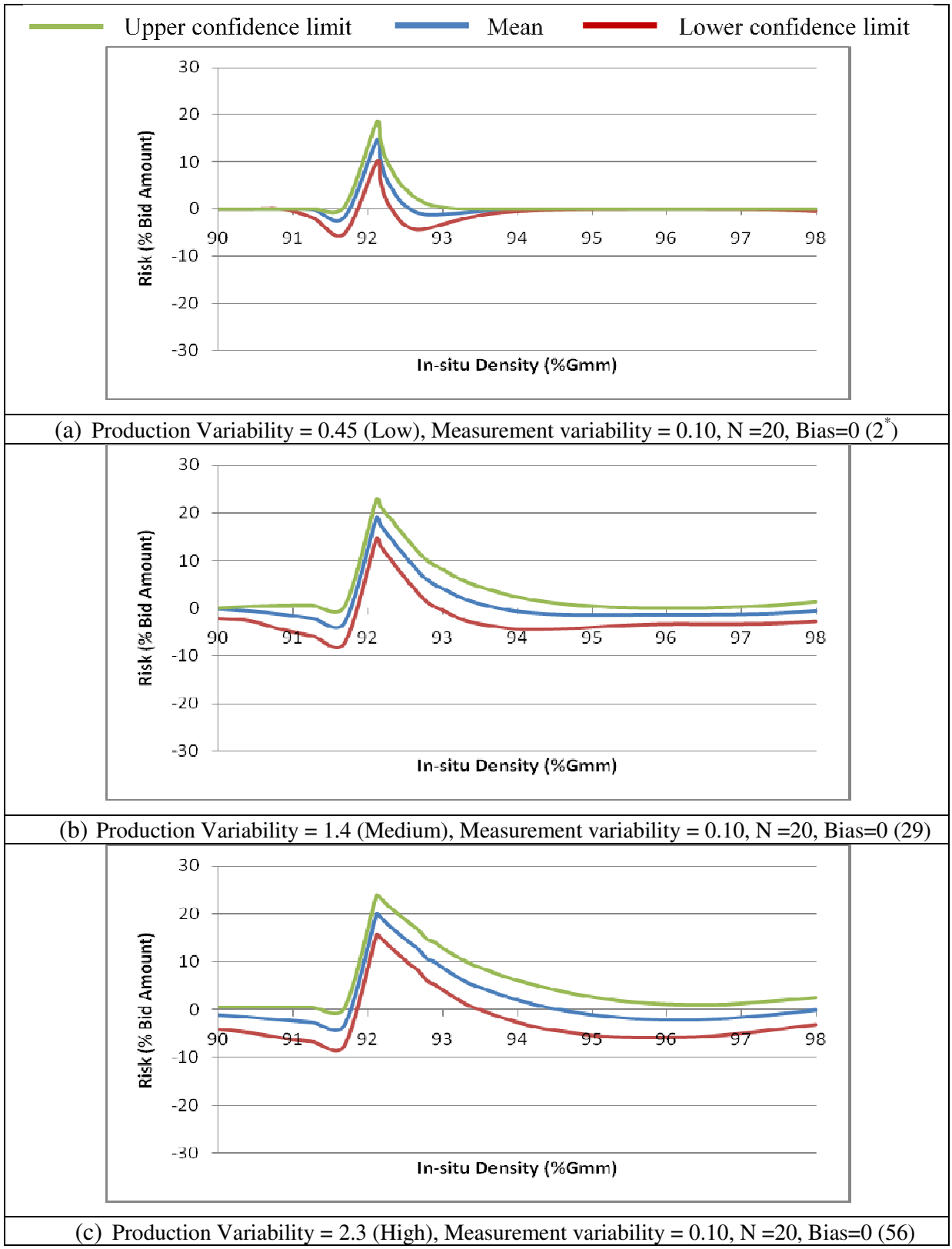
AMSim was run for all the 81 cases identified in the run matrix (Table 8.1). Figures 8.1 through 8.4 show sample results from the 81 runs performed with AMSim. It would be important to describe the concept of risk plot first. The x-axis in the risk plot has mean of the quality characteristic in the QA program which is being analyzed. The three risk curves shown in each of the figures correspond to the mean and upper and lower 90% confidence interval. Any point

on the mean risk plot will show a magnitude of payment risk with 50% likelihood if the mean of the quality characteristic achieved in a specific project is equal to the corresponding in-situ density value.

Figures 8.1 through 8.4 were selected to demonstrate some of the salient conclusions that can be derived through this analysis as listed below. This will be followed by analysis of variance and corresponding conclusions for the entire run matrix.

(1) The analysis presented in this chapter shows the effect of using pay formula instead of pay schedule. Since pay formula is a continuous function, risk is also smooth except for sharp points of inversion around the specification limits. The sharp inversion in the risk plot around the specification limits means that an error in the measured quality characteristic value putting it on one side of the specification limit would lead to a substantially different pay factor compared to that on the other side of the step for the same level of quality achieved. This is an undesirable feature of a QA program and efforts should be made to reduce the magnitude of risk and narrow the window in which inversion occurs even if it can be eliminated.

(2) Figure 8.1 shows the effect of production variability on risk. Plots a, b and c correspond to low, medium and high production variability with all other factors being the same.
    a. It is interesting to note that as the production variability increases from 0.45 to 1.3 the increase in payment risk is sharp.
    b. The increase in magnitude dampens for higher production variability values.
    c. With higher production variability the confidence interval around the mean risk widens considerably.
    d. More importantly there is appreciable amount of risk of overpayment or underpayment even when the contractor produces in the middle of the specification (allowable) window. In this case, the allowable window is 92% to 100%. This is highly undesirable. It has been observed that stricter allowable production variability on construction projects, within reasonable limits, can lead to lower production variability and thus reduction in this undesirable characteristic.
    e. It is important to note that risk is not equally balanced between the contractor and the agency. Especially, if the contractor's production mean is around the specification limit it is expected that the agency will overpay him. A good specification should have balanced risk. The advantage of balanced risk is that if pay factors are calculated on per lot basis, most of overpayment and underpayment would cancel each other for the aggregate payment for the entire project.



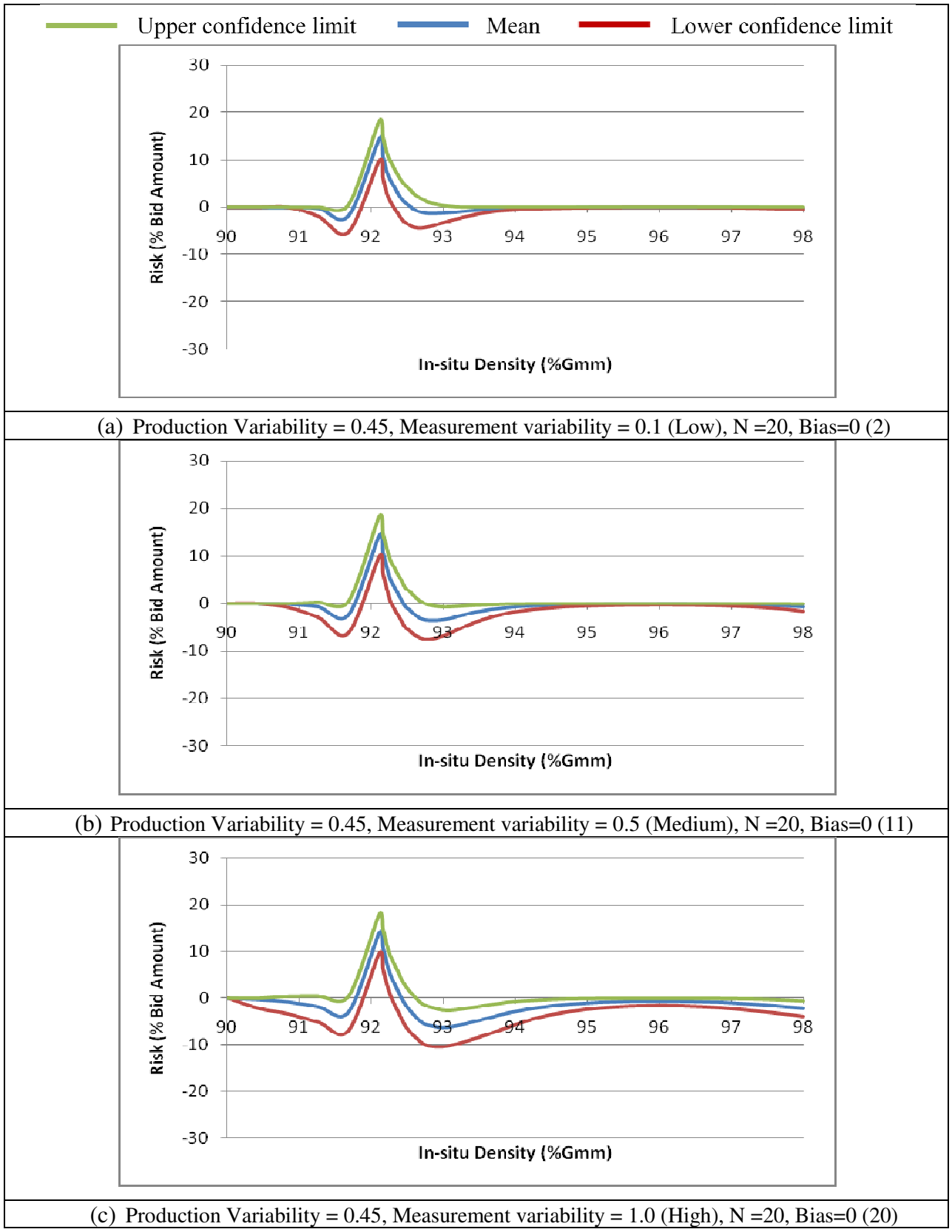
(a) Production Variability = 0.45 (Low), Measurement variability = 0.10, N =20, Bias=0 (2*)

(b) Production Variability = 1.4 (Medium), Measurement variability = 0.10, N =20, Bias=0 (29)

(c) Production Variability = 2.3 (High), Measurement variability = 0.10, N =20, Bias=0 (56)

Figure 8.1 Effect of production variability on risk for flexible pavements

*Run number for the case in the run matrix (Table 8.1)

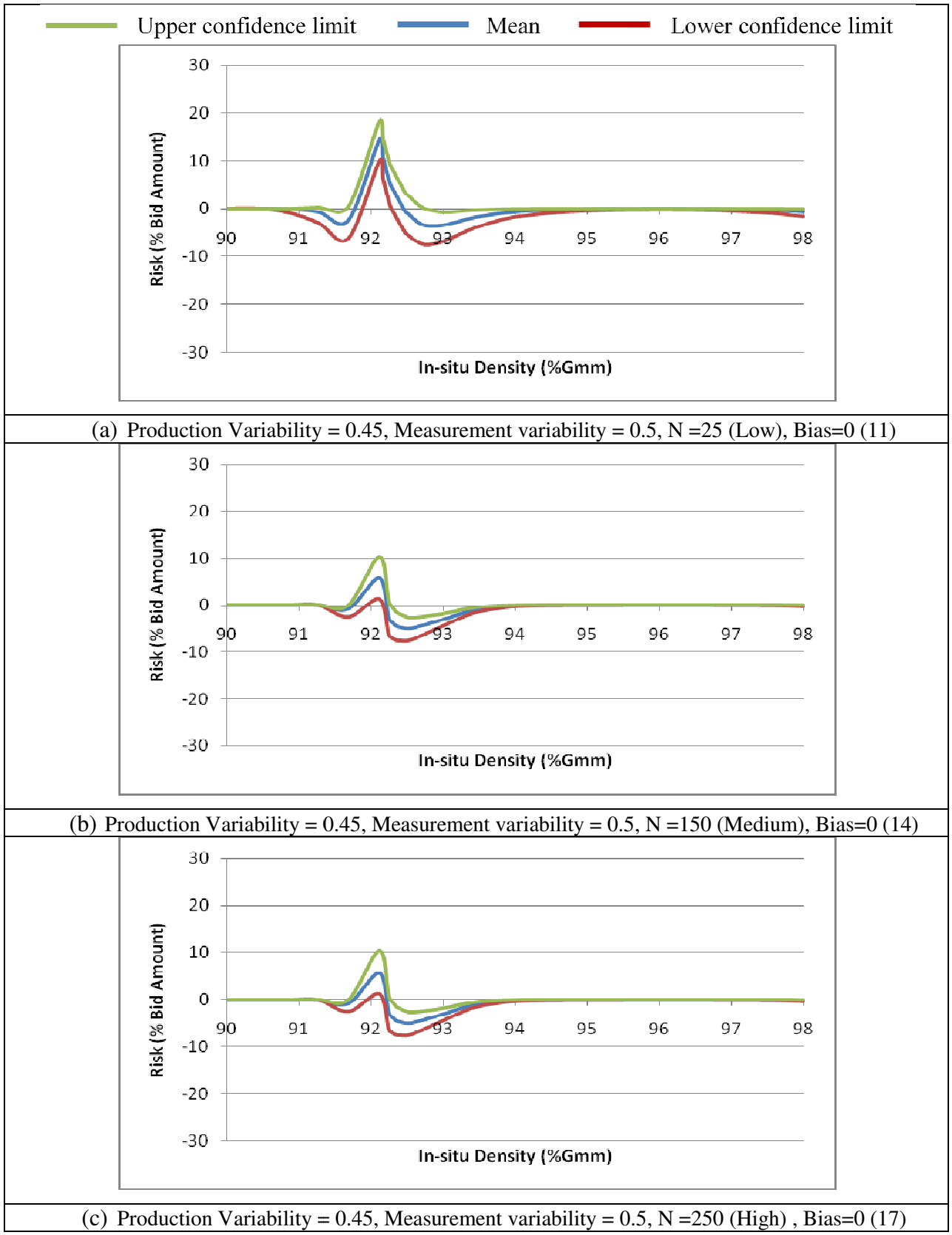


(a) Production Variability = 0.45, Measurement variability = 0.1 (Low), N =20, Bias=0 (2)

(b) Production Variability = 0.45, Measurement variability = 0.5 (Medium), N =20, Bias=0 (11)

(c) Production Variability = 0.45, Measurement variability = 1.0 (High), N =20, Bias=0 (20)

Figure 8.2 Effect of measurement variability on risk for flexible pavements

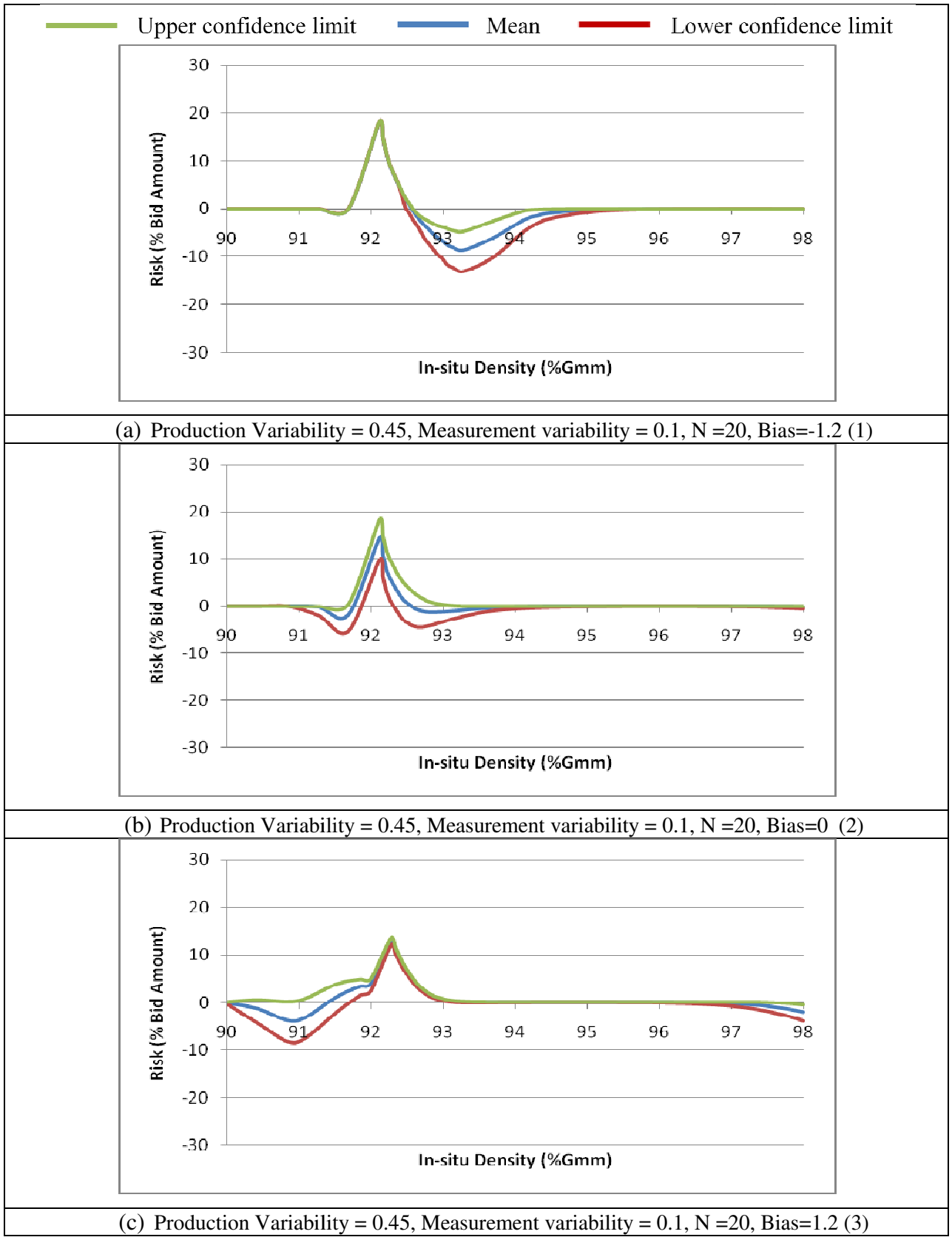


(a) Production Variability = 0.45, Measurement variability = 0.5, N =25 (Low), Bias=0 (11)

(b) Production Variability = 0.45, Measurement variability = 0.5, N =150 (Medium), Bias=0 (14)

(c) Production Variability = 0.45, Measurement variability = 0.5, N =250 (High) , Bias=0 (17)

Figure 8.3 Effect of sample size on risk for flexible pavements



(a) Production Variability = 0.45, Measurement variability = 0.1, N =20, Bias=-1.2 (1)

(b) Production Variability = 0.45, Measurement variability = 0.1, N =20, Bias=0 (2)

(c) Production Variability = 0.45, Measurement variability = 0.1, N =20, Bias=1.2 (3)

Figure 8.4 Effect of measurement bias on risk for flexible pavements

(3) The plots in Figure 8.2 show the effect of measurement variability on payment risk.
   a. Increase in measurement variability leads to an increase in payment risk (compare plot (a) to plots (b) and (c)).
   b. The increase in risk because of measurement variability increasing from 0.1 to 0.5 is not as high as that because of increase in production variability. This is not always true. In this case, the measurement variabilities used are that of density measurement using pavement cores. This process has higher level of accuracy than most of the other tests used in quality assurance programs. This highlights the fact that more accurate test methods would lower risk in payment.

(4) Sample size has significant influence on payment risk as is evident from the plots presented in Figure 8.3. In all the three cases measurement variability was 0.5% which generally leads to appreciable level of payment risk. However, as the sample size becomes larger risk goes down considerably. In addition to the lowering of the risk, an increase in sample size also leads some amount of redistribution and therefore balancing of risk between the agency and the contractor. Redistribution of risk can be higher in other cases as was found with thickness specifications for the rigid pavements (see chapter 17).

(5) Figure 8.4 shows the effect of measurement bias on risk. The first plot in Figure 8.4 corresponds to a bias of -1.2%, the second plot to no bias and the third to a positive bias of 1.2%. A bias of -1.2% means that the agency consistently measures thickness to be lower than what it would even if measurement variability were present. In other words, the mean of a large sample of thickness measurements would be lower than the actual value by roughly 1.2%. If there were no bias and only measurement error was present, the mean of such a large sample of thickness measurements would be very close to the actual thickness. The three plots in Figure 8.4 clearly shows that bias can not only affect the magnitude of the risk, it can completely change the sign of the risk as well. When the bias is -1.2% (negative), risk just right of the lower specification limit is high and the contractor is expected to be underpaid. However, when bias is 1.2% (positive) the risk in the same region collapses to nearly zero and increases in the region just left of the specification window. Therefore, bias must be controlled carefully and eliminated from measured value through proper testing and calibration in the initial lot.

8.7 ANALYSIS OF VARIANCE (ANOVA)

The comparisons in the preceding section among different cases from the run matrix show the effect of the four variables namely production variability, measurement variability, sample size and bias. Analysis of variance can not only help quantify the effect of these factors on payment risk but also it can give insight into the interaction effects of these factors. However, to be able to run ANOVA, the risk plots by themselves cannot be used. The concept of Composite Risk Index (CRI) was presented earlier in this chapter. CRI helps assign one index

value to a risk plot representing one case scenario considering several factors simultaneously. Without the use of such an index, thorough statistical analysis with such scenarios would be nearly impossible. Table 8.3 shows the values of CRI for all the 81 cases in the run matrix.

Table 8.4 shows the ANOVA table for CRI for all the 81 cases in the run matrix. Note that X1 through X4 represent the four factors being analyzed here and have been listed below the table. The following conclusions can be derived from the table.

(1) The p-values for main effects of the four factors show that all of them except production variability are statistically significant. This does not mean that production variability is not significant. One must look at interaction effects before making any conclusions.
(2) Looking at the main effects alone may indicate that sample size is relatively much more significant than the other three factors. Despite the fact that several interaction effects are also significant, high contribution to the total variance definitely indicate greater effect of that variable, sample size in this case, on the outcome of the experiment.
(3) Interaction effect between production variability and measurement variability leads to confounding results. For example, the solid line in Figure 8.5 shows that higher measurement variability leads to higher CRI which means higher risk when production variability is equal to 0.45 in this case. However, the trend is opposite when production variability is equal to 2.3 (high) while all other variables remain constant between the two cases.

Table 8.3 Calculated CRI Values for the Scenarios Identified in the Run Matrix.

Sample Size	Bias	Prod Var = 0.45			Prod Var = 1.4			Prod Var = 2.3		
		Meas var=0.1	Meas var=0.5	Meas var=1.0	Meas var=0.1	Meas var=0.5	Meas var=1.0	Meas var=0.1	Meas var=0.5	Meas var=1.0
20	-1.2	2.53	3.41	5.27	6.86	6.75	6.97	6.17	6.74	6.32
	0	0.92	1.79	3.36	3.51	2.51	4.96	6.00	5.65	5.55
	1.2	3.84	3.05	4.74	4.61	4.16	3.73	6.74	6.14	5.60
150	-1.2	2.98	3.99	10.55	4.53	8.32	11.60	6.98	8.25	9.67
	0	0.22	0.95	2.52	1.06	1.29	2.50	1.97	1.97	2.42
	1.2	3.75	3.98	4.11	4.12	4.49	4.14	4.91	4.76	4.81
250	-1.2	3.03	4.05	10.59	7.61	4.77	12.54	8.39	9.04	10.49
	0	0.16	0.99	2.18	0.94	1.08	2.28	1.48	1.70	2.59
	1.2	3.30	4.83	4.38	4.36	4.28	4.86	4.77	4.83	4.62

* Measurement Variability, in %

** Production Variability, in %

Table 8.4 ANOVA Table for CRI of all 81 Cases in the Run Matrix

Source	Sum Sq.*	d.f.	Mean Sq.	F	Prob>F
X1	8.6	2	4.3	1.6	0.22
X2	82.7	2	41.3	15.0	0.00
X3	56.4	2	28.2	10.2	0.00
X4	176.0	2	88.0	31.9	0.00
X1*X2	86.5	4	21.6	7.8	0.00
X1*X3	24.3	4	6.1	2.2	0.08
X1*X4	0.9	4	0.2	0.1	0.99
X2*X3	25.8	4	6.4	2.3	0.07
X2*X4	28.1	4	7.0	2.5	0.05
X3*X4	75.0	4	18.7	6.8	0.00
Error	132.4	48	2.8		
Total	696.6	80			

*Constrained (Type III) sums of squares.

X1: Production variability

X2: Measurement variability

X3: Sample size

X4: Bias

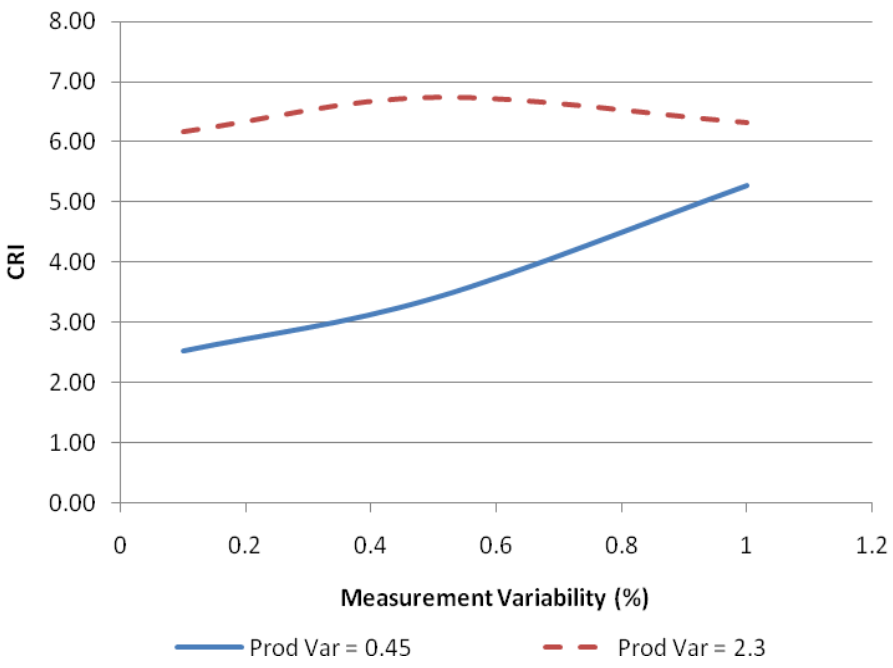


Figure 8.5 CRI as a function of production and measurement variability

(4) Measurement error can be controlled by the agency although it can most likely never be reduced to zero. If the measurement error is kept at minimum possible level, the risk in payment would go down while using the same QA program. Maintaining control over measurement variability is generally not too difficult. It would require that repeated measurements be taken in the beginning to assess the repeatability of the instrument/method and calibration be checked while doing the test section/initial lot in the beginning of the project.

(5) Bias seems to have the most drastic effect on payment risk because it can not only change the magnitude of risk but also alter the sign of risk. However, CRI does not catch this phenomenon because CRI treats the positive and negative risk as equally undesirable and does not discriminate between the two. This can be seen as a shortcoming of CRI. However, the authors have found through experience that it is very difficult, if not impossible, to design an index which is sensitive to the magnitude as well as sign of risk in the same plot. It means that most likely an accompanying risk index would have to be defined to cater to the needs of balancing risk between the agency and the contractor.

8.8 Conclusion

This chapter presents the details of the Monte-Carlo based simulation that was developed as part of this project to assess the current QA program of MDOT. The analysis conducted using the simulation showed that production variability, measurement variability, sample size and bias

have significant influence on the risk in payment to be made to the contractor. This knowledge leads to identification of ways to reduce payment risk. The simulation can be used to analyze all other variables of a QA program and thereby improve it to achieve a lower risk of overpayment or underpayment. The analysis also showed that if production variability is high despite very low measurement variability and mean production being in the middle of the specification window risk exists. Therefore, not only the contractor should produce right around the target he should be encouraged to maintain low variability in production quality. This is also significant from the point of view of pavement performance, as has been shown in chapter 7.

Generally the test methods and instruments are standardized and calibrated in the beginning of the construction project. For longer projects, the instruments may develop bias with continued use over several days. Bias has a very significant effect on payment risk. Such situations can lead to disputes and even law suits. Therefore, bias must be avoided through suitable inspection of the functioning of the test instruments.

Chapter 9: Feedback Process to Design for HMA Pavements

The aim of a quality assurance program for pavement construction is to assess the quality of the pavement constructed by the contractor and pay the contractor accordingly. It invariably involves testing for various quality characteristics. The data collected through this effort should therefore represent the quality of the end product in comparison with the quality targeted through the design process. Therefore, the QA program cannot only be used for determining the payment to be made to the contractor but also to provide feedback to the design process itself.

A feedback process is required primarily to check if pavement materials and layers are being produced according to the design plan and if the variability is within expected and acceptable limits. The as-constructed QA data can then be used to update the main statistics of input design variables (mean and standard deviation), which can be fed back into the design system to revise the expected performance. Figure 9.1 schematically shows the feedback process.

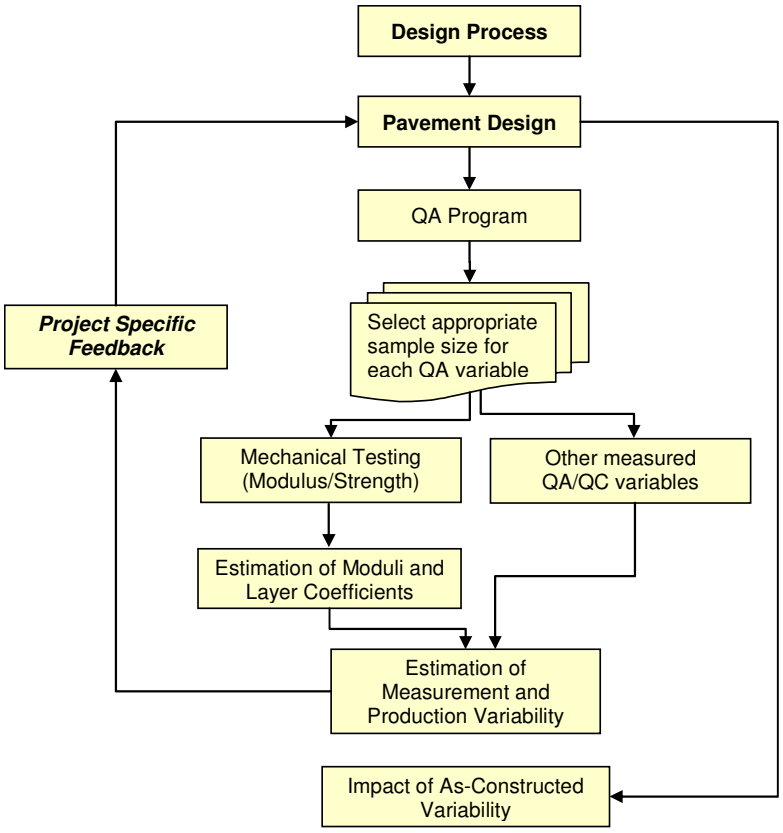


Figure 9.1 Flowchart showing feedback process for design

The following provides further description of the elements in this feedback process.

(1) *Selection of appropriate sample size:* One of the most important variables in the feedback process which needs to be optimized is the sample size. As the sample size becomes larger the confidence interval for a given design input (quality characteristic) tightens around the mean. The tighter the confidence interval the better the feedback process. This is discussed in detail in section 9.4.

(2) *Mechanical and/or material testing for modulus*: MDOT currently uses the AASHTO 1993 design guide for designing its pavements. Modulus values of the constructed pavement are required for the pavement structural design. The modulus values for the various layers can be indirectly measured through non-destructive testing in the field (e.g., FWD test) or directly measured through laboratory testing of cores obtained from the field. Alternatively, it can be estimated from material volumetrics and binder viscosity using established relationships (e.g., Witczak's E* equation). Note that the M-E PDG method also requires modulus testing in the form of dynamic modulus for the HMA layer and modulus of elasticity for the remaining layers.

(3) *Estimation of moduli and layer coefficients*: AASHTO 1993 uses the concept of layer coefficients (for asphalt pavements) to come up with the structural thickness. These layer coefficients need to be estimated from modulus values, which can be estimated from mechanical and/or material testing (see item (2) above). In the M-E PDG framework, the E* value is estimated through backcalculation using FWD test data (level 1) or from Witczak's equation (levels 2 and 3). Therefore, in the latter case, the feedback will consist of updating the input volumetrics and binder viscosity or dynamic shear modulus.

(4) *Use of other QA and QC data for design*: Quality characteristics data obtained through a QA program from pavement construction projects can be used as input for design either directly (in the MEPDG) or indirectly through correlations (in the current AASHTO 1993 design procedure). An effort can also be made to collect contractor's QC data as long as they are deemed comparable. This is described in sections 9.1 through 9.3.

(5) *Estimation of measurement and production variability*: The overall variability that construction data shows has two components, namely (a) measurement variability and (b) production variability. Production variability is the actual variability in the constructed pavement because of variability in material, construction practices and equipment, and climatic conditions. When various tests are used to determine the level of quality achieved, production variability gets masked with measurement variability because of error in the test equipment and/or process. However, only production variability affects pavement performance. Therefore, in the beginning of a construction project, measurement variability should be estimated for the various test methods to be used under the quality assurance program. This will help estimate actual production variability in the constructed pavement.

(6) *Impact of as-constructed variability*: Production variability will lead to variability in pavement performance. In the AASHTO 1993 design procedure, the loss/gain in design life (ΔPSI) can be directly back calculated using the design equation or by iteration using the Darwin design software. In the M-E PDG framework, the software can be used directly to predict the loss/gain in pavement life.

The following sections describe in detail the use of actual QC and QA data from MDOT projects to (1) validate the assumption of normality in QC/QA data, which is required for all the analyses used in this report, (2) demonstrate the applicability of using QC data in addition to QA data, and (3) investigate the effect of sample size on the feedback process to design.

9.1 MDOT Data Analysis

Michigan uses a standard format for collecting data for its QA program from pavement construction projects. All the data is stored in the form of Microsoft Excel Workbooks. An effort is also made to collect contractor's QC data as well in the same workbook. All the QA data from the projects contracted in 2008 were collected by the authors. There were 200 such files in total. Figure 9.1 shows a sample of the QA data from one of the project workbooks.

A typical pavement project is divided into lots which are further subdivided into sublots. Most of the sublots are composed of roughly equal amount of material used in the bound pavement layers. A specific number of samples is collected from each sublot for each quality characteristic. In the Excel worksheet shown in Figure 9.2, a set of rows are assigned for each sublot, and each row represents a sample within a sublot. The snap shot of the table represents lot number 1; each additional lot would be represented by a similar table. The worksheet has been designed in a way that one worksheet should be used for one mix type on any project. Most of the time, one mix type represents one HMA layer within the pavement. In those instances where the mix design formula is changed during the course of the project, separate worksheets may be used for different mix designs. However, it was found that the data entry in the actual worksheets was not as systematic as intended. For example, in many instances the same mix type for a project had data entered in more than one workbook with few lots worth of data in each of them. There is a possibility that not all of the data was documented in these sheets.

The first step in the analysis process was to consolidate and extract good data. Consolidation of data involved identifying all the data corresponding to one mix type of one project from different Excel files and putting them together in one place. Filtering involved eliminating duplicate data and in some cases those which seemed to be in error. This is a fairly arduous process since it needs to be done for each of the mix types and each of the quality characteristics separately, which means checking more than 1000 sets of data and moving them as required. At the end of the consolidation and filtering process a total of 127 mix worksheets from different projects were obtained. The next step was to analyze this set of data for various characteristics.

9.2 Assumption of Normality

This report documents the development of simulation methods to analyze MDOT QA specifications and provide for ways to improve it. An underlying assumption behind these simulations is that construction project data are normally distributed. This assumption is made based on observations made by other researchers in the past (Hall and Williams, 2002). However, this assumption is so crucial that it should be verified in every case as much as possible. Since the authors had access to MDOT QA data, this exercise was performed.

Figures 9.3 and 9.4 show plots from QA data corresponding to several mixtures from different projects to verify the assumption of normality. These plots have been presented for visual appreciation of the characteristics of QA data, which is very important from a simulation point of view. This will be followed by Chi-square goodness of fit test results for all the data being analyzed. In Figures 9.3 and 9.4 blue stars represent cumulative probability distribution functions for the actual QA data. The red dots represent the same for ideal normal distribution. If the two sets follow the same trend the data being tested can be assumed to be normally distributed.

MDOT HMA PWL Worksheet v1.10 - FY 2008 5/27/08
(effective for Contracts after Oct 07)

Control Section	13081	Binder Method	Back Calculated
Job Number	60527A	Plant Location	Kalamazoo
Testing Lab	MDOT/SW Reg	Plant ID	110-05
Lab Location	Kalamazoo	Layer	Top
Mix Type	3E50	Stress Level	Standard
Mix Design #	08MD160	Application	Mainline
Contractor	Aggregate Industries		

=Sublot RQL
=Lot RQL
=Density Outlier
=Action Limit
=Sus. Limit

	Binder	AV	VMA	Density
Low Spec	- 0.40 of JMF	3.00	13.00	92.00
High Spec	+ 0.40 of JMF	5.00	15.00	----------
Target	----------	4.00	----------	----------
AQL	90	90	90	90
RQL	50	50	50	50

							Target		MDOT QA					Contractor Random QC				MDOT QA Dispute Resolution ***Fill out only if there is a dispute***				
Sample Date	Mix Type	JMF	Production Type	Lot	Sublot / Core #	Sublot Tons	Binder Content	AV	QA Binder Content	QA AV	QA VMA	QA Density	QA Density Sublot Avg	QC Binder Content	QC AV	QC VMA	QC Density	Dispute Binder Content	Dispute AV	Dispute VMA	Dispute Density	Dispute Density Sublot Avg
5/8/2008	3E50	08MD160	UDP	1	1A	750.00			4.85	3.77	13.64	93.30		4.80	4.25	13.96	94.00					
5/8/2008	3E50	08MD160	UDP	1	1B							93.50					93.30					
5/8/2008	3E50	08MD160	UDP	1	1C							92.70										
5/8/2008	3E50	08MD160	UDP	1	1D							94.80	93.58									
5/8/2008	3E50	08MD160	UDP	1	2A	750.00			4.80	3.61	13.38	92.90		4.70	3.77	13.31	91.50					
5/8/2008	3E50	08MD160	UDP	1	2B							91.30					96.40					
5/8/2008	3E50	08MD160	UDP	1	2C							91.80										
5/8/2008	3E50	08MD160	UDP	1	2D							91.40	91.85									
5/12/2008	3E50	08MD160	UDP	1	3A	750.00			4.88	4.05	13.95	94.90		4.98	4.45	14.53	93.40					
5/12/2008	3E50	08MD160	UDP	1	3B							93.70					94.90					
5/12/2008	3E50	08MD160	UDP	1	3C							95.40										
5/12/2008	3E50	08MD160	UDP	1	3D							93.00	94.25									
5/14/2008	3E50	08MD160	UDP	1	4A	750.00			5.04	4.02	14.27	92.40		5.12	4.69	15.05	92.60					
5/14/2008	3E50	08MD160	UDP	1	4B							92.40					92.50					
5/14/2008	3E50	08MD160	UDP	1	4C		4.00	4.00				93.00										
5/14/2008	3E50	08MD160	UDP	1	4C							92.90	92.68									
5/14/2008	3E50	08MD160	UDP	1	5A	750.00			4.75	3.72	13.37	93.30		4.86	4.88	14.65	94.60					
5/14/2008	3E50	08MD160	UDP	1	5B							92.80					93.30					
5/14/2008	3E50	08MD160	UDP	1	5C							93.00										
5/14/2008	3E50	08MD160	UDP	1	5D							91.90	92.75									
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		
		08MD160	UDP	1																		

Figure 9.2 Sample MDOT QA data worksheet

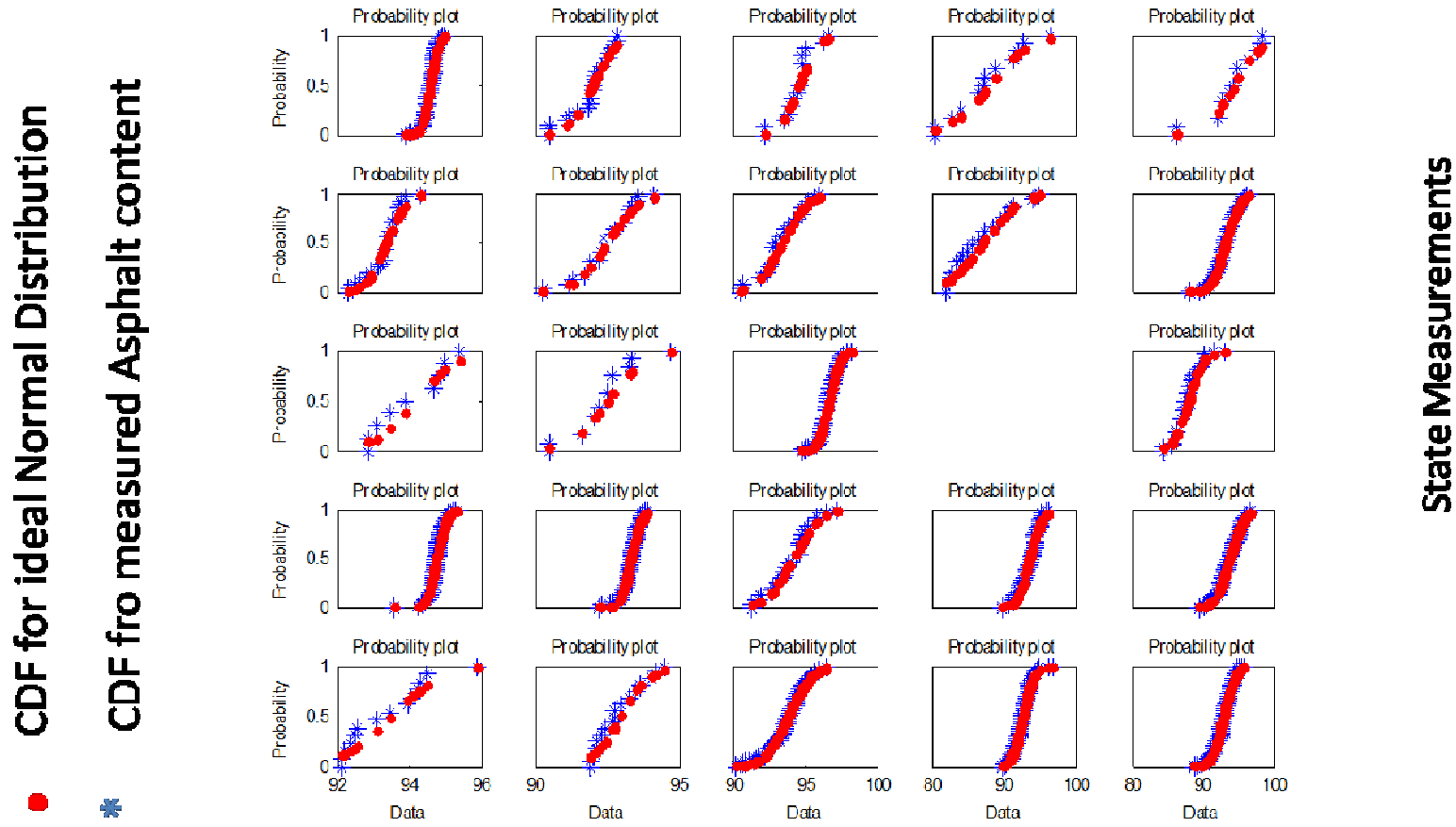


Figure 9.3 Test of assumption of normality – Asphalt content QA data

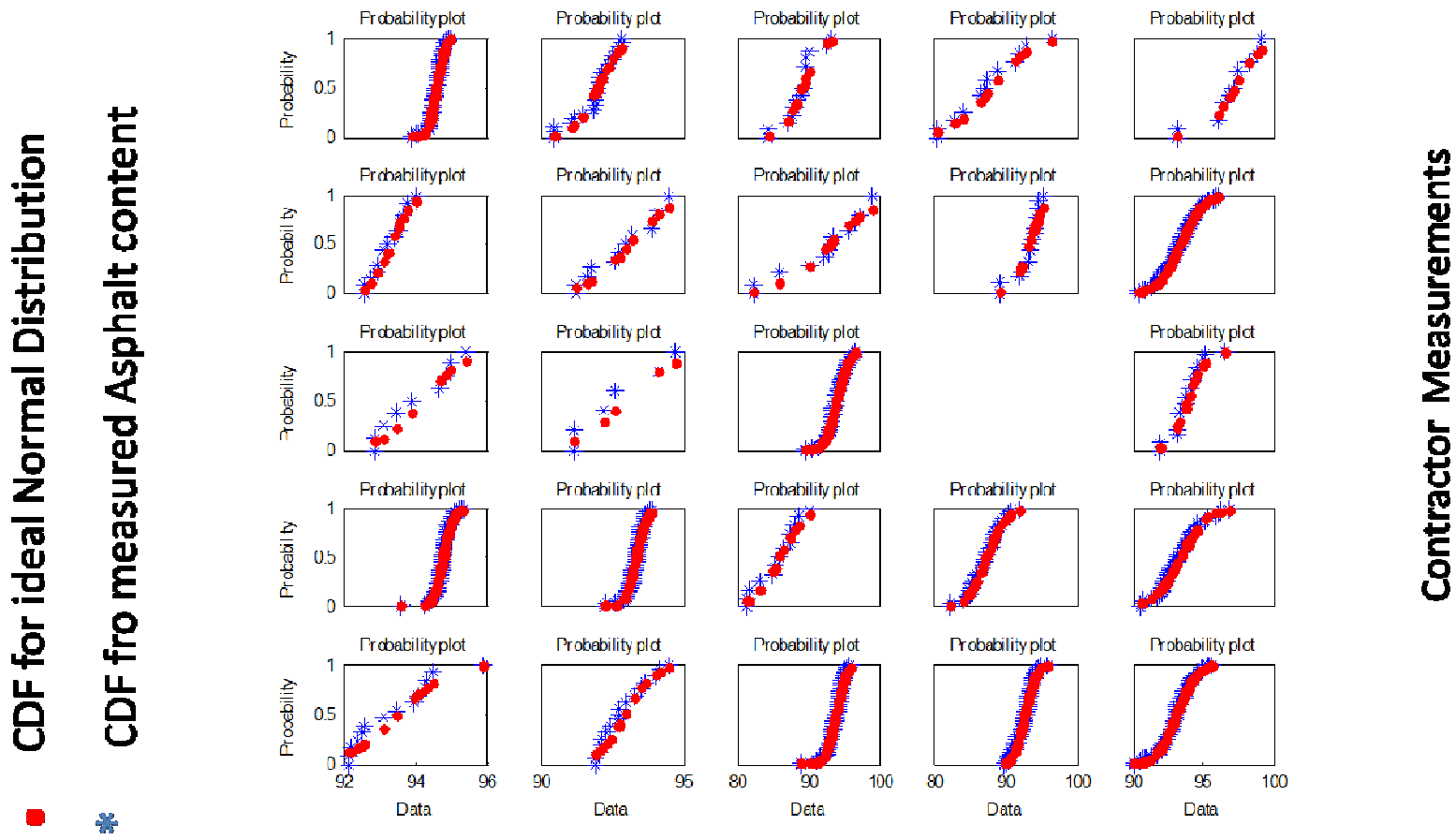


Figure 9.4 Test of assumption of normality – Asphalt content QA data (Continued)

An advantage of the visual representation is that one can see the parts of the data which deviate from normal behavior and account for that in the simulation so that the artificially generated data has the same characteristics as the real data. In the case of asphalt content, the QA data seems to be very close to ideal normal distribution in the majority of cases.

Figures 9.5 through 9.8 show results from Chi-square goodness of fit test for asphalt content, density, VMA and air voids QA data. The null hypothesis in this test is that the data are a random sample from a normal distribution with mean and variance estimated from the data, against the alternative that the data are not normally distributed with the estimated mean and variance. The result "h = 1" corresponds to p-value less than 5%, which means that the null hypothesis can be rejected at the 5% significance level. The result "h = 0" corresponds to p-value more than 5%, which means that the null hypothesis cannot be rejected at the 5% significance level. Therefore, an h-value of 0 means that the data is normally distributed at 5% significance level. It can be seen from the plots that the vast majority of the mix data are normally distributed in all the cases. For the sake of brevity cumulative density plots for other quality characteristics have not been included.

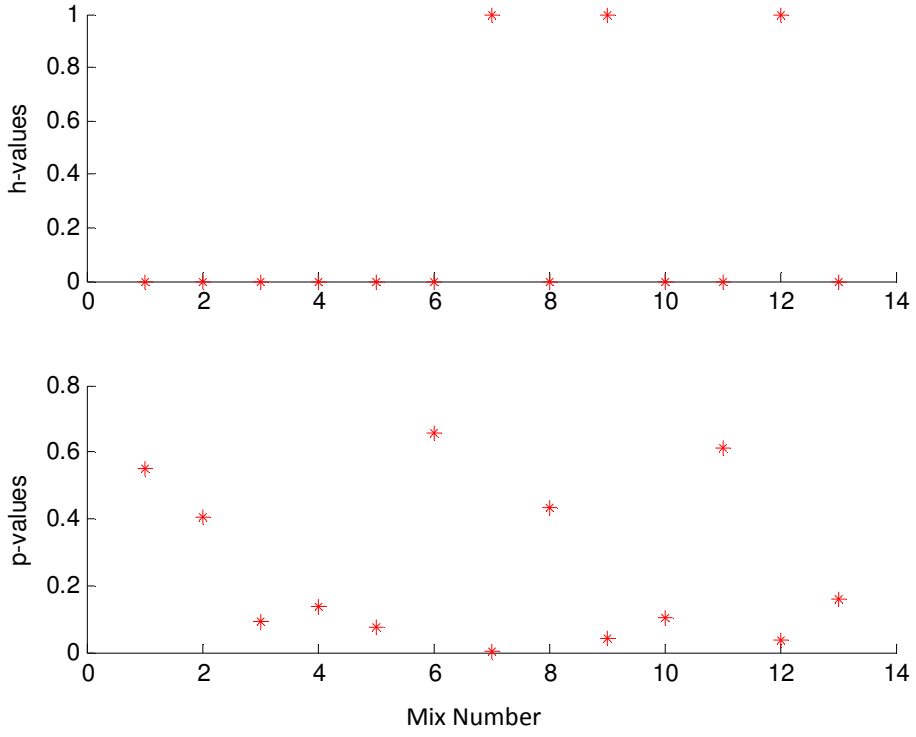


Figure 9.5 p and h – values from Chi-square goodness of fit for asphalt content QA data

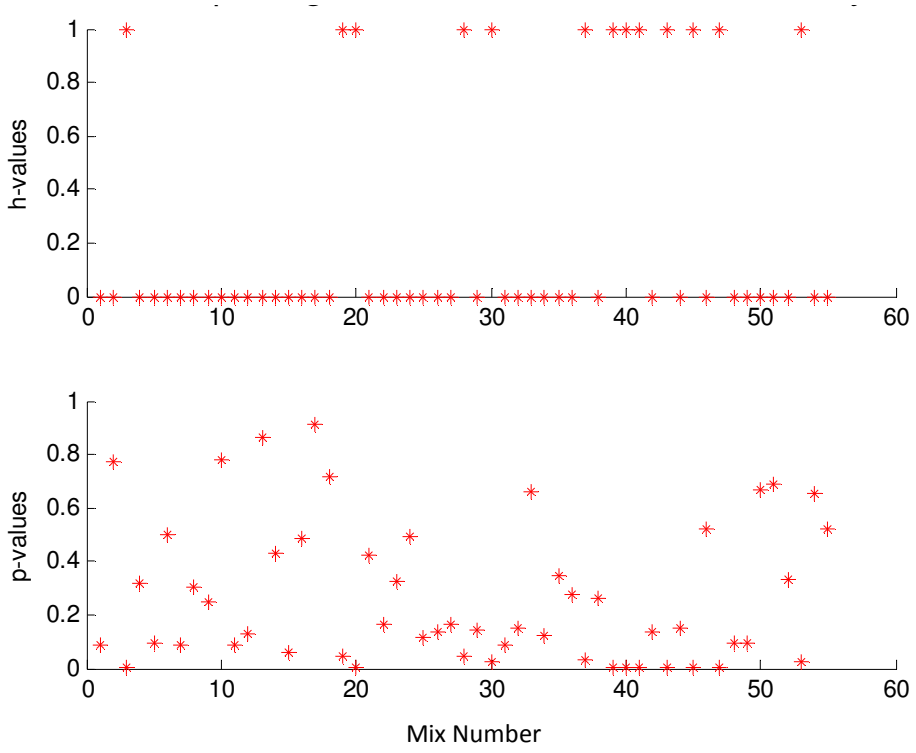


Figure 9.6 p and h – values from Chi-square goodness of fit for density QA data

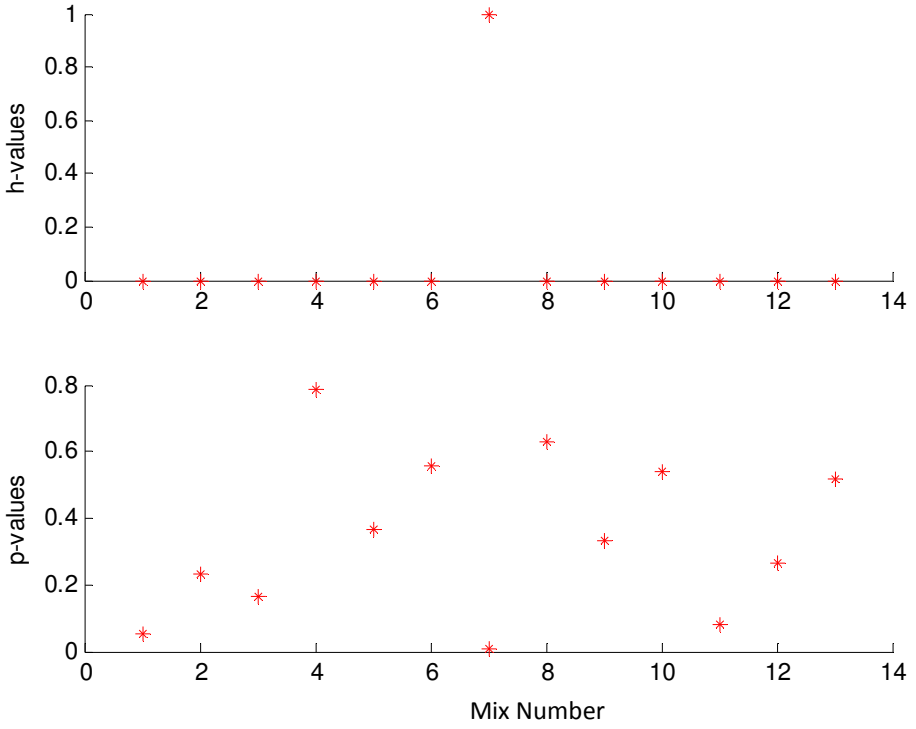


Figure 9.7 p and h – values from Chi-square goodness of fit for VMA QA data

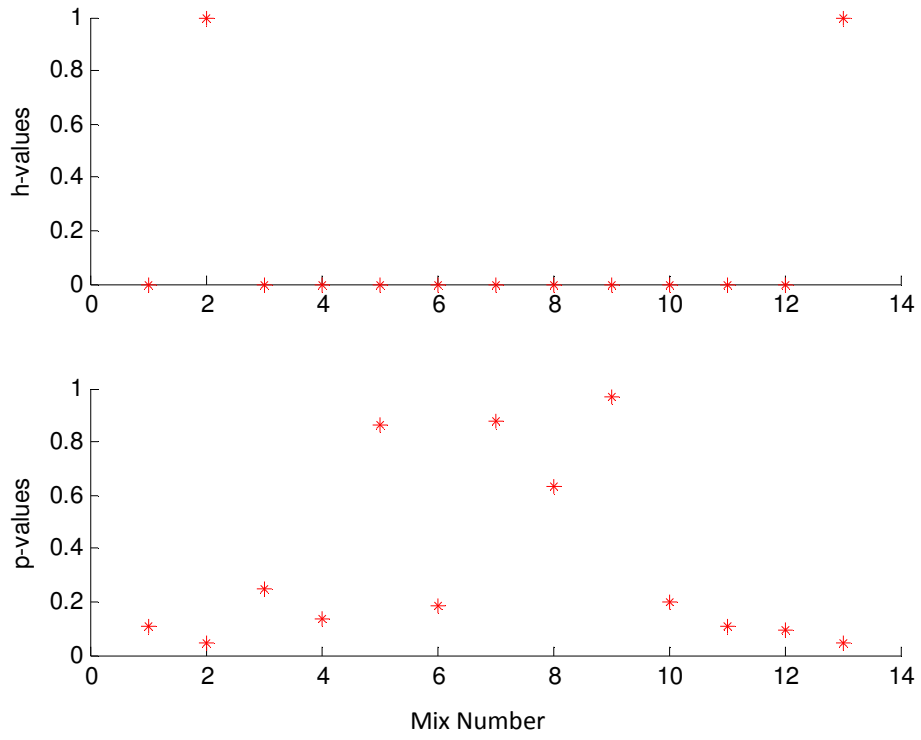


Figure 9.8 p and h – values from Chi-square goodness of fit for Air voids QA data

9.3 Comparison of QA and QC Data

The lot and sublot identification used for the QA testing remain the same for QC testing. Therefore, it is expected that the two sets of data should be similar, i.e. two samples should belong to the same population. Most of the times QC data is strictly used by the contractor only for quality control purposes and the state or the highway agency performs its own set of tests for QA. The state or highway agency determines payment for the contractor using QA data only. However, some states have procedures in place to check the correctness of QC data and use it to determine a pay factor. The main advantage of such a process is that the sample size used for pay factor calculation is larger and, therefore, can lead to lower risk in pay. MDOT does not use QC test data in pay factor calculation. An exercise was performed in this research to do a preliminary feasibility study to use QC test data (in addition to QA data) in pay factor calculation.

The first exercise in the feasibility study would be to test the assumption that QC and QA data are indeed two samples from the same population. If this assumption does not hold for any reason QC data cannot be used in pay factor calculation. Figures 9.9 through 9.12 present

comparisons between state and contractor data for the corresponding mixes to get a visual appreciation of the similarity or differences.

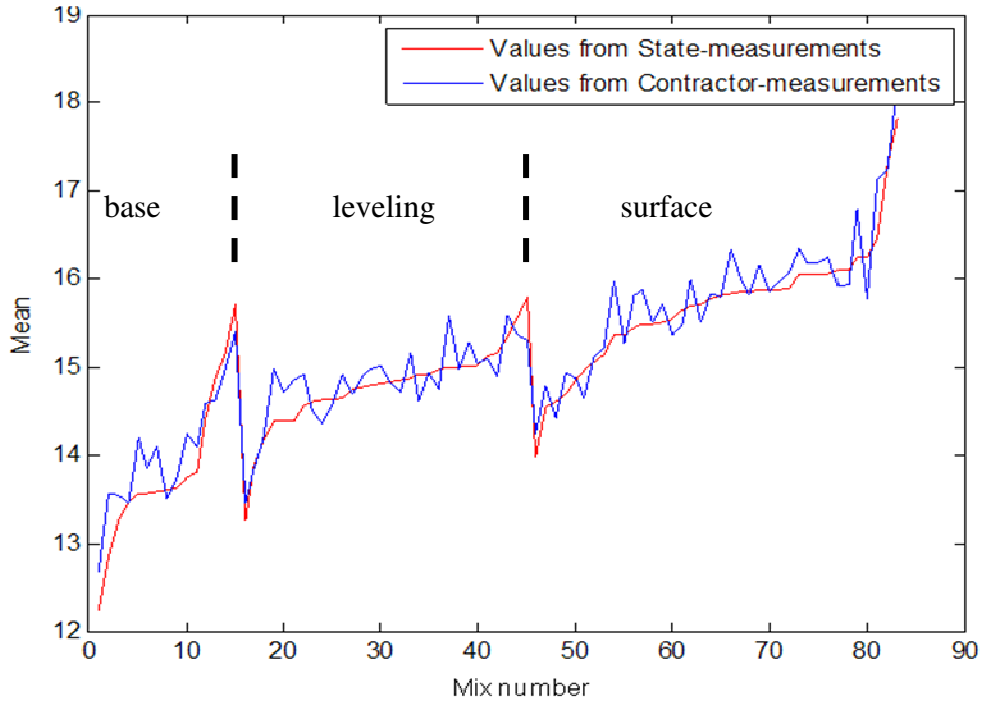


Figure 9.9 Comparison of state and contractor measured mean VMA for different mixes

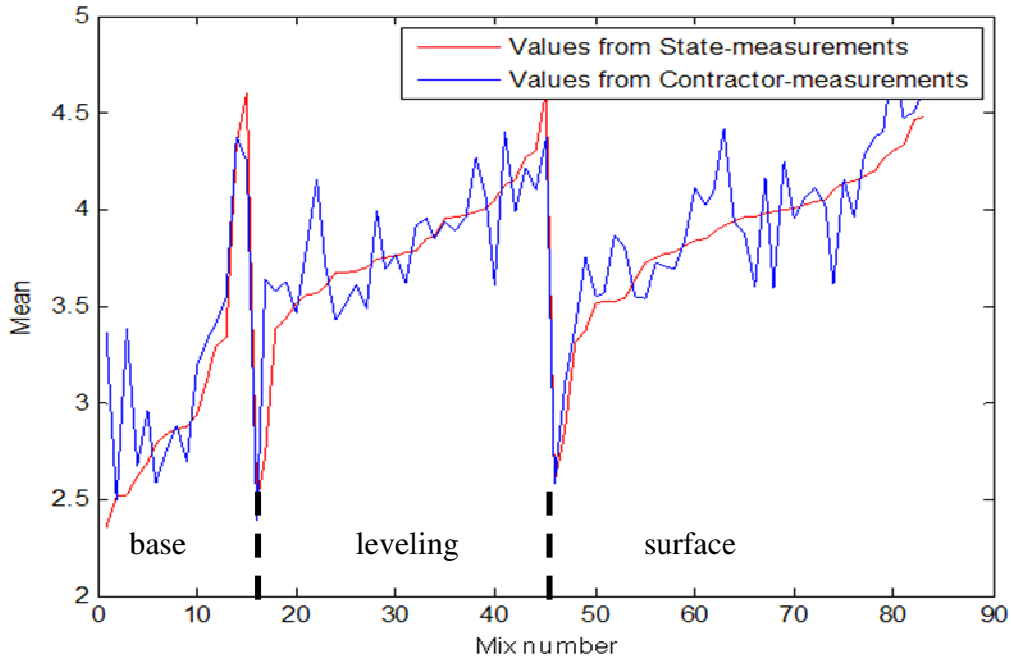


Figure 9.10 Comparison of state and contractor measured mean Air Voids for different mixes

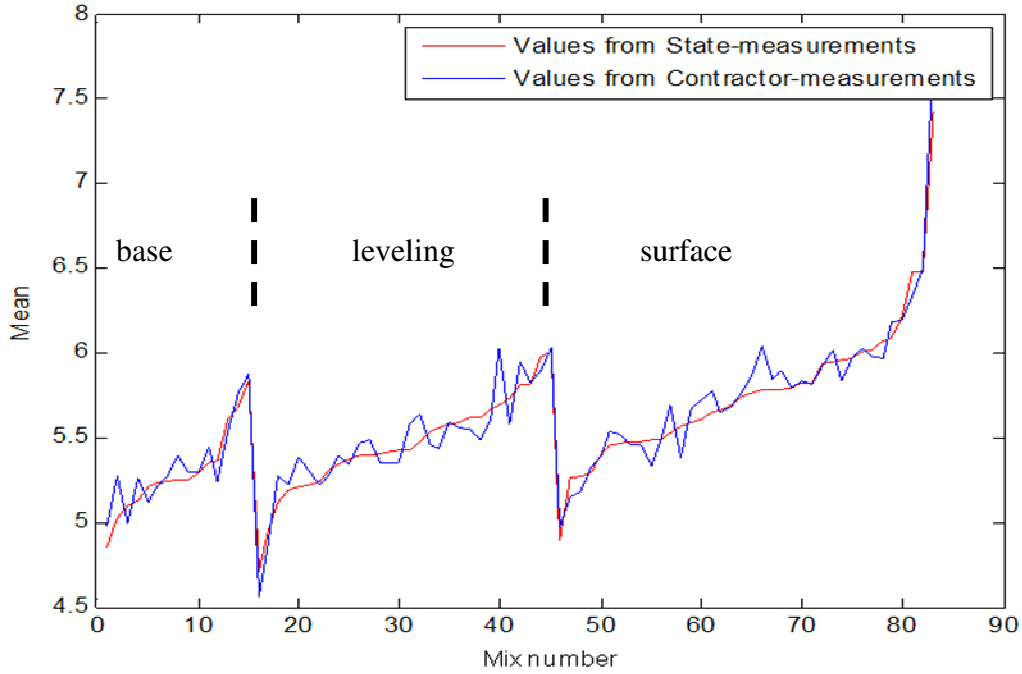


Figure 9.11 Comparison of state and contractor measured mean Asphalt Content for different mixes

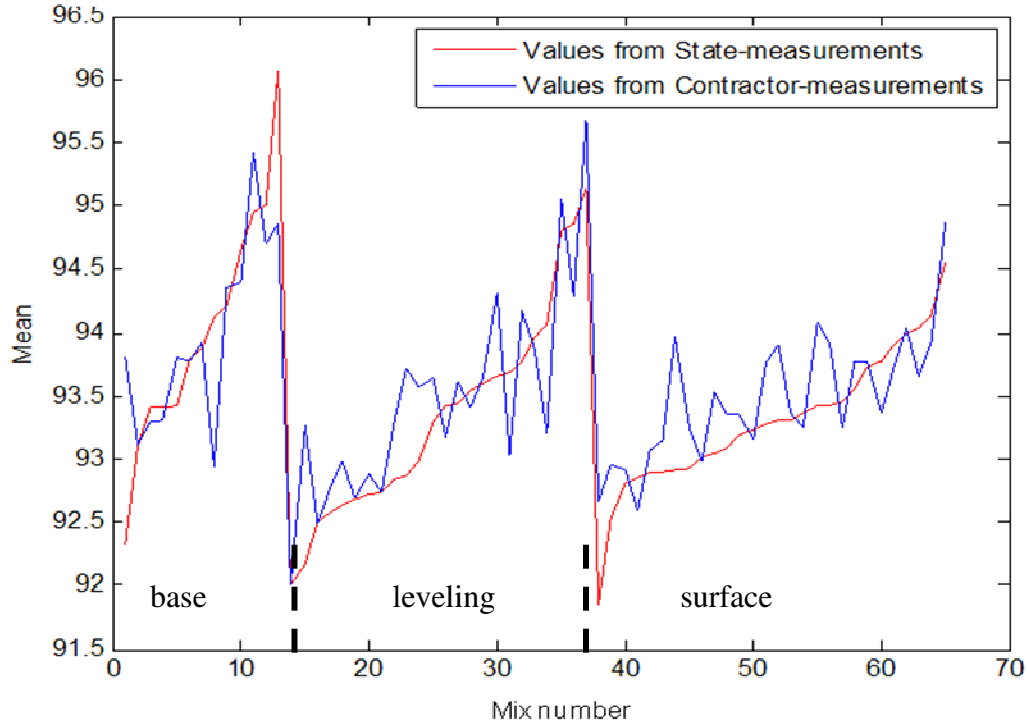


Figure 9.12 Comparison of state and contractor measured mean In-situ density for different mixes

The data used in this analysis were collected from different projects which were not related to each other. Therefore, there is no expected trend between any two projects or mixes. However, data for similar layers should be analyzed together. For example, one could put the base courses in one category, or lump the binder and surface courses in another category. If the state and contractor data corresponding to different mixes are plotted in any random sequence, both series will look more or less like the blue lines in the plots above. Because of the high fluctuation in values from mix to mix, it will be difficult to see any trend in comparing the two data sets. Therefore, the following steps were performed to obtain the above plots:

(1) Categorize data for the base course, leveling course and surface course (3 categories).

(2) Within each category sort the state measured data in ascending order and plot it along with the corresponding contractor data for that mix under the same mix number.

The plots for all four quality characteristics show the contractor data to be close to the corresponding state measured values. However, in some cases (e.g., relative in-situ density for the surface course, Figure 9.12), the contractor values seem to be higher. This is an example of bias.

Figures 9.13 and 9.14 show comparisons of standard deviation, 5th and 95th percentiles, and sample size between the state and contractor data. The purpose of presenting these plots is not only to compare the data but also to demonstrate the range of these statistics observed in the field. For example, in the case of air voids, more than 95% of the projects have a standard deviation less than 1%. Assuming an average standard deviation of 0.6%, if the QA specifications have an allowable window of ±1% (i.e., 1.67 times the average standard deviation) then percent-within-limits for an average project would be 90%, which translates into 100% pay. However, half of the projects which have higher standard deviation would attract penalty.

Sample size for the vast majority of mixes is less than 30. Section 9.4 in this chapter will show the impact of sample size on the error when estimating the variability and the resulting difference in pavement performance and error in pay factor calculation.

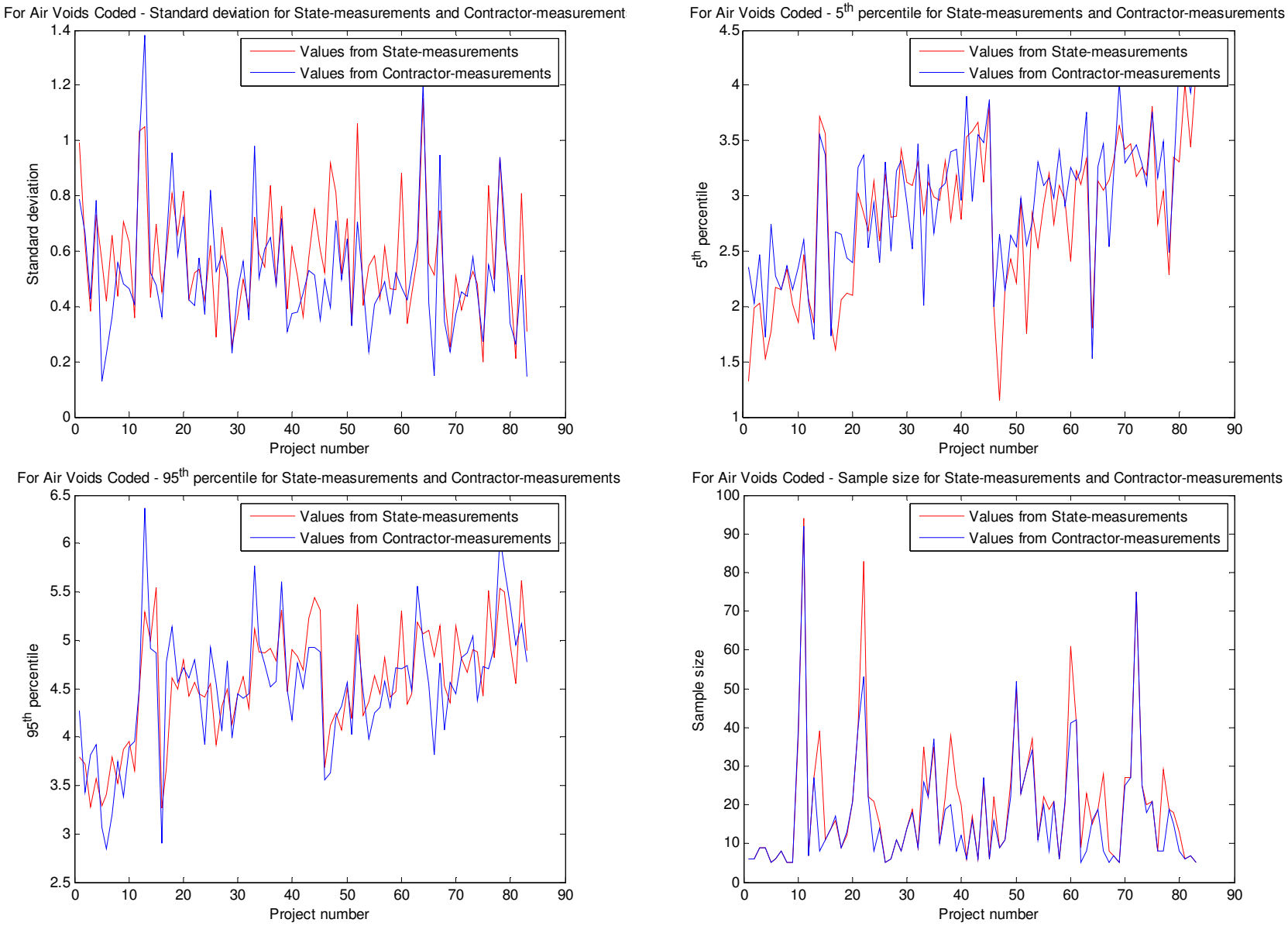


Figure 9.13 Comparison of state and contractor measured Air Voids for different mixes – standard deviation, 5th & 95th percentile and sample size

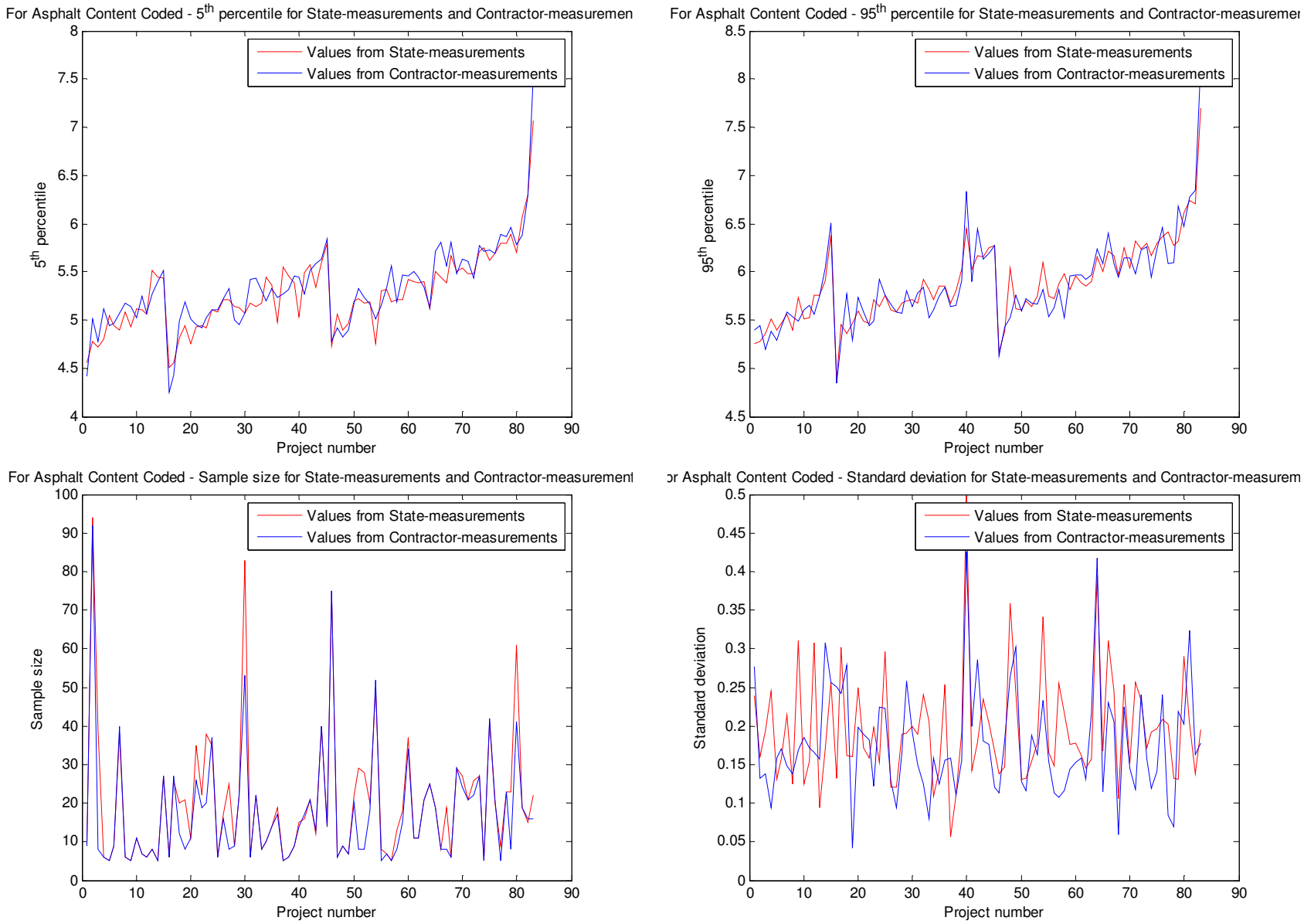


Figure 9.14 Comparison of state and contractor measured Asphalt Content for different mixes – standard deviation, 5th & 95th percentile and sample size

9.4 Effect of Sample Size on Feedback Process using Simulation

Chapter 8 of this report describes the development of Monte-Carlo based simulation to assess risk in payment to be made to the contractor in MDOT QA program. The same fundamental concept of simulation can also be employed to develop an optimal feedback process to design. Section 9.2 establishes the validity of synthetically generated data being similar to the actual field data collected from MDOT construction projects, both being normally distributed. The advantage of the synthetically generated data is that the error in the data is known a priori. Therefore, simulation using such synthetically generated data can be used to assess the extent to which data collected in the field represents true pavement quality compared to the design target. This section presents the details of this exercise.

One of the most important variables in the feedback process which needs to be optimized is the sample size. The feedback simulation developed in this project for flexible pavement construction was used to estimate the statistics enumerated below as a simultaneous function of sample size and mean of quality characteristic. Each scenario was simulated 10,000 times to identify the distribution of these statistics, allowing for a probabilistic study.

(1) Error in estimating the mean of a quality characteristic (density, air voids, asphalt content etc.) in a lot.

(2) Error in estimating the variability (standard deviation) in quality characteristics in a lot, and

(3) Risk in pay factor calculation for a lot.

All the above assessments were performed for a lot because MDOT QA program determines pay factor on a lot basis. Figures 9.15 through 9.17 show the above mentioned statistics as a function of sample size and mean quality characteristic (Q/C). The middle surface in Figure 9.15 represents the mean error in estimate of the mean quality characteristic. The surfaces that are above and below represent the 90% confidence interval for the error. The following observations can be made from this plot.

(1) The mean of the error is essentially equal to zero for all sample sizes and all values of mean Q/C

(2) It can be clearly seen that as the sample size becomes larger the confidence intervals tighten around the mean. The tighter the confidence interval the better the feedback

process would be. A tight confidence interval means that the estimate of the error lies within a small window, or in other words there is high probability that the error would be close to zero since the surface representing the mean of the error is essentially flat at zero level.

(3) The 90% confidence interval of error is tighter for higher mean Q/C, especially on the positive side of the mean. This happens because it would take much more compactive effort to reach 97% or higher density (density is being analyzed in this case).

The decision regarding optimal sample size for feedback will have to be made by MDOT. This is because defining the level of risk that MDOT is willing to take to save testing time (by not having a very large sample size) is a function of many considerations that only MDOT can weigh.

The plots presented in Figures 9.15 through 9.17 are helpful in understanding the trend in error, and therefore, how the optimal size should be selected. However, to be able to make this decision, MDOT would need a table or a plot showing a relationship between sample size and a metric tangible enough to make decisions (e.g., the width of the confidence interval).

Figure 9.16 shows the error in the estimate of variability (standard deviation) for different sample sizes and varying mean values of the quality characteristics. The overall behavior is similar to that observed in the case of the error in estimate of mean Q/C. The difference is quite noticeable when the sample size is small. For small sample size, the error is negative for all values of mean Q/C. In other words, a small sample size would lead to an underestimation of variability.

Figure 9.17 shows risk as a function of sample size and mean quality characteristics. The perspective view of this plot was chosen to be different from the preceding two plots because of the complex geometry of the surface. Therefore, the left horizontal axis which represented sample size in other plots has Mean Q/C and the right horizontal axis has Sample size instead of Mean Q/C. The effect of sample size on risk is similar to that for the estimate of mean and standard deviation.

Figures 9.18 through 9.20 were generated to get a better understanding of the magnitude of the effect of sample size on the three statistics being considered in this analysis.

Figures 9.18 and 9.19 have very similar trends and show that with increasing sample size the error in the estimate of mean and variability falls sharply in the beginning and then the reduction in error slows down. Therefore, MDOT will have to decide on the sample size beyond

which the reduction in error is not worth the extra effort of having a larger sample size to increase gain.

Figure 9.20 shows the reduction in risk across the entire range of the quality characteristic with increasing sample size. It is noticeable that for certain mean values of the quality characteristic (around the lower specification limit) an increase in sample size leads to a small decrease in risk whereas the reduction in risk is appreciably higher for mean quality characteristics away from the lower specification limit. This is because of the sharp increase in the magnitude of risk around the specification limits that was observed in risk analysis presented earlier.

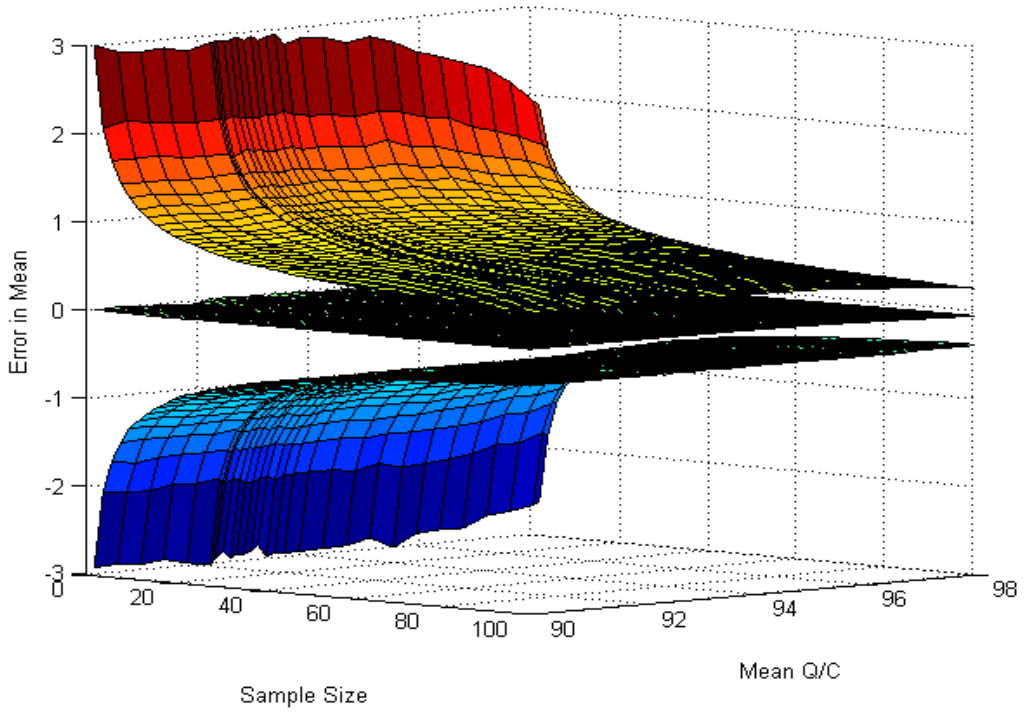


Figure 9.15 Error in estimate of mean quality characteristic with 90% confidence interval from feedback process

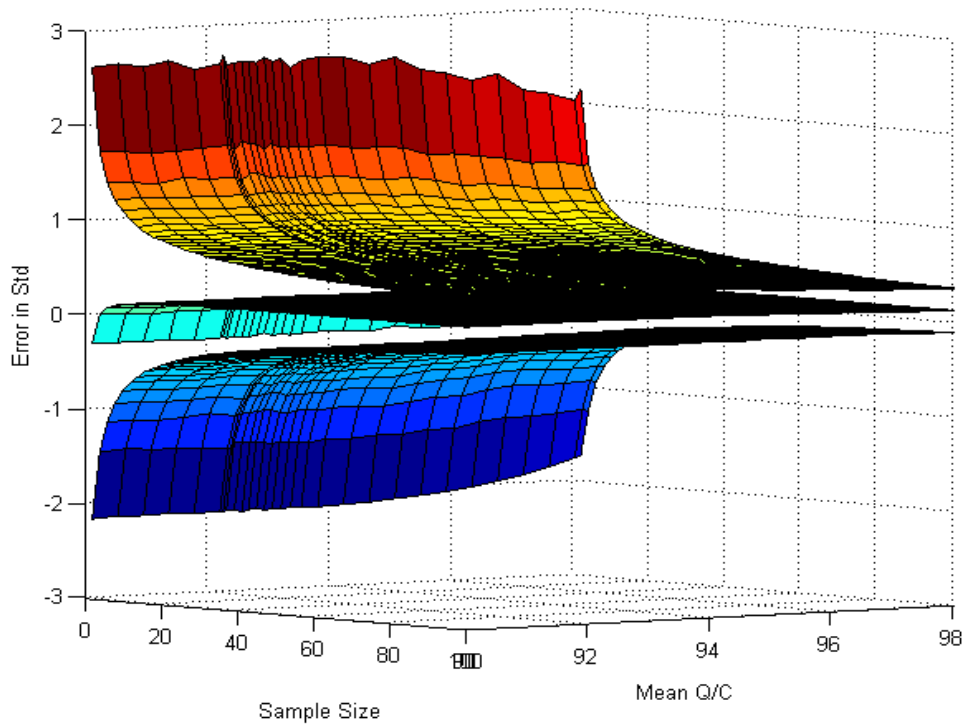


Figure 9.16 Error in estimate of variability in quality characteristic with 90% confidence interval from feedback process

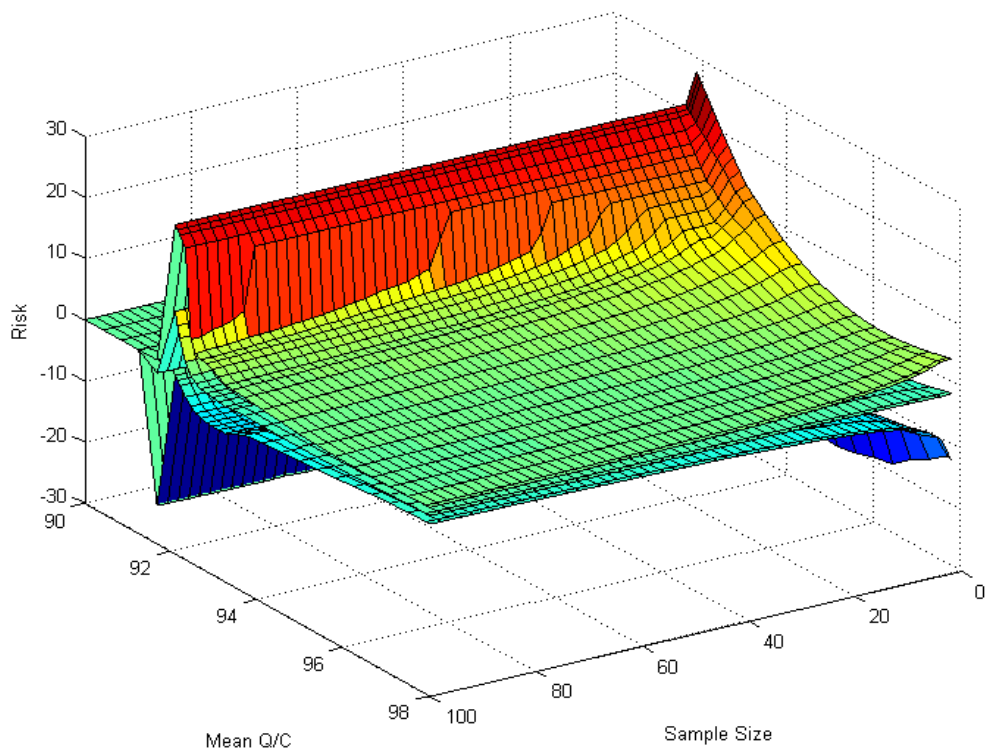


Figure 9.17 Error in estimate of payment risk with 90% confidence interval from feedback process

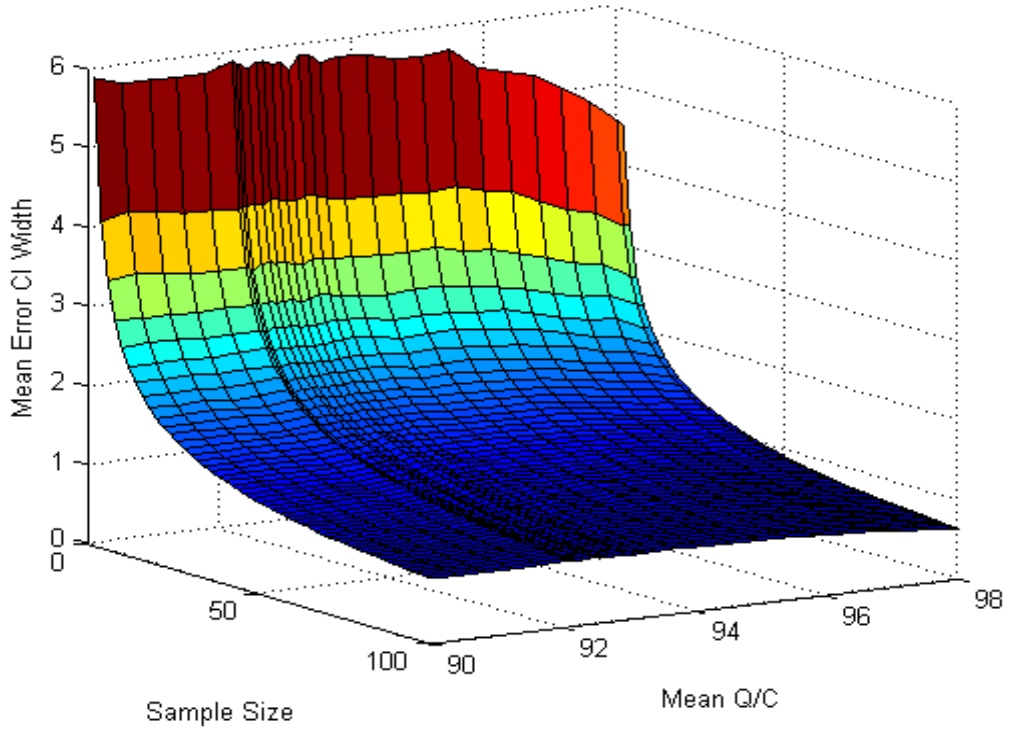


Figure 9.18 Width of 90% confidence interval in estimate of Q/C mean from feedback process

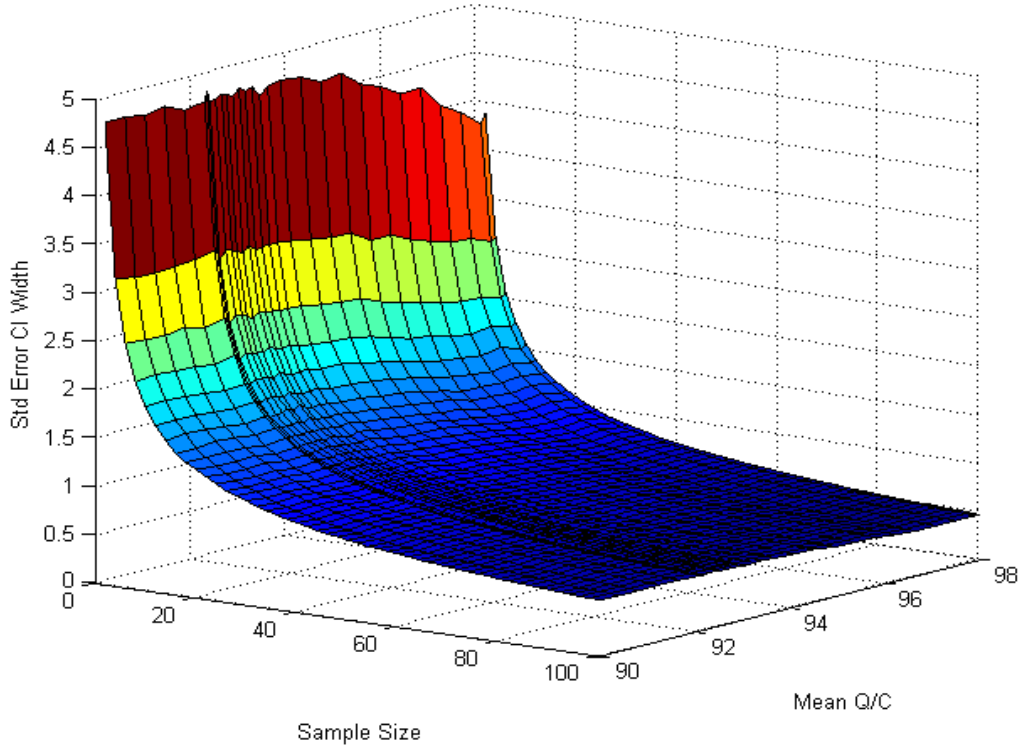


Figure 9.19 Width of 90% confidence interval in estimate of Q/C variability from feedback process

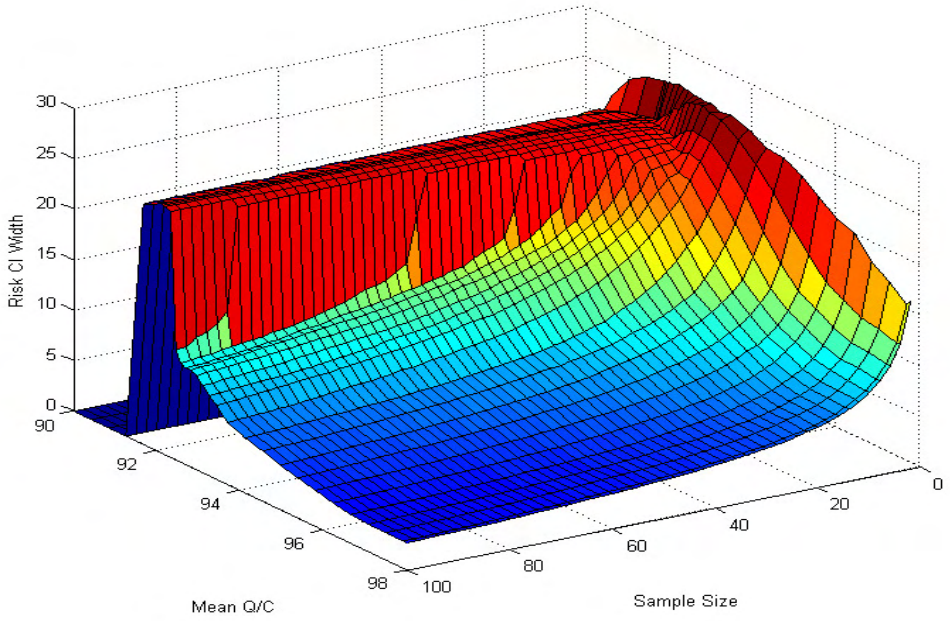


Figure 9.20 Width of 90% confidence interval in estimate of payment risk from feedback process

Tables 9.1 through 9.3 present the same information as the preceding three figures (Figures 9.18 through 9.20) but in tabular form to be able to see the magnitudes of confidence intervals, which would enable making decisions. One should consider the width of the confidence interval for the entire range of the quality characteristic for deciding on the sample size since different projects will have different values for the mean Q/C, although the sample size for the feedback process will probably have to be the same. It can be simplified one step further if we study the average width of confidence interval for the entire range of Q/C versus sample size. The last row in all these tables shows the average values. The sample sizes used in the analysis were varied from 2 to 100 with a step of 2. For the sake of brevity these tables present only a few selected values.

Figures 9.21 through 9.23 present the summarized form of the results obtained from this analysis. Figure 9.21 shows the maximum error in estimate of mean in 90% of the cases for different sample sizes. In other words, for example, if the sample size is 60 the maximum error in 90% of the cases will be lower than 1% (in-situ density in this case). However, if the sample size is only 8 the error in mean could be as high as 2.8%. If only two samples were collected the error in mean could be as high as 5.6% in ninety percent of the cases. This not only shows the benefit of having a larger sample size, but also quantifies the benefits in terms of reduction in error.

Table 9.1 Width of 90% confidence interval of error in estimate of mean quality characteristic for different sample sizes

Mean Q/C	Sample Size											
	2	**8**	**14**	**20**	**26**	**32**	**38**	**44**	**50**	**60**	**80**	**100**
90.0	5.8	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.1	1.0	0.9	0.8
90.4	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
90.8	5.8	3.0	2.2	1.9	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
91.3	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
91.7	5.8	2.9	2.2	1.8	1.6	1.5	1.3	1.3	1.2	1.1	0.9	0.8
92.1	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
92.2	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
92.2	5.8	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.1	1.1	0.9	0.8
92.3	5.8	2.9	2.2	1.9	1.6	1.5	1.3	1.2	1.2	1.1	0.9	0.8
92.4	5.8	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.2	1.0	0.9	0.8
92.5	5.9	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.1	1.1	0.9	0.8
92.6	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
92.7	5.7	2.9	2.2	1.9	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
92.9	5.9	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.2	1.0	0.9	0.8
93.0	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.2	1.1	0.9	0.8
93.2	5.8	3.0	2.2	1.9	1.6	1.5	1.3	1.2	1.2	1.1	0.9	0.8
93.3	5.9	2.9	2.2	1.9	1.6	1.5	1.3	1.3	1.2	1.1	0.9	0.8
93.5	5.9	2.9	2.2	1.8	1.6	1.5	1.3	1.2	1.2	1.1	0.9	0.8
93.9	5.8	2.9	2.2	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
94.3	5.8	2.9	2.1	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
94.7	5.7	2.8	2.1	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
95.2	5.5	2.8	2.1	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
95.6	5.6	2.8	2.1	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8
96.0	5.6	2.7	2.1	1.7	1.5	1.4	1.3	1.2	1.1	1.0	0.9	0.8
96.4	5.3	2.7	2.0	1.7	1.5	1.3	1.2	1.1	1.1	1.0	0.8	0.7
96.9	5.2	2.5	1.9	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8	0.7
97.3	5.0	2.5	1.9	1.6	1.4	1.2	1.1	1.1	1.0	0.9	0.8	0.7
97.7	4.6	2.3	1.8	1.5	1.3	1.2	1.1	1.0	0.9	0.9	0.7	0.7
98.0	4.5	2.3	1.7	1.4	1.3	1.1	1.0	1.0	0.9	0.8	0.7	0.6
Avg.	5.6	2.8	2.1	1.8	1.6	1.4	1.3	1.2	1.1	1.0	0.9	0.8

Table 9.2 Width of 90% confidence interval of error in estimate of variability in quality characteristic for different sample sizes

Mean Q/C	Sample Size											
	2	**8**	**14**	**20**	**26**	**32**	**38**	**44**	**50**	**60**	**80**	**100**
90.0	4.8	2.1	1.6	1.3	1.2	1.0	0.9	0.9	0.8	0.7	0.6	0.6
90.4	4.6	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.8	0.6	0.6
90.8	4.8	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.8	0.6	0.6
91.3	4.7	2.1	1.6	1.3	1.2	1.0	0.9	0.9	0.8	0.7	0.6	0.6
91.7	4.8	2.2	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.1	4.8	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.2	4.7	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.2	4.9	2.2	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.3	4.7	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.4	4.8	2.2	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.5	4.8	2.2	1.5	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.6	4.9	2.2	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
92.7	4.7	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.8	0.7	0.6
92.9	4.7	2.2	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
93.0	4.7	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
93.2	4.7	2.1	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
93.3	4.7	2.1	1.5	1.3	1.1	1.0	0.9	0.8	0.8	0.7	0.6	0.6
93.5	4.7	2.1	1.5	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.6	0.6
93.9	4.8	2.0	1.5	1.2	1.1	1.0	0.9	0.8	0.8	0.7	0.6	0.5
94.3	4.6	2.0	1.5	1.2	1.1	1.0	0.9	0.8	0.8	0.7	0.6	0.5
94.7	4.6	2.0	1.5	1.2	1.1	0.9	0.9	0.8	0.7	0.7	0.6	0.5
95.2	4.5	1.9	1.4	1.2	1.0	0.9	0.8	0.8	0.7	0.7	0.6	0.5
95.6	4.4	1.9	1.4	1.2	1.0	0.9	0.8	0.8	0.7	0.7	0.6	0.5
96.0	4.4	1.9	1.4	1.1	1.0	0.9	0.8	0.8	0.7	0.6	0.5	0.5
96.4	4.3	1.8	1.3	1.1	1.0	0.9	0.8	0.7	0.7	0.6	0.5	0.5
96.9	4.1	1.8	1.3	1.1	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.5
97.3	4.0	1.8	1.3	1.0	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.5
97.7	3.9	1.7	1.3	1.1	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.5
98.0	3.8	1.7	1.3	1.1	0.9	0.8	0.8	0.7	0.7	0.6	0.5	0.5
Avg.	4.6	2.0	1.5	1.2	1.1	1.0	0.9	0.8	0.8	0.7	0.6	0.5

Table 9.3 Width of 90% confidence interval of error in estimate of risk in payment for the quality characteristic for different sample sizes

Mean Q/C	Sample Size											
	2	**8**	**14**	**20**	**26**	**32**	**38**	**44**	**50**	**60**	**80**	**100**
90.0	15.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
90.4	17.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
90.8	22.2	20.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
91.3	25.3	22.7	22.6	22.0	19.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
91.7	26.0	23.5	23.7	24.0	24.0	24.1	24.1	24.1	24.1	24.1	24.0	23.8
92.1	26.6	23.6	23.9	24.2	24.3	24.3	24.4	24.4	24.5	24.5	24.5	24.6
92.2	26.9	23.7	24.0	24.1	24.3	24.3	24.4	24.4	24.4	24.5	24.5	24.5
92.2	27.2	23.7	24.0	24.1	24.3	24.3	24.3	24.4	24.4	24.5	24.5	24.5
92.3	26.8	23.6	23.9	24.1	24.2	24.3	24.2	24.3	24.3	24.3	24.4	24.4
92.4	26.5	23.6	23.8	24.0	24.1	24.1	24.2	24.1	24.2	24.1	24.1	24.0
92.5	27.0	23.4	23.7	23.8	23.9	23.8	23.8	23.8	23.8	23.7	11.5	9.9
92.6	26.8	23.5	23.6	23.6	23.5	23.5	23.4	17.7	14.7	13.0	10.9	9.6
92.7	26.9	23.1	23.4	23.4	23.2	18.7	15.5	14.1	13.3	12.3	10.5	9.6
92.9	26.7	23.0	23.1	22.8	18.0	15.7	14.4	13.5	12.7	11.6	10.4	9.6
93.0	26.8	22.8	22.5	19.1	16.4	15.0	13.7	12.9	12.2	11.2	9.9	8.9
93.2	26.6	22.3	22.3	18.0	15.9	14.4	13.3	12.3	11.6	10.6	9.2	8.2
93.3	26.4	22.5	20.8	17.1	15.1	13.7	12.5	11.7	10.8	9.9	8.5	7.4
93.5	26.6	22.6	19.0	16.0	14.0	12.6	11.3	10.6	9.9	9.2	7.7	6.7
93.9	26.0	21.2	16.7	13.8	11.8	10.5	9.6	8.8	8.1	7.3	6.3	5.6
94.3	24.9	19.8	14.5	11.6	10.2	9.2	8.5	7.8	7.3	6.6	5.7	5.2
94.7	24.8	16.3	12.2	10.2	9.1	8.3	7.6	7.0	6.6	6.1	5.2	4.7
95.2	24.1	13.9	10.8	9.1	8.1	7.3	6.7	6.3	5.9	5.4	4.7	4.2
95.6	22.8	12.9	9.8	8.3	7.5	6.7	6.1	5.6	5.3	4.8	4.2	3.8
96.0	21.9	11.6	8.8	7.4	6.5	5.9	5.4	5.0	4.8	4.4	3.7	3.4
96.4	21.3	10.5	7.9	6.7	5.8	5.3	4.9	4.6	4.3	3.9	3.4	3.0
96.9	19.1	10.1	7.5	6.3	5.5	5.0	4.6	4.3	4.0	3.6	3.1	2.8
97.3	18.3	9.6	7.3	6.1	5.4	4.9	4.4	4.1	3.9	3.5	3.0	2.7
97.7	17.0	9.5	7.3	6.0	5.3	4.8	4.4	4.1	3.8	3.5	3.1	2.7
98.0	16.0	9.3	7.1	6.0	5.3	4.8	4.4	4.0	3.8	3.5	3.0	2.7
Avg.	23.8	17.7	15.7	14.5	13.6	12.2	11.7	11.2	10.8	10.3	9.3	8.8

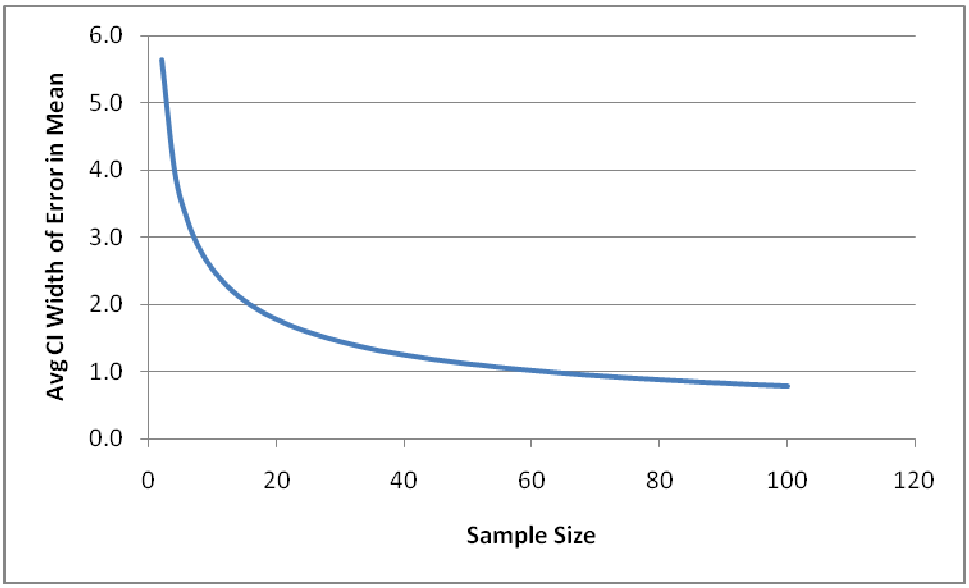


Figure 9.21 Average width of 90% confidence interval of error in estimate of mean

Figure 9.22 shows the maximum error in estimate of standard deviation in 90% of the cases for different sample sizes. Figure 9.23 is probably even more relevant because it shows how risk in payment to be made to the contractor goes down (in 90% of the cases) with increasing sample size for a lot. It is also more relevant because it includes the effects of errors in the estimation of the mean as well as variability. It should be noted that these errors are for each of the lots. A project will have several lots and the errors may cancel each other out, at least partially, when the payment is calculated for the entire project. However, a good quality assurance program should minimize risk in lot pay factors as well.

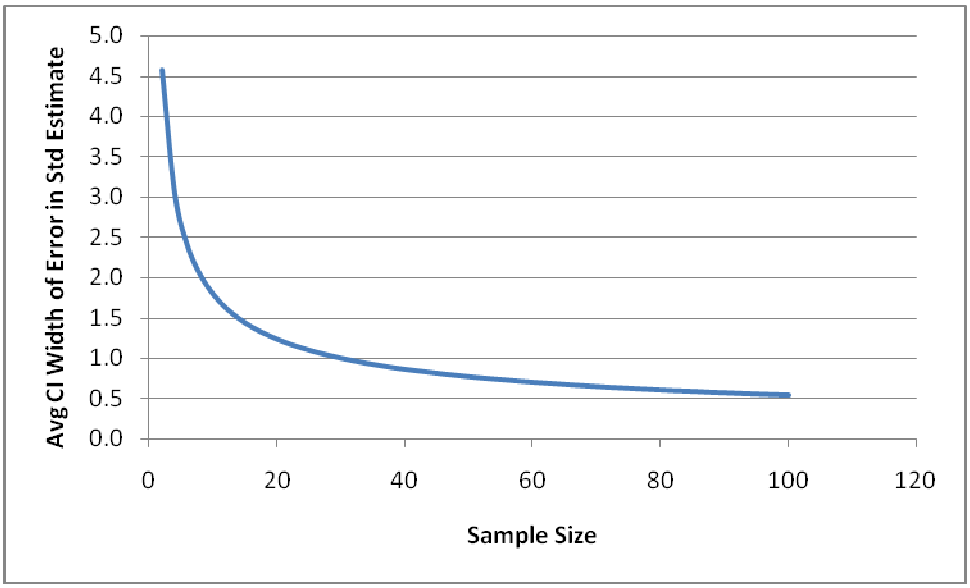


Figure 9.22 Average width of 90% confidence interval of error in estimate of variability

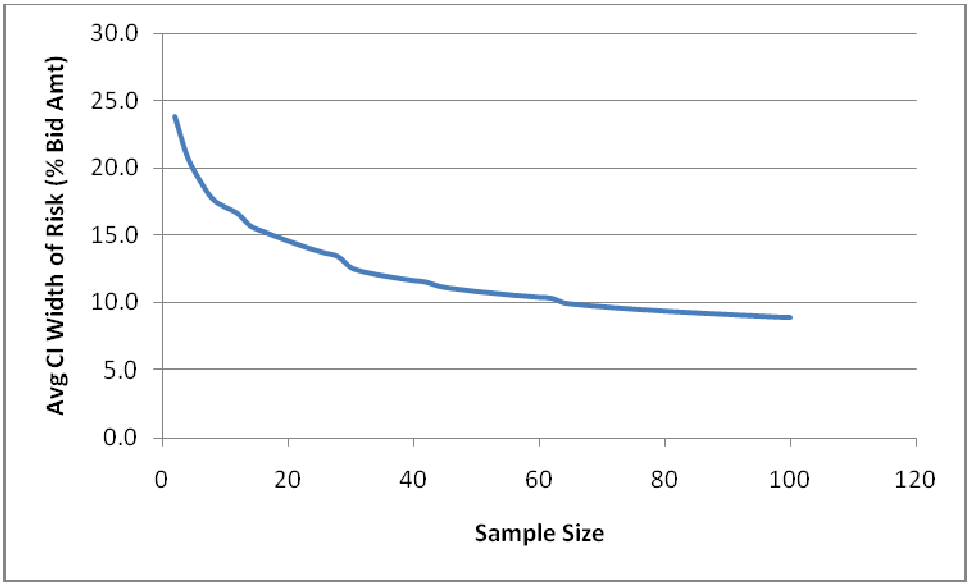


Figure 9.23 Average width of 90% confidence interval of risk

9.5 Conclusion

This chapter proposed a feedback process to design using QA and QC data. This chapter also presented the analysis that was performed using actual field data from projects constructed under MDOT QA program in the year 2008. This analysis helped in understanding the nature and magnitude of variability in MDOT projects and differences in these variabilities among projects depending on the strictness of quality control and other factors. This will help in making realistic assumptions in the design process. It was also established that most of the construction data follow a normal distribution to a large extent. This validated the assumption of normality that was made in developing the risk analysis simulation for MDOT QA program. Another simulation was developed to design optimal feedback process for design. The simulation helped in estimating the errors that can be expected depending on the sample size that is used in the feedback process and the associated probabilities. Finally, plots were developed to relate sample size to probabilistic error in estimation of the mean and standard deviation of the quality characteristic and payment risk. These plots can be used directly by MDOT to decide on the appropriate sample size (i.e., not too small to lead to higher errors and risk and not too large to be too costly or impossible to carry out).

Chapter 10: Conclusions and Recommendations for HMA QA Program

This chapter presents the overall strategy adopted in this research followed by the conclusions and recommendations that have been derived. A good quality assurance program should use the quality characteristics which can ensure pavement performance that meets or exceeds the design target. While there are several components to the QA program, identification of the suitable quality characteristics is the most important one.

10.1 Identification of Suitable Quality Characteristics for QA Program

Identification of suitable quality characteristics requires preparing an exhaustive list of potential characteristics which should be considered for inclusion in the QA program. The different sources for preparing this list are enumerated below.

(1) Quality characteristics being used in other states' QA programs: Different states use varying combinations of quality characteristics. Some of these quality characteristics are used in determining payment to be made to the contractor for any project, while others are used merely to provide feed back for proper construction.
(2) Quality characteristics being used in other states' QC programs: Any of the quality characteristics which are used in QC programs, i.e. is monitored by the contractor, but is not used in the QA program should also be considered.
(3) Quality characteristics used in the Mechanistic-Empirical Design Guide Software: The MEPDG software includes models to predict pavement performance from material, pavement structure, construction, traffic and environmental variables. The models used in the software are the result of studies carried out by many research teams after extensive testing and analysis to relate those variables to performance. Therefore, those variables or quality characteristics which are within the control of the contractor and which can be tested at the time of construction should also be included in the list.
(4) Quality characteristics studied in other research projects which have been shown to have impact on performance.

The second step is to shortlist those candidate QA variables which can be shown to affect pavement performance. These relationships can be established through one or more of the following options.

(1) Empirical data from Michigan: If empirical data can be obtained which establish that changes in levels of one quality characteristic leads to change in pavement performance, either individually or in conjunction with other quality characteristics, then it would be the most preferred way.
(2) Empirical data from other states: If Michigan data is not available or is not good enough to establish relationships mentioned in option 1 then empirical data from other states can be used. This is a slightly more indirect way of establishing whether a certain quality characteristic should be used in the Michigan QA program. This is because any other state may have climate, typical construction materials, construction practices and traffic different from those in Michigan. However, if some of these factors are matching with Michigan and/or if those parameters affect performance very significantly then they must be analyzed using such data.
(3) Analysis using MEPDG: MEPDG software puts together the best available response and performance models for flexible as well as rigid pavements. The models are generally mechanistic-empirical in nature. They have been developed and calibrated using empirical data. Therefore, analysis performed using these models simulates using empirical data but with more flexibility, although it also comes with prediction errors. Nonetheless, it should be recognized that simulation results are indicative rather than predictive. This means that these results should be used to guide decisions in general; but they cannot be used to provide direct predictions for a particular project. Therefore, relative differences in performance because of these quality characteristics are more relevant. MEPDG also allows for studying the effect of variability in these quality characteristics on pavement performance.
(4) Other research studies: Other research studies firmly establishing relationship of the candidate QA variables to performance can also be used to verify the findings from the above three options. In the case of variables for which none of the above options can be feasibly used for analysis this may be the only option.

The third and last step is to identify those variables which should be incorporated into the Michigan DOT QA program. The criteria for including those variables are that:

(1) They affect pavement performance either directly or in conjunction with other variables.
(2) They need to be tested individually and cannot be estimated or calculated from other significant QA variables already being used in the QA program.
(3) It is feasible to test for them within a reasonable amount of time during the construction.
(4) The testing for these candidate QA variables does not require very specialized or costly equipment.

It is possible for the contractor to control these variables through sound construction practices and tight quality control.

10.1.1 Comparison of MDOT QA programs with others in the US

The MDOT HMA QA program is similar to those being used by the states who have adopted end result specifications for their QA program. This similarity exists in (1) the tests that are used for verifying the quality of the constructed pavement, (2) the specifications limits that are used and (3) the quality characteristics, the statistical method and the pay formula used for calculating the payment to be made to the contractor. In other words there is nothing alarmingly different in the QA program being used by MDOT as compared to other "ERS" states.

10.1.2 QA Parameters Indentified by Other Studies

Based on literature review of studies that looked at the effect of pavement design and construction variables on performance, the following variables were identified as key QA parameters:

1. Asphalt Content
2. Air Void Content
3. Mixture Density
4. Aggregate Size and Gradation
5. Pavement Thickness
6. Binder Performance Grade

All the above variables can be input explicitly in the MEPDG software to study theoretical effects on performance.

10.1.3 Summary of Results from Empirical Analysis

The analysis using empirical data from the LTPP database showed that such data generally cannot be used to develop a performance model that explicitly relates specific QA variables to performance. Such analysis can at best show some trends as to how a particular quality characteristic can affect certain aspects of pavement performance.

The set of data extracted from LTPP database in this analysis showed that pavements with low (< 4%) or high (>8%) in-situ air voids have higher probability of rutting than those in which air voids fall in the range of 4 to 8%. This indicates that in-situ air voids immediately after construction is an important test for a quality assurance program. Contrary to intuition the

asphalt content data does not show any clear trend in its effect on rutting. This is most likely because of not having data points corresponding to very low or very high asphalt content compared to the corresponding design values. The trends observed in the case of fatigue cracking and longitudinal cracking were not very clear. This may most likely be because of the fact that several other factors are also affecting the fatigue performance and these cannot be controlled in the analysis with such data.

An argument was presented in chapter 6 as to how percent-within-limits for a particular quality characteristic should be related to pavement performance to rationally justify its use in the pay factor equation. The analysis did not show a clear relationship between PWL for in-situ density and fatigue cracking, longitudinal cracking and transverse cracking. Based on the plots showing PWL (plant air voids) versus cracking performance it can be said that while lower PWL does not necessarily mean poor performance, the probability of better performance certainly goes up with higher PWL. Further plots show that as PWL (plant air voids) become higher, i.e. air voids are within the specification limits to a greater extent, there is better longitudinal and transverse cracking performance. No observable relationship between PWL (plant air voids) and rutting was found in this analysis.

The conclusions derived from the empirical data analysis can at best be considered as mixed and not clear. This happens because of several factors discussed in chapter 6. These conclusions, therefore, indicate that a thorough mechanistic-empirical analysis needs to be performed to derive firm conclusions required for assessing quality assurance programs like the one being used by the state of Michigan.

10.1.4 Summary of Results from Mechanistic-Empirical Analysis

Based on the limited number of verifications, MEPDG seems to give reasonable predictions for different pavement distresses. Therefore, MEPDG can be used to explore all the aspects of mix and construction process which influence pavement performance. Pending verification through other methods, the other aspects of mix production or lay down that are not being accounted for in the current QA program can be identified though the strategy presented in this report. The proposed strategy can also be directly used to determine the targets and allowable windows for each of the quality characteristics being used in the QA program.

MEPDG can be used to simulate actual construction and resulting pavement performance. A preliminary example of that has been presented in this report. When extended with possible life cycle cost analysis, such a simulation process can lead to a rational pay factor formula/procedure.

Effect of varying air void distribution on average project performance

The results of MEPDG analyses corresponding to cases of varying in-situ air voids to achieve various PWL ranges, while fixing the asphalt content (at -0.4% of optimum), showed that on average, the maximum difference in fatigue was only about 2% between the worst PWL and the acceptable PWL (greater than 90%); a negligible difference. It was also noticeable that for very similar PWL values, different projects showed different amounts of fatigue cracking. This is because fatigue cracking in each project is estimated by sampling. As long as the range, mean and standard deviation of two samples are the same they would have the same PWL. But the average fatigue cracking can vary depending where each of the samples falls within the specifications. For example, if a project has all densities equal to 93%, PWL would be 100%. If the density of all the samples were 98%, PWL would still be 100%. However, fatigue performance for the 98% density project would be much better than that for the project with mean density of 93%.

Rutting performance estimated through the simulation showed that rutting performance improves as PWL approaches 100%. However, this improvement is very small. In other words rutting performance is not so sensitive to PWL (in-situ air voids) according to MEPDG results. The analysis also showed that IRI does not seem to be much affected by PWL (in-situ air voids).

It must be mentioned that all these results represent average project performance as opposed to a more realistic sublot by sublot performance, and they were only for one type of mix. It is important to study different types of mixes to assess how PWL values affect the performance for each one of them. It is quite possible that this effect may be much more pronounced in other types of mixes.

Effect of varying air void distribution and asphalt content on project sublot performance

The observed effects of PWL on average performance for a given project were not as pronounced as one would expect intuitively. The main reason for this is the attenuation of

distress levels because of the averaging of performance across sublots. Therefore, it is more realistic to show results in terms of performance by individual sublots. Results in terms of fatigue performance for various scenarios (three air void and five asphalt content distributions) showed that:

1. Fatigue performance of different sublots within a project can vary significantly;
2. Within a given asphalt content range, there are more underperforming sublots when the air void distribution moves towards the upper specification limit ;
3. As the asphalt content range moves dry-of-optimum, more sublots are underperforming in fatigue, even at relatively high combined PWL.
4. As the asphalt content range moves wet-of-optimum, more sublots perform better in fatigue, even at relatively low combined PWL. This may however lead to more rutting within the HMA layer.

The analysis also showed that cases where the combined percent-within-limits (PWL) values were very similar (combined PWL is generally used directly to calculate payment for the contractor), the average fatigue cracking at 30 years varied from 0.7% to 6.8%. This much of variation is very significant because they would be categorized as being in very good and nearly poor condition, respectively. The 75th percentile value for the same cases varies from 0.8% to 8.7% and the 95th percentile varies from 1.1% to 9.9%. This clearly shows that the candidate QA variable causing this variation, asphalt content in this case, has significant influence on fatigue performance.

Variations in fatigue performance as a result of change in in-situ air voids in these cases are also appreciable, although they are not as large as those because of asphalt content. When the air voids was increased from 3.5% to 8%, the 75th percentile for fatigue cracking increased from 13.9% to 17.9%. An advantage of percentile distresses is that it can possibly identify those cases where a few sublots perform very poorly while the rest of the sublots may have acceptable performance. Such a scenario should not be acceptable because the distressed sublots may force early repair work for the entire project.

For rutting performance, the results showed that the range of rutting values corresponding to the average, 75th and 90th percentiles for all the cases was rather small. It should be noted that in these runs the asphalt content was allowed to vary only up to 0.5% more than the optimal. If higher asphalt contents were allowed the increase in rutting would have been probably more

significant. However, the selected criteria, namely the average rutting at 30 years, the 75th and 90th percentile rut values are still good indicators of the effect of QA variables on performance.

The results also showed that even though different projects may have very similar combined PWL values, they may perform differently in rutting. However, this does not necessarily mean that PWL methodology does not work. Different factors may have relatively different influence on rutting performance and they may affect each others' effect as well. It is the goal for a good QA program to factor this in, to determine how combined PWL is calculated for the entire project.

Effect of varying HMA thickness on project performance by sublot

Based on the analysis and results for the effect of AC thickness variation on project performance, the following conclusions were drawn:

1. The performance of different sublots within a project can vary significantly.
2. HMA mat thickness variation has a significant influence on fatigue performance. The mean fatigue ranges from 25.3% to 40.6% for HMA mat thickness ranging from 7.1 in. to 8.1 in., respectively. Thus, the loss in fatigue performance can be significant with as little as 0.5-in. reduction in HMA mat thickness.
3. There seems to be almost 10% increase in fatigue distress for every 0.5-in. decrease in HMA mat thickness from 8.1 in. to 7.1 in.
4. The mean IRI ranges from 143 in./mile to 158 in./mile for thicknesses ranging from 8.0 in. to 7.1 in., respectively. It should be noted that in MEPDG, the IRI values are influenced by rutting and fatigue distresses.

MEPDG simulations provided valuable information to help us understand how performance varies within a project within such narrow ranges and variations in inputs, which is not possible to gain any other way.

10.2 Summary of Results from ERS Risk Analysis

A Monte-Carlo simulation was developed to assess the current QA program of MDOT. The analysis of the simulation showed that production variability, measurement variability, sample size and bias have a significant influence on the payment risk to be made to the contractor. This knowledge leads to identification of ways to reduce payment risk. The simulation can be used to analyze all other variables of a QA program and thereby improve it to achieve lower risk of overpayment or underpayment. The analysis also showed that, even though

measurement variability is very low, high production variability and mean production being in the middle of the specification window still creates risk. Accordingly, the contractor should (1) produce right around the target, and (2) maintain low variability in production quality. This is significant from the point of view of pavement performance.

Generally the test methods and instruments are standardized and calibrated in the beginning of the construction project. For longer projects the instruments may develop bias with continued use over several days. Bias has a very significant effect on payment risk. Such situations can lead to disputes and even law suits. Therefore, bias must be avoided through suitable inspection of the test equipments.

10.3 Feedback Process for Design

A feedback process to design using QA and QC data was proposed. In addition, an analysis was performed to (1) to understand the nature and level of variability in MDOT projects, and (2) to isolate the effect of strict quality control and other factors on the variability among projects. This will help in making realistic assumptions in the design process. The analysis was performed using actual field data from projects constructed under MDOT QA program in 2008. It was also established that most of the construction data follow normal distribution to a large extent. Consequently, the assumption of normality that was made in developing the risk analysis simulation for MDOT QA program was validated.

Another simulation was developed to design optimal feedback process for design. The simulation helped in estimating the errors that can be caused by the sample size in the feedback process and the associated probabilities. Finally plots were developed to relate sample size to the probabilistic error in estimating the mean and standard deviation of the quality characteristic and payment risk. These plots can be used directly by MDOT to decide about the appropriate sample size (i.e., not too small to lead to higher errors and risk and not too large to be too costly or impossible to carry out).

10.4 Use of Non-destructive Tests in QA Program

A detailed review of non-destructive tests is provided in Appendix B. The following are some of the relevant tests for use in QA programs of flexible pavements:

- Ground Penetration Radar (GPR) testing for thickness measurement
- Falling Weight Deflectometer (FWD) testing for modulus estimation
- Dynamic Cone Penetrometer (DCP) and/or lightweight FWD for modulus measurement of unbound layers
- Infra-red thermal imaging of HMA mats for checking uniformity and detection of temperature segregation

10.5 Overall Conclusion and Recommendations

Based on the review of Michigan and other DOT QA programs, it is concluded that the MDOT QA program is on par with ERS based QA programs used by the majority of the states. In other words there is nothing alarmingly different in the QA program being used by MDOT as compared to other "ERS" states. The results presented in this report confirmed the importance of HMA asphalt content, air void content and HMA thickness on long-term performance. It is recommended that HMA thickness be considered as an additional candidate QA parameter and that the interactions between QA characteristics (e.g., air voids and asphalt content) be incorporated in the pay formulae.

Because empirical analyses linking key characteristics to long-term performance were inconclusive (not enough data from the MDOT construction database and inconclusive results from the LTPP database), it is recommended that the mechanistic-empirical approach be adopted for this purpose. With the future possible adoption of the MEPDG by MDOT and other DOT's, it is suggested that the MEPDG be adopted for this purpose. The analyses conducted as part of this research study can serve as examples for such future efforts. The advantage of mechanistic-empirical approach is its ability to quantify the relative effects of deviations from the target on long-term performance and to include interactive effects between different QA characteristics. This allows for modifying/refining the pay formulae based on rational arguments.

Therefore, potential improvements to the QA program should focus on fine tuning the specification limits used and refining the pay formulae to minimize the risk associated with construction variability. In addition, accounting for interactions between certain QA construction quality characteristics (e.g., asphalt content and air voids) in these formulae should help in preventing extreme combinations that have drastic negative effects on pavement performance. Ideally, these refinements should be made based on mechanistic analyses. The pay formulae should be based on the final pavement performance such as pavement life, which can be predicted based on the MEPDG.

The QA data and the pavement surface distress data obtained by MDOT's PMS are the two most relevant data for evaluating the effectiveness of the current QA processes. Unfortunately, MDOT's QA data are either incomplete or missing. A good database system for storing QA data should therefore be developed.

The complexity of the QA processes increases as the number of characteristics is increased. If we rely on probability/statistics methods to investigate the impacts of acceptance sampling rules on the risks of accepting poor quality level of products and rejecting good quality level of products, it may suggest that there is a need to investigate how to reduce the number of characteristics for QA processes without affecting product quality level. However, if we use simulations based on mechanistic modeling, we can account for multiple QA characteristics and their interactions without the need for complex analyses.

Finally it is recommended that the QA data be used as part of the feedback process for design, as described in chapter 9 of this report. Results from probabilistic analyses like those described in chapter 9 can be used for the selection of optimal sample size for QA testing in order to minimize the error in estimating the mean and standard deviation of the quality characteristic and estimated pavement life.

The use of non-destructive testing to quantify as-constructed material properties should be a systematic part of the QA program. For AC pavements, GPR and FWD testing should be

conducted as they offer complementary information on the pavement structure and material properties/parameters. In addition, DCP and/or lightweight FWD testing for unbound materials should be conducted in reconstruction projects or in areas where pavement coring is done. Thermal imaging of the HMA mat should also be made part of the QA program in addition to the current density measurement requirements.

Part II : Rigid Pavements

CHAPTER 11: MDOT Quality Assurance Data for Rigid Pavements

Certain projects were identified to get a sample of construction and materials test data collected historically on highway construction projects. Although this project was meant to evaluate quality *assurance* procedure, data from quality *control* procedures were also collected as part of the data collection effort. The rationale behind this decision is that different quality assurance procedures were followed at different times in the history of highway construction in Michigan. However, if suitable data were collected under the quality control procedures, they could be used with the current quality assurance procedure, or a slightly modified version of it (like a shadow specification), to assess how the construction would be rated according to those specifications. This can then be related to the actual performance of those pavements. In other words, any material test data, whether under QC or QA, can be evaluated for their possible relationship to pavement performance. If any of such test results show strong correlation to performance, then they can be considered for quality assurance.

11.1 Selection of Sample Projects

Two sources of data were identified:

(1) A list of projects for which performance data had already been processed by MDOT was used to identify the first set of sample projects for data collection.

(2) Actual construction documents from those projects stored in boxes in the record center of Michigan Department of Transportation and microfilms stored in C & T office of MDOT, Lansing.

Table 11.1 lists the projects which were selected as sample projects. All the data available at MDOT related to these projects were collected. This was done in order to understand the types of data that would be available for analysis in these projects.

Table 11. 1 Sample rigid pavement projects selected for data mining.

REGION	ROUTE	CS	JN	BMP	EMP	LET	LOCATION
University	I-96 EB	47065	28215	5.671	9.200	6/7/1996	Chilson Road to Dorr Road
University	I-96 WB	47065	28215	5.623	9.223	6/7/1996	Chilson Road to Dorr Road
Southwest	I-94 EB	11017	32516	0.905	5.603	5/12/1995	E of I-196/US-31 Interchange to W of M-140
Southwest	I-94 WB	11017	32516	0.888	5.886	5/12/1995	E of I-196/US-31 Interchange to W of M-140
Metro	WB/I-27	63191	36003	0.000	2.230	5/12/1995	Ramp from WB I-96/NB I-275 to WB I-96
		61131	3036				

Key: CS: Control Section
JN: Job Number
BMP: Beginning Mile Point
EMP: Ending Mile Point

11.2 Sample Data Gathering

All the boxes and microfilms (if available) were searched for data related to the sample projects listed in Table 11.1. Table 11.2 summarizes the number of concrete pavement projects with test data available. Appendix A gives a more detailed list of the types of data available.

Table 11. 2 Summary of test results collected from sampled concrete pavement projects.

Material Property	Cores	Stockpile	Job Site
Compressive Strength	2		
Gradation		2	2

It can be seen that, for concrete pavements, there is very little test data available in the project construction files. There is a possibility that this data had been collected, but was either misplaced or is being used by some third party. Other quality characteristics used in contemporary QA programs are also very scarce.

Considering the fact that there is scarcity of data required for determining the influence of various quality characteristics on pavement performance, it was deemed necessary to consider other alternatives rather than relying on Michigan data. The alternatives were to collect relevant data from the Long Term Pavement Performance (LTPP) database and to conduct mechanistic-empirical analysis using MEPDG and other performance models.

CHAPTER 12: Review of MDOT QA Program and Design Process for PCC Pavements

12.1 MDOT PCC Pavement Quality Assurance Program

Portland cement concrete work is investigated during all phases of production through placement for adequate and acceptable quality. Sampling, testing, and inspection should meet or exceed the minimum rates specified by the Michigan Department of Transportation. The following are the current concrete Quality Assurance (QA) sampling and testing requirements:

- Concrete Temperature
- Concrete Slump
- Concrete Air Content
- Concrete Yield
- Concrete Strength
- Concrete Pavement Thickness

Random sampling and testing methods, which assure all material being produced have an equal chance of being selected for testing, is specified by the engineer. The materials quality assurance manual describes such acceptable methods.

- **Concrete Temperature**: It is required that, at the time of placement, concrete temperatures be between 45° F and 90° F. Steam or hot water coils, live steam, and indirect hot air is usable for heating concrete ingredients, e.g. water and/or aggregates, if necessary, to meet the minimum placement temperature. Also, if the mean daily air temperature is expected to remain below 45° F during the curing period, the engineer may allow or require the use of concrete accelerators, e.g. additional cement or an admixture, to accelerate the rate of strength gain.

- **Concrete Slump**: Slump reading should not exceed 3.0 inches or the contractor's approved mix design. Samples for slump measurements are taken and tested in accordance with ASTM C 143, Standard Test Method for Slump of Hydraulic-Cement Concrete.

- **Concrete Air Content**: At the time of placement, existence of 6.50 ± 1.50 percent entrained air is required. Samples for determination of air content for freshly mixed concrete are taken

in accordance with MTM 207 and tested by ASTM C 231, Standard Test Method for Air Content of Freshly Mixed Concrete by the Pressure Method, or by ASTM C 173, Standard Test Method for Air Content of Freshly Mixed Concrete by the Volumetric Method.

If freshly mixed concrete contains slag or other highly absorptive coarse aggregate, air content of the mixture is determined according to ASTM C 173, Standard Test Method for air content of freshly mixed concrete by the volumetric method.

- **Concrete Yield**: As concreting operation starts and specified slump and air content is attained, unit weight of each mix design should be immediately determined. Average unit weight from three different batches is used for determination of actual mixture yield. That is, in the progress of concreting, if yield based of a single unit weight determination differs from theoretical value (adjusted for differences in air content) more than ±2 percent, two additional unit weight determination needs to be made.

- **Concrete Strength**: At least once every 200 cubic yards for a specific mix design of one day's production, concrete strength against flexure and compression should be determined for two cylinder or beam samples. Specified procedures for concrete strength sampling, curing, and testing are ASTM C 31, ASTM C 39, ASTM C 78, and ASTM C 293, respectively.

- **Concrete Pavement Thickness**: Final acceptance of concrete pavements is based on thickness and, if required, depth of reinforcement below the concrete pavement surface. In accordance with MTM 201, cores are taken and determination of pavement thickness is conducted. If thickness or location of steel reinforcement exceeds acceptable tolerances, samples are classified by their deficiency and the contract unit price will be adjusted accordingly. These adjustments will be applied cumulatively to the pavement unit being evaluated.

Figure 12.1 shows the flowchart for acceptance testing procedure for PCC pavements, with corresponding tables 12.1 through 12.3. Table 12.4 show the sampling rates for the concrete QA program.

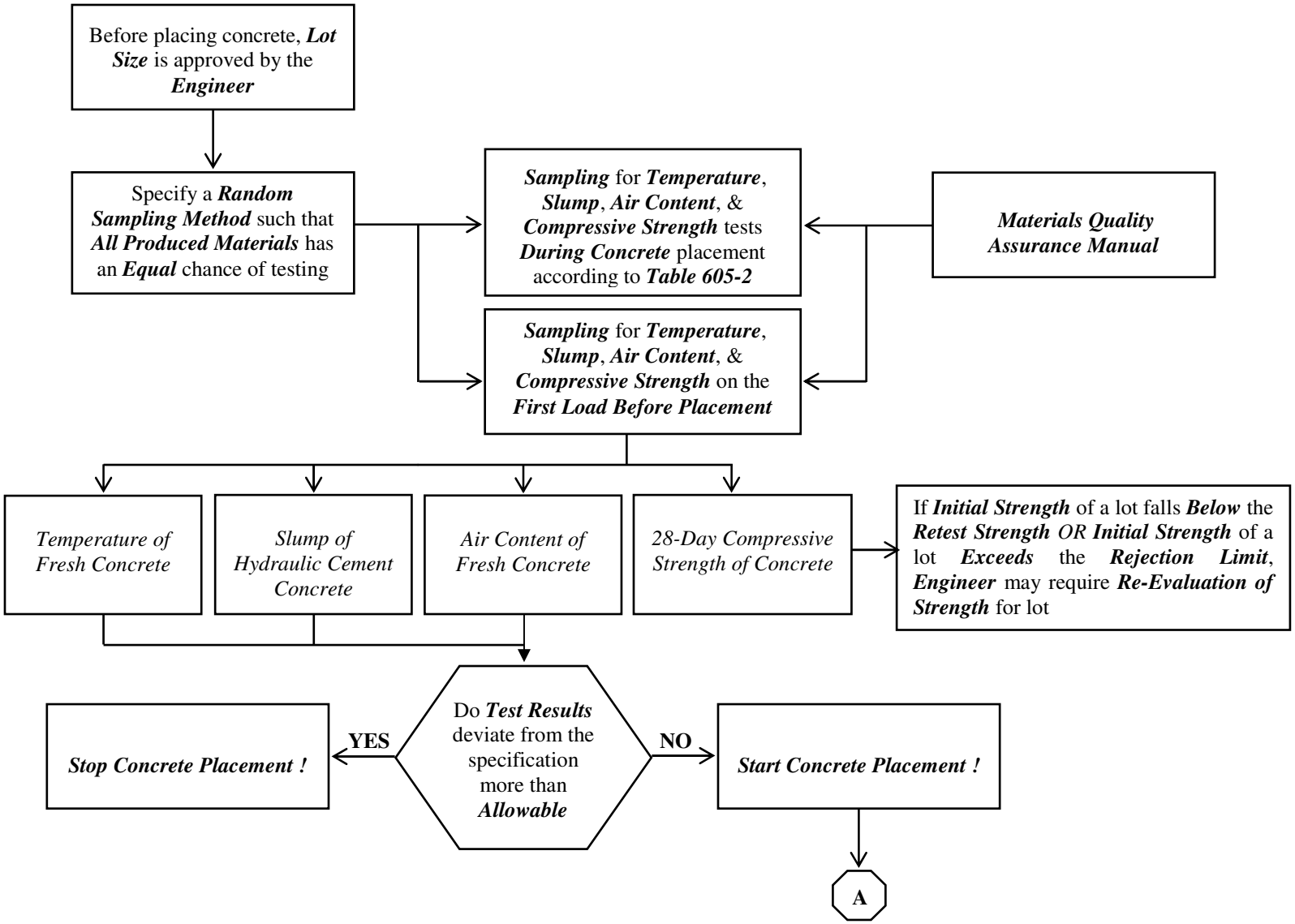


Figure 12.1a Flowchart for Acceptance Testing Procedure for PCC Pavements

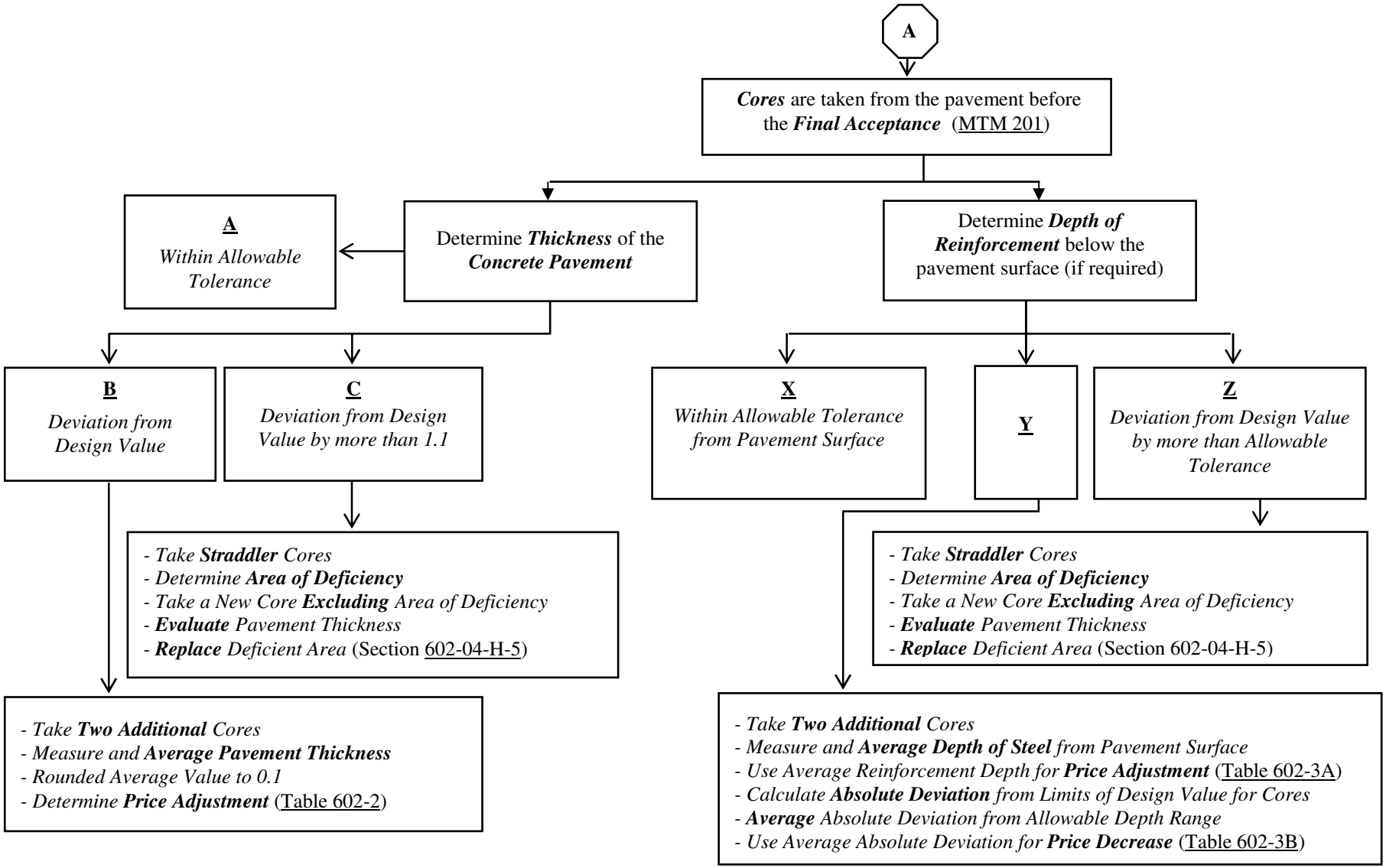


Figure 12.2b Flowchart for Acceptance Testing Procedure for PCC Pavements

Table 12.1 Price Adjustment for Concrete Thickness Deficiency (MDOT Table 602-2)

Initial Core Type	Deficiency in Thickness (Inch)	Price Adjustment (Percent)
A	0.20 or Less	0
B	0.30	-5.0
B	0.40	-15.0
B	0.50	-25.0
B	0.60 To 1.0	-50.0
C	1.10 and Over	-100 [a]

a. *Corrective Action* up to and including *Remove* and *Replace* pavement

Table 12.2 Price Adjustment for Depth of Steel from Pavement Surface (MDOT Table 602-3A)

Initial Core Type	Tolerance on Depth of Reinforcement (Inch) [a, e] For Uniform Plan Thickness (Inch) [c]					Price Adjustment (Percent)
	7.75 - 8.50	8.75 - 9.50	9.75 - 10.50	10.75 - 11.50	Shoulder	
Z	0.0 – 0.9	0.0 – 0.9	0.0 – 0.9	0.0 – 0.9	0.0 – 0.9	-100 [d]
Y	1.0 – 1.9	1.0 – 1.9	1.0 – 1.9	1.0 – 2.4	1.0 – 2.4	-25 [d]
X [b]	2.0 – 4.0	2.0 – 4.5	2.0 – 5.0	2.5 – 5.5	2.0 – 4.0	0
Y [b]	4.1 – 4.8	4.6 – 5.4	5.1 – 6.0	5.6 – 6.6	4.1 – 5.0	-25
Y [b]	4.9 – 6.4	5.5 – 7.2	6.1 – 8.0	6.7 – 8.8	5.1 & Over	-50
Z [b]	6.5 & Over	7.3 & Over	8.1 & Over	8.9 & Over	-	-100 [d]

a. When a pavement is specified to reinforce with two layers of reinforcement, only the top layer of steel will be measured for the proper depth.
b. When a core length measures 0.20 inches or more over the plan thickness, the maximum depth range will be increased by one-half the excess core length over the plan thickness. For each core, the increase will be rounded off to the nearest tenth of an inch according to AASHTO R 11 and then added to the range shown.
c. Pavement base course
d. Corrective action up to and including remove and replace pavement
e. Use the same depth range used for pavement thickness that the shoulder is tied to. Use average shoulder thickness, if tapered.

Table 12.3 Adjustment for Deviation of Depth of Steel from Design Range ((MDOT Table 602-3B)

Initial Core Type	Allowable Average Absolute Deviation from Design Depth of For Uniform Plan Thickness (Inch) [c]					Price Adjustment (Percent)
	6.50 - 7.50	7.75 - 8.50	8.75 - 9.50	9.75 - 10.50	11.0 - 13.0	
X [d]	0.0 – 0.5	0.0 – 0.5	0.0 – 0.5	0.0 – 0.5	0.0 – 0.5	o
Y [d]	0.5 – 1.0	0.5 – 1.0	0.5 – 1.0	0.5 – 1.0	0.5 – 1.0	-10
Y [d]	1.0 & Over	1.0 & Over	1.0 & Over	1.0 & Over	1.0 & Over	-25
Design Range	2.0 – 4.0	2.0 – 4.0	2.0 – 4.0	2.0 – 4.0	2.0 – 4.0	

a. When a pavement is specified to be reinforced with two layers of reinforcement, only the top layer of steel will be measured for the proper depth.
b. Use same depth range used for pavement thickness that the shoulder is tied to. Use average shoulder thickness, it tapered.
c. Pavement or base course.
d. When a core length measures 0.20 inches or more over the plan thickness, the maximum depth range will be increased by one-half of the excess core length over the plan thickness. For each core, the increase will be rounded off to the nearest tenth of an inch according to AASHTO R 11 and then added to the range shown.

Table 12.4 Sampling Rates, Re-Sampling Rates, & Rejection Limits for Concrete QA

	Grade of Concrete				
	D	S1	T	S2 / P1	S3 / P2
Critical Concrete QA Items					
Initial Sampling Rate (per lot)					
Lot Size 0 – 100 Cubic Yards	3	3	3	3	3
Lot Size Over 100 Cubic Yards	6	5	4	5	4
Retest Strength (PSI)					
Lot Size 0 – 100 Cubic Yards	4,500	4,000	3,500	3,500	3,000
Lot Size Over 100 Cubic Yards	4,000	3,500	3,000	3,000	2,500
Non-Critical Concrete QA Items					
Initial Sampling Rate (per lot)	3	3	3	3	3
Retest Strength (PSI)	4,500	4,000	3,500	3,000	2,500
All Concrete QA Items					
Rejection Limit (Percent)	10	10	10	10	10
Re-Sampling Rate (per lot)	6	6	6	6	6

12.2 MDOT PCC Pavement Design Process

An effective pavement design is highly dependent upon performing an adequate investigation of the existing pavement structure. Therefore, prior to construction/reconstruction of a pavement some investigations such as:

- reviewing as-built plans
- reviewing and analyzing existing pavement distress condition
- determining causes of pavement surface distresses
- evaluating pavement ride quality
- reviewing pavement remaining service life
- evaluating drainage system
- evaluating subgrade

should be conducted. A comprehensive investigation of the pavement structure not only ensures that the Engineer employs the proper reconstruction or rehabilitation strategies, but also aids the Designer in the selection of appropriate input values for pavement design.

MDOT uses the pavement design methodology recommended by the American Association of State Highway and Transportation Officials, AASHTO:

- 1993 AASHTO Guide for Design of Pavement Structures
- AASHTO Pavement Design Software DARWin

The following summarizes typical values recommended for design inputs :

- New / Reconstruction Design Life ……..…… 20
- Accumulated ESAL's Years ……. 20
- Initial Serviceability ……. 4.5
- Terminal Serviceability …… 2.5
- Reliability Level ……... 0.95
- Overall Standard Deviation …….. 0.39

- 28-Day Mean PCC Modulus of Rupture ……. 670 psi
- 28-Day Mean Elastic Modulus of Slab ……... 4,200,000 psi
- Load Transfer Coefficient (J) …….. 2.7 (Tied Shoulder/Widened Lane)
  ……... 3.2 (Untied Shoulders)

- Mean Effective k-Value: 50 – 200 psi/in (Typical Range)

Use AASHTO's chart for estimating composite modulus of subgrade reaction and for the correction of effective modulus of subgrade reaction for potential loss of subbase support (Figures 3.3 and 3.6 in AASHTO's 1993 Guide for Design of Pavement Structures).

- Effective Existing Pavement Thickness (Condition Survey Method): Pavement management condition data are used as an aid but a site review of the existing pavement and the planned amount of joint work to be done prior to the concrete overlay must be obtained.

- Overall Drainage Coefficient ……1 to 1.05

CHAPTER 13: Survey of QA Programs for Rigid Pavements in the United States

Performance-related specifications for PCCP are not used by as many agencies as for the flexible pavement; however their use is increasing more rapidly than the HMA pavements. The vast majority of the states use the same tests for QA as for QC.

Table 13.1 gives the number of agencies using particular attributes for concrete QC and QA. Forty agencies had responded to the survey from which this data has been extracted (Hughes, 2005).

Table 13.1 Most commonly used quality characteristics for QC and QA on PCC pavements (Hughes, 2005)

Quality Characteristic	No. of Agencies	
	QC	QA
Air content	25	38
Thickness	14	36
Slump	24	33
Cylinder strength	18	31
Gradation	25	26
Beam strength	14	18
Water-cement ratio	12	16
Ride quality	1	15
Aggregate fractured faces	7	6
Sand equivalence	0	3
Permeability	0	3
Core strength	0	2

Table 13.1 and figure 13.1 show the attributes that are most often used for acceptance of PCC pavements: Air content, used by 38 agencies; thickness used by 36; and slump used by 33. Thirty one agencies accept PCC structures based on cylinder strength, and 26 accept gradation. Michigan also uses the first four of these quality characteristics in its QA program. The lesser-used acceptance attributes are aggregate fractured faces, sand equivalence, permeability and core strength.

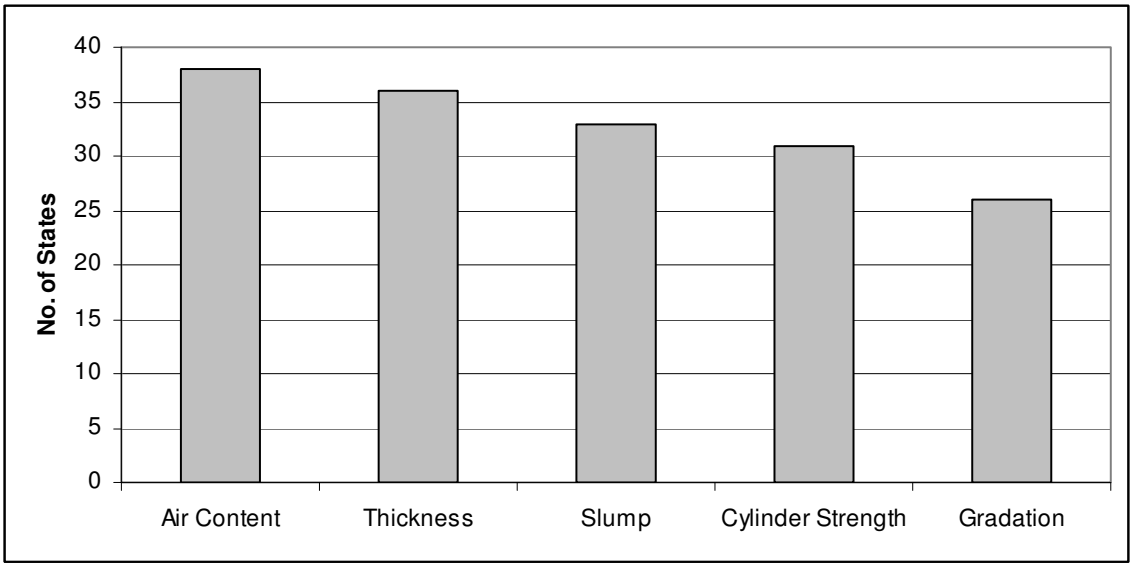


Figure 13.1 Attributes most often used for QA of PCC (Hughes, 2005)

13.1 Regional Level Analysis of the QA Practices

We have listed the QA practices in various states on the basis of their regional location in US.

The quality assurance parameters and the corresponding criteria for rigid pavements in southern states are given in table 13.2 and compared to those used in Michigan. We can observe that there is lot of variability in the criterion followed across the states. For instance, the sample and testing procedure for compressive strength is AASHTO T 22 in Arkansas, but in Georgia, GDT-22 is used for the same quality characteristic. In Texas, the tolerance for deviation from the QA standards for compressive strength is 20% of the mean while Arkansas requires a minimum compressive strength of 4000 psi.

The QA program for rigid pavements in Michigan is done on characteristics such as concrete temperature, slump, air content, yield, strength and pavements thickness. It is required that at the time of placement, the concrete temperature should be in the interval 45° F to 90° F. The concrete slump metric should not exceed 3.0 inches. The sample for slump measurements are taken in accordance with "Department Methods" and tested by ASTM C 143.

At the time of placement, 6.50 ± 1.50 percent entrained air is required. Samples for determination of air content for freshly mixed concrete are taken in accordance with MTM 207 and tested by ASTM C 231 or C173. The final acceptance of concrete pavement is based on

thickness. In accordance with MTM 201 cores are taken and determination of pavement thickness is conducted.

Other QA procedures for concrete pavements are summarized in Table 13.3 (Western states) and 13.4 (other states).

Table 13.2 Quality Assurance Practices in Southern States versus Michigan

	Arkansas	Georgia	Texas	Michigan
Sample & Test				
Gradation	AASHTO T 27			
Compressive strength	AASHTO T 22	GDT-22	Tex-448-A Tex-418-A	ASTM C 31
Air content	AASHTO T 152	GDT-26	Tex-414-A	MTM 207, ASTM C 231
Slump	AASHTO T 119	GDT-27	Tex-415-A	ASTM C 143
Sample location				
Compressive strength		Paver		
Air content		During Pouring Operation at Roadway		
Slump		During Pouring Operation at Roadway		
Frequency				
Gradation	One test per 750 Cubic meters of mix	Two per One Day of Production		
Compressive strength	Test specimens for compressive strength determined by cores will be obtained according to AASHTO T 24	Two per One Day of Production		3 to 6 per lot
Air content	Sampled after placement on grade, but before consolidation by paver or vibrators	Two per One Day of Production		3 to 6 per lot
Slump	Sampled after placement on grade, but before consolidation by paver or vibrators	Two per One Day of Production		3 to 6 per lot
Tolerance				
Gradation	±7.0			
Compressive strength	4000 psi (min)		20% of the mean	2,500 to 4,500 for diff grades
Air content	4.0%-8.0%		± 1%	6.50 ± 1.50
Slump			± 1%	± 3 inches

Table 13.3 Quality Assurance Practices in Western States versus Michigan

	Colorado	Utah	Michigan
Sample & Test			
Thickness	CP 68	AASHTO T 24	ASTM C 31
Compressive Strength		AASHTO T 22, 24	MTM 207, ASTM C 231
Air Content	T 141, 152		
Sample Location			
Thickness	Finished Concrete	Finished Pavement	
Compressive Strength		Grade	
Air Content	Mixer Discharge		
Frequency			
Thickness	In accordance with Subsection 412.21.	1 per 12,000 Square Feet	
Compressive Strength		Smaller of One Strength Test Each 700 Square Yards or one Day's Placement	
Air Content	1 per 2500 sq. yards. Or one per day if less than 2500 sq. yards. are placed in a day tests.		
Tolerance			
Thickness		0 - ⅛ ⅛ - ¼ ¼ - ½ ½ - ¾ >¾	
Compressive Strength		(Below 4,000 psi) 1 - 100 101 - 200 201 - 300 301 - 400	

Table 13.4 QA Practices in Some of the States for PCC Pavements

(a) Alabama

Quality Characteristic	Sampling Method	Sample Size	Sampling Location	Testing Frequency
Compressive strength	AASHTO T 23 & AASHTO T 141	Set of 2 Cylinders	ALDOT 210	1 per Pavement Testing unit
Air Entraining	AASHTO T 141	Minimum of one	ALDOT 210	Minimum 1 per Pavement Testing unit
Slump	AASHTO T 141	Minimum of one	ALDOT 210	Minimum 1 per Pavement Testing unit
Thickness	AASHTO T 24	One core	ALDOT 210	1 per Pavement Testing unit

(b) Indiana

Quality Characteristic	Sampling Method	Test Standard	Frequency	Tolerance
Air Content	ITM 802	AASHTO T 152 or ASTM C 173	One per sub lot	– 0.8% to +2.4% of the 6.5% Target
Flexural Strength		AASHTO T 97	Two beams per sub lot	
Thickness of PCCP		ITM 404	Two per sublot	
Unit Weight and Relative Yield		AASHTO T 121	One per sublot	± 3.0%
Water-Cement Ratio		ITM 403	Once per week	± 0.030

(c) Maine

Quality Characteristic	Sample & Test Standard	Sample Location	Tolerance
Gradation	AASHTO T30	Paver Hopper	
PGAB Content	AASHTO T308	Paver Hopper	Target +/-0.4%
%TMD (Surface)	AASHTO T269	Mat behind all Rollers	
%TMD (Base or Binder)	AASHTO T269	Mat behind all Rollers	
Air Voids at N_d	AASHTO T 312	Paver Hopper	4.0% +/-1.5%
%VMA at N_d	AASHTO T 312	Paver Hopper	
Fines to Effective Binder	AASHTO T 312	Paver Hopper	0.6 to 1.2
%VFB	AASHTO T 312	Paver Hopper	

CHAPTER 14: Literature Review – Rigid Pavements

Several research projects were reviewed to verify the significance of candidate QA variables identified in this project for pavement performance. These variables are listed in the following sections.

14.1 Slab Thickness

PCC slab thickness is obviously the basic parameter in the design of rigid pavements. In support of such a statement, it has been experimented that higher tensile strength is not necessarily required for excellent long-term performance in any climate because slab thickness is an effective parameter in controlling the fatigue resistance (Hansen et al., 2001). However, quality of subgrade and compressive strength of the concrete affect the PCC slab thickness design.

In general, variation in thickness affects pavement performance such as IRI, faulting, and cracking. For instance, predicted 20-year IRI decreases as concrete slab thickness increases (Khanum et al., 2005). Among pavements built on fine-grained soils, those with 8-inch PCC slab have higher changes in IRI than those with 11-inch PCC slab (Chatti et al., 2005). Faulting performance improves in pavements with thicker slabs, (Khanum et al., 2005).

In addition, it has been reported that increasing slab thickness from 8 to 12 inches decreased pavement cracking from 90 percent to 0.3 percent.(Khanum et al., 2005). The study of LTPP data has also proved that occurrence of longitudinal and transverse cracking among pavements with thinner PCC slabs (8 inches) is higher than among those with an 11-inch thick PCC slab (Chatti et al., 2005).

14.2 Water-Cement Ratio

The water-cement ratio is defined as the mass of mix water divided by the combined mass of the cement and any additional cementitious admixtures. This ratio is important in determining the overall strength of the mix as well as other mechanical properties including