Technical Report Documentation Page

1. Report No. FHWA/TX-14/0-6675-1	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Revised Pay Adjustment Factors for HMA and Concrete Pavements		5. Report Date February 2014; Published November 2014 6. Performing Organization Code
7. Author(s) Prasad Buddhavarapu, Andre Smit, Jorge Prozzi, Wei Fan, Zegeye Gurmu		8. Performing Organization Report No. 0-6675-1
9. Performing Organization Name and Address Center for Transportation Research The University of Texas at Austin 1616 Guadalupe Street, Suite 4.202 Austin, TX 78701 College of Engineering and Computer Science The University of Texas at Tyler 3900 University Boulevard, Tyler, TX 75799		10. Work Unit No. (TRAIS) 11. Contract or Grant No. 0-6675
12. Sponsoring Agency Name and Address Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, TX 78763-5080		13. Type of Report and Period Covered Technical Report: September 2011–February 2014 14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

16. Abstract

The current pay adjustment system for HMA production, placement, and ride consistently rewards contractors, but does not necessarily result in improved performance of constructed HMA pavements and longer service life. The current system needs to be changed in order to improve the quality of pavements in Texas and provide performance-related incentives. A database framework was developed incorporating TxDOT's SiteManager QC/QA database and network-level performance data in the Pavement Management Information System (PMIS) database, yielding a large dataset comprising more than 600 pavements across Texas with available QC/QA data and performance records spanning 3 to 10 years. The research team evaluated the influence of variations in the construction QC/QA parameters on pavement performance. Advanced statistical modeling of these relationships using econometric approaches was conducted to establish the significance, sensitivity, and consistency of these parameters in regard to pavement performance. The statistical models provided the tools necessary to evaluate the current pay adjustment system with an eye to developing new performance-related specifications. This report provides recommendations for revising the production and placement pay adjustment factors for HMA pavements and revised pay adjustments for the ride quality of HMA and concrete pavements.

17. Key Words	18. Distribution Statement
Pay adjustment factor, pavement performance, HMA production and placement, bonus & penalty, ride quality.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161; www.ntis.gov.

19. Security Classif. (of report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	190	

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

Revised Pay Adjustment Factors for HMA and Concrete Pavements

Prasad Buddhavarapu
Andre F. Smit
Jorge A. Prozzi
Wei Fan
Zegeye Gurmu

CTR Technical Report:	0-6675-1
Report Date:	February 2014
Project:	0-6675
Project Title:	Evaluation of Bonus/Penalty Pay Adjustment Systems for HMA and Ride Specifications of Concrete and Asphalt Pavements
Sponsoring Agency:	Texas Department of Transportation
Performing Agency:	Center for Transportation Research at The University of Texas at Austin The University of Texas at Austin

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.

Center for Transportation Research
The University of Texas at Austin
1616 Guadalupe, Suite 4.202
Austin, TX 78701

http://ctr.utexas.edu/

Disclaimers

Author's Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation.

Patent Disclaimer: There was no invention or discovery conceived or first actually reduced to practice in the course of or under this contract, including any art, method, process, machine manufacture, design or composition of matter, or any new useful improvement thereof, or any variety of plant, which is or may be patentable under the patent laws of the United States of America or any foreign country.

Engineering Disclaimer

NOT INTENDED FOR CONSTRUCTION, BIDDING, OR PERMIT PURPOSES.

Research Supervisor: Jorge Prozzi

Acknowledgments

The authors would like to thank Tomas Saenz from the El Paso District, who served as Project Director. The authors are also grateful for the assistance provided by Maghsoud Tahmoressi and Dale Rand from Pavetex Inc., who served as consultants on the project.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Background

The Texas Department of Transportation (TxDOT) currently uses a pay adjustment factor (PAF) system for production and placement of hot mix asphalt (HMA) and ride quality of HMA and concrete pavements; this system has been in existence for almost a decade. Under the current PAF system, the reward or penalty to a contractor is based on whether a contractor exceeds or fails to meet the average contractor level of performance, which was calculated based on historical data reflecting an "average" contractor's capabilities. As contractors have gotten increased experience with TxDOT projects, their average performance has also increased. Consequently, current PAFs are skewed, resulting in PAFs that are mostly greater than 1.0, with contractors being rewarded more often than penalized. A sound and rational pay adjustment system should aim to enforce strict quality control in pavement construction projects and should be independent of historical data. Furthermore, such a system should be developed based on the relationship between measurable parameters in a construction project and expected performance of the constructed facility such that the bonuses or penalties applied can be economically justified. It is necessary to revise the current pay adjustment system to ensure that bonuses awarded to contractors don't exceed the benefits to the highway agency, and that the penalties levied on contractors don't fall short of the potential losses incurred by the agency due to reduced service life as a result of poor workmanship. TxDOT needs to maintain a harmonious relationship with local construction contractors but at the same time assure consistently superior products. A rational performance-based pay adjustment system improves the contractual relationship by balancing the risks between TxDOT and contractors to benefit both parties in the long run.

The fundamental purpose of any construction specification is to ensure the expected long-term performance of the final product. A construction project that fails to meet the required quality level should always result in reduced payment to the contractor in order to recover the future costs incurred by TxDOT for additional maintenance. On the other hand, a project with superior quality must be rewarded based on actual savings to the agency corresponding to the improvement in performance due to higher construction quality. A performance-related specification (PRS) incorporates the economic implications associated with superior (or inferior) pavement performance since it translates to extended (or reduced) pavement service life. The main goal of any PRS is not to improve the quality of construction, but to improve the specifications by determining what best reflects the quality of the product and to create a contractual framework that maximizes the cost effectiveness. Implementing a PRS in TxDOT's highway construction specifications requires an understanding of the relationship between measurable quality control parameters and long-term performance, which is essential to assess the quality of any construction job.

Quality control in any pavement construction can be implemented at three levels. The first level is controlling the variability in the material properties that results from the production-related fluctuations. It is important to identify the material properties that may possibly influence HMA performance and to have a good understanding of the relative magnitude of such effects on future field performance. A balanced production pay adjustment system should therefore aim at rewarding superior-performing mixes and penalizing poor-performing mixes. Secondly, it is important to control the variability resulting from improper placement in the field. Quality control during construction of a pavement structure is just as important as controlling the quality of the plant material. Thirdly, a minimum ride quality requirement is also essential for controlling the overall level of service for a pavement facility. A PRS specification should at least address the financial implications due to deviations in the above three aspects from their target values.

1.2 Research Objectives

The primary goal of this research was to evaluate and modify (if justified) the existing pay adjustment system in TxDOT regarding the aforementioned three essential quality issues: production and placement of HMA and ride quality of HMA and concrete pavements. The project accomplished the goal by sequentially addressing the following four objectives:

1. Evaluate the validity of the existing pay adjustment system for ride quality of HMA and concrete pavements (Specification item 585).
2. Evaluate the validity of the existing PAFs for HMA production (Specification item 341).
3. Evaluate the validity of the existing PAFs for HMA placement (Specification item 341).
4. Modify the existing pay adjustment system based on performance considerations.

In addition to the above-mentioned main objectives, the research also addressed other objectives considered essential for the overall success of this project. These objectives include the development of a database that will be populated with project-level information on production and placement pay factors along with volumetric properties of the mixture. This database was concatenated with the respective performance information upon integrating with Pavement Management Information System (PMIS) databases. Furthermore, a gap analysis was conducted based on the requirements from the study to develop an experimental design that will include test sections for validation of the revised PAFs based on pavement performance. Additionally, an evaluation of the current sample sizes (i.e., sampling frequency) and sampling methods used by TxDOT was also conducted to determine its adequacy and recommendations were developed.

1.3 Report Outline

A brief outline of this research report, including a short summary of individual chapters, is presented below.

Chapter 1 describes the motivation for this research project and introduces the rationale behind pay adjustment systems. The overall goals and major objectives of this research project are also provided.

Chapter 2 provides a comprehensive review of the earlier literature on pay adjustment practices across the US and other countries. Special emphasis is placed on reviewing PRS practices in other states within the US. A review of the various sampling methods during production and placement quality control is also included.

Chapter 3 describes the exercise of building the essential datasets for achieving the overall objectives of this research project. The relevant TxDOT databases are described in detail and a methodology for the integration of these databases is provided. Also provided is a preliminary data exploration, including distributions of quality control parameters and historical PAFs awarded to the contractors.

Chapter 4 describes the econometric model building process that is essential for establishing a relationship between the quality control parameters and pavement performance. This section reports several interesting empirical findings that provide a basis for validating the existing pay adjustment system and proposes revisions thereof.

Chapter 5 describes a methodology for performing sensitivity analysis based on the statistical model developed in Chapter 4. This chapter identifies the performance-sensitive quality control parameters for production, placement, and ride quality. A methodology for modifying the existing PAF system is also provided. The chapter ends with a discussion of the revised proposed PAFs for production and placement of HMA, and ride quality of HMA and concrete pavements.

Chapter 6 addresses the evaluation of the current TxDOT sampling frequency and sampling methods. The chapter provides a description of the work done as part of the study that included a plant mix collection exercise followed by laboratory testing of several hot mix specimens. A statistical analysis was conducted to evaluate the existing sampling frequency and to identify the optimal sampling frequency.

Chapter 7 identifies the need for validating the proposed pay adjustment system using field performance data. Such a validation requires a number of years of performance data, which is beyond the scope of this research project. Nevertheless, an experimental design is proposed for validation of the methods proposed as part of this research. Several HMA and concrete sections were identified across Texas. A brief summary of various important features of the identified test section is provided.

Chapter 8 summarizes the major research findings at different stages of this research project. A discussion of the proposed pay adjustment system and recommendations for implementation thereof are provided.

Chapter 2. Literature Review

This chapter provides a review of pay adjustment systems that are currently being used in other states and countries. It also identifies mixture properties and other variables considered in the calculation of bonuses and penalties for pay adjustment. The review critically evaluates the relevant literature and summarizes the most significant aspects thereof. To date, most of the work on the development of performance-based PAFs has been based on laboratory studies. For this reason, during the literature review process, special emphasis was placed on identifying research based on field performance, full-scale testing, and validation studies. The chapter initially describes the history of PAFs and the development of PRSs. Subsequently, a discussion of various studies across the country regarding the implementation of PRSs is provided. The discussion focuses on the quality control parameters in HMA production and placement, followed by a review of the pay adjustment systems concerning ride quality. The chapter ends with a review of sampling frequency and methods for quality control of HMA production and placement operations.

2.1 History of Pay Adjustment Systems

The tradition of contracting for construction of public roads in the US dates back to the 19th century. Construction specifications evolved as an essential component of contracts. The earliest specifications involved only a prescription of materials and required construction methods—the so-called traditional specifications (Chamberlin, 1995). In the early 1950s, with the onset of the Interstate System and the burgeoning capital expenditure for highway infrastructure development after World War II, larger companies started contracting highway projects. As a result, several new sophisticated construction technologies emerged, most of which were developed by contractors. Traditional method specifications stifled developments in the industry. Moreover, the non-uniformity and non-applicability of general construction methods for project-specific requirements caused several deficiencies in construction quality. In the early 20th century, the American Association of State Highway Officials (AASHTO) helped develop relatively uniform specifications across different US states in conjunction with the American Society of Testing and Materials (ASTM). Subsequently, in the early 1960s, the AASHTO road test demonstrated the inherent variability in highway construction processes and highlighted the rare possibility of 100% compliance. Consequently, statistical end result specifications evolved and by the 1970s replaced traditional specifications at the construction site.

The fundamental intention of any construction specification is improvement of the long-term performance of the finished product. A construction project that deviates from the required quality level should always result in a reduced contractor payment to recover the costs incurred by the agency for additional future maintenance costs. On the other hand, a project with superior quality must be rewarded based on actual savings to the agency, using a figure corresponding to improvements in performance due to higher construction quality and subsequent extended service life. With the advent of sophisticated data models and high quality control procedures in the highway industry, it is clear that traditional and statistical end result specifications do not always ensure the desired quality. Knowledge of the relationship between construction variables and long-term performance is essential in assessing the quality of any construction job. Consequently, research communities identified the necessity for more restrictive PRSs in the highway industry. It is important for any highway agency to establish a harmonious relationship with the local construction contractors, while at the same time assuring consistent superior products. Hence, contractors must be rewarded or penalized based on the actual benefits or losses to the highway agency, which PSRs make possible. The main goal of any PRS is not to improve the quality of construction but to improve the specifications by determining what best reflects the quality and to create a contractual framework that maximizes cost effectiveness (Chamberlin, 1995).

Material and construction quality characteristics (such as initial smoothness, pavement thickness, field air voids, asphalt content, and the strength of concrete cores) are related to fundamental engineering properties that directly or indirectly influence performance. PRSs allow the evaluation of these key characteristics, influencing the long-term performance of the finished product. Performance models in a PRS framework can be used to identify the quality characteristics that influence future performance but also to quantify the corresponding effects.

PRS models may be classified into two broad types: performance-prediction models and maintenance-cost models. The performance-prediction models forecast the time required for the appearance of a pavement distress above the acceptable level, whereas maintenance-cost models predict the changes in future maintenance costs incurred by the agency due to an inferior or superior quality of construction. A typical PRS can be used to a) establish a relationship between measurable quality characteristics and product performance, b) identify the optimum quality level required to maximize the performance for the incurred cost, c) develop a rational basis for contractor’s pay adjustments, and d) integrate construction processes with future pavement management.

The above-mentioned benefits motivate the implementation of PRS by TxDOT to replace the existing pay adjustment schedules, which are based on absolute deviation of quality variables from their target values as established using historical data. The existing pay schedules do not necessarily align with expected performance of pavement projects, leading to either a benefit or loss to the agency. New pay schedules based on PRS models developed using existing TxDOT pavement performance databases are therefore recommended to develop a rational pay adjustment scheme to share the risks between agency and contractor.

2.2 Performance-Related Specifications

2.2.1 Earlier Developments

Research into the development of relationships between construction quality measures and performance was not implemented until the early 1980s (Chamberlin, 1995). The National Cooperative Highway Research Program (NCHRP) synthesis report on statistically oriented end result specifications was published in 1976. Subsequently, the fundamental concepts behind PRS and pay adjustment schedules were developed and well documented in the late 1970s to early 1980s (Irick, 1988; Welborn 1984; Majidzadeh et al., 1984; Von Quintus et al., 1985; Shah, 1987; Weed, 1989). In 1980, a new research program was initiated by the Federal Highway Administration (FHWA) with two major objectives: 1) identify the existing specifications that relate to performance and develop a complete system for PRSs for both flexible and concrete pavements (Mitchell, 1981), and 2) provide a rational basis for pay adjustment plans. The FHWA study found a lack of models to predict the performance of roads from existing material and construction variables. Consequently, in 1985, AASHTO sponsored NCHRP project 10-26 with the aim of identifying the variables measured during quality control that influence the future performance of constructed pavements. The project highlighted the inadequacy of databases at the time for developing the required PRS models. Consequently, NCHRP 10-26 concluded that further research on PRS should focus on developing a general framework that provides multi-stage derivation of the required PRS relationships (Chamberlin, 1995). It highlighted the lack of primary relationships that directly relate material and construction variables to the performance indicators. Irick (1988 and 1990) also identified the necessity to develop secondary relationships that relate the material and construction variables measured during construction with performance indicators.

In 1986, NCHRP project 10-26A, "Performance-Related Specifications for Hot-Mix Asphalt Concrete," concluded that development of PRSs is an implementable goal. A pay adjustment scheme based on the difference between as constructed life cycle costs (LCCs) and target or design LCCs was also included in the refined framework developed during the project. It was recommended to develop additional or refined performance models before replacing the existing specifications with PRS. The laboratory and field performance data from the ongoing Strategic Highway Research Program (SHRP) and NCHRP projects as well as the Long-Term Pavement Performance (LTPP) program were used to develop performance relationships. The FHWA continued to refine primary and secondary relationships for asphalt pavements with additional laboratory and accelerated field studies at a test track facility (Shook et al., 1992; Weed, 1982).

In 1987, the FHWA initiated another NCHRP project for developing PRSs for concrete pavements. Similar research was carried out in developing primary and secondary relationships. However, two additional approaches for developing pay adjustment schedules were included: a method based on liquidated damages, originally developed by the New Jersey DOT (NJDOT), and LCC analysis models as presented in the AASHTO guide.

In 1990, an FHWA study identified 16 high priority research and development needs for the management of highway construction engineering. Out of these, the development of PRS ranked first with a proposal of 60% funding allocation (Chamberlin, 1995). During the same year, a prototype PRS was developed for portland cement concrete (PCC) pavements based on the NJDOT research. These new specifications included measurement of thickness, concrete strength, air content, and roughness of the pavement.

2.2.2 A General Conceptual Framework for PRS

In 1988, Irick proposed a conceptual framework for the development of PRS. Irick (1990) and Anderson (1990) proposed detailed PRS frameworks for concrete and asphalt pavements, respectively. The fundamental concepts in developing a PRS are broadly divided into two sections (Chamberlin, 1995):

1. Design, construction, and performance variables:
    a. Primary dependent variables
        i. Stress indicators
        ii. Distress indicators
        iii. Performance indicators
        iv. Cost indicators
    b. Primary independent variables (stress and distress prediction factors)

        i. Traffic factors
        ii. Environmental factors
        iii. Structural factors
    c. Secondary independent variables
        i. Material and construction surrogate factors for primary prediction factors
        ii. Material and construction control factors
    d. Design criteria
        i. Distress performance criteria
        ii. Reliability criteria
        iii. Time and applications criteria
    e. Uncontrolled independent variables or error term
        i. Uncontrolled deviations from specified levels
        ii. All remaining uncontrolled independent variables
2. Aspects to be considered when developing performance-related material and construction specifications:
    a. **Primary relationships:** Identify and/or derive all primary prediction equations. The primary independent variables, which may or may not be amenable for measuring in the field, are used as independent variables for modeling primary dependent variables.
    b. **Material and construction candidate variables:** Identify all the significant independent variables in the primary relationships. Identify the variables that are easy to measure at the site during the construction process with reasonable cost even though they are not significant.
    c. **Secondary relationships:** Identify the secondary prediction equations. These equations relate the measurable material and construction variables and performance of the pavement.
    d. **Material and construction specification:** Develop algorithms for using primary and secondary equations integrated with design and cost optimization criteria.

The performance models developed in the above steps are used to estimate the expected life of the pavement for any given level of construction quality. The pay schedule is calculated based on calculated expected life estimates. The following three methods for calculation of pay schedules are mentioned in an NCHRP synthesis by Chamberlin (1995):

1. Considering the difference in expected life of as-constructed and as-designed pavement structures as a measure of quality differential is used in New Jersey. This approach does not consider maintenance costs and user operating costs.

2. An FHWA research project developed an approach which calculates the pay schedules based on the difference in life cycles costs associated with designed and constructed products. This approach includes maintenance costs but it does not include user operating costs.
3. NCHRP Report 332 discusses another approach based on estimated economic life period of the project, which is defined as the age at which minimum annual cost occurs. Both maintenance and user operating costs are included in this approach.

2.2.3 PRS Development at Westrack

In 1994, the FHWA funded the Westrack project to further develop PRSs. As part of the project, accelerated testing of 34 experimental HMA pavement test sections was carried out over a 2.5-year period on a closed loop test track facility 1.75 miles in length. Westrack is located approximately 160 miles southeast of Reno in the Nevada desert. It produced adequate data for the refinement of existing performance relationships in order to better account for the influence of off-target values for different material and construction variables. Several factors were considered in the experimental design of the project, including coarse aggregate type, aggregate gradation, asphalt binder type, asphalt content, air void content, and thickness. Driverless trafficking technology was used for loading of the test sections in order to minimize human error. NCHRP Report 455 (Epps et al., 1999) explains the track geometric features and further information detailing the design of the test sections.

Epps et al. (1999) summarized various activities during 4-year project and highlighted major findings. They briefly discussed the development of both regression and mechanistic performance models used for fatigue and permanent deformation of the asphalt surface layers. Fatigue cracking was found to be dependent on asphalt content, aggregate gradation, and compaction based on both laboratory and field results, whereas permanent deformation was dependent on asphalt content, compaction level, and pavement temperature.

Epps et al. (1999) also described a methodology for the calculation of pay factors. Initially, Monte Carlo simulations were carried out to establish the distribution for equivalent single axle loads (ESALs) to failure in terms of both fatigue cracking and rut depth using the corresponding performance models. In this process, random material and construction characteristic values were obtained from their probability distributions. The mean and standard deviation required for building such distributions were obtained from the lot's quality control/quality assurance (QC/QA) data. The expected number of ESALs required for failure was obtained for different random material and construction property combinations as part of the Monte Carlo simulation. Consequently, the distribution for ESALs to failure was obtained. Similarly, the probability distribution for the target number of ESALs was obtained. Thus, the estimated ESALs required for failure were calculated for both on-target construction activity (as-designed) and off-target constructed activity (as-designed) from the corresponding distributions. A ratio between the estimated ESALs for target and as-constructed is calculated (RP). The expected life (OTY) as a result of any off-target material and construction value is calculated as Equation 2.1:

$$\mathrm{OTY} = \frac{\ln(1+\mathrm{RP}\lfloor(1+g)^{\mathrm{TY}}-1\rfloor)}{\ln(1+g)} \tag{2.1}$$

where:

g = annual traffic growth rate expressed as a decimal,

TY = number of years of pavement life resulting from on-target construction.

The extended/reduced duration of pavement expected life causes a movement of future rehabilitation projects in the time scale. The net present value of such movements can be assessed based on the cost model (Epps et al., 1999) in Equation 2.2:

$$\Delta PW = 100\left(\frac{1+d}{1+r}\right)^{TY}\left(\frac{1+r}{1+d}\right)^{OTY} - 100 \qquad (2.2)$$

where:

r = annual rate of construction-cost inflation,

d = annual discount rate.

Epps et al. (1999) also mentioned that the maximum bonuses for superior construction were set at 50% of the agency savings, while the penalties were set at 100% of the added agency costs. They also explained various stages in the development of the PRS software HMAspec (alpha version released in February 2000). The software provides a platform to produce a PRS based on several performance models as developed at Westrack, Monte Carlo simulations, and LCC models. Hand et al. (2004) summarized a case study including PRS in HMA based on four field sections constructed on I-80 east of Reno, Nevada. HMAspec software was used in calculating the pay factor in these field projects. Hand et al. (2004) reported that all the pay factors were very close to 1.0 using the HMAspec software for that particular construction job; however, the study highly recommended refinement of the existing performance models used in the HMAspec software.

In conclusion, the Westrack project highlighted the importance of combining full-scale testing with laboratory testing for the development of PRSs. It is interesting to note that the PRSs were not completely developed during this project (Epps et.al, 1999). However, Westrack did provide other useful information apart from the partial development of PRS, in areas such as pavement construction; QC/QA during construction; vehicle operations; materials specifications; and pavement rehabilitation. Also, the project demonstrated the possibility of achieving close tolerances in terms of asphalt content, aggregate gradation, and compaction during construction (Epps et al., 1999). The Westrack project recommends a field implementation of the new PRS system as a “shadow specification” followed by a trail specification in order to evaluate the impact of new specification systems. It also highlights the additional need for further field and laboratory testing in other environments for further calibration of the developed performance models.

2.2.4 Material and Variables Related to Performance

In 1999, NCHRP project 9-15 was initiated with an objective to identify construction-related quality characteristics and as-produced HMA quality characteristics influencing long-term pavement performance. Based on test results obtained during field experiments (from different projects in Colorado and Illinois) and literature review and surveys, five quality characteristics for HMA pavements were identified for successful incorporation in PRS based on their importance in determining the overall performance of HMA pavements. These are ride quality, in-place density, in-place permeability, longitudinal construction density, and segregation. It is to be noted that, out of these five, TxDOT uses in-place air voids and ride quality for pay factor calculations. Also, the test methods and preliminary threshold values for the above quality characteristics are recommended and described in detail in the final NCHRP project report (Killingsworth, 2001). However, the threshold levels recommended are not validated using long-term performance data; instead, it is assumed that these levels are adequate for achieving the required performance levels. The feasibility of the equipment used to measure quality in the field, the contractor's ability to employ this equipment for the measurement of these quality characteristics, and the variability of each device were also investigated in this project. Since the test protocols were developed based on limited field data, further refinements and validations are recommended. Table 2.1 summarizes the field equipment used and test parameters for each quality characteristics evaluated.

Table 2.1: Field equipment and test parameters for different QC characteristics

Quality characteristic	Field equipment	Test parameter
Segregation	Road Surface Analyzer	Estimated Texture Depth
Initial smoothness	Lightweight Profiler	International Roughness Index
In-place material density	Pavement Quality Indicator	Density
Longitudinal joint density	Pavement Quality Indicator	Density
In-place permeability	NCAT Field Permeameter	K-Value

A survey revealed that substantial number of agencies use the practice of the contractor controlling the quality and the agency performing acceptance (Hughes, 2005). Gradation, asphalt content, volumetric properties and compaction are the most frequently used QC attributes. The performance-based approach considers the mean and variance of material and construction variables rather than considering percent within limits (PWL), the method currently used by many agencies. Note that TxDOT is still using the absolute deviation method, which does not account for variability in the test data.

2.2.5 Probabilistic PRS Approach Based on the Mechanistic-Empirical Pavement Design Guide

Most of the present research efforts recommend calculating pay factors based on the change in LCCs due to a superior or inferior construction quality. However, this approach requires several assumptions and uncertainties regarding the selection of a time period during which the acceptable quality characteristics would perform adequately. As an alternative, NCHRP project 9-22 considered incorporation of the Mechanistic-Empirical Pavement Design Guide (MEPDG) into a PRS.

In 2000, NCHRP project 9-22 was initiated by the FHWA to address some of the recommendations from the Westrack project. The first objective of the project was to improve the HMAspec software to an implementable level. It found the capability of Westrack PRS software was inadequate for general use across the US. Consequently, the project considered the possibility of incorporating the MEPDG in an HMA PRS system. However, the MEPDG approach was found too complex for implementation. The research team initiated another simple approach which relates the HMA dynamic modulus with pavement distresses using closed form solutions. This final version of the HMA PRS was developed as Quality-Related Specification Software (QRSS). The QRSS is a stand-alone program that calculates predicted HMA pavement performance in terms of both fatigue and permanent deformation using material and construction properties as input variables for both as-designed and as-constructed pavements. The stochastic predictions are implemented using Monte Carlo simulations in order to account for construction variability in all the material and construction properties. The new software also accounts for project-specific climatic, structural, traffic, and desired level of service.

El-Basyouny and Jeong (2010) presented the PRS framework developed as part of NCHRP project 9-22. The Witczak predictive equation is employed to calculate the dynamic modulus using the material and construction variables. Subsequently the concept of effective temperature was used to consider the climatic effects on HMA dynamic modulus and to calculate an effective dynamic modulus. The effective dynamic modulus and several material- and construction-related variables were inputs for closed form solution distress prediction models to calculate distress levels. The estimated distresses were further converted into pavement life. A Monte Carlo simulation was run using the mean and variance of the material and construction variables to obtain the probability distribution of the predicted life for as-designed and as-constructed pavements. Subsequently, the cumulative probability distributions were obtained to calculate the predicted life difference (PLD) for each lot considered. The pay factor is estimated based on such PLD values from each lot and a weighted average of all such pay factors from different lots is calculated based on the individual lot tonnage. Finally, the pay factor is further adjusted according to the initial roughness value of the pavement. Examples for pay factor calculation are shown in Figure 2.1 and Figure 2.2. The boundary pay factor values (minimum and maximum) are to be determined before implementing this framework, which can be agency-specific.

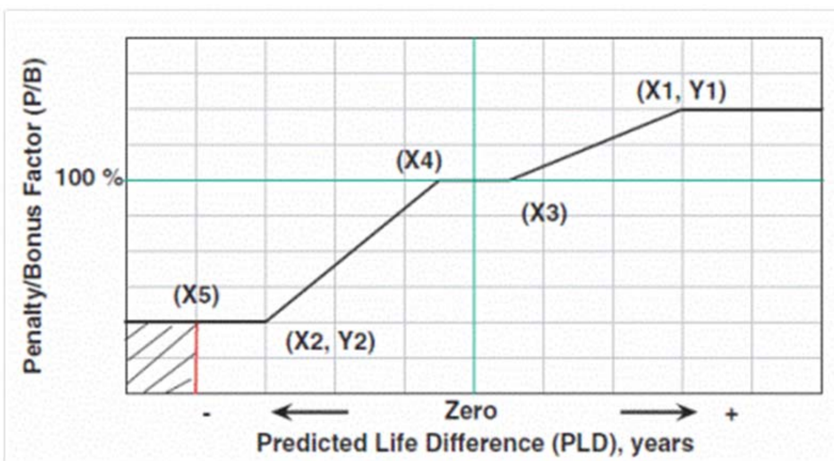

Figure 2.1: Pay adjustment versus PLD relationship for rut and fatigue-cracking (after El-Basyouny and Jeong, 2010)

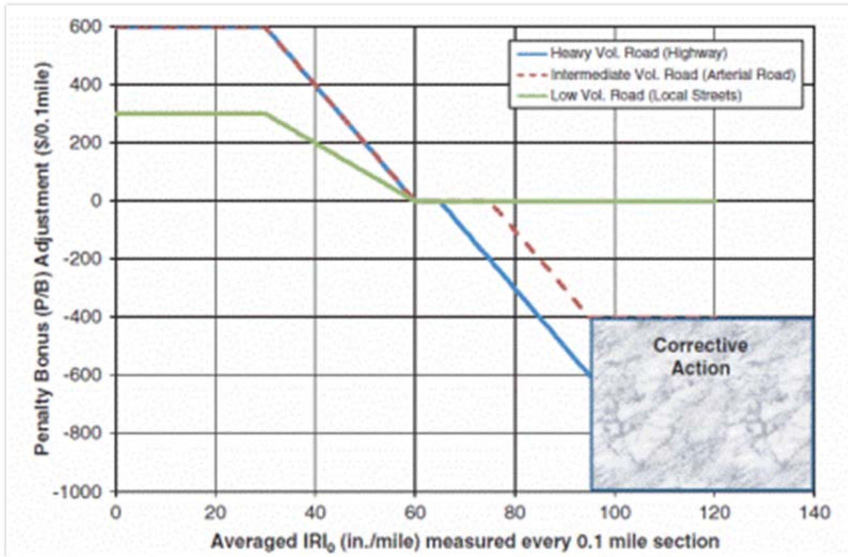

Figure 2.2: Pay adjustment versus PLD relationship for International Roughness Index (after El-Basyouny and Jeong, 2010)

2.2.6 PMIS in PRS—A Wisconsin Study

Choi et al. (2004) highlighted the use of PMIS databases in developing PRS for Wisconsin DOT (WisDOT) projects. They recommended the development of performance models using the PMIS databases to better reflect the local condition (environmental, traffic, etc.) in the performance predictions for a PRS. A brief summary of the PRS development for a specific project in Wisconsin was presented as an example. Figure 2.3 shows various stages in the PRS development process. Initially, the formulation of the following performance models was achieved using linear regression on several material and construction variables (such as air voids, asphalt content, and percent passing the #200 sieve) and age of the pavement with 34 sample data points from 8 different HMA projects in Wisconsin (Equations 2.3 and 2.4):

$$IRI = 5.66 - 5.22P_{200} + 3.82AV + 0.644AC * P_{200} - 0.924AC * AV + 0.426P_{200} * AV + 0.00151Age \quad (2.3)$$

$$PDI = 461 - 137P_{200} + 12.5AC * P_{200} - 17.6AC * AV + 18.0P_{200} * AV - 0.0427Age \quad (2.4)$$

where,
AC = asphalt content measured during construction
P_{200}= percent passing No.200 sieve measured during construction
AV = air voids measured after construction
Age = days after opening to public traffic

The above models were used to predict the performance as a function of age of the pavement for a given set of construction quality data. Subsequently, the performance predictions are converted into life cycles costs using the critical performance levels for WisDOT. A Monte Carlo simulation procedure using Excel spreadsheets was developed to calculate the distributions for the life cycles costs for both as-designed and as-constructed values. Information regarding as-target mean and standard deviation of material and construction variables is collected from the existing specifications and as-constructed material and construction variables from construction QC data. The pay adjustment is developed based on the difference between the LCCs between as-constructed and as-designed values as shown in Figure 2.3. Choi et al. (2004) recommends a separate contract support system for each construction project to implement PRS. Such a system should be based on existing databases consisting of construction data, performance, QC/QA information, traffic data, and environment data.

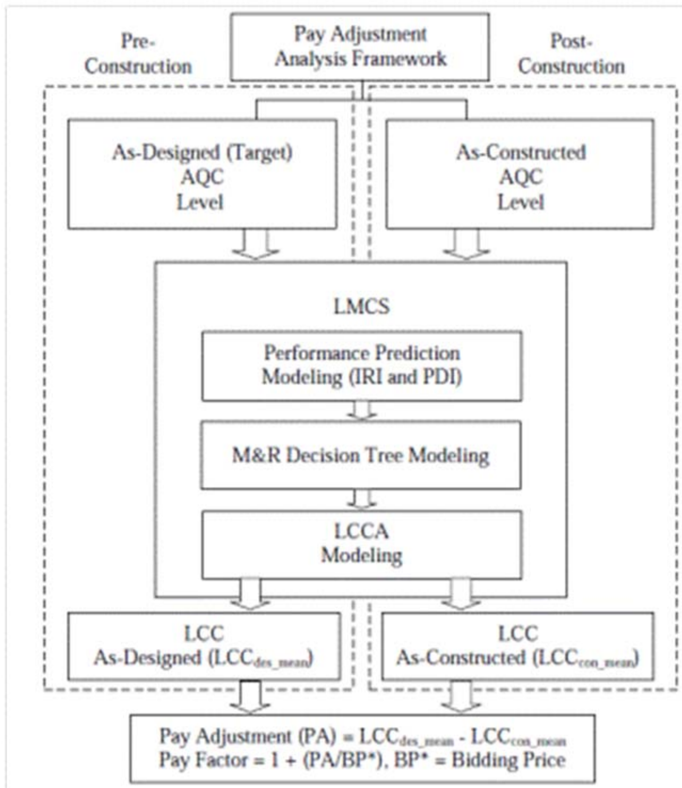

Figure 2.3: Analysis framework for determining pay adjustment at state DOTs (after Choi et al., 2004)

2.2.7 PRS in California

The California DOT (Caltrans) is one of the few agencies actively developing performance-based pay factors. A case study at the University of California (Popescu et al., 2006) demonstrated a procedure for developing performance-based pay adjustment schedules using performance prediction models. These models were developed based on mechanistic-empirical pavement analysis and SHRP-developed laboratory test data. The procedure is applicable for different types of HMA mixes with adjustments for fine graded mixes. Popescu et al. (2006) mentioned that air void content, binder content, and aggregate gradation were important for rutting performance, while air void content, binder content, and thickness of the HMA layer were significant for fatigue cracking. Popescu et al. (2006) also reported that the existing CAL-ME (an empirical and mechanistic pavement design guide for Caltrans) can be used for the calculation of the ESALs for failure. The ESALs required for failure is calculated corresponding to each random combination of material and construction variables in a Monte Carlo simulation process. Consequently, a distribution for ESALs is obtained for each distress mode from Monte Carlo simulation methods. Distributions are established for both as-constructed and as-designed conditions. A ratio between mean ESALs is calculated as Relative Performance (RP). The least RP value is selected among different RP values obtained for various distress modes. Finally, cost models are used to calculate the pay adjustment for a given combination of the material and construction variables. An Excel spreadsheet-based "pay factor calculator" is also attached in a report by Popescu et al. (2006) for the benefit of Caltrans. They recommend the necessary steps to implement the PRS in daily contracting operations by initiating pilot and shadow projects. A comparison between experience-based pay factors from more than 80 projects from 1997 to 2000 and calculated performance-based pay factors using material and construction data collected from these projects is presented in Figure 2.4.

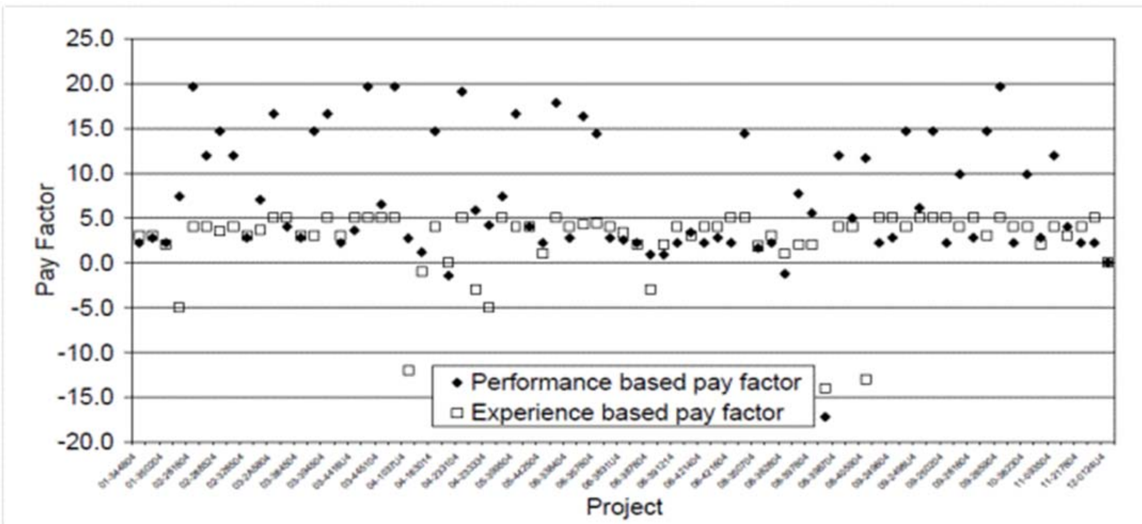

Figure 2.4: Comparison of performance-based and experience-based pay factors for a set of QC/QA projects in Caltrans during the period 1997–2000 (Monismith et al., 2004)

It was found that pay factors based on rutting performance were lower than those of the experience-based approach used by Caltrans. On the other hand, the pay factors calculated based on cracking performance were generally higher than those of the experience-based approach. Monismith et al. (2004) also mentioned that it is important to link material- and performance-related databases for a gradual transition from the experience-based pay factor calculation. They also highlighted the importance of the PRS in revealing the relative importance of several material and construction variables for both agency and contractors, which may enhance the awareness among contractors and agencies to improve the uniformity in important quality characteristics with respect to performance.

2.2.8 Transition to PRS

Buttlar et al. (1998) described the efforts in Illinois to develop end-result and performance-based specifications. They indicate that the transition from experience-based method specifications to performance-based specifications requires considerable time. A few steps are also recommended for any highway agency in order to foster such transitions:

1. Make an initial move to statistical QC/QA.

2. Develop comprehensive end result specifications to consider all relevant quality characteristics.
3. Monitor and support the development of primary and secondary prediction relationships.
4. Develop performance-related pay factors.
5. Compare the performance-related pay factors with the end result based on statistical pay factors.
6. Periodically repeat steps 3, 4, and 5 for a gradual transition from a statistical-based approach to the performance-based approach.

Buttlar et al. (1998) also described a pilot field study at Edgewood, Illinois, for enhancing end-result specifications. Continued monitoring of this project is expected to produce important inputs for the development of performance specifications in Illinois.

2.3 Production and Placement PAF

2.3.1 Current TxDOT Production and Placement PAF

PAF calculation is divided into two components: the production PAF and the placement PAF. The production PAF is based on the laboratory-molded density using the engineer's test results. PAFs are determined for each sublot using the average absolute deviation (AAD) from the target laboratory-molded density. A lot refers to the quantity of HMA representing one day's production, typically about 2,000 tons. The number of lots varies depending on the size of the project. The plant material is sampled from each sublot (typically 500 tons) during a normal day's production and the lab-molded density is determined for each sublot. The final production PAF for completed lots is the average of the PAFs from the four sublots sampled within that lot.

The placement PAF is based on in-place air voids using the engineer's test results. It is determined for each sublot and requires in-place air void measurement. The placement PAF for completed lots is the average of the placement PAFs for each of the four sublots within that lot. The production and placement PAFs for dense-grade mixes as provided in TxDOT's specifications are illustrated in Figure 2.5 and Figure 2.6. Similar PAFs exist for other specification items.

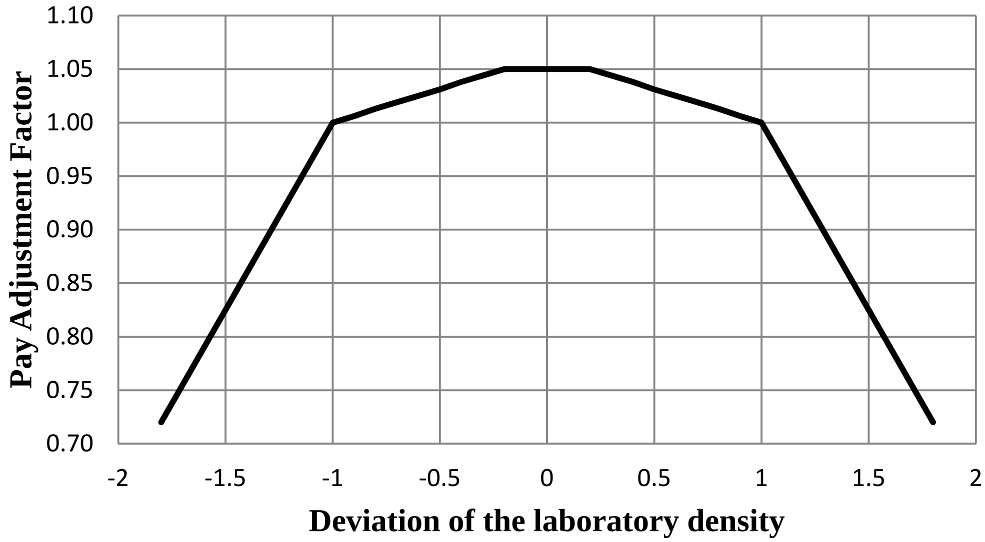

Figure 2.5: Production PAFs for Item 341 (TxDOT, 2004)

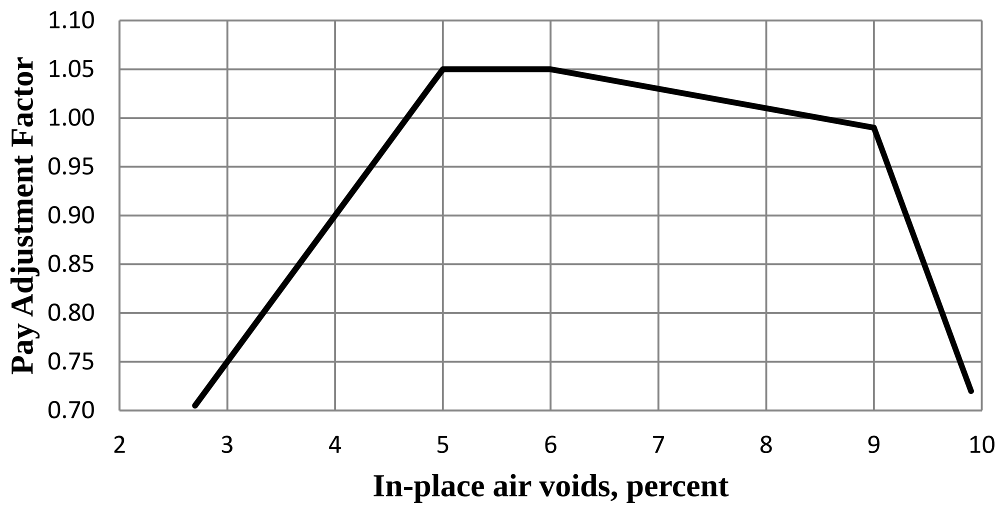

Figure 2.6: Placement PAFs for Item 341 (TxDOT, 2004)

If the production or placement PAFs for three consecutive lots falls below 1, production is suspended until further test results confirm to the satisfaction of the Engineer that the next material produced or placed will result in pay factors of at least 1. The total adjustment pay (TAP) is based on the applicable PAFs for production and placement for each lot calculated as follows in Equation 2.5:

$$TAP = (A + B)/2 \tag{2.5}$$

where:

A = Bid price × production lot quantity × average PAF for the production lot
B = Bid price × placement lot quantity × average PAF for the placement lot + (bid price × miscellaneous quantities × 1.000)

2.3.2 Current State of the Practice of Other State DOTs

In a research study conducted for Caltrans, Deacon et al. (1997) developed an efficient PAF to award or penalize contractors in the construction phase of new flexible pavement construction in which the performance model recommended was based on a mix analysis and design system. The PAF used was based on fatigue distress and specifically took into account the means and variances of asphalt content, air-void content, and asphalt-concrete thickness. Although the previous cost model presented to Caltrans in the summer of 1996 considered the time of first rehabilitation, the overlay thickness, and the timing of second and subsequent rehabilitation activity based on the AASHTO performance equations, the new cost model presented accounted for only the time to the next rehabilitation activity. In addition, Deacon et al. (1997) pointed out that a penalty it should be equal to the full amount of added cost to agency. In contrast, if a bonus is given, it should be some fraction (such as 50% as suggested) of the full added benefit to the agency. Deacon et al. (1997) recommended that a desirable pay schedule should incorporate average and standard-deviation based on field measurements, and be as simple as possible. In this effect, increments of 5% in the bonuses/penalties were considered appropriate.

Another study in California (Popescu et al., 2006) recommends using performance models for fatigue and rutting to determine the pay factors for asphalt and concrete pavement construction. It is also suggested that a fixed weighting scheme be implemented to take into account the relative effect of the different parameters that influence the performance of the HMA. It should be noted that currently Caltrans uses a fixed weighing factor of 0.3, 0.4, and 0.3 for the asphalt content, air void percentage, and aggregate gradation controls, respectively. This method provides for a full bonus for superior construction and a full penalty for inferior construction. In order to compare the proposed method and the current system, Popescu suggests that the agency select a series of QC/QA construction projects and determine pay factors by both the current procedure and the proposed performance-based approach towards evaluating the efficacy of implementing a performance-based approach.

The Washington DOT's (WSDOT) specification uses a variable sampling plan to measure in-place density, asphalt content, and aggregate gradation. Payment is determined by calculating a PWL, then a series of parabolic pay equations are applied, depending on sample size. The PWL methods used by WSDOT balances the risk between the contractor and the WSDOT well; yet there is a risk that a random sample will not be representative of the material as a whole, and will thus result in an incorrect estimate of material quality. This could mean a quality product could be rejected as unsatisfactory or vice versa. This results in two issues that differ from similar specifications: 1) the expected pay for material produced at acceptable quality level (AQL) is greater than 1.0, and 2) WSDOT's AQL is 95 PWL, but contractors seem to consistently produce material near 90 PWL (Mahoney et al. 2001). To help compensate for poor quality work or to reward superior work, a pay factor is used. Pay factors relate quality of work to actual pay. Put simply, a pay factor is a multiple applied to the contract price of a particular item. In general, most plans apply a pay factor to the contract price based on the calculated quality (expressed as percent defective [PD] or PWL) of a particular quality characteristic. Theoretically, material produced at AQLs receives a pay factor of 1, material produced at a rejectable quality level (RQL) is rejected, material produced between AQL and RQL receives a pay factor less than 1, and material that performs in excess of AQL receives a pay factor greater than 1. Pay factors usually range from a high between 1 and 1.12 down to a low between 0.5 and 0.75 (Mahoney and Backus, 2000). These simple relationships do not always hold true for the following two reasons: 1) expected pay is different than contractual pay, and 2) material produced at AQL may not receive a 1 pay factor. Figure 2.7 shows the difference between the specified or contractual pay factor and the expected pay factor for the WSDOT specification.

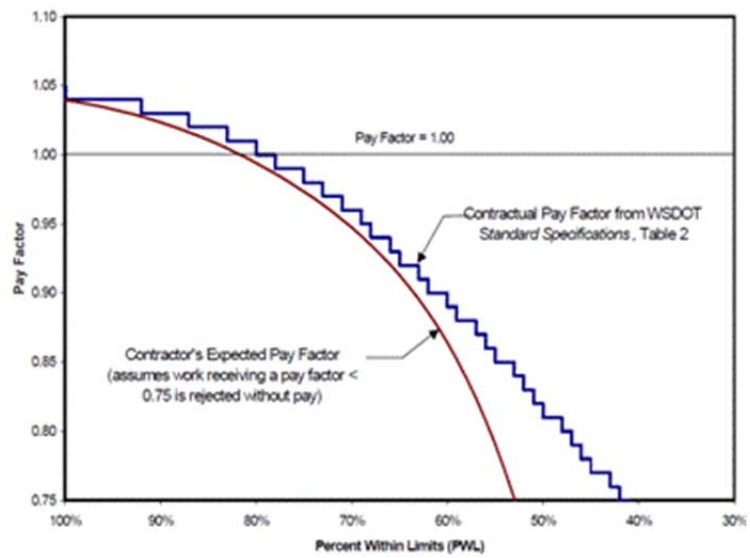

Figure 2.7: Difference between the specified, or contractual, pay factor and the expected pay factor for WSDOT specification (after Mahoney et al., 2001)

Kim et al. (2006) developed a methodology for the determination of price-reduction factors (PRF) for density deficient asphalt concrete mixes in North Carolina. The study included two different mixes with nominal maximum aggregate sizes of 19 mm and 9.5 mm. The samples were compacted using the Superpave gyratory compactor and tested for fatigue cracking and permanent deformation. The authors used the indirect tension (IDT) fatigue test to determine the fatigue life of the mix. The plastic deformation of the mix was determined on the triaxial repeated load permanent deformation test. The authors also evaluated the rutting and fatigue performance of the mix using the Model Mobile Load Simulator (MMLS3), which is essentially a unidirectional vehicle load simulator. The laboratory results were used to determine the PRFs based on the number of cycles to failure for each of the two distress mechanisms. The regression analysis of the laboratory data resulted in the following models (Equations 2.6 and 2.7):

For fatigue cracking:

$$PRF = -0.215 \times (\%AV) + 2.72 \tag{2.6}$$

For rutting:

$$PRF = 2446.8 \times e^{(-0.9753 \times (\%AV))} \tag{2.7}$$

where:

%AV represents the Percentage of Air Voids

Kim et al. (2006) recommend using the PRFs provided in Table 2.2 for density-deficient asphalt mixes.

The authors highlighted the finding that the PRFs that were determined for each of the two distress mechanisms were not affected by the testing program (IDT versus MMLS3), but they differed for each of the two distress mechanisms under consideration. It was also noted that the PRF determined for fatigue cracking was in general in good agreement with current North Carolina DOT guidelines. However, the PRF determined for rutting was in general much lower than existing standards. The authors also recommended that the PRF shown in Table 2.2 should be considered as the lower limit, as they are of the opinion that the deficiency in the density may only be applicable for a certain depth in the entire pavement structure. Therefore, adopting the recommended values may overly penalize the contractors.

Table 2.2: Recommended PRF values for different air void contents (Kim et al., 2006)

Price Reduction Factor				
	Indirect Tension Fatigue Test	**Triaxial Repeated Load Permanent Deformation Test**	**MMLS3**	
% AV	Fatigue	Rutting	Fatigue	Rutting
8	1.00	1.00	1.00	1.00
9	0.83	0.31	0.79	0.38
10	0.68	0.10	0.57	0.14
11	0.55	0.03	0.36	0.05

Based on a survey by Russell et al. (2001), Table 2.3 lists attributes from pay factors from 40 state specifications. Table 2.4 provides several linear pay factor equations from these specifications, including the test property, sample size, and RQL assigned to the equations. Based on their analysis of the survey data, Russell et al. (2001) presented three pay factor equations shown in Table 2.5, classified by severity of financial penalty: lenient, moderate, and severe.

Table 2.3: Pay adjustment attributes for 40 states (after Russell et al., 2001)

Attribute	Number of States Specifying
Type of Adjustment	
Factor	36
Fixed Rate	4
Bonus	21[a]
Aggregate Gradation Sieve Sizes	
12.5mm (1/2")	15
9.5mm (3/8")	15
4.75mm (#4)	17
2.36mm (#8)	18
2.07mm (#10)	10
1.18mm (#16)	7
600 m (#30)	10
450 m (#40)	10
300 m (#50)	10
75 m (#200)	25
Asphalt and Mixture Properties	
Asphalt Content	31
Air Voids	16
Voids in Mineral Aggregate	7
Stability	2
Voids Filled with Asphalt	1
Asphalt Penetration	1
Anti-Strip Additive	1
Moisture Content	1
Theoretical Maximum Density	1
Density	
Percent Theoretical Maximum Density	29
Percent Test Strip Density	6
Percent Laboratory Maximum Density	4
Smoothness	
Profile Index	16
Rolling Straightedge	1
Profilometer/Mays Meter	1
Method of Combination	
Weighted[b]	25
Minimum[c]	12

[a]Bonus provision is contained within the Factor or Fixed Rate.
[b]Weights summing to 1.0 are multiplied to each property then summed.
[c]Minimum individual pay factor of all measured properties is used.

Table 2.4: Pay factor equations (after Russell et al., 2001)

State	Pay Equation	Test Property	Sample Size, n	RQL, PWL
New Jersey	PF= 102 – 0.2×PD PF = 10 + 1.0×PWL[a]	Density	5	50
New Mexico	PF = 55 + 0.5×PWL	AG, AC, AV, Density	3 (minimum)	60
New York	PF = 21.7 + 0.833×PWL (PWL≥94) PF = 57.8 + 0.499×PWL (PWL<94)	Density	4	5[b]
South Dakota	PF = 55 + 0.5×PWL	AG, AC, AV, VMA, Density	5	60
Vermont	PF = 83 + 0.2×PWL	AV	3 (minimum)	50
Virginia	PF = 55 + 0.5×PWL	AC, AV, VMA	4	40

AG = Aggregate Gradation
AC = Asphalt Content
AV = Air Voids
VMA = Voids in Mineral Aggregate
[a]Equation given as an example in the specification only.
[b]Remove and replace for material PWL < 5.

Table 2.5: Pay factor used in data analysis (after Russell et al., 2001)

Pay Factor (PF) Severity	PF Equation
Lenient	PF = 83 + 0.2 PWL
Moderate	PF = 55 + 0.5 PWL
Severe	PF = 10 + 1.0 PWL

2.3.3 Relationship between Production & Placement Quality and Performance

Smit and Prozzi (2008) demonstrated how it is possible to track the network-level performance of surface mixtures using TxDOT databases. The authors indicate that the traffic load, climatic conditions, the past service life of the mixture, design and construction properties, as well as the underlying structure on which the surface was paved are all contributing factors to the life of the pavement. Their findings include the following:

1. Traffic is associated with increased roughness (the more traffic, the rougher the road).
2. Maintenance activities (in dollars) are also associated with increased roughness. This can be interpreted as follows: the rougher the road, the more TxDOT has to spend on it.
3. The wet-warm region in Texas appears to be the worst in terms of roughness, while central Texas (mixed/moderate environment) was in the best condition.
4. The Interstate Highway system (IH) is maintained in better condition than is the US Highway system, which in turn is better maintained than the State Highway system (SH), followed by the Farm to Market (FM) system.

The authors felt that it was also necessary to expand their analysis to include factors such as binder performance and grade, and some HMA mixture-related properties such as asphalt content, voids in the mineral aggregates (VMA), density, and lift thickness.

Patel (1995) conducted research on the development of fatigue-based payment adjustment factors for use in PRSs for full-depth asphalt pavements. In his research he stated that reduction of structural integrity due to fatigue cracking may exacerbate the occurrence or severity of other distresses such as rutting and thermal cracking. Therefore, the analysis focused primarily on examining the quality characteristics that influence AC fatigue cracking. The quality characteristics that were used were asphalt content, voids content, fines content, and AC thickness. These characteristics were investigated to also influence other distresses, but the correlations were too weak in comparison to fatigue cracking. The suggestion given was to base the PAFs on the deviation from the target nominal value rather than from the target standard deviation. Of the quality characteristics analyzed, surface thickness was determined to have more influence on fatigue cracking than any other characteristic. A maximum pay factor of 105% and a minimum factor of 80% were justified based on changes in fatigue life of ±50%.

Deacon et al. (2000) provided an approach to develop performance models from the results of the Westrack accelerated pavement test program for use in a PRS for asphalt concrete mixes. Flexural fatigue tests and a shear test were conducted to determine fatigue response and rutting characteristics. The results were then used to develop performance models. Based on the performance models, it was determined that aggregate gradation, asphalt content, and air-void content should be considered when determining pay factors for rutting.

Monismith et al. (2000) applied an approach to quantitatively establish penalties and bonuses for asphalt concrete construction using performance models for asphalt concrete obtained from the California Accelerated Pavement Testing Program for fatigue and Westrack for rutting. Specifically for rutting, the system considers the means and variances of asphalt content, air-void content, and aggregate gradation. For fatigue, the means and variances of asphalt content, air-void content, and asphalt thickness are included. Monismith et al. (2000) then calculates the cost by using a cost model that takes into account the agency cost consequences of delaying or accelerating the time to the next surfacing or rehabilitation activity. The resulting pay factor is that associated with the shortest life determined for the two distress modes.

Yu (2005) proposed using the Linear Mixed Effects Model (LMEM), also referred to as the longitudinal model, to predict future conditions of a specific pavement section by a weighted combination of the deterioration trends of the family average and that of the specific pavement. The LMEM's distinctive feature is that the mean response is modeled as a combination of population characteristics that are assumed to be shared by all individuals, and subject-specific effects that are unique to a particular individual. Yu (2005) proposed using pavement conditions such as traffic loading, thickness, climates, and pavement condition prior to the last treatment as models to be used in the LMEM. The results of the LMEM have shown better accuracy of predicting these conditions than other methods previously used. Yu also stated that roughness and distress-based ratings of asphalt were the two most important variables in determining the life of a road; however, he did not specify the important material parameters that affect these variables. Yu stated that prior knowledge of the factors that affect performance is essential to developing reasonable models. Materials, traffic loading, and pavement structure are the major factors that should be considered in prediction models.

Vazquez et al. (2010) wanted to establish the relationship between how the operational tolerances affect the expected performance of HMA, and also to determine the effects of variability in key mix design factors, such as asphalt content, gradation, and density. The authors found that the 1) fatigue life of HMA samples tested decrease with increased tensile strain, 2) fatigue life increases with increased asphalt content, 3) the effect of density in performance was not found to be statistically significant, and 4) permanent deformation under the Hamburg wheel tracking device increased with increased asphalt content. For the tested mixtures, the critical combination influencing performance was high asphalt content with the fine gradation. They indicate that "the additional cost resulting from the PAF could be offset by the additional benefits in terms of extended performance and the savings resulting from minimizing disruption and user's costs, such as the delay costs incurred during maintenance and rehabilitation activities" (Vazquez et al., 2010).

In 2007, TxDOT funded a research project to investigate the effect of variations in key volumetric properties within the tolerance limits on HMA mixture performance (Vazquez et al., 2010). These volumetric properties included binder content, aggregate gradation, and air voids. The effects of these variables were tested on two of the most popular dense graded mixes in Texas: Type C and Type D from a rutting and fatigue cracking perspective. The permanent deformation of the mixtures was evaluated using the Hamburg wheel tracking device while the fatigue performance was evaluated using the four-point bending beam test. In addition, the fracture resistance of the mixture was evaluated using the Texas Overlay tester. Subsequent statistical analyses of the laboratory data showed that the binder content and the filler to binder ratio has a statistically significant effect on the rutting, fatigue, and fracture performance of the mix (Banerjee et al., 2011). The scope of the research project included investigating the effect of variations in the volumetric properties within the tolerance limits on the overall performance of the mix. The sensitivity of the mix to the binder content and filler/binder ratio, as determined in this study, is provided in Table 2.6.

Table 2.6: Summary of performance statistics (Banerjee et al., 2011)

Mixture Type	Optimum Asphalt Content (%)	F/B*	Rut Depth (mm)	Rutting Sensitivity		Fatigue Sensitivity	
				Asphalt Content (%)	F/B	Asphalt Content (%)	F/B
Limestone Type C	4.6	0.75	5.63	330	1.16	0.462	-0.594
Limestone Type D	5.3	0.50	5.58				
Limestone SMA-D	6.2	1.23	9.37				
Gravel Type C	5.6	0.79	4.16				
Gravel Type D	5.9	0.75	6.42				

*Filler-binder ratio

In the authors' opinion, a rational approach towards revising existing PAFs should meet the following requirements at the minimum: it should commend better quality and consistent product so as to promote the design and construction of superior performing mixtures. In this way, the additional cost resulting from the PAFs could be offset by the additional benefits in terms of extended pavement life. The development of performance-based PAFs should be conducted based on the results of laboratory performance tests and then validated with short- and long-term field performance. The mix variables that could be considered for developing PAFs should include density, gradation, and binder content. The findings of this research indicate the important effect of binder content and aggregate gradation on performance; it is therefore suggested that due weights should be assigned to these two variables for calculation of the PAFs. Table 2.7 provides interim recommendations that should be evaluated in conjunction with project-specific conditions and other practical mix production constraints.

Table 2.7: Recommended PAFs based on binder-content and filler-binder ratio (Banerjee et al., 2011)

Recommendation for Adjustments to Performance-Based PAF based on Binder Content				
Deviation from Target	**Relative Effect on Rutting Performance (mm)**	**Effect on Fatigue Life**	**Effect on Fracture Performance (Overlay Tester Cycles)**	**Revised PAF**
+0.6%	1.98	+89.3%	+38	1.00
+0.5%	1.65	+70.2%	+31	1.00
+0.4%	1.32	+53.1%	+25	1.05
+0.3%	0.99	+37.6%	+19	1.10
+0.2%	0.66	+23.7%	+13	1.10
+0.1%	0.33	+11.2%	+6	1.10
+0.0%	0.00	+0.0%	0	1.10
-0.1%	-0.33	-10.1%	-6	1.00
-0.2%	-0.66	-19.2%	-13	0.90
-0.3%	-0.99	-27.3%	-19	0.75
-0.4%	-1.32	-34.7%	-25	0.60
-0.5%	-1.65	-41.3%	-31	0.40
Recommendation for Adjustments to Performance-Based PAF based on Filler-Binder Ratio				
+0.5	0.58	-49.5%	Decreased Fracture Performance*	Remove & Replace
+0.4	0.46	-42.1%	Decreased Fracture Performance*	0.60
+0.3	0.35	-33.7%	Decreased Fracture Performance*	0.75
+0.2	0.23	-23.9%	Decreased Fracture Performance*	0.85
+0.1	0.12	-12.8%	Decreased Fracture Performance*	1.00
+0.0	0.00	0.0%	No effect*	1.10
-0.1	-0.12	+14.7%	Increased Fracture Performance*	1.10
-0.2	-0.23	+31.5%	Increased Fracture Performance*	1.10
-0.3	-0.35	+50.7%	Increased Fracture Performance*	1.05
-0.4	-0.46	+72.8%	Increased Fracture Performance*	1.05
-0.5	-0.58	+98.2%	Increased Fracture Performance*	1.05

*Effect on the fracture performance of the mix could not be calculated as it is related to the absolute value of the filler fraction in the aggregate blend rather than the filler-binder ratio. However, the filler content of the mix has an inverse relationship with the fracture performance which implies that high filler content will have a higher filler-binder ratio provided the binder content is held constant which will eventually compromise with the mixture's fracture properties.

The authors of the study recommended that the final PAF should be assessed based on the combined performance edge the agency can gain from the product that is being delivered by the contractor. It should be noted that a weighted sum of the bonuses based on the binder content, filler-binder ratio, and laboratory compacted density in the ratios of 0.3, 0.3, and 0.4, respectively, will provide a PAF that is geared more towards improved performance than just a consistent end product. A higher weight was provided to the laboratory-compacted density in order to give credit to the current pay adjustment schedule as it has proven to produce consistent and durable mixes for construction and maintenance projects throughout Texas.

Tong (2009) stated that a robust relationship between pavement performance and construction QC information should be established. It was recommended that a new PAF system based on the PWL format should be implemented in Texas. He also stated that performance models should be developed for pavement fatigue and rutting that incorporate quality characteristics such as asphalt content, asphalt mix air voids, and aggregate gradation. Tong (2009) also studied the effect of different volumetric properties on the International Roughness Index (IRI). The study reported that neither laboratory compacted density nor the in-place air voids has a significant effect on IRI. It is known that TxDOT has an aggressive maintenance plan which does not allow severe signs of key distresses to manifest. Thus, higher PAFs were not seen to have any kind of correlation with IRI or rut depths. It should be noted in this context that the performance data being referred to here corresponds to TxDOT's PMIS distress data, which is collected at a network level. It was also reported that current TxDOT PAFs are not based on expected pavement performance but rather on historical data and the contractor's capabilities. Moreover, the lack of any correlation with IRI fails to indicate if higher PAFs resulted in any benefit to TxDOT. Because PWL is widely used in other states and has a well-documented basis, Tong's recommendations included developing a new PAF system based on the statistical acceptance plans and pavement performance models. The author also suggested that the current pay factor calculation (which is assessed on the basis of laboratory-molded density) leads to PAFs of 1.05 for most contractors. It was also noticed that although the in-place air voids tend to vary quite a bit, it was still possible to control the variability within the bonus range.

Figure 2.8 and Figure 2.9 demonstrate that contractors in general consistently earn bonuses in hot mix jobs across Texas, particularly with dense-graded Type C and D mixes. The composite PAF distribution is clearly biased towards the right, which implies that in most cases the contractor will earn a pay bonus. This result was highly unusual considering that the sample size used for this study included more than 30,000 data points. Thus, in Tong's opinion, simply controlling the laboratory-molded density and the in-place air void will not yield an effective payment adjustment system. In fact, Tong has even pointed out that, in extreme cases due to market pressure, the contractors may frequently bid slightly lower on hot mix jobs with the assumption that they will receive bonus pay. Similar observations have also been noticed elsewhere (Mahoney et al., 2001). While this method of competitive bidding has tended to eliminate a substantial portion of the overpayment, it also fails to commend a good quality job. This further underscores the importance of a complete evaluation of the current pay adjustment schedule and an overhaul of the current practice.

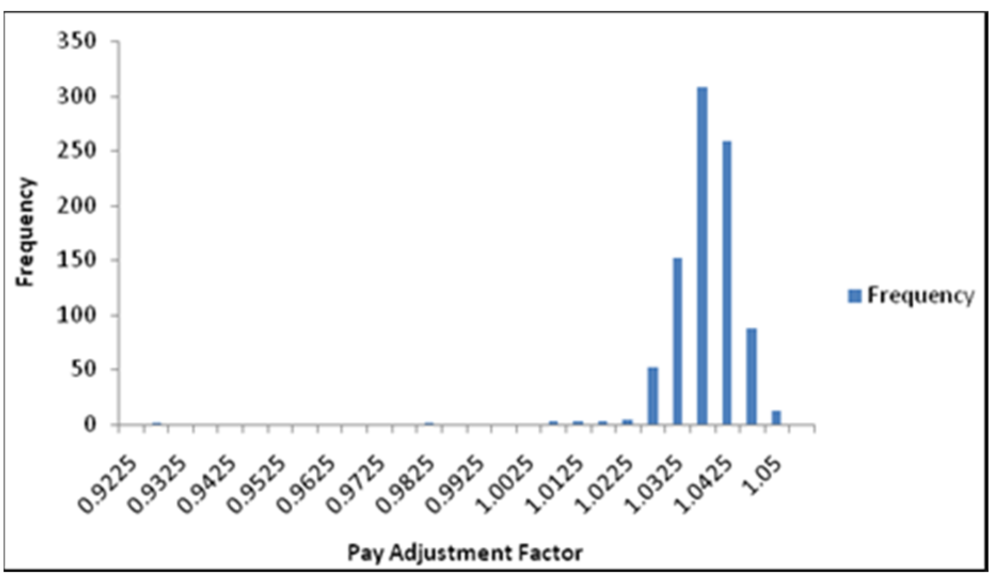

Figure 2.8: Calculated composite PAF for all Type C projects (Tong, 2009)

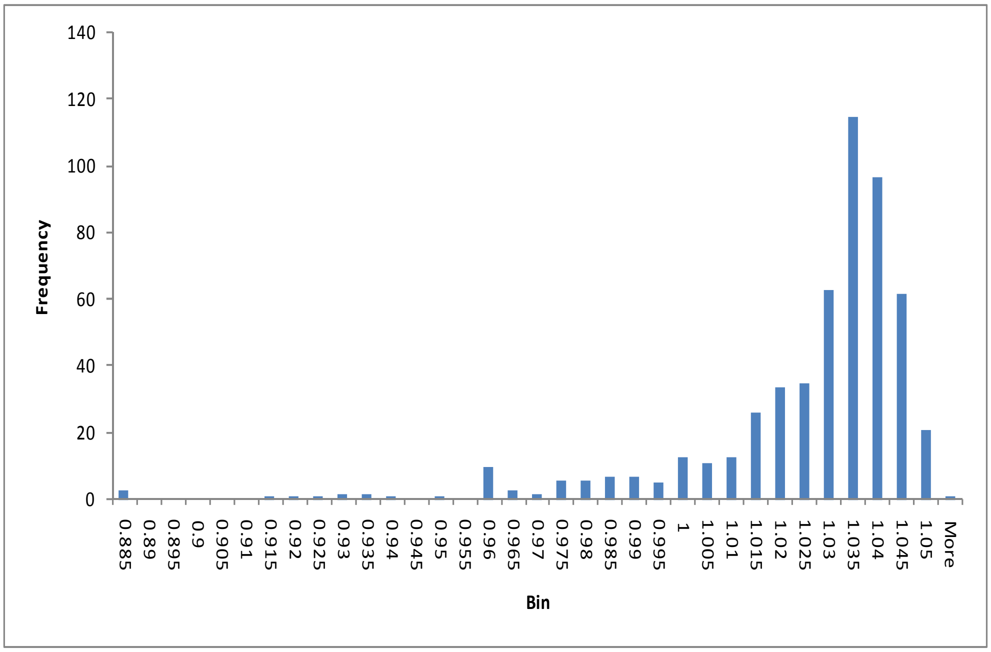

Figure 2.9: Calculated composite PAF for all Type D projects (Tong, 2009)

2.4 Ride Quality PAF

2.4.1 Current TxDOT Ride Quality PAF

TxDOT uses a pay adjustment schedule to commend construction projects based on ride quality. The construction division provides the necessary guidelines for selection of the appropriate pay schedule. The procedure takes note of the existing IRI, facility type, posted speed, the number of smoothness opportunities, and other mitigating factors before identifying the pay adjustment schedule that fits the profile of the specific job (TxDOT, 2004), as shown in Table 2.8.

Table 2.8: Guidance for selecting pay adjustment schedules (after TxDOT, 2004)

<table>
<tr><th colspan="4">Project Description</th><th>Recommended PAS*</th></tr>
<tr><td rowspan="3">New construction or major rehabilitation (IH, US, multilane highways)</td><td rowspan="2">Rigid pavements</td><td colspan="2">Continuously reinforced concrete pavement</td><td>2</td></tr>
<tr><td colspan="2">Jointed concrete pavement</td><td>3</td></tr>
<tr><td colspan="3">Flexible pavements with a total HMA thickness > 1.5”</td><td>1</td></tr>
<tr><td rowspan="7">Overlays or minor rehabilitation</td><td colspan="3">Rigid pavements (bonded and unbonded concrete overlay)</td><td>3</td></tr>
<tr><td colspan="3">Flexible pavements with a total HMA thickness < 1.5”</td><td>3*</td></tr>
<tr><td rowspan="5">Flexible pavements with a total HMA thickness > 1.5”</td><td colspan="2">Posted speed < 45 MPH</td><td>3*</td></tr>
<tr><td rowspan="2">More than 2 smoothness opportunities</td><td>Other than 2-lane undivided highways</td><td>1*</td></tr>
<tr><td>2-lane undivided highways</td><td>2*</td></tr>
<tr><td rowspan="2">1 smoothness opportunity</td><td>Other than 2-lane undivided highways</td><td>2*</td></tr>
<tr><td>2-lane undivided highways</td><td>3*</td></tr>
</table>

*Pay adjustment schedule

TxDOT uses the guidelines provided as part of Item 585 specifications (TxDOT, 2004) to determine the pay adjustment schedule for ride quality requirements for HMA and concrete pavements. A roughness index of less than 60 inches/mile is considered excellent while anything beyond 95 inches/mile will require some kind of corrective action according to TxDOT guidelines. It is also known that the roughness index for a given pavement structure can be reduced approximately by 50% with each lift of HMA until it gets below 60 inches/mile, where it reaches a point of diminishing return. When determining a pay adjustment schedule, the existing condition of the pavement, previous experiences with similar projects, the ability of a contractor to improve the existing ride with the number of smoothness opportunities specified, and the need for higher ride quality should all be taken into consideration. The payment adjustment schedules that are currently in practice are shown in Figure 2.10.

The figure indicates that, in general, the requirements are more stringent for new construction as well as scenarios where there is more than one smoothness opportunity (for example, diamond grinding and placing a single lift of asphaltic concrete). Figure 2.10 demonstrates that the penalty schedule is most severe for Schedule 1, less severe for Schedule 2, and Schedule 3 has no penalty.

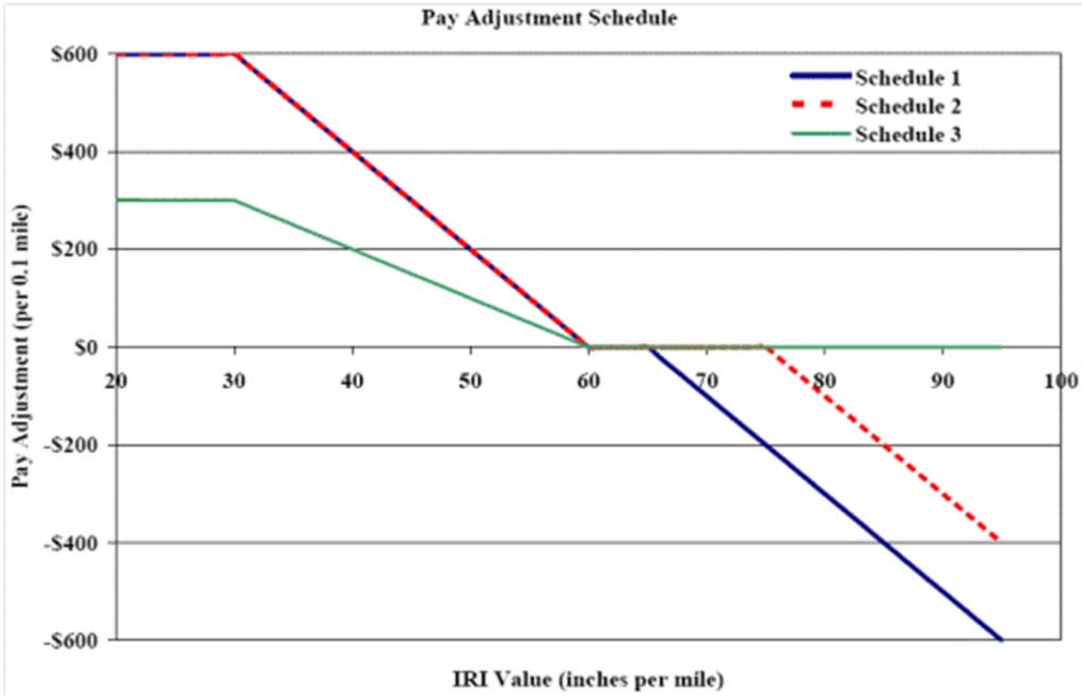

Figure 2.10: Graphical illustration of payment adjustment schedules

2.4.2 Current State of the Practice of Other State DOTs

State Highway Agencies (SHAs) in the US commonly use smoothness specifications to ensure that the public is provided with quality roads. Based on a survey conducted by the FHWA in 1995, smoothness of ride was found to be one of the most important factors in increasing public satisfaction with the highway system. Therefore, to help encourage contractors to build smoother roads, monetary incentives and disincentives are used based on the initial roughness values used by the SHAs. To justify the extra costs associated with smoothness specifications, Ksaibati and Mahmood (2002) aimed to demonstrate that smoother roadways do indeed stay smoother over time. To investigate this, a large number of test sections from the LTPP database were included in the study. After analysis of all the statistical data was preformed, it was found that (1) the IRI values clearly increase over time in a linear fashion, (2) initial roughness values do affect future roughness values of pavements, (3) a pavement section built with a smoother surface will remain smooth with time, and (4) asphalt pavements with high initial IRI show a higher increase in future roughness than the sections with low initial IRI. Table 2.9 and Table 2.10 show the Connecticut DOT pay factor specifications for concrete and asphalt pavements.

Table 2.9: Pay factor of Connecticut DOT for asphalt pavements

IRI (meters per kilometer)	Percent adjustment (PF)
<0.789	10
0.789–0.947	63.29 (0.947 - IRI)
0.948–1.262	0
1.263–1.893	39.68 (1.263 - IRI)
>1.893	-50

Table 2.10: Pay factor of Connecticut DOT for concrete pavements

Profile Index (mm/km)	Percent Paid
0–40	105
41–80	104
81–120	103
121–160	102
161–180	101
181–200	100
200+	grind

Chou and Pellinen (2005) used artificial neural networks to develop time-dependent roughness prediction models for different types of pavements, including PCC pavements, asphalt overlay on cement concrete pavements, and full-depth asphalt pavements. It was mentioned that the best way to determine pay factor limits was to use actual zero blanking band smoothness measurements. The FHWA recommends a terminal IRI value of 170 inches/mile as acceptable for highway systems (FHWA, 1998). The authors considered typical scenarios for Indiana and evaluated the service life of pavements until a terminal IRI of 170 inches/mile was reached. In the process, they also assumed different initial IRI values to account for systematic differences that exist between higher and lower facility types (IHs versus FM roads). Table 2.11 summarizes the service lives obtained from the roughness progression curves.

Table 2.11: Service lives for various initial smoothness limits (Chou and Pellinen, 2005)

Type of Pavement	Pay Factor (%)	Initial Profile Index (inches/mile)	Service Life (yr.)	Life Increase (yr.)	Life Increase (%)
HMA	96	34.66	5.8	-	-
	100	29.71	7.1	1.3	22.4
	102	23.77	8.2	1.1	15.5
	104	13.87	10.4	2.2	26.8
	105	9.90	11.1	0.7	6.7
Overlay	96	34.66	7.2	-	-
	100	29.71	7.6	0.4	5.6
	102	23.77	8.2	0.6	7.9
	104	13.87	9.6	1.4	17.1
	105	9.90	10.5	0.9	9.4

The authors evaluated the relative gain in the service life of a particular facility due to superior construction that resulted in lower initial smoothness limits and higher pay bonuses for the contractors. It was pointed out that the increments in the pay bonuses did not follow a systematic trend. For example, a hike in the pay factor from 0% to 2% did not yield a significant improvement in the service life of a composite pavement as compared to a pay hike from 2 to 4%. Therefore the authors proposed that the smoothness limits for both HMA and PCC pavements should be rather based on the cumulative probability density curves (CDF) of the zero blanking band from the California Profilograph output traces. In the author's opinion, the zone between the 50th and 90th percentiles on the CDF should correspond to the 100% pay factor for

PCC and that for HMA the zone should be between the 75th and 95th percentiles. The revised pay factors recommended as part of this research study are summarized in Table 2.12.

Table 2.12: Revised initial smoothness limits (Chou and Pellinen, 2005)

Pavement Type	Percentile	Initial Profile Index (inches/mile)	Initial IRI (inches/mile)	Pay Factor (%)
HMA	10th	5.94	32.33	104
	75th	16.84	55.78	102
	95th	24.76	72.89	100
	99th	29.71	83.67	96
PCC	10th	9.90	47.54	104
	50th	16.84	63.39	102
	90th	29.71	92.54	100
	95th	36.64	108.39	96

2.5 Expected Pavement Life and Economic Value

The following section discusses research efforts relating the estimation of expected pavement life using the PWL approach.

2.5.1 Estimating Performance Life

One of the first tools used and a major finding in the development of the PRS system was the realization that “the life cycle cost (LCC) of the as constructed pavement can be used as the overall quality characteristic to be controlled” (Darter, 1993). The LCC can then be related to all the distress factors that, in turn, are a function of the various material and construction factors measured during construction. Seeds et al. (1997) related this to the PRS in that the pay adjustment for a given lot can be calculated based on the difference between the as-constructed LCC pavement and the as-designed pavement. Figure 2.11 describes the anticipated process for calculating the LCC for both the as-constructed and as-designed pavements. The remainder of this section is organized into the following three subsections based on the methods used to estimate the performance life using PWL: linear model; polynomial model; and exponential model.

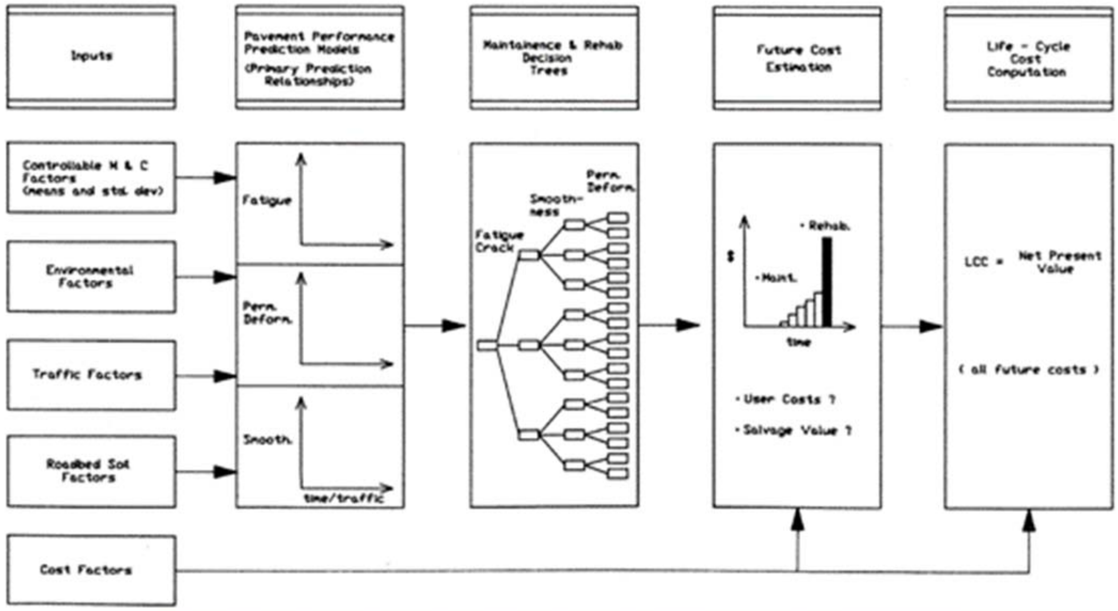

Figure 2.11: Calculation of life-cycle cost for a given lot (Seeds et al., 1997)

Linear Model

Seeds et al. (1997) derived a pay-factor equation by executing a computer model over a wide range of inputs and applying standard statistical procedures, including analysis of variance and regression analysis. The resulting equation took the following linear form in Equation 2.8:

$$\mathrm{PF} = \mathrm{a_0} + \mathrm{a_1} * \mathrm{PD_1} + \cdots + \mathrm{a_i} * \mathrm{PD_i} + \cdots + \mathrm{a_{ij}} * \mathrm{PD_i} * \mathrm{PD_j} + \cdots + \mathrm{error} \tag{2.8}$$

where:

PF = pay factor (ratio of the final payment to the contractor's bid price),
PD_i= percentage defective for each key (significant) material and construction factor,
a_i, a_{ij} = equation coefficients established from regression analysis.

Palise et al., (1998) describe the pay factor system used by the NJDOT. Since the NJDOT has been successful in the past with a composite pay equation for Portland cement concrete pavement based on three quality characteristics simultaneously (thickness, strength, and smoothness), Palise et al. (1998) suggested using similar characteristics for asphalt. Consequently, the three acceptance parameters decided on were in-place air voids, thickness, and smoothness. In order to turn the parameters into a composite pay equation, Palise et al. (1998) decided to compute the pay factor for each quality characteristic and then compute the overall PAY FACTOR for the lot as the arithmetic average of the three individual values of pay factor, given that no individual pay factor is zero. If any pay factor is determined to be zero then the overall pay factor for the lot is zero. The formula for the pay factor lot (PF_{LOT}), which is applied to the bid price of the surface layer, is given by Equation 2.9:

$$\mathrm{PF_{LOT}} = (\mathrm{PF_v} + \mathrm{PF_T} + \mathrm{PF_s})/3 \tag{2.9}$$

where:

PF_v= air-voids PF (Percent)
PF_T= thickness PF (Percent)
PF_s= smoothness PF (Percent)

However, since the additive process allowed a surplus in one quality measure to offset a deficit in another, this could be a potential problem. Although this equation may not be appropriate in all cases, it is not believed to take away from the practicality of this particular approach. In order to illustrate how the equation operates over a wide range of quality levels, Palise (1998) prepared Table 2.13, which shows that the pay factor values calculated using the equation are fairly appropriate for the many different combinations of quality levels for the individual characteristics that may occur in the real world.

Table 2.13: Pay factor values of various combinations of quality levels (Palise, 1998)

PERCENT DEFECTIVE (PD)			PAY FACTOR
AIR VOIDS	THICKNESS	SMOOTHNESS	(PERCENT)
0	0	0.0	102.3
0	10	0.0	102.0
0	0	1.3	101.7
10	0	0.0	101.0
0	10	1.3	101.3
10	10	0.0	100.7
10	0	1.3	100.3
10	10	1.3	100.0
30	30	2.0	98.3
50	30	2.0	89.0
70	30	2.0	79.7
30	50	2.5	78.0
50	50	2.5	68.7
70	50	2.5	59.3
30	70	3.0	57.7
50	70	3.0	48.3
70	70	3.0	39.0
80[a]	70	2.5	45.3
70	80[a]	3.0	34.3
80[a]	80[a]	3.0	29.7
100[a]	50	2.5	0.0[b]
50	100[a]	2.5	0.0[b]
50	50	3.5[a]	0.0[b]

[a]Remove-and-Replace Option; [b]Zero-Pay Provision

Polynomial Model

Weed (2000a) presented a new method for statistical construction specifications based on a single quality measure that is a composite of individual quality measures to make appropriate acceptance decisions. Weed mentioned that previous specifications were based on multiple quality characteristics. These specifications used pay equations that include a separate term for each of the quality characteristics, so that the resultant pay adjustment is a function of the combined effect of all quality measures. The new method was used to accomplish the same purpose and was believed to simplify the procedure and offer several practical advantages. When considering a particular construction item, several different types of decisions must be made and different types of tests must be performed. Thus, to design an acceptance procedure that is fair, effective, and free from inconsistencies, a complex equation is developed, as shown in Table 2.14, which provides an example of a rejection provision for air voids and thickness in HMA.

Table 2.14: Pay factor values of various combinations of quality levels (Palise, 1998)

	QUALITY LEVELS		
CASE	AIR VOIDS	THICKNESS	REJECTABLE?
1	PD = 0 (Excellent)	PD = 75 (RQL)	Yes
2	PD = 75 (RQL)	PD = 0 (Excellent)	Yes
3	PD = 74 (Almost RQL)	PD = 74 (Almost RQL)	No

The RQL for both air voids and thickness has been defined as a PD value of 75 or greater. If either characteristic exhibits this level of quality, the lot may be declared rejectable. But, clearly, an instance or case may arise where the RQL provision is not triggered at all. To allow for this inconsistency, a composite RQL provision can be derived as a joint function of the two quality measures.

According to Weed (2000a), Equation 2.10 and Figure 2.12 illustrate an RQL provision of this type derived from basic performance considerations on the basis of joint quality:

$$1.273PD_{Voids} + 1.2373\ PD_{Thick} - 0.0109\ PD_{Voids}PD_{Thick} \geq 100 \quad (2.10)$$

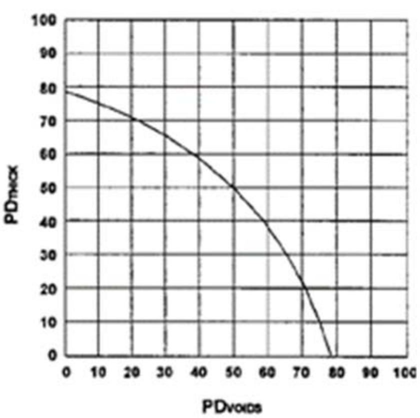

Figure 2.12: Graph of RQL provision (Weed, 2000a)

All points on the curve in Figure 2.12 were judged to be equally detrimental in terms of pavement performance and any combination of PD_VOIDS and PD_THICK that fall on or above the curve would be judged rejectable. While this equation was developed specifically to make an appropriate rejection decision on the basis of joint quality measures of air voids and thickness, upon further development it becomes a useful tool for the retesting and pay adjustment decisions that must also be made. To illustrate how this can be done, Weed (2000a) presented a general performance model in which EXPLIF represents the expected life of the pavement in years, the C terms are coefficients to be determined, and PD voids and PD thick are the measures of air voids and thickness quality, as represented in Equation 2.11.

$$\text{EXPLIF} = C_0 + C_1(PD_{Voids}) + C_2\,(PD_{Thick}) + C_3\,(PD_{Voids}\,PD_{Thick}) \qquad (2.11)$$

To determine the four unknown coefficients in the above equation, four known points that span a wide range of quality are needed. These may be obtained from any valid source, such as agency experience, established performance models, or field experiments, and the values presented in Table 2.15 were obtained from a combination of these sources (Weed, 2000a).

Table 2.15: Performance values used to solve the unknown coefficients (Weed, 2000a)

	EXPECTED LIFE (YEARS)	
	PD_{THICK} = 10	PD_{THICK} = 90
PD_{VOIDS} = 10	20	10
PD_{VOIDS} = 75	10	5

After substituting the values from Table 2.15 into Equation 2.11, and solving the simultaneous equations, the performance model in Equation 2.12 is obtained:

$$\text{EXPLIF} = 22.9 - 0.163(PD_{Voids}) - 0.135(PD_{Thick}) - 0.000961(PD_{Voids}\,PD_{Thick}) \quad (2.12)$$

Since the EXPLIF equation provides a direct estimate of performance as a function of the individual quality measures, it will form the basis for the performance-related pay schedule, and ultimately can be used to develop a single, composite quality measure that will simplify the overall acceptance process and the various decisions that have to be made. Weed (2000a) concluded that the usage of a composite quality measure provided a practical and effective means to make three different types of acceptance decisions: retest, reject, and pay adjustment.

Exponential Model

Weed (2002) proposed a method in which empirical performance data can be combined with logical assumptions about mathematical form and boundary conditions to develop quantitative models sufficiently accurate for use with statistical construction specifications.

Graphical representations for IRI values tend to have an upward curve, which means that the IRI becomes progressively rougher with time, and in some cases this is supported by actual data. One explanation for this could be that the increased roughness will result in a greater dynamic loading because of the bouncing of heavy vehicles, and thereby accelerating the rate of deterioration. Using this information, one such model is given by Weed (2002); in which t is the time (in years) since initial construction and B is the maximum number of years that the service life of the pavement will be extended if the initial IRI is zero (Equation 2.13).

$$\text{IRI} = \text{A}(t + B)^c \qquad (2.13)$$

The appropriate indicator of pavement riding quality was determined to be the IRI. A model can be created using a two-step procedure that describes how IRI typically increases with time, and then that curve is used to derive a model that gives expected life as a function of the initial IRI. Table 2.16 is an excerpt from Swanlund (2000) and provides three separate estimates of the increased service life for HMA pavement as a function of a reduction in roughness.

Table 2.16: Performance values used to solve the unknown coefficients (Weed, 2000 a)

	Average Percent Increase in Service Life of HMA Pavement for Selected Levels of Reduction of Roughness		
	10 Percent	25 Percent	50 Percent
State #1	8	20	39
State #2	3	9	18
State #3	5	11	23
Average	5.3	13.3	26.7

The model decided upon by Weed is presented in Equation 2.14.

$$\text{EXPLIF} = 15e^{-0.3278\,IRI^{1.284}} \tag{2.14}$$

where:
EXPLIF = expected life of HMA overlay (years),
IRI = as-constructed IRI (m/km).

From the above equation, it is clear that expected pavement service life can be given as a function of initial ride quality and can be combined with life-cycle cost analysis principles to compute an appropriate level of pay adjustment.

Weed (2003) synthesized earlier work to recommend an exponential model for performance measures, shown in Equation 2.15:

$$y = Ae^{-Bx^{C}} \tag{2.15}$$

where:
y = performance measure,
x = quality measure, and
A, B, C = constants to be determined.

This exponential model creates a sigmoidal shaped curve that properly recognizes that often a point of diminishing returns occurs both for extreme good quality and for extreme poor quality. Weed (2003) stated that this model ensured a measure of realism that was not always present with other models. Determining points for this model were the maximum, the AQL, and the RQL. Weed further stated (2003) that it was desirable to make *y* the expected life and the variable *x* a quality measure such as PD or PWL. These two quality measures are complements of each other. The multiple-parameter model is shown in Equation 2.16.

$$\text{EXPLIF} = Ae^{-(B_1 PD_1{}^{C_1} + B_2 PD_2{}^{C_2} + \cdots + B_k PD_k{}^{C_k})} \tag{2.16}$$

Weed evaluated these equations by controlling the three quality characteristics and the expected life, thus resulting in a final equation for the expected life (Equation 2.17):

$$\text{EXPLIF} = 13.8e^{-(0.0126PD_{VOIDS} + 0.0107PD_{THICK} + 0.00924PD_{SMOOTH})} \quad (2.17)$$

The coefficients in the above equation indicate that air voids have the largest influence on the expected life, whereas smoothness has the least effect. Obviously sufficient quality in all three characteristics would increase the expected life of the overlay. However, Weed (2003) found that extra quality in one characteristic can offset deficient quality in others. Superior quality in a characteristic cannot mask the extremely poor quality of another. Therefore, it is essential to ensure a balanced quality on all performance related characteristics.

2.5.2 Translating Estimated Performance Life into Economic Value (PAF)

To achieve consistency in the magnitude of pay adjustment, a method is required to relate pavement quality to expected life, which in turn should translate into an economic value (Weed, 1998). This is important to fully recoup the real costs incurred by highway agencies as a result of defective work, as suggested by the AASHTO *Guide for Design of Pavement Structures* (1993). In order to relate quality to a monetary value, it is first necessary to relate quality to expected life. Provided the pavement design procedure is based on some type of fatigue relationship, such as the procedures for either rigid or flexible pavement in the AASHTO *Guide for Design of Pavement Structures* (1993), the same procedure can be used in reverse to estimate the load-carrying capacity from the as-built values measured in the field. To determine pavement life, Weed (1998) derived the following formula (Equation 2.18):

$$L_E = \frac{\ln\{\left[\frac{N_E}{N_D}\right][(1+R_{TRAF})^{LD}-1]+1\}}{\ln(1+R_{TRAF})} \quad (2.18)$$

where:

L_E = expected life (years);
L_D = design life (years);
N_E = expected load-carrying capacity (ESALs) calculated from as-built measurements;
N_D = design loads (ESALs) calculated from design values; and
R_{TRAF} = annual traffic growth rate (decimal).

However, after several tests of the formula, it was found that in most cases the annual traffic growth rate can be assumed to be 0; therefore, as a practical expedient, the traffic growth rate was ignored in the procedure. The degree of stability also suggested that a generic pay-adjustment schedule can be based on PD as the quality measure but that some accommodation may have to be made when applied to pavement layers of substantially different thicknesses. This problem occurred because the thicker layer comprised a greater percentage of the total Structural Number value of the pavement structure so that, when this layer was defective, the detrimental effect on pavement performance was more pronounced. Next, Weed gave formulas for finding the future cost in terms of the present cost, inflation rate, and the present worth in terms of the future cost and interest rate (Equation 2.19 and 2.20):

$$C_n = C_0(1 + R_{INF}/100)^n \quad (2.19)$$

where:

C_n= future cost after n years;
C_0= present cost; and
R_{INF}= annual inflation rate (percent)

$$W_0 = C_n/(1 + R_{INT}/100)^n \tag{2.20}$$

where:
W_0= present worth;
C_n= cost *n* years in the future; and
R_{INT}= annual interest rate (percent).

These two equations were simplified to a single expression, by summing the geometric series containing the R term, because the long-term inflation rate was always less than the long-term interest rate, R < 1. Therefore, the series converged to a finite sum given in Equation 2.21:

$$W_o = \frac{C_o(R^{L_D} - R^{L_E})}{(1 - R^{L_O})} \tag{2.21}$$

where:
W_o= present-worth cost of rescheduling of future overlays;
C_o= current total project cost;
L_D= design life of pavement (years);
L_E= expected life of pavement (years);
L_o= typical expected life of overlay (years);
$R = (1 + \frac{R_{INF}}{100})/(1 + \frac{R_{INT}}{100})$;
R_{INF} = annual inflation rate (percent); and
R_{INT} = annual interest rate (percent).

Weed (1998) concluded that the pay schedule could be expressed in the traditional way as a percentage of the bid price of some component of the paving project or in a possibly more desirable form as dollars per unit area. Table 2.17 was an example of appropriate pay adjustments given based on the pavement thickness quality level.

Table 2.17: Pay adjustment as a function of thickness quality level (Weed, 1998)

PERCENT DEFECTIVE (PD)	CORRESPONDING EXPECTED-LIFE RATIO[a] (L_T / L_0)	APPROPRIATE PAY ADJUSTMENT[b] ($/yd²)	($/m²)	EQUIVALENT PERCENT PAY ADJUSTMENT[c]
5	1.06	+0.99	+1.18	+20
10 (AQL)	1.00	0.0	0.0	0
20	0.94	-1.14	-1.36	-23
30	0.89	-1.94	-2.32	-39
40	0.85	-2.50	-2.99	-50
50	0.82	-3.27	-3.91	-65
60	0.79	-3.86	-4.62	-77
70	0.76	-4.56	-5.45	-91
80	0.72	-5.18	-6.20	-104
90	0.67	-6.24	-7.46	-125

[a] Based on Table 3 results for pavement thickness deficiencies
[b] Computed with Equation 6 using a total project cost of C_0 = $500,000, equivalent unit cost of $15/yd² ($17.94/m²), paving area of 500,000/15 = 33,333 yd² (27,866 m²), L_D = 20 years, L_O = 10 years, and R = (1 + 4/100)/(1 + 8/100) = 0.963
[c] Based on bid price of surface layer of $5/yd² ($5.98/m²)

In order to convert the expected life into a value, Weed (2003) determined Equation 2.22 was appropriate:

$$\text{PAYADJ} = C(R^{DESIGN} - R^{EXPLIF})/(1 - R^{OVLIF}) \qquad (2.22)$$

where:

PAYADJ = appropriate pay adjustment for pavement or overlay
C = present total cost of resurfacing
DESIGN = design life of pavement or overlay
EXPLIF = expected life of pavement or over lay
OVLIF = expected life of successive overlays
R = (1 + INF)/(1 + INT), where INF is the long-term annual inflation rate in decimal for, and INT is the long-term annual interest rate in decimal form. This equation justifies large pay adjustments that reflect real costs to the highway agency when the quality of the resurfacing differs substantially from the design level. It was described that many agencies may choose to have an RQL provision supercede the pay schedule for extremely low values of expected life, which would give an option to remove and replace at the time of construction.

When trying to develop a pay schedule, Weed (2003) stated that bonuses must be limited due to budget limitations and because of the possibility that a pavement might fail to achieve the expected extended life as a result of some condition not accounted for. The solution to this is to have a compound pay schedule to be less harsh on contractors who only marginally deviate from the desired quality, and to ensure large pay reductions for substantially undesirable work. The compound pay schedule resulting consisted of Equations 2.23 and 2.24:

For an EXPLIF < 5 years:

$$\text{PAYADJ} = 2{,}600(\text{EXPLIF}) - 26{,}000 \qquad (2.23)$$

For an EXPLIF > 5 years:

$$PAYADJ = 15{,}400(EXPLIF) - 90{,}000 \quad (2.24)$$

where:

PAYADJ is per lane kilometer.

Table 2.18 provided forms of the pay schedule, yet Weed mentions that these may be slightly altered so long as the intersecting point is at 5 years. This procedure was most effective for quality characteristics that had no correlation between them. Because correlation between the variables did not provide any additional useful information and could greatly complicate the risk analyses.

Table 2.18: Range of values computed with the above equations (Weed, 2003)

INDIVIDUAL QUALITY LEVELS			EXPECTED	PAY ADJUSTMENT[b]
PD_{VOIDS}	PD_{THICK}	PD_{SMOOTH}	LIFE (Years)[a]	($/Lane Kilometer[c])
0	0	0	13.8	+25,484
5	0	5	12.4	+16,517
5	5	5	11.8	+12,527
10	10	10	10.0	0
0	0	45	9.1	−6,590
0	45	0	8.6	−10,349
45	0	0	7.9	−15,732
25	15	30	6.5	−26,935
40	15	30	5.4	−36,162
40	30	30	4.6	−43,117
40	30	55	3.6	−52,112
65	30	55	2.7	−60,502
65	75	55	1.6	−71,152
90	90	90	0.7	−80,202

[a] Computed with Equation 10
[b] Computed using Equation 11 and associated constants
[c] $/Lane Kilometer x 1.61 = $/Lane Mile

2.6 Review of Sampling Methods

The pavement performance compliance rating measure is critical in any pay adjustment system. Russell et al. (2001) presented five different measures that were used to determine specification compliance: (1) average; (2) quality level analysis (i.e., PWL or PD); (3) AAD; (4) moving average, and (5) range. Table 2.19 provides the characteristics of these compliance measures along with supporting equations. Kvasnak et al. (2005) also mentioned that PWL, absolute deviation method, and probability-based approach are widely used in the United States. The PWL algorithm is explained in Figure 2.13.

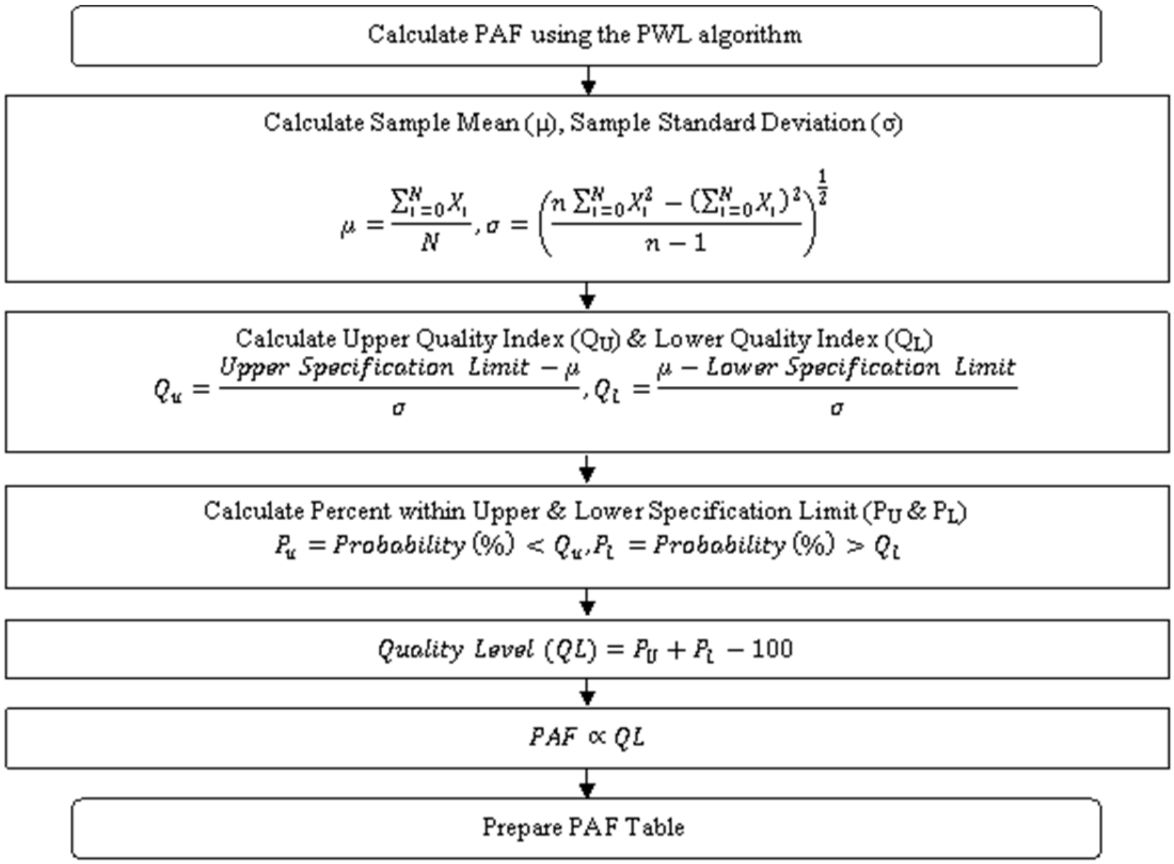

Figure 2.13: Determination of the PAF using the PWL methodology

Table 2.19: Characteristics of compliance measures (adapted from Russell et al., 2001)

Compliance Measure	Definitions and Characteristics	Equation
Sample Mean	• Arithmetic average of tests • Variation must be known. • A confidence interval should be constructed.	$C.I. = z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$ Where, C.I.=Confidence Interval of mean; $z_{\alpha/2}$=standardized statistic; σ^2 = known variance; and n = number of tests.
PWL or PD	• Represents the percent of a normal population that falls either within or outside specification limits; • Incorporates the sample mean and standard deviation. • Quality indexes for the upper and lower specification limits are first calculated (Q_u and Q_l) then applied to statistical tables to determine the estimated PWL or PD.	$Q_U = \frac{(USL - \bar{X})}{s}$ $Q_L = \frac{(\bar{X} - LSL)}{s}$ Where, USL=Upper Specification Limit LSL=Lower Specification Limit $\bar{X}$ = sample mean; and S = sample standard deviation PWL = $Q_u + Q_l - 100$
AAD	• Average of the individual deviations from the target value. • Allows greater cumulative deviations from the target for smaller sample sizes.	$\Delta = \frac{\Sigma\,\lvert X - T\rvert}{n}$ Where, Δ = average absolute deviation; X = individual test result; T = Target value; and n = number of tests
Conformal Index (CI)	• Squares of the individual deviations from the target value. Similar to the standard deviation. • Discourages mid lot process adjustments by not allowing positive and negative deviations to cancel out.	CI = $\sqrt{\frac{\Sigma(X_i - T)^2}{n}}$ Where, CI = Conformal Index; X_i = Individual test result; T = Target value; and n = number of tests
Moving Average	• Measures the arithmetic moving average of several consecutive tests. • Evaluates changes or trends in the moving average relative to target values or specification limits.	$\bar{X} = \frac{\sum X_i}{n}$ Where, $\bar{X}$ = sample mean; X_i = individual test result; and n = number of test.
Range	• Measures the arithmetic range of tests. • Compares the range of values to specification limits, but does not compute the distribution of this range.	Range = Max - Min where, Max = Maximum test value; and Min = Minimum test value.

Kvasnak et al. (2005) stated that the Michigan DOT (MDOT) uses the PWL method while TxDOT uses absolute deviation. The probabilistic approach employs a distribution based on the data collected and then uses this distribution to calculate the probability of meeting the job mix formula (JMF) requirements. Thus, it accounts for variances in the material, collection and location encountered in the analysis. Several agencies have modified their existing systems to account for variations arising due to sampling procedures. However, the Federal Highway Regulations published in 1998 do not provide any specific guidelines related to sampling for QC/QA testing and thus there have been reports of conflicting results obtained by the owner agencies and contractors. The following findings and recombination were made by the authors.

1. AAD presents inherent mathematical inconsistencies that weaken its usefulness as a quality measure. It is variably sensitive to both the shift of the mean away from the target value and the variability of the population itself.
2. The conformal index (CI) is somewhat more consistent than AAD as a quality measure; however, it has a similar weakness in that markedly different combination of mean and standard deviation can produce the same CI value.
3. Although AAD and CI can be made to work for two-sided specifications for which there is a specific target value, they are not well suited for single-sided specifications for which a single, specific target value cannot be defined.
4. The PD/PWL approach works reasonably well for both single-sided and double-sided specifications. It has a similar shortcoming as AAD and CI in that it cannot distinguish between widely different distributions at PD = PWL = 50.
5. The problems could be overcome by basing the pay equation on the mean and standard deviation computed from the sample.
6. These results may make it possible to improve existing highway specifications to make them more effectively oriented toward actual performance.

Burati and Weed (2006) used computer-simulated studies to evaluate the accuracy and precision of the estimators for three quality measures: PWL, AAD, and CI. For each estimator, the variability decreased as the sample size increased. Both PWL and AAD were found to be unbiased estimators, yet CI appeared to be a biased estimator that consistently underestimated the true population CI. PWL has become the preferred measure since it incorporates both the sample mean and the variability in an efficient statistical way. Yet the decision to choose PWL or AAD will need to be based on factors other than the bias and precision of their respective estimators.

Karimi (2009) conducted research for the Maryland State Highway Administration on the durability of HMA pavements in which he evaluated and assessed individual and composite pay factors as well as the risks taken by the agency and contractor. Operating characteristic curves were created to estimate the risks to the agency and contractor. Karimi evaluated the three main quality indicators: PWL, AAD, and CI. When assessing the average differences of simulated lots and actual population values, he concluded that AAD and PWL are both accurate whereas CI showed to be a biased estimator. The recommended quality indicator by AASHTO is the PWL; therefore, Karimi used PWL in his analyses.

Burati et al. (2003) described that if a quality characteristic is to be used for payment determination, the quality measure to be related to the payment must be decided upon. Among the measures evaluate in this study were PWL, PD, AAD, moving average, and CI. They indicate that PWL proves to be a good measure of quality since it is easy to see that the more material within the specification limits, the better quality of the HMA. It was also stated that when evaluating the AAD quality measure, it was found to encourage a contractor to manipulate the lot characteristics in order to achieve a more suitable result. Since the variability of the lot is not measured, the sample means and sample standard deviations vary. For this reason it is not recommended that this quality measure be used for QA acceptance plans. Some agencies have used the moving average. This process is illustrated in Figure 2.14.

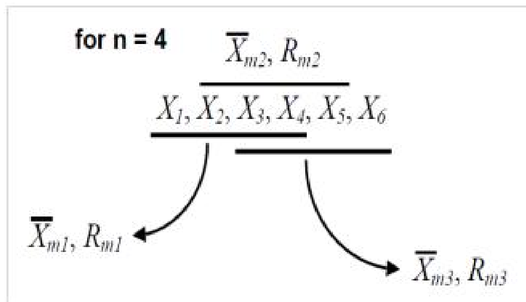

Figure 2.14 Illustration of the moving average (Burati et al. 2003)

The measure is calculated by taking four samples in each lot within the project. The first moving average is calculated from the first four samples. The second moving average then uses its fifth value instead of the first value. For the third moving average, the sixth value replaces the first value in the previous moving average, and so on. Since this quality measure uses moving averages between lots, Burati et al. (2003) recommended it not be used when determining pay adjustments on a lot-by-lot basis.

Burati et al. (2003) recommended that the PWL quality measure be used for QA acceptance plans. Although the PWL measure was chosen, it also has drawbacks. One problem in particular is the fact that a given PWL can represent many different populations. The PWL quality measure was shown to have fewer drawbacks than that of the other measures, and is able to be used with one-sided and two-sided acceptance properties in which the others cannot. It was stated that PD is equally as suitable as PWL for determining acceptance. Among the other measures evaluated, AAD seemed to be the most effective. Yet the drawback of having variability when converted to a pay adjustment yields the quality measure to be inefficient.

Weed (1999) stated that highway agencies have used several different statistical processes to rate quality:

1. Sample mean (x)
2. PD or its complement, PWL
3. AAD
4. CI

It was found that the pay equations based on the mean and the standard deviation calculated from the QA sample can be tailored to closely match the value of the actual constructed product estimated by life-cycle cost techniques. Weed (1999) also discussed the potential problems with existing QA measures, stating that, If the desirable properties of a quality measure were to be enumerated the following would be considered among the most important:

1. Performance related
2. Consistent, and
3. Effectiveness

A performance-related quality measure is one that is strongly correlated with the performance of the constructed product. A consistent quality measure is one that always increases (or always decreases) as the level of performance increases. And finally, an effective quality measure is one that is sufficiently sensitive so that it is capable of detecting significant differences in expected performance (Weed, 1999).

Put simply, in order for a quality measure to be suitable for assurance specifications, it should meet the three basic requirements nearly every time, be unbiased, and finally be relatively precise. Three methods for measuring a desired quality are CI, ADD, and PD or PWL. CI is the root mean square of the departures from the target value; AAD is the average departure (without regard to direction) of the quality parameter from the target value; PD or PWL represents the percent of a normal population that falls either outside or within specification limits, respectively. The CI method is the least popular and hardly used, while the AAD is used by a few agencies. By far, the PD or PWL is the most popular and widely used (Weed, 1999).

From Figure 2.15, Weed (1999) concluded that AAD is determined entirely by the population spread in the former case and entirely by the population shift in the latter case. Even though this may be less evident in larger sample sizes, it still creates doubt as to the consistency of AAD and its use as a quality measure.

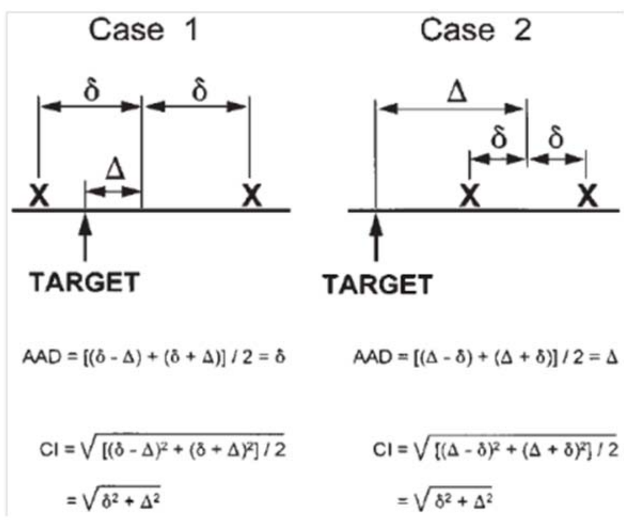

Figure 2.15 Comparison of mathematical properties of AAD and CI for sample size of N = 2. Δ represents population spread within itself. Δ represents shift of population away from target (Weed, 1999)

Weed (1999) described another problem with AAD and CI: both are based on a target value (usually the midpoint between the upper and lower spec. limit), and consequently, are not well suited for a one-sided specification for which a single, specific target value cannot be defined. However, PD/PWL is suitable for both single-sided and double-sided specifications, as illustrated in Figure 2.16. However, PD/PWL has a potential drawback when the PWL value is around 50. When PWL > 50, a decrease in the standard deviation (with no change in the mean) causes the PWL value to increase, but when PWL < 50 just the opposite effect occurs. Even though under normal conditions and with good QC this problem may not be as pronounced, Weed (1999) recommends seeking better methods that are less prone to these shortcomings.

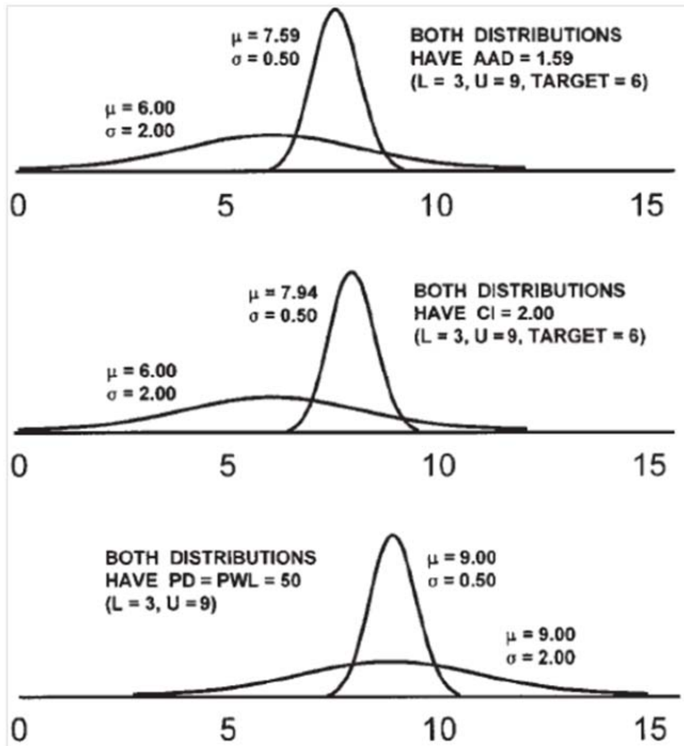

Figure 2.16 Potential weaknesses of common statistical measures of quality (Weed, 1999)

Weed offered an approach based on the mean and the standard deviation individually. This allows greater flexibility to tailor the pay equation to suit a specific performance model. Some general forms of the pay equations are as follows (Equations 2.25–2.27):

Single lower limit: $PF = 100 + A[\{x - B(s) - LIMIT\}/LIMIT]$ (2.25)

Single upper limit: $PF = 100 = A[\{ LIMIT - x - B(s)\} / LIMIT]$ (2.26)

Double Limit: $PF = PF_{max} - A[\{ABS(x - TARGET) = B(s)\} / TARGET]$ (2.27)

where:

PF = pay factor (percent),
PF_{max}= maximum pay factor for double-limit specification (Percent),
A, B = equation coefficients,
X = sample average
s = sample standard deviation,
LIMIT = limit for single-limit specification,
TARGET = target value for double-limit specification,
ABS = absolute value operator.

Weed (1999) indicates that the coefficient B in the above equations is dependent on sample size, and that in Equation 2.27, the positive and negative departures from the target provides greater payments as the sample average moves farther above a single lower limit and as the standard deviation becomes smaller. Highway construction acceptance procedures must be designed to control the characteristics that contribute to a successful performance product. In order to do this, they must properly reward the contractor who conscientiously observes good quality practices, and they must be capable of recouping the economic loss to the highway agency when the failure to follow good practices results in a less-than-satisfactory product.

Burati et al. (2004) completed a report for the FHWA to develop a comprehensive QA manual containing step-by-step procedures and instructions for developing effective and efficient QA specifications. Within this project, Burati et al. (2004) evaluated different quality measures to assist in determining a PAF. During this research, PWL (otherwise known as PD) was determined to be the most effective and efficient quality measure since it combines both the sample mean and standard deviation. Among the other quality measures evaluated were the AAD and the CI. When quality measures were first being introduced in evaluating HMAC, the typical measures were the mean and standard deviation. However, problems arose with the validity of these methods. The mean alone did not consider the variability in the mix.

Quality measures provide an estimate for a true population value, and it is important to select a measure in which the estimator provides an unbiased estimate for the total population, stated by Burati et al. (2004). It is known that if the estimator is unbiased that it is accurate; such is the case with PWL. Among the important factors is that the measure provides for variability where a low variability is represented as precise. A quality index, such as Q, is used to provide an estimate in PWL. The Q index consists of an upper and lower index, demonstrated in Equations 2.28 and 2.29 provided by Burati et al. (2004).

$$Q_L = \frac{\bar{X} - LSL}{s} \quad (2.28)$$

$$Q_U = \frac{USL - \bar{X}}{s} \quad (2.29)$$

where:

Q_L =quality index for the lower specification limit
Q_U =quality index for the upper specification limit
LSL = lower specification limit
USL = upper specification limit
$\bar{X}$ = sample mean for the lot
S = sample standard deviation for the lot

Q_L is used when there is a one-sided lower specification limit, while
Q_U is used when there is a one-sided upper specification limit.

For two-sided specification limits, the PWL value is estimated as shown in Equation 2.30:

$$PWL_T = PWL_U + PWL_L - 100 \quad (2.30)$$

where:

PWL_U= percent below the upper specification limit
PWL_L= percent above the lower specification limit
PWL_T= percent within the upper and lower specification limits

Burati et al. (2004) used a computer simulation to analyze the accuracy of results obtained from the PWL quality measure method. The computer simulation was used to generate samples of various sizes from known populations, from which the PWL was calculated, and then were evaluated to measure the accuracy and precision of this method. It was found that the PWL measure provided an unbiased estimate of the population mean, and it was determined that the amount of variability in the PWL decreased as the sample size increased.

Figure 2.17 provides a representation to prove the fact that as the sample size increases, the variability decreases. Provided in the figure are three different lines: a sample size of 3, 5, and 10. The variability is described by the standard deviation on the y-axis, while the PWL is shown on the x-axis. The difference in variability is best seen at PWL = 50. At this point the height of the curve for a sample size n = 10 has the lowest standard deviation, and therefore the smallest amount of variability.

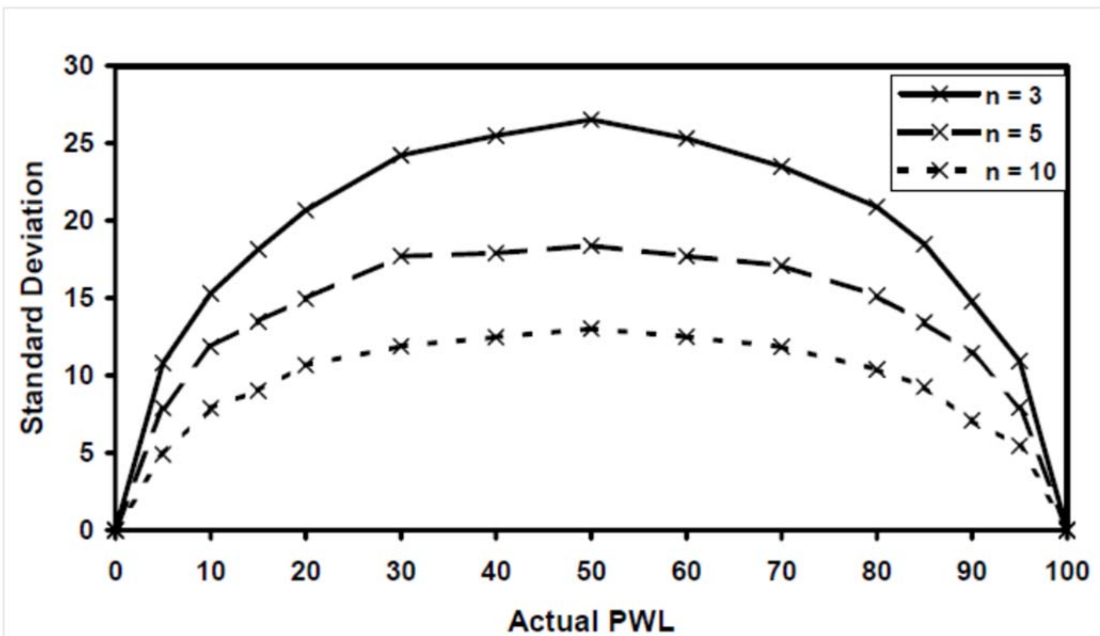

Figure 2.17 Standard deviation vs. actual PWL with different sample size (Burati et al. 2004)

Although widely used in the past, the AAD quality measure has some faults that lead to the assumption that it is not accurate. Burati et al. (2004) showed that this measure allows the contractor to manipulate a lot during production to provide a more suitable deviation. If a target value is defined, as in the case of AAD, the contractor can increase the mean of lots after finding that the previous lots were below the target value. This would result in the average for the entire project to hew more closely to the target value than if the same process was being used throughout the whole project. To account for this manipulation, the absolute value is used. The equation used for the AAD quality measure is as follows in Equation 2.31:

$$AAD = \frac{\sum|X_i - T|}{n} \quad (2.31)$$

where:

X_i= individual test results
T = target value
n = number of test per lot

As with the PWL quality measure, Burati et al. (2004) used a computer simulation program to evaluate the accuracy and precision of AAD values. As the population mean strays from the target value, the AAD value increases. The computer simulation revealed that the AAD measure yields an accurate representation of the population being that it is an unbiased estimator. As with the PWL measure, the deviation from the target decreases as the sample size increases. Although this method was proved to be an unbiased estimator, the fact that the results can be manipulated by the contractor shows that this quality measure is not the ideal method.

The CI quality measure is very similar to the AAD measure, yet the CI uses the squares of the individual deviations from the target instead of the average of the deviations from the target as in the AAD measure. Burati et al. (2004) stated that similar to the AAD measure using absolute value to discourage mid-lot process adjustments, the CI uses the squares to not allow for positive or negative skews. The equation used to calculate the CI is as follows in Equation 2.32:

$$CI = \sqrt{\frac{\sum(X_i - T)^2}{n}} \quad (2.32)$$

where:

X_i = individual test results
T = target value
n = number of tests per lot

Burati et al. (2004) used another computer-simulated evaluation to determine the accuracy and precision of the CI quality measure. This evaluation revealed that the CI quality measure is slightly biased because the average of the CI estimates is always lower than those of the population CI. The bias does, however, decrease as the sample size increases, but the presence of the bias still makes this quality measure unsuitable for use. Since both the PWL and AAD yielded unbiased results and the CI proved to be a biased estimator, the CI was eliminated from the possibility of being an effective and efficient quality measure.

To determine which quality measure, AAD or PWL, was most effective and efficient, Burati et al. (2004) conducted more analysis by comparing the skew, bimodal distribution, and as they relate to PAFs. When comparing the skew and bimodal distribution of each, AAD seemed to function just as well as the PWL measure. Yet when comparing how each measure converts into a pay adjust factor, the PWL measure exceeded the performance of the AAD measure. The reason for this is that the AAD measure yielded a greater payment variability as opposed to PWL, and showed signs of bias once converted to the payment. Also, since the AAD measure does not directly measure the lot variability, a given lot AAD could come from a number of different populations. Burati et al. (2004) revealed that PWL contains fewer drawbacks since it incorporates both the sample mean and standard deviation. Another attribute of the PWL is that it can be used with both one-sided and two-sided acceptance properties, which makes it more versatile than the AAD method, which cannot be applied to one-sided limits. Burati et al. (2004) conclude that the PWL quality measure is the best measure for use in QA specifications based upon its effectiveness and efficiency.

Based on the extensive literature review of the state of the practice on the PAFs as discussed above, most have showed the benefits of using PWL over other compliance measures to estimate expected pavement life.

2.6.1 Sampling Methods and Their Potential Impacts

QC is the contractor's responsibility during construction. QA is ensuring the project is built according to the acceptance criteria established for the contract. These methods can be carried out at the supplier's plant or at the project site. QA starts with the designers and receives constant reviews during project development. This can carry over to design assistance during construction and be combined with value engineering proposals to ensure superior quality of the work. These methods ensure the quality by requiring target values for the in-place product characteristics such as pavement thickness, rideability, strength, and compaction.

The QC/QA processes depend on reliable test results. It is important to note that up to 60% of all variability measured during testing of certain properties can be attributed to sampling and testing reproducibility and repeatability problems and not to the materials and construction practices (Deacon et al. 2001). Taute et al. (2007) pointed out that all necessary precautions should be made to minimize the variations and obtain a true result from the tests carried out.

Taute et al. (2007) indicated some specific areas that should be focused on, such as test location, sample size, lot size, and variation and repeatability. Testing on binder and aggregate should be carried out at the source (i.e., the plant) because larger samples can be used and the binder content does not have to be extracted. In order for the sample to be representative of the property being tested, larger sample sizes are preferred because the potential for error is decreased. It is recommended that coring or slab samples be used when testing binder content at a specific location on the road. Testing done on a lot should be as representative of a uniform lot as possible. If there is visible variation within a lot, then multiple samples should be taken. Taute et al. (2007) state that all sampling and testing of materials should be performed by suitably qualified materials technicians in laboratories that are properly equipped and are suitably accredited. Figure 2.18 provides a graphical description of the processes and quality management.

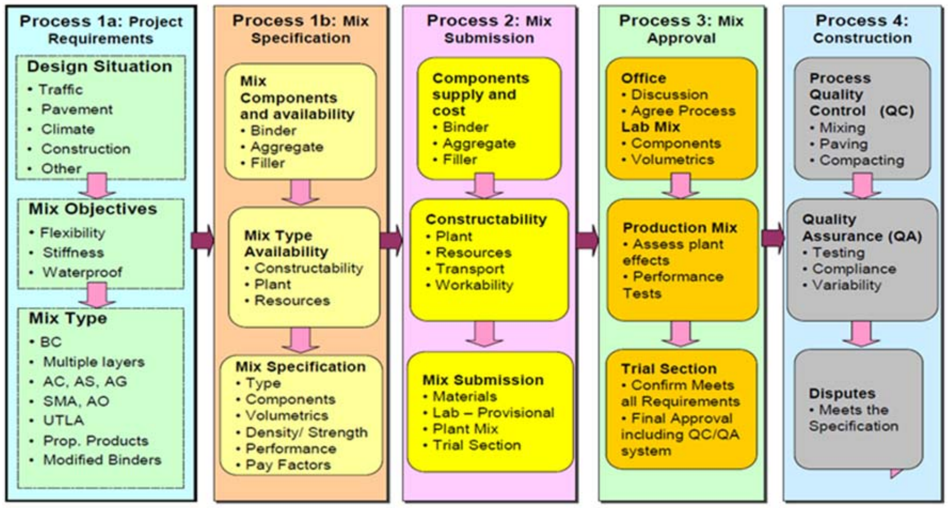

Figure 2.18: Processes and quality management (Taute et al. 2007)

Most US states require the contractor to perform mix designs (i.e., 36 of 42) and provide QC plans (i.e., 34 of 40) (Russel et al. 2001). The QC plan is used to provide the agency and contractor with a list of the tests and inspections that will be conducted during the construction of HMA. Russel et al. (2001) stated that a typical QC plan contains the types and frequencies of tests and inspections, methods for material storage and handling, a list of personnel responsible for various QC functions, and methods to ensure that testing equipment complies with testing standards.

Lot Size

Patel (1996) defined a population (or lot) as a quantity of material that has essentially the same characteristics. Most lots are produced according to production time, most often one per day or one per portion day. Patel (1996) stated that a statistical QC and quality acceptance plan must be based on the evaluation of the population by the use of a random sample. Parker (1995) also discussed that randomized sampling was essential in order to represent the true population and avoid biased results.

Many factors should be taken into consideration when determining lot size, which includes variability of the material/product and economics of consequences of accepting a poor quality product (Patel, 1996). During the process of HMA paving, unavoidable changes of materials, rates of production, or weather conditions may have a large effect on the quality of the material/product from day to day. In terms of the economics of consequences, a large lot could be rejected and the consequences to the producer could be substantial; if a large lot is accepted when it should have been rejected, the costs to repair it in the future are going to be much larger. Therefore, in the author's opinion, it would be best to choose a lot size that best reflects the population while keeping in mind not to make it too large or small.

One major issue in determining sampling levels is the identification of an efficient sample size and the accommodating lot size to gather information on the properties of the HMA within the lot. In order to determine lot size, a sample size must be decided upon. One method for determining lot size is to evaluate the amount of time it takes to take four samples, then basing the sublot size on that time. Another is to determine the lot size for the quantity of asphalt by multiplying the number of samples by the sampling frequency. When determining lot size, four basic considerations must be accommodated: practical, economical, statistical, and equitable constraints. Practical constraints may consist of the amount of time, people, and equipment available to complete the task. All operations within the sampling field require time and money; therefore, the economic efficiency of the different considerations must be evaluated. Statistical issues include the ability to characterize the test results using statistical principles. Lastly, a fair and equitable approach must be taken to provide for the well-being of the agency's and contractor's business. Russel et al. (2001) stated that another factor to take into consideration when determining lot size is whether to have a time-based or quantity-based approach. Time-based lots are determined by setting a number of lots by the period of production; for example 1 per day. Quantity-based lots are determined by specifying a number of lots per amount of tonnage, length, or area of roadway produced. In the quantity-based approach, lots can sometimes overlap between days depending on the tonnage, length, or area specified. States can specify different lot sizes for different characteristics or properties being measured.

Patel (1996) also conducted a review of other states' practices, as shown in Table 2.20. It was stated that lot size and sampling frequency for controlling and accepting paving mixtures are highly variable. He showed that some specify length or area, and others use a lot size based on production or tonnage. The most common lot size was 1 day's production, which was then subdivided into two to five sublots. Control sampling frequencies varied from 1 per 2000 tons to 1 per 500 tons, whereas acceptance sampling frequency is much less. This table also includes the states' specifications on acceptance testing.

Table 2.20: Specifications of mixture lot size and testing frequency (Patel 1996)

Mixture Lot Size and Testing frequency			
State	**Lot Size**	**Contractor Testing Frequency**	**Acceptance Testing frequency**
Alabama	One day's production	Production hours — Sets of Marshal Samples/day 0–3.00 — 1 3.01–6.00 — 2 6.01–9.00 — 3 9.01–12.00 — 4 12.01–15.00 — 5	1 per lot for verification. Uses contractor test for pay adjustment.
Arizona	One shift's production	1 per 500 tons for aggregate gradation, 1 per 100 tons for AC, mix gradation, voids, 10 per compaction lot.	Mixture: 1 per sublot/4 per lot.
Colorado	Based on quality level of moving average of five consecutive values	1 per 500 tons for AC and density and 1 per 1000 tons for gradation under "green" condition	Under "green" condition ac: 1 per 2500 tons Gradation:1 per 3000 tons Density: 1 per 500 tons.
Georgia	One day's production	Not specified	Up to 8 per lot
Indiana	Binder or base: 4000 tons Surface: 2500 tons	1 per sublot/5 per lot (1 per 800 T sublot for binder/base; 1 per 500 T sublot for surface).	1 per sublot/4 per lot (1 per 1000 T sublot for binder/base; 1 per 625 T sublot for surface).
Iowa	One day's production	4 per day or 1 per 750 T whichever is greater	1 per day (used for verification).

Mixture Lot Size and Testing frequency			
State	**Lot Size**	**Contractor Testing Frequency**	**Acceptance Testing frequency**
Kentucky	One day's production	2 per day for AC, gradation and moisture.	2 per day for AC, 6 per day for voids, VMA, per 2 days for gradation (after start-up).
Michigan	Initial production: 2000 tons; next 4000 tons: 4000 tons; thereafter 6000 tons	1 per sublot/5 per lot	1 per day (used for verification).
Mississippi	One day's production for densities and smoothness, mixture properties on individual tests	1 per ½ day for gradation, moisture, voids; 3 per day for AC.	1 per day for gradation: 1 per 3 hr production for AC and voids, 5 per lot for density.
New Jersey	For mixture properties: 3000 tons (600 T sublots); for density 5000 yd^2 (1000 yd^2 sublot); for thickness 15000 yd^2 (5000 yd^2 sublot)	Not specified	1 per sublot/5 per lot for AC, voids, stability, gradation, and density; 15 per lot for thickness.
North Carolina	One day's production	After uniformity has been established: Daily Production (tons) — Samples/day 80–1000 — 1 1001–2500 — 2 2501–4000 — 3 4000+ — 4	Min 10 % of process control testing for verification. Pay adjustments based on contractor's tests. Acc test must be within 2 days of sampling.
Ohio	Density: generally one day's production Mixture: 2000 T	Generally 1 per 4 hrs production.	Verification on split samples, if acceptable use contractors test for acceptance.
Oklahoma	4000 tons or 1 day's production whichever is smaller (sublot≈1000T)	1 per sublot/4 per lot for AC and Gradation: 3 per sublot averaged as one for air voids and stability, and density.	1 per sublot/4 per lot
Texas	One day's production	1 per sublot/4 per lot for AC, gradation, moisture, 2 per sublot averaged as one/4 per lot for air voids	1 per 12 sublot for AC, gradation, moisture; 2 per 12 sublot for air voids
West Virginia	None for moisture; density: 1000 ft; smoothness: 1 mi; and thickness: 2000 ft.	No minimum requirements	1 per 1000 ft for density and smoothness; 5 per lot of 2000 ft averaged as one for thickness.
Wisconsin	Not specifically stated, mix per project	Based on production per day: 20–600 T: 1 per day; 601–1500 T: 2 per day; 1501–2700 T: 3 per day; 2701–4200 T: 4 per day; add 1 per 1500 t thereafter.	Verification on split samples, if acceptable use contractors test for acceptance

Mixture Lot Size and Testing frequency			
State	**Lot Size**	**Contractor Testing Frequency**	**Acceptance Testing frequency**
AASHTO	Mix per project	Temp of mix, mat: 1 per hour; Gradation, AC, Compaction: 1 per 500 T	Verification on split samples (min 10%), if acceptable use contractors test for acceptance
WASHTO	Mix per project	Sand Eq. agg grad, fractured faces: 1 per 1000 T; Mix Gradation, AC, Compaction: 1 per 500 T; Thickness: as needed.	Grad, VTM, VMA, AC: 1 per 1500 T; Compaction: 1 per 500 T.
Demo Project.	One day's production	Sand Eq: 1 per 5000 T; Agg grad: 4 per day; moisture: 2 per day; AC: 1 per 1000 T.	MTD,VTM,VMA,AC: 4 per day; Compaction: 1 per 1000 T.

Sublots and lots are used to accurately distribute samples that represent the entire population of characteristics. The lot is composed of several equal-size sublots, to allow efficient sampling under often changing construction conditions. For example, four 1,000-ton sublots could be used to create one 4,000-ton lot.

Sample Size and Frequency

Patel (1996) stated that two issues govern how the sample size is determined: 1) the amount of risk the buyer and seller are willing to accept when accepting a product and 2) the time and cost of the sampling and testing. When a larger number of samples are taken, the time and cost to take the samples increases significantly. Yet there still must be a median to meet at because increasing the amount of samples reduces the risk to the buyer and seller. Therefore, the agency must balance the risk with the costs of sampling and testing. Table 2.21 shows the sampling locations and frequencies used by the Illinois DOT (IDOT).

Table 2.21: Sampling specifications of IDOT (Patel 1996)

Characteristic		Sampling Location	Frequency
Aggregate	Sand Equivalency	Just prior to mixing with asphalt cement	1 per 5000 tons but no less than 1 per week
	Gradation	Plant cold fee or hot bins	4 per each day of production
	Moisture Content	Plant cold feed belt	2 per each day of produciton
Asphalt Content		Feed line to plant pugmill	1 per 1000 tons but not less than 1 per day of produciton

Table 2.22 shows a comparison of the minimum testing frequencies between IDOT and the suggestions from AASHTO. All of the AASHTO specifications are based on quantity (i.e., 500 tons). IDOT chooses to use time for gradation and asphalt content (i.e., 1 per half-day production), and length for density (i.e., 1 per 800 m).

Table 2.22: Comparison of contractor's QC guidelines (Patel 1996)

Parameter	AASHTO	IDOT
Gradation	1 per 500 tons	1 per half-day production.
Asphalt Content	1 per 500 tons	1 per half-day production.
Density	1 per 500 tons	1 per 800 m (1/2 mi.) for ≤ 75 mm (3 in) 1 per 400 m (1/2 mi.) for > 75 mm (3 in)
Thickness	As needed to control operations	not specified
Smoothness	As needed to control operations	not specified

Russel et al. (2001) conducted a data analysis that found that between-day variation was present when collecting samples. They stated that contractors' samples should be collected and tested within each day of production, instead of high frequency testing on certain days. It was recommended that three (four if time permits) hot mix tests be conducted during each day of construction. They found that this sample size yielded the most reliable estimate for the distribution of differences across a project. The number of samples within a lot can affect risk levels for both the agency and contractor; therefore Russel et al. (2001) recommended that states specify PWL as a compliance measure to evaluate sample sizes for the lot and acceptable levels of risk.

Table 2.23 provides the relative advantages and disadvantages when specifying either time or quantity for sampling size and frequency. A time-based sampling frequency has the advantage of evaluating a given production time, where material and project characteristics may be better understood within time periods. A typical lot size from time-based sampling is 1 per day. With quantity-based sampling, material from different periods is assembled into one lot—the materials from these periods may have different production features. Quantity-based sampling has several advantages over time-based sampling. A given quantity of material can be traced through plant mixing and laydown operations, allowing for a discrete evaluation of the material throughout construction. Mix storage times may be an additional reason for selecting production as a testing frequency denominator, because material in storage for an extended period may not be tested with a time-based sampling frequency. An important characteristic of quantity-based sampling is that both small and large contractors must sample material at the same rate—this produces a testing specification that is fair to all parties.

Table 2.23: Advantages/disadvantages of time- and quantity-based sampling (Russel et al. 2001)

Advantages	Disadvantages
Time	
• Evaluates an isolated period where process conditions can change between periods. • Separates production into time-based sequences where process levels can be understood from a daily perspective. • May be easier to account for work using days, rather than tonnage from multiple days.	• Amount of material in each sublot and lot can vary creating an unequal evaluation of material quantities. • Smaller contractors producing a smaller tonnage will require more testing than larger contractors producing a larger tonnage. • If an insufficient number of samples are collected prior to a work stoppage, it may be difficult to collect the remaining samples in a random manner.
Quantity	
• Same amount of material is used with each acceptance decision. • Testing frequency easily related to sublot and lot size. • Lot and sublot sizes can be managed to ensure samples are collected correctly. • Small and large contractors sample at the same rate.	• Material from multiple days may have different characteristics and confound the statistical properties of the lot. • Combining days with unequal means and variances may provide incorrect statistical estimates. • Specification limits should acknowledge the possibility of added variation from combining material from different days.

The six states selected for the study conducted by Russel et al. (2001) specified both time and quantity for sampling frequency. Table 2.24 provides a summary of the lot and sublot sizes used by the state when sampling for density. The sampling frequency was identified interchangeably with sublot size for all states.

Table 2.24: State specifications for density sampling (Russel et al. 2001)

State	Sampling Method	Sublot Size (Sampling Frequency)	Lot Size
Kentucky	Cores	1 per 2,500 feet	1 day (total varies on length paved)
Arizona	Cores	No specified size	10 per day
Florida	Cores	1 per 300 meters	5 per 1,500 meters
Wisconsin	Nuclear Gauge	No specified size	5 per 750 tons[a] and 7 per 750 tons[c]
Minnesota	Cores	No specified size	2 per 300 tons (lot increases for increased daily tonnage)
Ohio	Cores	No specified size	10 per day
[a] Agency performs acceptance testing. [c] Contractor performs acceptance testing.			

Time-based and quantity-based methods are also used to determine sampling frequency, which are normally correlated with the sublot size as Table 2.24 indicates. It is also common that time and quantity are used together to determine frequencies and sizes. Kentucky, for example, specifies a lot size of one per day and a sublot size/sampling frequency of one per 2,500 feet.

For QC of binder content, Taute et al. (2007) suggested using a sampling frequency of one per 100 tons of asphalt produced, whereas a frequency of one for every 200 ton is suggested for sampling of aggregates. The sampling for QA should be governed by the lot size and could possibly be higher than that of QC sampling depending on the degree of control of mix variables at the plant.

Sampling and Testing Methods

In past decades, various techniques have been suggested to properly collect HMA samples. This includes samples taken either at the plant or on the roadway behind the paver. Testing methods have also been suggested for different HMA attributes. The basis to develop specifications is having a sound random materials sampling and testing procedure. In a study by Buttlar (1998), a lot is defined as one day of mixture production for the control of as-produced quality characteristics, which were then divided into four sublots. For the control of as-constructed quality characteristics, a lot was defined as 800 meters of paved roadway, which was then subdivided into four 200-meter sublots. The first sampling method used for density readings was the existing IDOT's method of taking five transversely-aligned density measurements at one randomly selected location in each lot. The second method consisted of a single density measurement taken at a random location in both the transverse and longitudinal directions within each sublot (200-m section). It was found that the second method gave a more representative assessment of overall lot density. This is illustrated in Figure 2.19 by the lower standard deviation of the moving average.

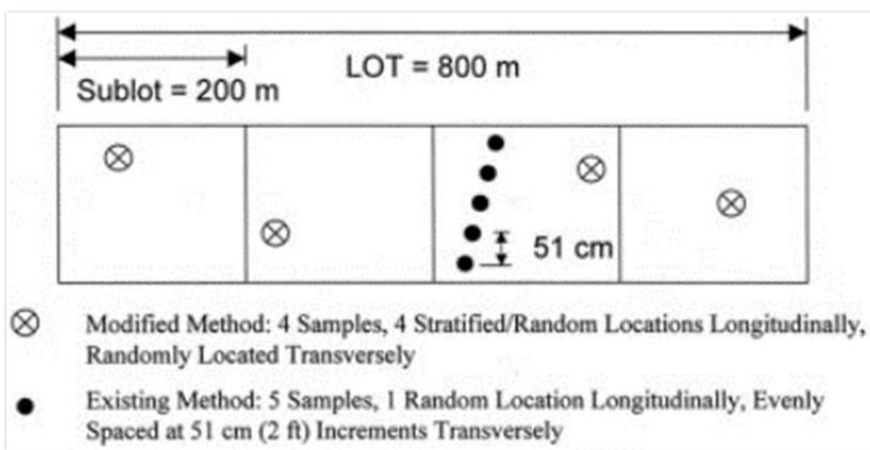

Figure 2.19: Random Sampling Techniques in Sublots (Buttlar 1998)

Most states typically test three different mix properties: aggregate gradation, asphalt content, and mix volumetrics (such as air voids and VMA). Russel et al. (2001) surveyed the 42 states within their study to develop Table 2.25, which shows the attributes specified by these states.

Table 2.25: Mix property acceptance attributes for 40 states (Russell et al. 2001)

Attribute	Number of States Specifying		
	Aggregate Gradation	Asphalt Content	Mix Volumetrics
Sublot Size			
1 per 500 tons to 1 per 900 tons	12	17	9
1 per 1,000 tons to 1 per 2,000 tons	12	10	15
1 per 3 hours	2	6	5
4 samples	1	1	0
Variable	2	2	2
Lot Size			
1 per 500 to 5 per 6,000 tons	13	17	15
1 to 4 per day	7	11	7
4 Sublots	2	2	1
Project	1	1	0
Total per Mix Design	4	4	3
Variable	2	2	2
Cumulative	0	1	1
Continuous	0	1	0
Sampling Location[a]			
Coldfeeds or Hot Bins	17	0	0
Plant Discharge	4	7	4
Truck	15	19	15
Windrow	1	2	1
Volume Analysis	1	2	0
Mat	9	11	9
Asphalt Content Testing Methods[b]			
Extraction	-	20	-
Nuclear Gauge	-	20	-
Ignition Oven	-	18	-
Plant Record	-	11	-
Tank Stickings	-	9	-
Specific Gravity	-	4	-
Compliance Measure[c]			
Quality Level Analysis	13	14	10
Absolute Average Deviation	7	8	7
Moving Average	6	7	6
Average	5	6	2
Range	3	3	3

[a]States may specify multiple locations for aggregate gradation and asphalt content.
[b]States may specify multiple testing options for asphalt content.
[c]One or more compliance measures may be specified within a state (i.e., may vary by property being tested).

As Table 2.25 indicates, most states use tonnage to define sublot sizes which range from one sample per 500 tons to one sample per 2,000 tons. Another way to measure lots is by time. These sublots typically are based on increments of 3 hours. The lot sizes can be quantity-based, which range from one per 500 tons to five per 6,000 tons, or time-based, which typically consists of one per day. The most popular sampling location among the states for mix properties is cold feeds (or hot bins) and the truck box. Cold feeds are used to store asphalt at the plant, and 17 states specify cold feeds as their preferred sampling location for aggregate gradation. On the other hand, 15 states specify the truck as the preferred location for sampling aggregate gradation. The truck box is also the most popular sampling location for asphalt content and mix volumetrics.

Table 2.26 provides the acceptance attributes for density among the 40 states surveyed. Similar to plant-produced mix properties, more of the states are using tonnage for sublots (19 states) and lots (17 states). Sublot sizes range from one test per 80 tons to one test per 1,500 tons. Other specified sublot and lot sizes include length, time, and area. Sublot length varied from 330 to 660 yards, time for sublots and lots ranged from one to five per day, and area included a 2,000 yd² (1,672 m²) sublot and 5,000 yd² (4,200 m²) lot.

Table 2.26: Density acceptance attributes for 40 states (Russell et al. 2001)

Attribute	Number of States Specifying
Sublot Size	
1 per 80 to 1 per 1,500 tons (1,360 Mtons)	19
330 to 660 yards (300 to 600 meters)	5
1 to 5 per day	4
Square yards	1
Square meters	1
Variable	1
None	9
Lot Size	
1 per 400 to 1 per 6,000 tons	17
5 to 10 per day	11
330 to 1760 feet (300 to 1,500 meters)	5
Total Per Mix Design	4
1 per shift	1
Cumulative	1
Variable	1
Sampling Method	
Nuclear Gauge	16
Core	15
Nuclear Gauge corrected to Core[a]	9
Reference[b,c]	
Theoretical Maximum Density	32
Laboratory Maximum Density	9
Test Strip	8
Compliance Measure[b,d]	
Quality Level Analysis	20
Average	8
Range	4
Moving Average	3
Average Absolute Deviation	3

[a]Number of cores for correcting nuclear readings ranged from 3 to 12.
[b]When the total number is less than 40, this means that some agencies did not provide a response.
[c]States may specify multiple options as a density reference.
[d]States may specify multiple options for compliance.

Table 2.27, adapted from Russel et al. (2001), lists the acceptance attributes for pavement smoothness for 26 states. Sampling methods observed consisted of the most popular California Profilograph (12 states), regular straightedge (6 states), profilometer (3 states), rolling straightedge (2 states), and Mays ride meter (1 state). The California Profilograph is a very similar method to that of the rolling straightedge. It consists of a metal frame on wheels that is rolled on the surface while a wheel in the center measures bumps in the road.

Table 2.27: Pavement smoothness acceptance attributes for 40 states (Russel et al. 2001)

Attribute	Number of States Specifying[a]
Sublot Size	
1 per 0.1 mile	9
1 per 100 feet (100 meters)	3
1 per day	1
10 per day	1
1 per 200 tons	1
Section	1
Lot Size	
Total Project	12
1 per day	3
1 per mile	2
1 per 2,500 feet	1
1 per 160 feet	1
1 per 1,000 tons	1
Sampling Method[b]	
California Profilograph	12
Regular Straightedge	6
Profilometer	3
Rolling Straightedge	2
Mays Ride Meter	1
Compliance Measures	
Profile Index	16
Surface Variation	6
Quality Level Analysis	1

[a]When the total number is less than 26, this means that some agencies did not provide a response.
[b]States may use multiple methods for sampling.

Table 2.28 demonstrates that different methods are used to sample pavement density. Included in these are core samples (15 states), nuclear density readings (16 states), and correcting the nuclear density readings to core samples (10 states). Table 2.29 lists the common methods used to measure pavement smoothness in 26 states.

Table 2.30 and Table 2.31 show examples of states specifying different attributes discussed above for acceptance. It is worth noting, however, that the listing in the tables is not exhaustive.

Table 2.28: Sampling locations for mix property acceptance

Attribute	Aggregate Gradation	Asphalt Content	Mix Volumetrics
Sampling Locations			
Cold Feeds or Hot Bins	KS, NC, OR, CO, ID, IL, IA, MD, MI, MT, NE, NJ, NV, ND, OH, PA, WY	-	-
Plant Discharge	Not found	MD, MI, NE, ND, PA, KS	NC, OR, MD, MI, NE, ND, PA
Truck	LA, ME, MN, VA, WA, WI, LA, MS, NC, OH	AL, CT, FL, GA, HI, IL, LA, KY, ME, MN, MS, NV, NJ, NY, NC, OH, OR, RI, SC, TN, TX, VA, WA, WV, WI, WY	FL, IA, KY, LA, ME, MN, OH, OR, VA, WI, MN, MS, NC, KS
Windrow	MN	MN	MN
Volume Analysis	Not found	Not found	Not found
Mat	NV, KS, AK, AR, IA, AZ, GA, PA, SD	CT, KS, MT, NV, KS, AK, AR, CO, IA, MD, GA, AZ, ID, IN, SD, TN	CT, KS, NE, NJ, AR, AZ, MD, PA, GA, SD, TN

Table 2.29: Asphalt content testing methods for mix property acceptance

Asphalt Content Testing Methods	Examples of states specifying
Extraction	KS, AZ, CA, CT, DE, FL, GA, HI, ID, IN, LA, MD, MA, MI, MN, NE, NH, NJ, NY, NC, OK, PA, RI, SC, TN, TX, UT, VT, VA, WI
Nuclear Gauge	AL, AK, AR, CA, CO, HI, IL, KY, ME, MD, MI, MS, MO, NV, OH, OR, TX, UT, VT, WA, WV
Ignition Oven	UT, AL, IN, AZ, MO, WI, MD, ME, GA
Plant Record	IA, KS, MO, NM, ND, SD, WY
Tank Stickings	Not found
Specific Gravity	Not found

Table 2.30: Compliance measures for mix property acceptance

Attribute	Aggregate Gradation	Asphalt Content	Mix Volumetrics
Compliance Measure[a]			
Quality Level Analysis (PD/PWL)	AK, CA, FL, KS, ME, NY, ID, IL, MD, OK, OR, WA, SD, NM, IN, DE, CO	AK, CA, FL, ME, NY, ID, IL, MD, MO, OR, WA, SD, VA, NM, IN, DE	FL, KS, NJ, PA, CT, LA, ME, MO, NY, ID, SD, VT, VA, NM, AZ, IN
AAD	GA, NC	AL, GA, NC, KY	AL, NC, KY
Moving Average	Not found	Not found	Not found
Average	Not found	Not found	Not found
Range	Not found	Not found	Not found

[a]One or more compliance measures may be specified within a state (i.e., may vary by property being tested).

Table 2.31: Sampling methods and compliance measures for density acceptance

Attribute	Examples of States Specifying
Sampling Method for Density	
Nuclear Gauge	AL, AZ, DE, KS, MD, ME, MA, MI, NH, NC, PA, WI, VA, WA, WV, WY
Core	IN, MD, IA, FL, AK, AR, CT, FL, GA, HI, KY, MO, MA, MN, NH, NJ, NM, ND, OH, OK, UT, VT
Nuclear Gauge Corrected to Core[a]	MN, MT, MS, NE, OR, TX
Compliance Measure[b]	
Quality Level Analysis (PWL/PD)	CA, CO, CT, FL, LA, ME, IN, KS, NJ, NY, NM, SD, DE, WA, AK, SC, MO, OK
Average	not found
Range	not found
Moving Average	not found
AAD	AL, NC

[a]Number of cores for correcting nuclear readings ranged from 3 to 12.
[b]States may specify multiple options for compliance.

Two methods are available to verify test results between the contractor and agency: (1) split-sampling and (2) independent-sampling. Many states are using split-samples for verification (i.e., 29 of 42), as Table 2.32 indicates, and a near-equal number are using independent samples (i.e., 20 of 42) (Russel et al. 2001). The benefit of split-sampling is the prevention of unnecessary project effects when comparing tests, such as the effects of sampling from different locations within a truck box or mat and sampling at different times during production. Independent-sampling allows the contractor and agency to conduct the sampling process independent of each other. Russel et al. (2001) recommended that the agency perform split-sample testing, rather than independent-sample testing for agency verification of contractor acceptance test. A split sample test is an experimental method in which a sample is divided into random sub-samples which are treated differently. In this case a sample is divided into two samples in which one is tested by the contractor and the other by the agency, which helps to ensure accuracy. Split-sample verification testing reduces the number of comparison tests and removes unnecessary project effects during the verification, such as materials, production, and sampling variation. Table 2.32 shows the amount of agencies that require a particular sampling attribute.

Table 2.32: Agency requirements for sampling attributes (Russel et al. 2001)

Attribute	Yes	No	Total Responding[a]
Contractor Requirements			
Technician certification required	36	6	42
Contractor provides mix design	36	6	42
Contractor QC Plan required	34	6	40
Project Resources			
Required time and cost determine testing levels	5	30	35
Staffing determines testing levels	11	31	42
Waive testing for small tonnage (< 500 tons)	22	8	30
Adjust testing levels during production	10	31	41
Acceptance Testing			
Contractor tests used for acceptance	27	15	42
Split-samples used for verification	29	13	42
Independent-samples used for verification	20	22	42
Pay adjustments	39	2	41
Dispute Resolution System	33	9	42

[a]When the total number responding is less than 42, this means that some agencies did not provide a response.

Mahoney et al. (2001) stated that WSDOT's specification uses a variable sampling plan to measure in-place density, asphalt content, and aggregate gradation. Acceptance sampling is just one broad form of an acceptance procedure used to decide whether work should be accepted, rejected, or accepted at a reduced payment. This method of acceptance/rejection relies heavily on random sampling to draw conclusions about a large amount of material or lot. If samples are not random, then the statistical basis for evaluating them and drawing conclusions about an entire lot is invalid. HMA construction acceptance sampling uses a modified version of random sampling that satisfies the random sampling assumption. In this version, in order to avoid sample clustering, each lot is divided into several equal-sized sublots in which each are randomly sampled. When sampling for in-place density, WSDOT divides its lots into five equal-sized sublots and takes one random sample from each sublot. Three basic rules must be followed when using this version:

1. The same number of samples is taken from each sublot.
2. Sublots are of equal size.
3. Samples are selected randomly from within sublots.

The two basic types of acceptance sampling are (1) attribute sampling and (2) variable sampling. Both attribute and variable sampling are used in HMA construction; however, variable sampling is more prevalent (Mahoney et al. 2001). Attribute sampling uses a basic pass/fail method rather than taking an actual measurement of the sample. An inspection of the sample is compared with a set standard, which then receives a pass or fail (accept or reject). An example of this would be WSDOT's asphalt concrete aggregate fracture test. In variable sampling, the quality characteristic being analyzed is actually measured, which in turn retains more information per sample than that of attribute sampling. For this reason, most HMA statistical acceptance plans use variable sampling. Acceptance sampling is a powerful audit tool because it allows reasonably accurate estimates of lot quality to be made based on test results from a relatively small number of random samples within the lot (Mahoney et al. 2001).

Martin (2003) stated that currently TxDOT samples and approves asphalt materials at the source. These materials are then used in highway projects without consideration of possible changes in properties that may occur between production and use during construction. Historic concern and limited recent data indicate that binder properties do change, after production contributing to construction and operation difficulties as well as poor performance. In Martin's study, the current TxDOT QA program for binders was evaluated and recommended revisions toward improving quality were made.

The initial strategy for evaluating the TxDOT binder QA program was to validate and further examine differences in properties between corresponding supplier and field samples and identify factors responsible for these changes (Martin, 2003). Difficulties in obtaining corresponding samples due to poor sample identification and lack of an easily accessible database resulted in an alternative approach, and an extensive laboratory experiment relying on supplier samples and simulation of storage conditions and contamination was designed. The researchers demonstrated the use of cluster analysis or CART as a methodology to analyze binder data to identify suppliers with a historical record of specification compliance or noncompliance by product. This type of analysis shows promise for estimating field sampling frequencies by supplier/product combination to reduce resource requirements to a reasonable level within current budget limits.

Martin (2003) recommended some changes that need to be done to the current binder QA program:

1. Manager shall be appointed and all employees need to be trained on the revised QA program.
2. The binder QA program by TxDOT is to be used as a tool.
3. Data shall be stored in a user friendly database.
4. Data need to be organized and analyzed frequently to detect problems.
5. Asphalt cements and emulsion residue should be uniformly tested using performance-related parameters.

Kvasnak et al. (2005) also reported that sampling represents a key component of a QC/QA program, especially one that involves determination of the PAF. Historical evidence showed that the choice of sampling methodologies can lead to conflicting results between the contractor and the agency. In recent times, this has contributed to the growth in popularity of split-sample testing, which typically involves preparing three different samples from a single field sample, each tested individually by different parties including the contractor, the agency, and a third neutral party for dispute mitigation. The authors also highlighted instances where they found significant differences in the material quality characteristic, each sampled using different techniques. For example, they observed that the measured asphalt binder content was most precise with the plate sampling method. On the other hand, the plate and rings method produced repeatable samples. The plate and shovel method was found to be ideal for low traffic levels. The authors also reported that the measured pay factor system was more precise than the actual pay factor system when using a probabilistic approach. It is therefore essential to recognize the importance of the role played by the sampling technique as it will govern the accuracy of the measurements that are recorded on the sample.

Karimi (2009) discusses a number of issues pertaining to sampling with regards contractor vs. agency data, plant vs. behind the paver data, and impact of sample size. The study found from statistical *F* and *t* analyses that samples gathered by the contractor contained significantly different results than those of the DOT's, which indicates that the data are not from the same population. The possibility of defining transfer functions between mix parameters using the QA and QC data was examined but it proved impossible to develop acceptable relationships. A major difficulty in conducting the QA and QC data analysis was to pair the observations from material in the plant (QC) and behind the paver (QA). Thus, a better material identification and tracking techniques is recommended.

Sampling Location

Parker (1995) stated that sampling for mat density should only take place on the roadway, which would also be the ideal location after compaction for other mix properties because it will most accurately reflect pavement quality. Although the roadway location for sampling is recommended, there are still some advantages to sample mix at the plant, such as ease of sampling loose mix instead of compacted mat, ease of preparing test specimens from hot loose mix instead of cold compacted cores or slabs, and proximity of sampling points to laboratories. Sampling at the roadway requires more resources: time, money, personnel, and transportation.

Russel et al. (2001) stated that a typical QC plan contains the types and frequencies of tests and inspections, methods for material storage and handling, a list of personnel responsible for various QC functions, and methods to ensure that testing equipment complies with testing standards. Within each sublot, samples are chosen using the randomized sampling method. Most states require the contractor to perform mix designs (i.e., 36 of 42) and provide QC plans (i.e., 34 of 40). The QC plan is used to provide the agency and contractor with a list of the tests and inspections that will be conducted during the construction of HMA.

Russel et al. (2001) provides a list of advantages and disadvantages in Table 2.33, indicating the effect of sampling location.

Table 2.33: Advantages/disadvantages between sampling locations (Russel et al. 2001)

Attribute	Plant	Laydown Operation
Material Characteristics	• Material may segregate within the truck box at the sample face and introduce bias. • Sample may not represent AC absorption found immediately before compaction.	• Segregation possible in paver hopper or on mat. • Sample has time to absorb AC and is considered to be more representative for compaction.
Resources	• Requires a reduced amount of resources. • Less time for overall sampling cycle.	• Requires more resources (technician time and vehicle to transport sample). • More time for overall sampling cycle.
Interaction of Construction Operations and Sampling	• Possibility of changing plant settings immediately before or after sampling (over-conscientious or over-eagerness to sample), creating dependency between plant operations and the sampling process. • Selecting a "more representative" truck in "trying to be fair" can introduce bias.	• Trucking and paving operations are independent of sampling (bias of sampling process is minimized). • Removes (1) opportunity to change process immediately before or after sampling and (2) any dependency between plant operations and sampling process.
Safety	• Climbing into the truck box to obtain a representative sample presents a safety problem. • Unsafe sampling environment may influence the ability to obtain a representative sample.	• Improved safety because sampling is made at ground level. • Safety concerns exist from moving equipment and projects paved under traffic

MDOT reported that the sampling of the material is as important as everything else associated with the determination of PAFs awarded to the contractors (MDOT, 2002). MDOT recommends plate and shovel sampling from behind pavers prior to compaction as they believe it provides a more representative sample. MDOT uses asphalt content, air void percentage, maximum theoretical density, and the VMAs to determine the pay factor for a specific construction job. However, MDOT uses the sample mean and standard deviation to compare against the JMF characteristic targets to calculate the incentives or penalties that are handed over to the contractors.

According to AASHTO T-168-03, samples for acceptance testing must be obtained from the roadway behind the lay down machine prior to compaction when sampling for volumetric tests, and after compaction for determining compacted density. Russel et al. (2001) recommends the following with regards sampling locations, frequencies, and sizes:

- Choose a location that provides the best opportunity to collect a representative sample of the work. Consider the four sampling attributes (material characteristics, resources, interaction of construction operations and sampling, and safety) and other state-specific attributes during this decision.

- Hot mix acceptance samples should be collected at the laydown operation to provide the most representative sample.
- Perform all sampling in strict accordance with randomization principles to obtain unbiased statistical estimates.
- Agencies should either collect their own samples or witness the contractor collect and split the sample and have the agency field representative immediately take possession of their portion of the split-sample.
- Review the relative advantages and disadvantages of time-based and quantity-based sampling.
- Determine sampling frequency using estimated times to complete each test.
    - Time-based: sampling rate, longest testing time
    - Units=samples-per-hour
    - Quantity-based: sampling rate, maximum production rate
    - Units=samples-per-tons (convert to area or length as desired)
- Collect samples with the presumption they will be tested to comply with randomization principles. If not all samples are tested, remove them in accordance with randomization principles.

Chapter 3. Database Development and Data Description

In order to better address the main research objectives, this study required the establishment of a data warehouse that will contain the PAFs related to mixture production and placement as well as the ride quality at the project level. Much of this information is already routinely collected by TxDOT and stored in various databases. A deeper understanding of the available data fields in each database is essential for linking multiple databases and extracting the required information. This chapter describes a methodology to gather project-level information from various data sources available within TxDOT and familiarizes the reader with the data warehouse developed by the research team by including a comprehensive discussion of the databases employed and a database integration procedure. The newly developed data warehouse, a project-level database, comprises the following information:

1. Type of project and location;
2. Production and placement bonuses/penalties awarded to the contractor;
3. Bonuses/penalties awarded to contractors based on ride quality;
4. Performance history in terms of key distress mechanisms that include distress, ride and condition score, surface rutting, cracking and direct roughness measures. This information is collected every year from the time of construction or rehabilitation; and
5. Volumetric properties of the surface mixture on the project including air voids, binder content, gradation, VMAs, voids filled with asphalt (VFAs), and other relevant variables as identified in the literature review.

The latter part of this chapter includes a detailed discussion on the distributions of the project-level volumetric properties, ride quality, and awarded bonuses and penalties. This exercise provides an insight into contractor tendencies and the variability of the hot mix production, placement, and ride quality across Texas.

3.1 TxDOT Databases

As part of the data extraction process, TxDOT's Design and Construction Information System (DCIS), SiteManager (SM), and PMIS databases were employed. Letting and budget-related information such as quantities and pricing of highway projects is generally included in the DCIS database. The SM database consists of asphalt mixture properties such as air voids, binder content, maximum theoretical density, gradation, ride quality measurements immediately after construction, and other relevant routinely collected information as part of QC/QA. It also contains information on the placement, production, and ride-quality-related bonuses/penalties awarded to the contractor. TxDOT collects network-level performance measures covering most of its highway centerline mileage across the state annually, which are stored in PMIS. The importance of linking these databases for developing performance-based pay factors was highlighted by the researchers as part of a previous TxDOT research project (Smit, 2005). For the current study, a database was populated by integrating the above three databases with project-level information, including production, placement, and ride quality pay factors, QC/QA data (for example, volumetric properties of the asphalt mixtures), and performance measures from PMIS. A comprehensive description of each database is given below. Each database consists of multiple tables including various data fields. The definitions and the importance of the individual fields are discussed in detail.

The Texas Cartographic Information Technology System (TxCIT) database, which was created as part of a TxDOT inter-agency program, provided the framework for the development of this study's data warehouse. TxCIT establishes a link between the SM and PMIS databases by using Texas Reference Marker (TRM) information obtained from DCIS and a geographical TRM database developed by TxDOT, as shown in Figure 3.1. TRM information missing from DCIS must be identified using a manual procedure and TxCIT provides a web-based geographical procedure whereby mapped project coordinates are used to extract TRM information as illustrated in Figure 3.2.

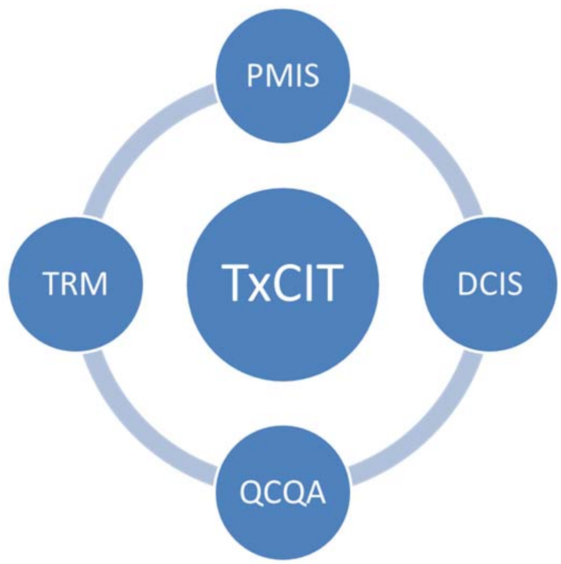

Figure 3.1: TxCIT database framework

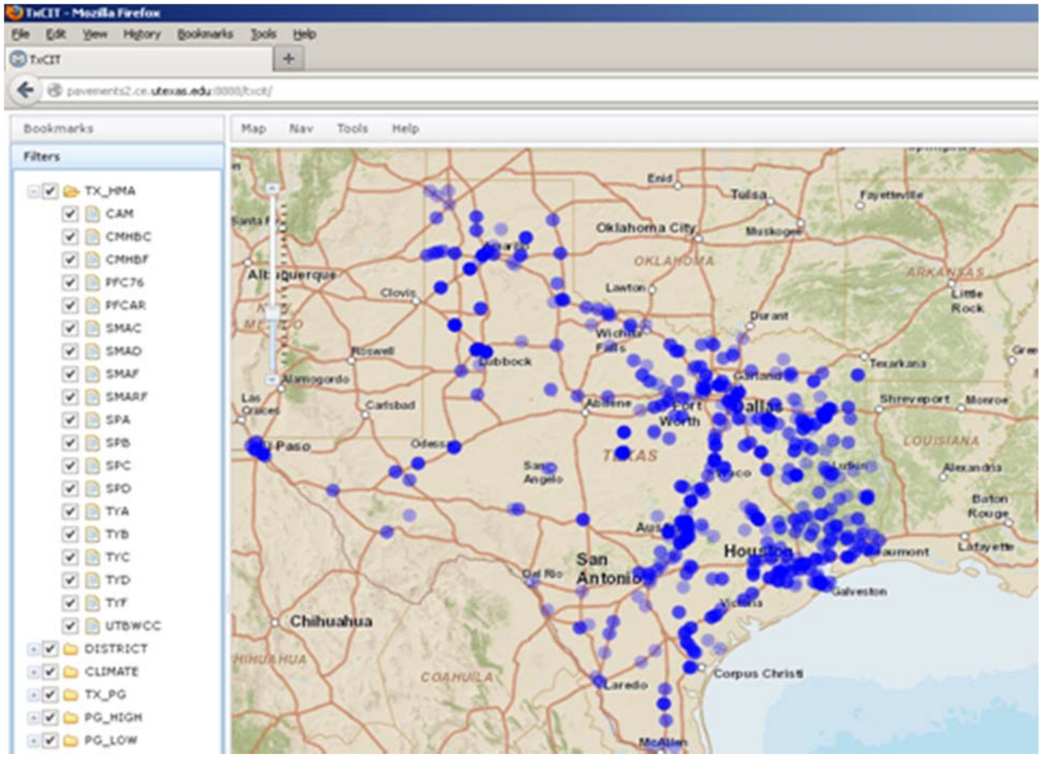

Figure 3.2: A screenshot of TxCIT database

3.1.1 Design and Construction Information System

The DCIS database includes five tables: 1) Bid, 2) Project, 3) Proposal, 4) Item List, and 5) Districts and County. A summary of the various tables and the corresponding data fields is shown in Table 3.1. Additionally, the format and a brief description of each data field are also indicated in the table. The role of these data fields in the context of developing a new database for the current project is discussed below.

Table 3.1: DCIS database description—tables and fields

Table Name	Field	Format	Description
Bid	CONTID	varchar(15)	Project number (control-section-job [CSJ])
	VENDOR	varchar(14)	Vendor Number
	ITEM	varchar(13)	TxDOT specification number
	BTUPRICE	decimal(13,5)	Pricing per unit quantity
	BTOQTY	decimal(12,3)	Unit quantity
Project	CONTID	varchar(15)	Project number (CSJ)
	PJDESC1	varchar(60)	Project description
	ISPECYR	varchar(2)	Specification year
	COUNTY	varchar(4)	Project county number
	PJDISTR	varchar(5)	Project district number

Table Name	Field	Format	Description
	PJROADNM	varchar(60)	Road (route) name
	PJBTERMI	varchar(10)	Beginning termini
	PJETERMI	varchar(10)	Ending termini
	PJBSTATN	varchar(20)	Beginning Station
	PJESTATN	varchar(20)	Ending Station
	PJLENGTH	decimal(9,4)	Project Length
	PJXCOORD	int	Longitude of midpoint
	PJYCOORD	int	Latitude of midpoint
Proposals	CONTID	varchar(15)	Project number (CSJ)
	ISPECYR	varchar(2)	Specification year
	CNDISTR	varchar(5)	Primary district
	COUNTY	varchar(4)	County number
	CNRDSYS	varchar(4)	Road system
	CNROUTE	varchar(20)	Route
	CNDTLET	Date	Letting date
	CNDTSTRT	Date	Estimated starting date
	CNDTCPE	Date	Estimated completion date
	VENDOR	varchar(14)	Contracted vendor
	UNITSYS	varchar(4)	Measurements system (English or metric)
Item List	ITEM	varchar(13)	Item number
	ISPECYR	varchar(2)	Specification year
	IDESCR	varchar(40)	Item description
	IUNITS	varchar(4)	Quantity unit

Bid table includes CONTID, a unique number assigned for each construction contract approved by TxDOT. A contract generally engages only one independent contractor/vendor and is identified by the field VENDOR in the database. The bid table consists of all individual vendors involved in the bidding process for a particular project. It is important to filter out the contracted vendor to which the contract was actually assigned. The ITEM field logs the individual items planned as part of the project. A comprehensive description of individual items is given in the TxDOT specifications. Since specifications change over time it is necessary to link the specification year along with the item number (ITEM). Additionally, the bid table also includes the planned quantities and the offered bid prices for each individual construction item.

Project table logs the project-specific information such as description of the project and its location information, including county; district; route name; and beginning, mid, and ending stations. However, some of these location-specific fields are not populated in the DCIS table. The ISPECYR data field indicates the year of the specification book corresponding to individual items listed in the bid table.

Proposal table reports information for all the TxDOT proposals and needs to be filtered for the contracted vendor (as in the case of the bid table). The table indicates an estimated start and completion date for the construction project. These dates may be used as an estimate of the project beginning and ending dates. However, more reliable estimates can be obtained from the SM database, as discussed later. The Counties and Districts are represented by number and corresponding textual descriptions can be obtained by linking to District and County tables.

Item table lists the description and relevant specification year for each individual item listed in the bid table. It also includes IDESCR field that describes the individual items within a given project. Lastly, the quantity units are provided by IUNITS field.

These tables are connected for extracting meaningful information using primary data fields or primary keys such as CONTID and ITEM.

3.1.2 SiteManager Database

QC/QA commonly involves laboratory testing of asphalt mix samples on a lot basis. A lot is defined as the quantity of asphalt mix produced in the hot mix plant on a single day, which is typically about 2,000 tons. Four sublots are randomly obtained from each lot for laboratory testing. Each sublot is used for regulating both production and placement processes during construction. Production QC/QA consists of three laboratory-molded samples (a total of twelve for four sublots), which are prepared using the loose mix collected from the plant or hot mix truck. On the other hand, field cores are obtained at any two randomly chosen locations (a total of eight for four sublots) for laboratory testing as part of the placement QC/QA process. The described framework is adopted by both TxDOT (as part of QA) and contractor (as part of QC), which doubles the aforementioned number of test samples. All QC/QA information is collected via spreadsheet templates filled by both TxDOT personnel as well as the contractor. The spreadsheet stored data is uploaded into the SM database server. It has been logging all the QC/QA information from various construction projects across Texas since 2003. The database comprises two tables: 1) CCSJ table and 2) QC/QA table. A discussion of utility and importance of each data field included as part of these two tables is given below. A summary of the various tables and the corresponding data fields is shown in Table 3.2.

CCSJ table contains data fields required for identifying a project, its executed items, and identification numbers for all the samples collected as part of the QC/QA. Table 3.2 lists various data fields listed in the table. CONT_ID is a unique identification number for an individual contract whereas PRJ_NBR indicates the corresponding project. Note that each contract assigned to any individual contractor may consist of more than one project. Hence, a project can be uniquely identified by a combination of the CONT_ID and PRJ_NBR. In other words, a two-variable primary key comprising CONT_ID and PRJ_NBR is essential. The laboratory samples and field cores are grouped together and sent over to the TxDOT/contractor laboratories in multiple batches. These sets of batches may comprise samples belonging to distinct lots. However, a unique SMPL_ID is given for all the samples in any selected set. Hence, SMPL_ID is not very useful for assigning specimens to the corresponding lots and sublots. However, it is a crucial element that facilitates integration of the CCSJ and QC/QA tables as described below.

Table 3.2: SM database description—tables and fields

	Field	Format	Description
CCSJ table	CONT_ID	nvarchar(15)	Control section job number
	PRJ_NBR	nvarchar(13)	Project control number
	LN_ITM_NBR	nvarchar(4)	Line item number
	SMPL_ID	nvarchar(18)	Sample identification number
QC/QA table	SMPL_ID	nvarchar(18)	Sample identification number
	TST_METH	nvarchar(10)	Test method applied
	SMPL_TST_NBR	nvarchar(10)	Sample test number (lot number)
	FLD_NBR	decimal(5)	Field reference number
	FLD_VAL	nvarchar(18)	Field value

QC/QA table consists of the data fields necessary for storing information collected during construction. The key for connecting this table with the previously described CCSJ table is the SMPL_ID data field, which is recorded in both tables. As mentioned earlier, the QC/QA information is stored in Excel spreadsheet templates. Multiple templates are used for collecting QC/QA information depending on the construction year, which asserts the need for a unique identifier. TST_METH indicates the spreadsheet template used during QC/QA data collection. Furthermore, FLD_NBR identifies each individual cell of any specific spreadsheet template. These identification numbers (FLD_NBR) are different for different QC/QA spreadsheet templates. Besides, FLD_VAL identifies the input value entered either by TxDOT personnel or the contractor into a specific spreadsheet cell. Thus, a combination of TST_METH, FLD_NBR, and FLD_VAL refers to a unique cell value in a specific QC/QA table. SMPL_TST_NBR indicates the lot number; which distinguishes the QC/QA information of any two samples from dissimilar lots. The production and placement QC variables are averaged at the lot level, as is discussed later.

SM stores the field values extracted from the TxDOT/contractor-filled spreadsheet templates. However, up until 2008 it did not store the calculated field values. As part of a previous research project (0-5496), the present research team recommended storing the calculated fields, including the raw data. The recommendation was implemented and the SM database was revised to store the calculated fields. For example, the objective is to store specific gravity information for a given laboratory specimen. The older version of the SM database stores the dry, submerged and saturated surface dry weights in three different fields, whereas the newer SM database (post 2008) stores the specific gravity value in addition to the three different weights of the specimen. Indeed, the specific gravity of the specimen is of prime importance for further analysis rather than the unprocessed information (weights). Since the revised SM database was executed after 2008, only 4 years of performance data is available for these projects. Since this is not sufficient to evaluate performance over the service life of pavements, it was necessary to extract information from the older SM database to calculate the information required for the study.

Unfortunately, the calculation of the required volumetric properties using the older SM database was cumbersome. This issue was also highlighted in a previous research report by the present research team (Smit et al., 2005). Researchers devised a procedure to calculate the required data fields (such as laboratory density, in-place air voids, etc.) using Visual Basic macros and SQL queries. This procedure calculated the required volumetric properties using the existing raw information from the SM database. Finally, a database including all the calculated production and placement characteristics was developed towards the end of this exercise.

The SM database stores various other construction characteristics apart from the QC/QA properties that may potentially be used in the context of developing performance-based pay factors. For instance, it logs the placement date of each lot, so the date of the final lot can be used as an accurate estimate of the project completion date. Such a date is essential for tracking post-performance of the project, a critical element of the performance-related pay adjustment system development.

3.1.3 Pavement Management Information System

The PMIS database maintains road condition information that is collected annually across the Texas road network. Two tables are being used in the present research project: PMIS data collection section and PMIS condition summary.

PMIS data collection section table consists of 39 different data fields describing the section characteristics on which the performance information is collected. The table includes section location information such as district, county, and route number. Traffic information such as 18-kip ESAL value, annual average daily traffic (AADT), maintenance cost, and pavement type are some of the noteworthy data fields. The number of 18-kip ESAL values is represented using the CURRENT_18KIP_MEAS data field. It is obtained from the TxDOT TRM database for the data collection section and represented in thousands. AADT is defined as the average daily estimate of the number of vehicles on all the lanes in a single direction for divided facilities; it includes traffic on all the lanes in both directions for frontage and undivided roads. AADT_CURRENT data field reports the maximum AADT value published for a given stretch of test section. The cost of the maintenance on the main lanes during the previous year of the data collection is reported as MAINTENANCE_COST_AMT. This value is calculated from the TxDOT maintenance management information system (MMIS). Pavement type of the predominant travel lane during the data collection year is reported as the pavement type for the test section. Geometric features such as the number of through lanes, total surface roadway width and shoulder width are also included. The fiscal year and location information such as beginning and ending TRMs are included in both of the above-mentioned tables. Hence, these data fields are primarily used to link the two tables.

PMIS condition summary table includes 47 different fields each describing performance of the test section in terms of specific distresses. Some of the most important data fields are described below based on TxDOT's PMIS dictionary:

1. Fiscal year
2. Signed highway roadbed ID
3. Beginning and Ending reference marker numbers
4. Beginning and Ending reference marker displacements
5. Distress score
6. Condition score

7. Ride score
8. Left and right IRI
9. Shallow rut
10. Deep rut
11. Severe rut
12. Block cracking
13. Alligator cracking
14. Longitudinal cracking
15. Patching
16. Transverse cracks
17. Visual lane code

Fiscal year is the year in which the data collection was conducted. *Signed highway roadbed ID* concatenates route name, route number, and road bed. Route name includes two letters and a list of various route names is given in Table 3.3. Route number is a four-letter string variable. Road bed ID is a code identifying the roadbed constituting a highway section as shown in Table 3.4. *Beginning and ending Reference marker numbers* are the nearest TRM for beginning and ending of the section respectively. The *beginning and ending Reference marker displacements* are the displacements from the beginning and ending points of the test section to their corresponding nearest TRMs.

Table 3.3: PMIS highway systems

Highway	System
IH	Interstate Highway
US	US Highway
UA	US Alternate
UP	US Highway Spur
SH	State Highway
SA	State Highway Alternate
SL	State Highway Loop
SS	State Highway Spur
BI	Off Interstate Business Route
BU	Off US Highway Business Route
BS	Off State Highway Business Route
BF	Off Farm or Ranch to Market Road Business Route
FM	Farm to Market Road
RM	Ranch to Market Road
RR	Ranch Road
PR	Park Road
RE	Recreation Road
FS	Farm to Market Road Spur
RS	Ranch to Market Road Spur
RU	Ranch Road Spur
RP	Recreation Road spur
PA	Principal Arterial Street System (PASS)
MH	Metropolitan Highway

Distress score describes the amount of surface distress on the data collection section. It is calculated by multiplying utility values for each distress evaluated on a pavement type and varies between 0 (worst) and 1 (best). *Ride score* is a measure of overall ride quality of the pavement section and varies between 0.1 (roughest) to 5.0 (smoothest). *Condition score* represents the overall condition of the road including surface distress and ride quality and ranges from 1 (worst) to 100 (best). The condition score integrates ride and distress scores. *Left and right IRI* describes the average IRI value on the left and right wheel paths, respectively. *Shallow rut* represents the percentage of the shallow rutting (ranges between 0.24–0.49 inches) in a given data collection test section. Similarly, *deep and severe ruts* represents the percentage of the deep (between 0.24–0.49 inches) and severe (from 1.00 to 1.99 inches) rutting in a given data collection test section.

Block cracking is the percentage of the lane area in the data collection test section with block cracking. Block cracking consists of interconnecting cracks that divide the pavement surface into approximate rectangular pieces (ranging from 1 foot x 1 foot to 10 feet x 10 feet). *Alligator cracking* is the percentage of the wheel path length in the data collection section with alligator cracking. Alligator cracking consists of interconnected rectangular blocks that arise out of fatigue failure. *Longitudinal cracking* is the average length of the longitudinal cracking per 100-ft section. Longitudinal crack lengths are measured in 100-ft-long sections and averaged across the data collection section. *Patching* is the percentage of lane area with patching in the rated lane of data collections section. *Transverse cracks* is the number of transverse cracks per 100-ft lane length of the data collection section.

Visual lane code identifies the lane of the data collection section in which the visual distress data is collected. Note that the performance information is only collected on a single lane, visually the most distressed (typically the outside lane carrying heavier traffic) for a given road and direction. Lanes in both undivided and divided highways are identified using a combination of a letter code and lane number. The letter codes of the road beds are shown in Table 3.4. Lanes are numbered from 1 to 5 in the direction of the increasing TRM and 6–0 in the direction of decreasing TRM. Figure 3.3 and Figure 3.4 demonstrate the lane identification for both divided and undivided highways.

Table 3.4: PMIS lane convention

Roadbed ID	Description
K	Single main lane road
A	Right frontage/service road
R	Right main lane road
X	Left frontage/service road
L	Left main lane road

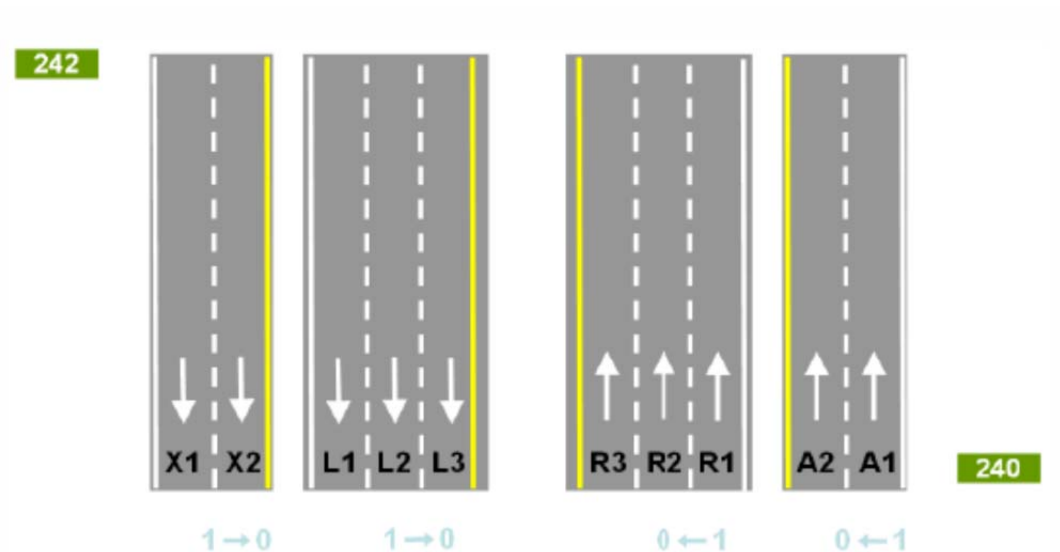

Figure 3.3: Divided lane identification

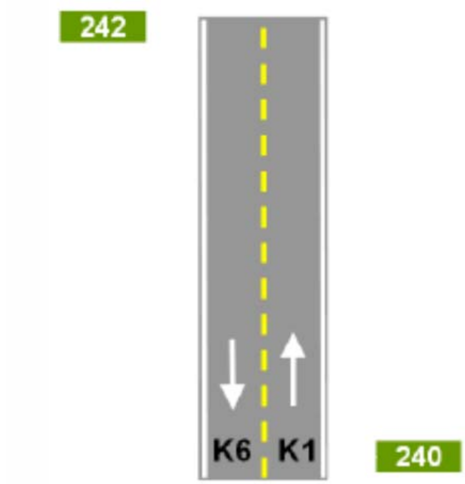

Figure 3.4: Undivided lane identification

3.2 Database Integration

The current research project aims to review the existing pay adjustment scheme and develop a new performance-based pay system if justified. Pay factors are typically calculated based on various parameters describing material and construction quality. Identifying the performance sensitivity of these parameters is essential for building a performance-based pay adjustment system. A relationship must be established between the existing QC/QA parameters and the corresponding field performance of the projects for the sensitivity analysis. The three aforementioned databases (DCIS, SM, and PMIS) are linked using various primary data fields in order to develop an integrated database that caters to the needs of the current research project. Such an integrated database may possibly be used to develop relationships between QC/QA parameters and the corresponding project field performance. Indeed, regression-based empirical models may be developed for attributing the variation in field performance to corresponding QC/QA parameters.

A construction lot comprises a day's asphalt mix production and consecutive placement in the field. The total number of lots in a project is dependent on its length. Also, each lot is divided into four sublots in a typical QC/QA framework. QC/QA parameters are collected in each sublot and averaged across each lot. These lot-level QC/QA parameters are related to field performance of the pavement section corresponding to that specific lot rather than the entire project. TRM location is essential for extracting performance of any pavement section using the PMIS database. Unfortunately, the location of pavement sections corresponding to an individual lot is not recorded in the SM database. The unavailability of the lot specific field performance requires a project-level approach. A brief description of the approach used is given below.

Firstly, it is assumed that all the lot-level averages which are calculated based on the corresponding four sublot values in a given project belong to the same population. Thus, project-level QC/QA parameter estimates are obtained by averaging the lot-level means. Laboratory-molded density, in-place air voids, asphalt content, VMA, theoretical maximum density and gradation properties were used in the present study as the QC/QA parameters. It may be argued that the mechanism of deterioration differs with asphalt mix type and therefore the influence of the above-mentioned properties may not be similar for all the mixes. For instance, raveling is predominant in a porous friction course. Thus, the properties that control the deterioration of mixes may be very different. In order to account for the mix type, it was decided to perform the analysis specific to the mix type which will be further explained later. A project-level estimate is calculated for each of these six properties and logged into the database using SQL queries. Data validation was carried out at each stage to ensure the reliability of the results. The SM database is not validated and consisted of several unexpected entries which demonstrate the need for data validation prior to the analysis. In summary, a table was prepared consisting of project-level information along with the corresponding project-level estimates of QC/QA volumetric parameters.

In the context of performance tracking, project-level performance trends are estimated from the PMIS database. As mentioned earlier, the PMIS data is collected typically on the outside lanes in both directions for divided facilities consisting of multiple lanes; it is collected only in one direction in the case of undivided facilities. Thus, the performance information for a given stretch of a pavement section is only based on a single lane irrespective of the number of existing lanes. Note that only the beginning and ending TRMs are available for locating a project. Since it is difficult to identify the exact lanes paved as part of any particular project, a few assumptions are necessary while extracting and attributing the performance information to that specific project. It is assumed that overlaying/reconstruction of any pavement section on an undivided facility always includes all the existing lanes. The assumption is justified because overlaying of single lane/direction may introduce level difference on these undivided roads, which is unacceptable—indeed, dangerous—at high speeds. Thus, the performance data collected on any of the existing lanes on a given pavement section shall be relevant to the corresponding overlay/reconstruction project.

On the other hand, divided facilities comprise multi-lane roads in two directions with a median. The presence of the median avoids the above-mentioned problem of the level difference and thereby allows for the independent construction of main lanes in different directions. The exact project location remains ambiguous with the available beginning and ending TRMs. The collected performance data cannot be averaged across the two directions unless it is ensured that, the corresponding project involves overlay/reconstruction of lanes in both the directions. In order to address this problem, the performance information collected on both directions was compared. A threshold value was selected for the difference between those performance measures. The similarity in the performance trends in two directions indicates the possibility of the project being executed in both directions; i.e., the same design was used on the opposing carriageways. The analysis comprises only those projects including the overlaying or reconstruction of lanes in both the directions on a divided highway. Averaged IRI (left and right wheel paths) was used as the performance indicator as part of this project. IRI has been one of the most consistent performance measures in the PMIS database. Rutting and cracking measures in the PMIS database were also evaluated but were found to be poor performance indicators compared to IRI. Performance measurement based on PMIS rutting and cracking data potentially mislead the overall performance of the project upon combining few years of data. This is primarily due to the highly noisy network-level rutting and cracking measurements. Also, rutting data was stored in different categories representing the intensity of the rutting; the definitions of these categories have changed recently and aggregating the rutting data over time needs to account for the respective transformations. The researchers excluded the rutting and cracking measurements from the analysis in order to ensure adequate reliability levels, despite the rigorous data cleaning and data collection efforts.

The performance information spanning a few years is summarized based on the following procedure. The researchers developed a ride-quality (in IRI) based deterioration index (rate of deterioration) for assessing pavement performance. Other such indexes could be equally used. The deterioration index is defined as the rate of change of averaged IRI across the pavement project over time. The rate is estimated using linear regression of the available annual ride quality measurements (IRI) corresponding to the post-construction period with time (in years). The slope of the estimated regression line, which primarily depicts the expected change in IRI per year, is taken as the deterioration index. The intercept of the regression line—i.e., estimated value of the IRI during the year of the construction—is utilized as the IRI immediately after the construction (or initial IRI). Figure 3.5 demonstrates the calculation of the performance measures as discussed. The construction year is identified manually by monitoring sudden drops in IRI or sudden improvement in condition score data in PMIS. In Figure 3.5 the construction year is assumed as 2009. The aforementioned regression includes the performance data starting from the construction year and thus, the accuracy of the construction year is critical. The aforementioned estimated initial IRI is utilized purely due to unavailability of the actual IRI measured immediately after pavement constructions in the relevant database. It is important to note that the time-span of the available performance data (i.e., analysis period) differs across individual projects due to distinct years of construction.

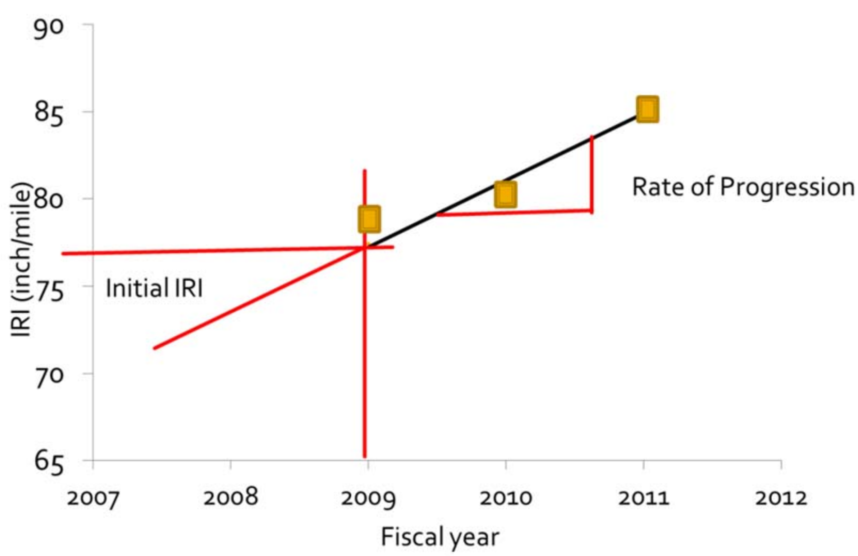

Figure 3.5: Calculation of performance measures

In summary, the database populated using the above-explained methodology includes project-level information including location, averaged QC/QA estimates (volumetric properties), initial IRI, and field performance (until 2012). The integrated dataset is fully described along with detailed descriptive statistics below. Note that most of the performance information collected in the PMIS database includes only surface-related distresses. Hence, this research study only included surface projects in the analysis towards the development of performance-based pay factors.

3.3 Descriptive Statistics

An integrated database was developed including project-level information comprising averaged QC parameters (volumetric properties of mixture, initial ride, etc.) and field performance (rutting, cracking, and rate of deterioration). The surface projects that comprised of paving Type C, Type D, stone matrix asphalt (SMA), coarse matrix high binder (CMHB), and Superpave (SP) are included in the database. However, a limited number of projects are available for SMA, CMHB, and SP categories due to their lower frequency of usage. The database contains information pertaining to more than 1,400 surface projects constructed after 2002. However, this study used information from 824 projects mainly because about 500 projects are less than 3 years old and thereby disqualified for the current analysis, and another 76 projects were filtered out during a manual data QC exercise for containing erroneous performance data (PMIS data) and missing QC/QA information. Projects containing irregular and unexpected performance trends (such as a decrease in performance after construction or unacceptable noise) were removed from the analysis. For instance, Figure 3.6 shows an example of a "good" project with an expected performance trend. The effort in plotting these performance trends for each project to ensure reliable deterioration rate estimates should be emphasized; it also suggests the need for data QC interfaces (noise filtering interface before inputting into TxDOT data servers) for PMIS. The aforementioned performance trends from more than 800 hot mix projects are documented on the attached CD-ROM as an addendum to this report. A detailed description of the dataset including various interesting characteristics and descriptive statistics is provided below. Table 3.5 and Table 3.6 provide a list of variables included in the analysis and reports important descriptive statistics.

Figure 3.6: Deterioration of a typical hot mix project (vertical line indicating the construction year)

Deterioration rate is estimated using a linear regression approach combining the first few years of ride quality data (IRI) as described earlier. Deterioration rate is censored at zero for all the projects with minor deterioration rate (close to zero). Thus, HMA projects were categorized into two groups: 1) those exhibiting almost no deterioration in the subsequent analysis period, and 2) those exhibiting a considerable level of deterioration within the analysis period. Thus, deterioration rate takes either a zero or a non-zero value. Figure 3.7 shows a histogram of the deterioration rate for all 824 projects included in the analysis. As indicated in Table 3.5, 146 pavements (out of 824 HMA projects) did not exhibit any signs of deterioration in terms of ride quality. The deterioration rate varied from 0 to about 60 inches/mile/year with an average rate of deterioration of 4.2 inches/mile/year. About 5% of the pavement sections deteriorated faster than 15 inches/mile/year. The distribution of the deterioration indicates that a minor portion of these new pavements did deteriorate rapidly.

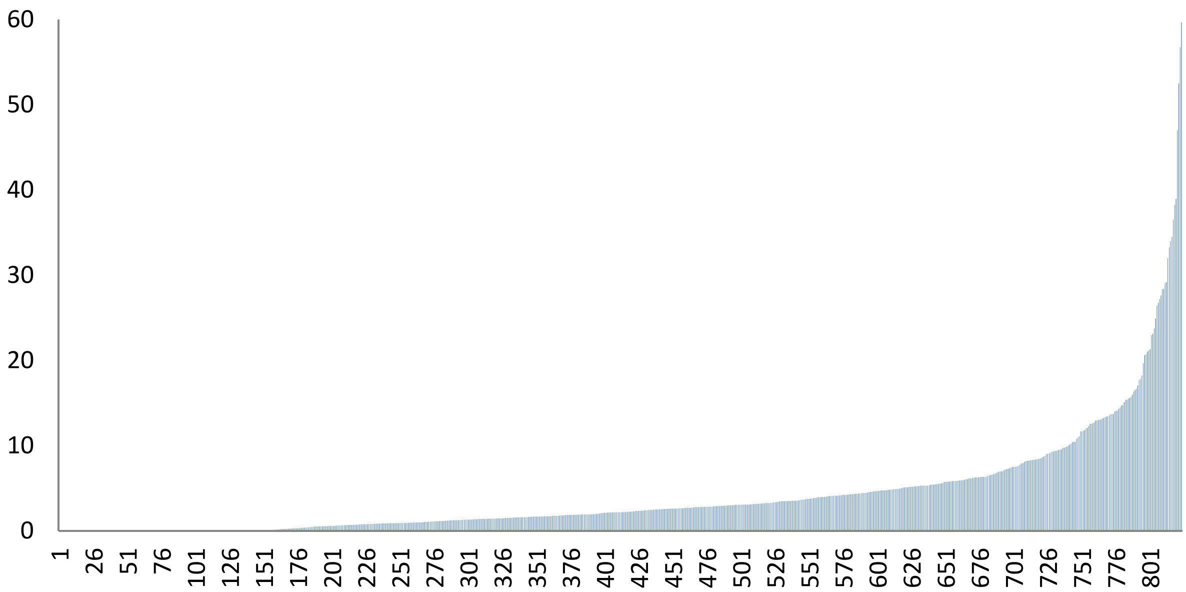

Figure 3.7: Distribution of deterioration rate for different projects

A comprehensive description of the characteristics of explanatory variables is also important to better understand the wide range of the hot mix projects included in the analysis. The analysis included both shorter and longer projects ranging from about 1-mile long to 23-mile long with an average length of 4 miles; about 34% of the projects are less than 2-mile long, while about 17% are as long as 7 miles. Approximately 57% of the HMA projects are associated with two-lane roads. About 5% of the pavements are constructed on routes with a maximum speed limit of 45 mph, while 49% of the pavements are on routes with more than a 65 mph speed limit. About 46% of the hot mix projects are located in rural areas. The projects selected for the analysis provide moderate coverage across all the facility types. The dataset included approximately 18% on Interstate Highways, 28% on US highways, 25% on State Highways, 27% on FM roads and the remaining being other facility types. Local environmental conditions markedly influence the performance of a pavement. In order to effectively account for the weather condition, the dataset included about 7% belonging to dry/cold region, 21% belonging to dry/warm region, 29% belonging to wet/warm region, and 21% belonging to wet/cold region, with the remaining located in a mixed-type weather. Both left and right shoulder width ranged from "no shoulder category" to about 28 ft and 19 ft respectively. Traffic is included in the dataset in four different measures: 1) 18-kip ESAL count, 2) AADT, 3) truck traffic percentage, and 4) estimated daily average of the ten heaviest wheel loads traveling a particular traffic section. The dataset included hot mix projects in low, moderate and high traffic regions. The descriptive statistics of the traffic counts are reported in Table 3.5.

Table 3.5: Project location features

Variable	Mean	Std. Dev	Min	Max
Deterioration	4.23	6.62	0	59.65
Project Length	4.07	3.54	0	22.94
Indicator Variable: Small Project	0.34	0.47	0	1
Indicator Variable: Medium Project	0.49	0.50	0	1
Indicator Variable: Longer Project	0.17	0.37	0	1
Indicator Variable: Two lane Road	0.57	0.49	0	1
Indicator Variable: Lower Speed Limit	0.06	0.23	0	1
Indicator Variable: Medium Speed Limit	0.46	0.50	0	1
Indicator Variable: High Speed Limit	0.49	0.50	0	1
Indicator Variable: Rural Area	0.46	0.50	0	1
Indicator Variable: Facility-IH	0.18	0.38	0	1
Indicator Variable: Facility-US	0.28	0.45	0	1
Indicator Variable: Facility-SH	0.25	0.43	0	1
Indicator Variable: Facility-FM	0.27	0.44	0	1
Indicator Variable: Facility-Other	0.03	0.16	0	1
Indicator Variable: Weather-Dry/Cold	0.07	0.26	0	1
Indicator Variable: Weather-Dry/Warm	0.21	0.41	0	1
Indicator Variable: Weather-Wet/Warm	0.29	0.46	0	1
Indicator Variable: Weather-Wet/cold	0.21	0.41	0	1
Indicator Variable: Weather-Mixed	0.21	0.40	0	1
Left Shoulder Width	6.72	2.81	0	28.00
Right Shoulder Width	7.72	2.73	0	19.20
18 Kip ESALs	10,137	13,443	40	86,717
Annual Average Daily Traffic	17,743	22,061	78	149,944
Truck Traffic Percentage	15.62	9.96	1.50	72.66
Traffic Load Estimate	150.11	24.06	89.00	195.67

This research is primarily interested in identifying the relationship between as-constructed QC information and the corresponding project performance. A wide range of volumetric properties, ranging from off-target to on-target values, is essential to identify the influence of these volumetric properties on performance. Unfortunately, the SM database contains limited projects with off-target volumetric properties and is dominated by projects with QC parameters within the acceptable thresholds of the existing specifications. The analysis conducted as part of this study therefore reflects changes in pavement performance due to a change in the QC parameters that are within the existing specification limits. It is important to realize that this research only accounts for the sensitivity of these quality parameters to performance within a narrow range of the specification limits. Indeed, it is important to identify such sensitivities for the development of PRSs.

Approximately 49% of the hot mix surface projects included in the dataset placed Type C mix (a popular ½" dense-graded mix in Texas) and about 45% of the surface projects placed Type D mix (a popular fine 3/8" mix). The remaining hot mix projects contained either SMA or Superpave mixes. About 19% of the HMA projects used performance-graded (PG) binder with a high temperature grade of 64, while 39% used a high temperature grade of 76.

The initial ride quality after construction significantly affects future deterioration. The dataset contained hot mix projects with an initial IRI value varying between 30 inches/mile to 160 inches/mile with a mean of 78 inches/mile. It is important to control for the initial ride quality during any type of performance modeling; initial ride quality affects the deterioration rate. The deterioration rate is estimated using the available performance data corresponding to post-construction period, which results in dissimilar analysis periods for individual projects depending on the respective years of construction. The dataset included analysis periods ranging from 3 to 10 years, with a mean analysis period of 5 years. Table 3.6 also includes descriptive statistics (mean, standard deviation, maximum, and minimum) for both production- and placement-related as-constructed quality parameters. The production-related quality parameters include plant density, plant VMA, plant VFA, asphalt content, maximum specific gravity (Rice density), and dust content of the plant mix. The placement-related quality parameters include in-place air voids, in-place VMA, and in-place VFA, which were measured using field cores.

Table 3.6: As-constructed QC parameters

Variable	Mean	Std Dev	Min	Max
Type C	0.49	0.50	0	1
Type D	0.45	0.50	0	1
Indicator variable: PG 64	0.19	0.39	0	1
Indicator variable: PG 70	0.41	0.49	0	1
Indicator variable: PG 76	0.39	0.49	0	1
Initial IRI	78.2	23.9	29.1	159.2
Analysis period	5.05	1.79	3.00	10.00
Laboratory density	96.34	0.48	94.75	98.82
Laboratory VMA	14.61	1.06	7.63	18.11
Laboratory VFA	74.67	3.83	54.49	90.58
Asphalt content	4.77	0.44	3.54	7.75
Maximum specific gravity	2.46	0.05	2.27	2.80
Dust content	3.18	1.60	0.00	12.27
In-field air voids	7.26	0.99	3.72	14.27
In-field VMA	17.81	1.41	9.63	24.34
In-field VFA	59.24	4.50	35.95	77.54

3.4 Exploratory Data Analysis: Distributions

This section details the distributions of the QC/QA parameters: 1) laboratory-molded density, 2) in-place air voids, 3) asphalt content, 4) VMA, and 5) as-constructed ride quality and awarded pay adjustments for production, placement, and ride quality. Subsequently, they are compared to most relevant theoretical distributions. Such plots reveal the contractors' rate of achieving the assigned target QC/QA parameters as well as their overall variability. Additionally, they assist in the selection or removal of a given QC/QA parameter for the newly developed performance-based pay factor system.

3.4.1 Laboratory Density

TxDOT currently controls the production quality of hot mix using laboratory density measurements. A tolerance limit of ±1% deviation from the respective target density is allowed for most of the hot mix items as per the specification book (2004). Three laboratory-compacted specimens are prepared using the loose asphalt mix specific for each sublot. The averaged density of these specimens is reported as laboratory-molded density corresponding to the respective sublot. A project-level laboratory density measurement was necessary for understanding the relationship between the laboratory density and corresponding pavement performance. A project-level laboratory density estimate is calculated by averaging the individual lot averages across a given hot mix project. Figure 3.8 shows a distribution of the project-level laboratory density. It indicates that laboratory density is approximately normally distributed and ranges from 94.5% to 97.7% with an average value 96.3%. The range of the laboratory densities is typical for Type C and Type D mixes, which are predominant in the current database.

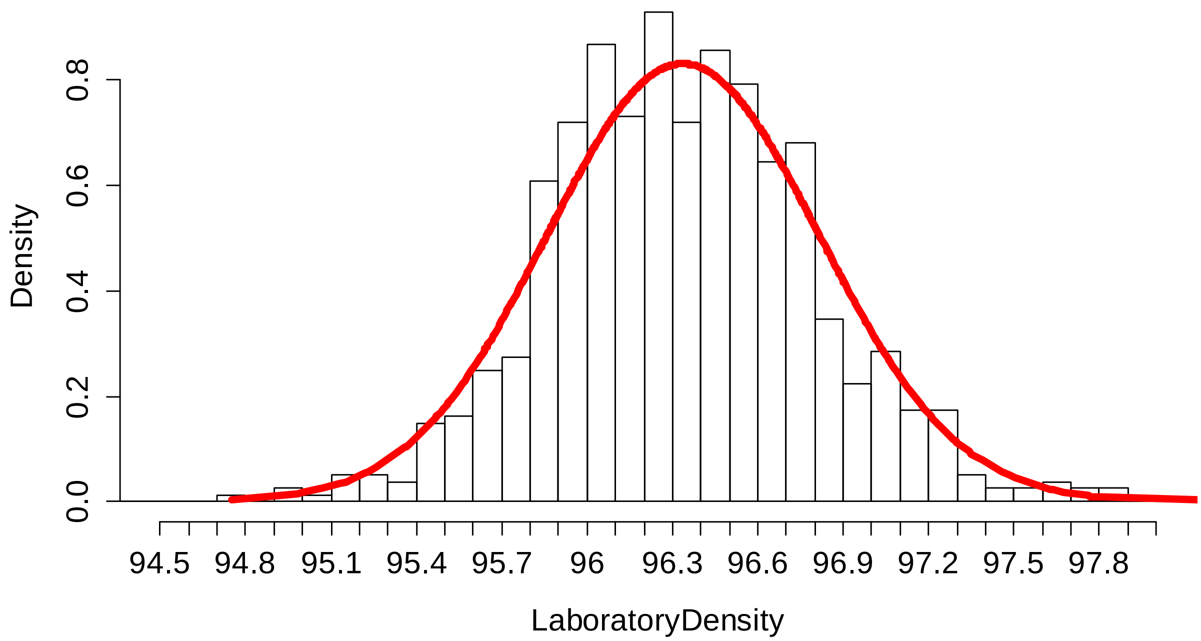

Figure 3.8: Actual and modeled distributions of laboratory density

Lot-level averaged deviation is used for the calculation of production-related PAFs. A look-up table is provided in the TxDOT specification book (2004) that lists the pay factors corresponding to various absolute deviation values. Figure 3.9 shows a distribution of the awarded production-related pay factors from the previous projects. It can be observed that a major portion of contractors received a production-related bonus. It indicates that nearly 100% of the sample population is within the tolerance limits. In other words, almost all the contractors are able to achieve the target value and may receive a bonus. Historically, contractors have developed the expertise and technology in achieving accurate laboratory-molded density. This behavior highlights the need for revising the existing pay factor system in the context of laboratory-molded density to "raise the bar" as most of the contractors are currently being awarded a bonus.

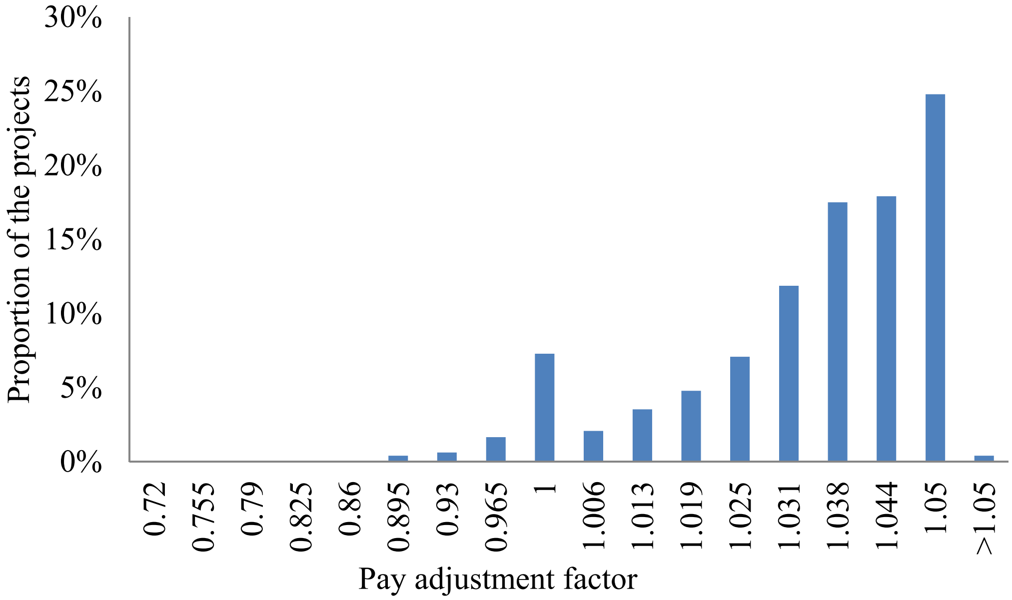

Figure 3.9 Awarded production pay factors

3.4.2 In-place Air Voids

TxDOT currently controls the production quality of hot mix using laboratory density measurements. Two cored field specimens are obtained from randomly selected locations within the paved section corresponding to each sublot. Air voids are calculated in the laboratory for these eight specimens and averaged across each sublot. Averaged compaction quality of a given hot mix project is obtained by averaging the lot-level compaction densities. Figure 3.10 indicates that the in-place air voids averaged at the project level follow a normal distribution. It is interesting to see a larger variability of in-place air voids spanning from 2.7% (over-compacted) to 9.9% and above (under-compacted). Figure 3.10 shows the actual in-place air voids but not deviations from the targeted in-place air voids as in the case of laboratory density. Lot-level averaged in-place air-void content is used for the calculation of placement-related PAFs. A look-up table is provided in the TxDOT specification book (2004) that lists the pay factor corresponding to various in-place air-void contents. Figure 3.11 shows a distribution of the awarded placement pay factors from the previous projects. Note that a major portion of contractors received a placement-related bonus, which is very similar to the production-pay factor.

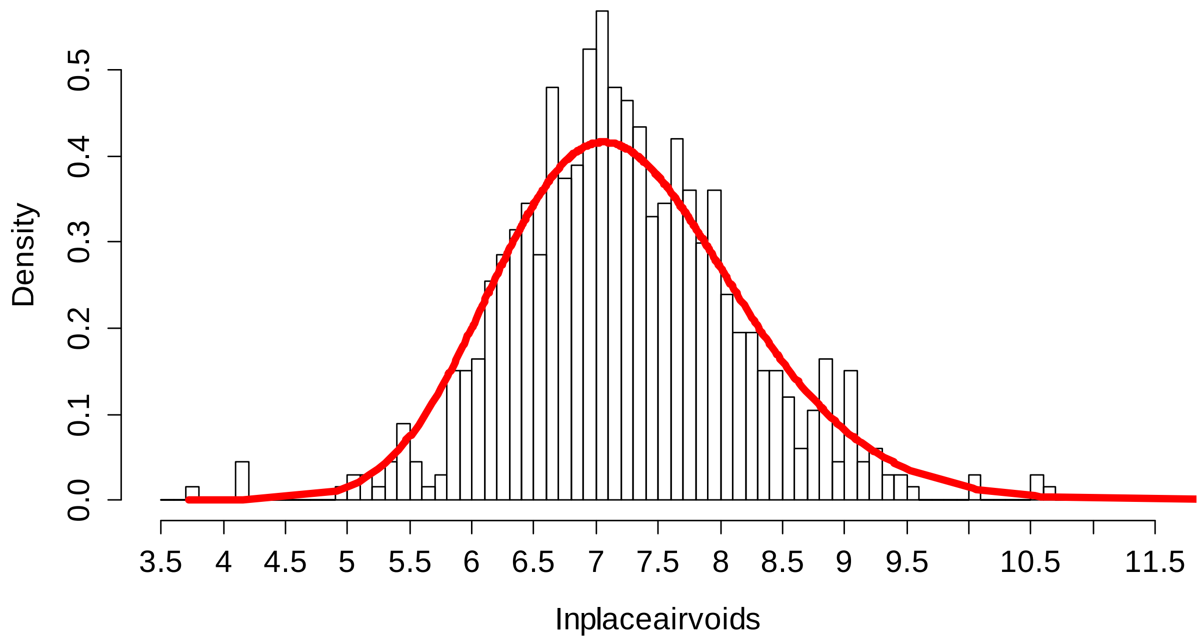

Figure 3.10: Actual and modeled distributions of in-place air voids

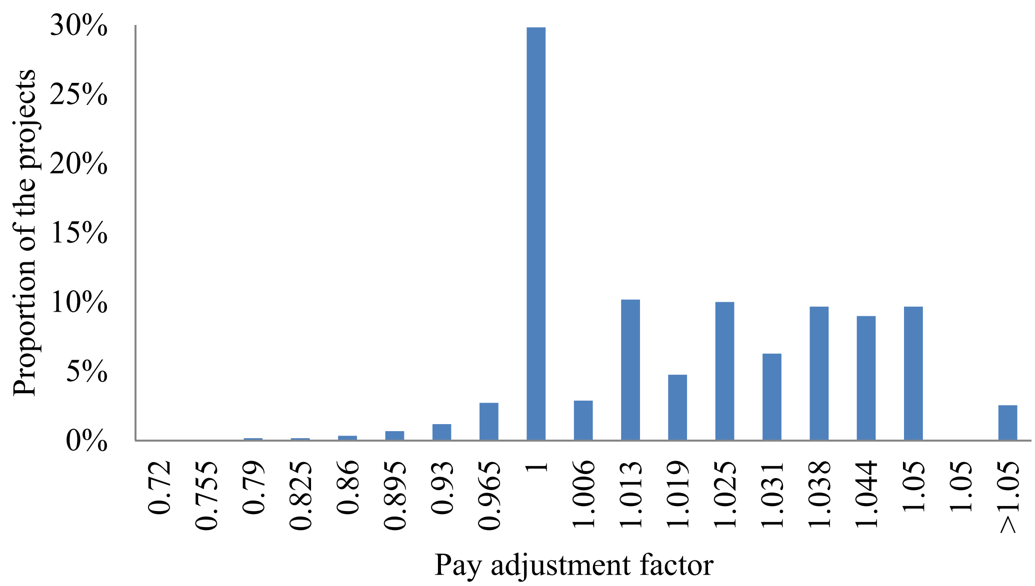

Figure 3.11: Awarded placement pay factors

3.4.3 Asphalt Content

The current specification restricts the deviation from the target asphalt content using upper and lower limits (±0.3%). However, asphalt content is not included in the present pay adjustment system. Asphalt content is measured during the production process using the ignition oven method in individual sublots. An average asphalt content of a given hot mix project is calculated by averaging the lot-level average asphalt content measurements. Figure 3.12 shows the distribution of the project-level averages of asphalt content. The asphalt content generally varied between 3.5% and 6%, which is typical for a dataset with predominantly Type C and Type D mixes.

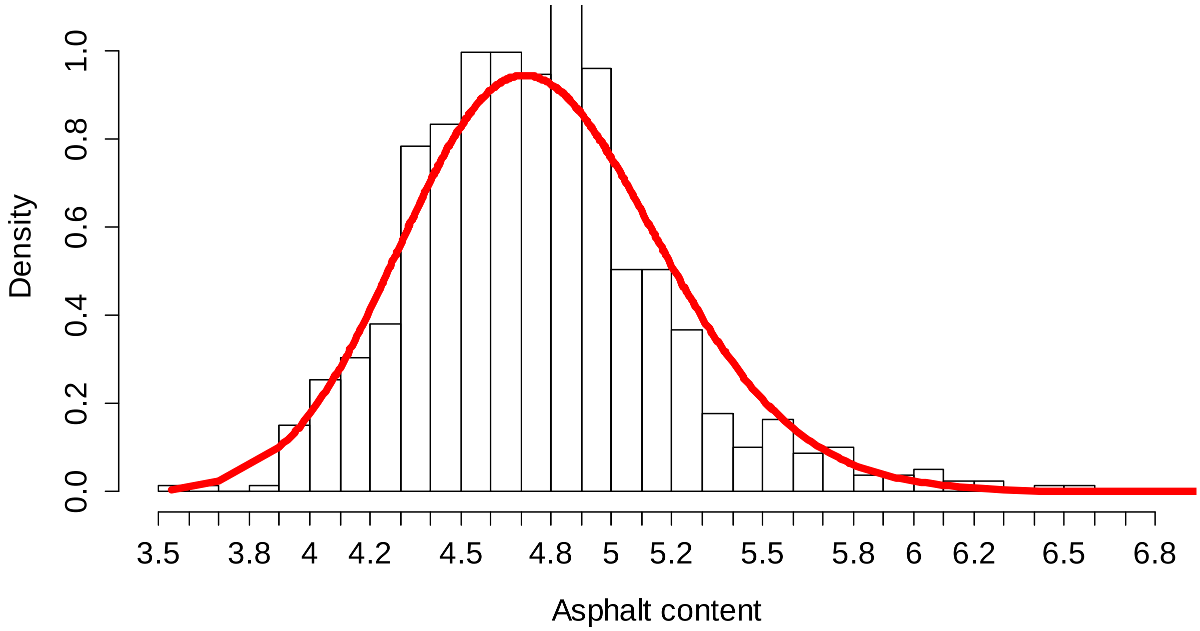

Figure 3.12: Actual and modeled distributions of asphalt content

An earlier TxDOT research project (Vazquez et al., 2010) emphasizes the influence of asphalt content on the fatigue performance of Type C mixes. Figure 3.13 shows a plot relating the asphalt content with number of cycles to failure in the bending beam apparatus for three different gradations confirming to the Type C specification (Item 341). The existing specification is providing a window of 0.6% asphalt content as a tolerable level. Figure 3.13 suggests a significant reduction in number of fatigue cycles corresponding to a reduction of 0.6% of asphalt content. This suggests that it is important to penalize or award contractors based on asphalt content as it significantly influences fatigue performance. Although a relationship is evident between asphalt content and laboratory cracking performance, a validation is necessary to ensure its correlation with field performance.

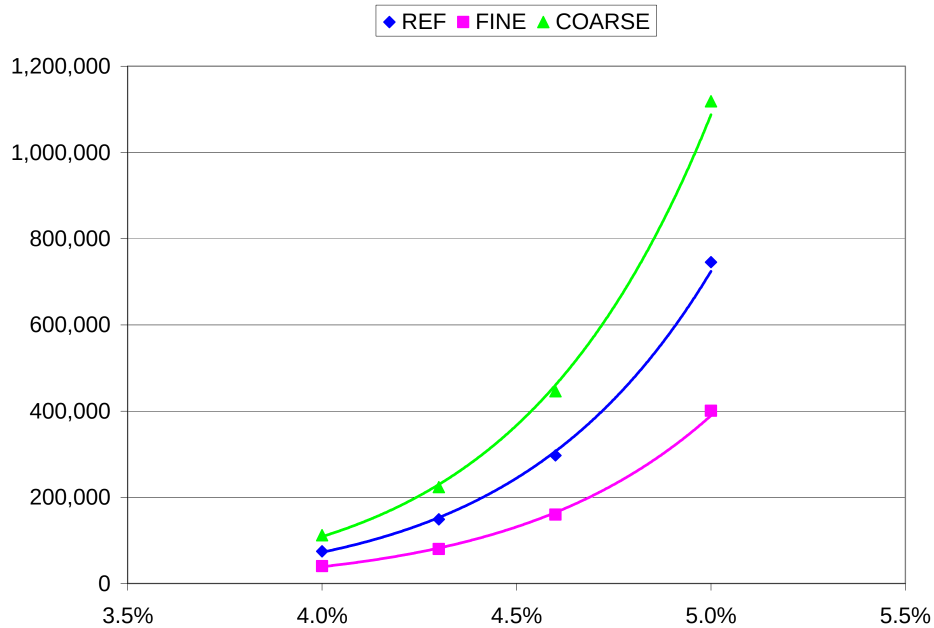

Figure 3.13: Relationship between asphalt content and fatigue life (Vazquez et al., 2010)

3.4.4 Voids in the Mineral Aggregate

VMAs are controlled during the production process, although it is not included in the current pay-adjustment system. The researchers explored the option of utilizing VMA of the field cored specimens as a surrogate for in-place air voids during placement QC. Because the asphalt content of these field-cored specimens is not available, the VMA of these specimens is calculated using the asphalt content measurements obtained during the production process; it is reasonable to assume that the asphalt contents in both field cores and the plant mix are similar if not identical. Thus, the VMA of the field-cores is a calculated by combining the production-specific asphalt content and the placement-specific in-place air voids. Note that the VMA of the field cores is merely a proxy for the in-place air voids for a given asphalt content.

Figure 3.14 shows the distribution of the project-level averages of the in-field VMA. Field VMA generally varied between 14.5% and 20.5%, which is typical for a dataset with predominantly Type C and Type D mixes.

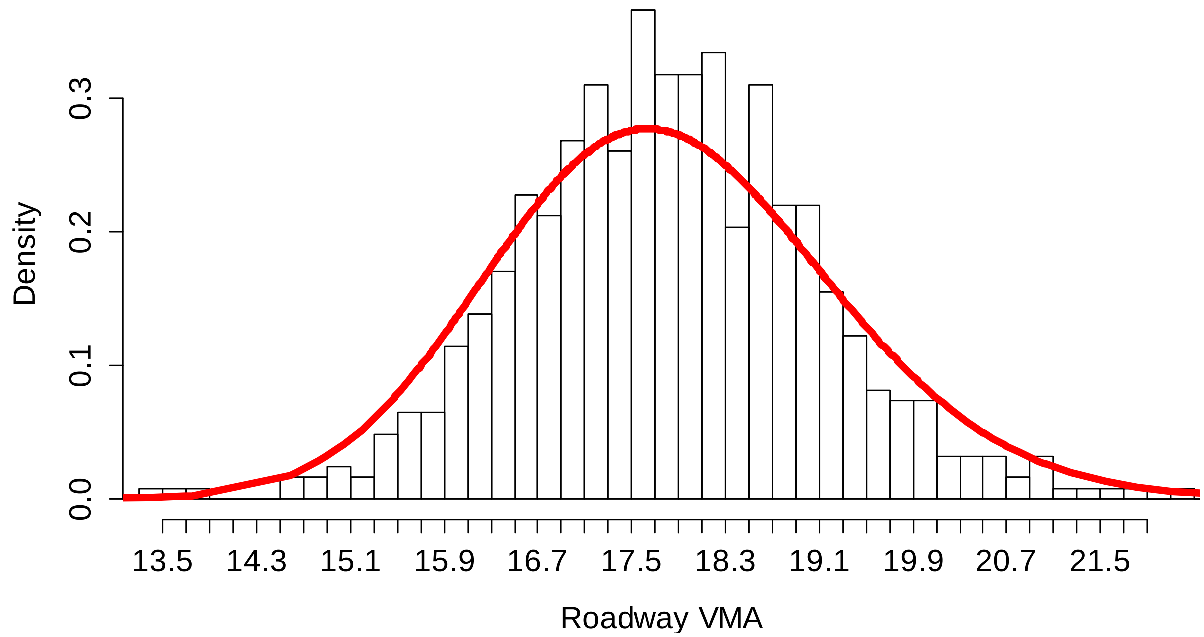

Figure 3.14: Actual and modeled distributions of in-field VMA

3.4.5 As-constructed Ride Quality (Asphalt Pavements)

TxDOT currently uses a pay adjustment schedule for concrete and HMA ride quality that aims at rewarding smoother surface finish. The calculation of ride quality pay factor is based on an immediate IRI measurement after construction. A look-up table is provided in Item 585 of the current TxDOT specification book (2004), which lists a fixed dollar amount corresponding to an IRI value on a 0.1-mile section. The IRI value on each 0.1-mile segment of the project is obtained and the corresponding dollar amounts are added to determine the ride quality bonus/penalty. All three schedules (1, 2, and 3) for ride quality use a linear relationship between ride quality and the bonus/penalty that is awarded to the contractor. According to TxDOT standards, an IRI measure of 95 inches/mile or more is considered unacceptable and, thus, requires some kind of maintenance intervention. It is therefore imperative to realize that the service life of a pavement is defined by the time it takes for the IRI to reach the terminal IRI value from its initial value after construction, until rehabilitation or maintenance occurs. According to the Item 585 specifications, TxDOT does not differentiate between an initial IRI value between 60 and 75 inches/mile for Schedule 2. However, it can make a significant difference as the time taken to reach the terminal IRI will be quite different in these two cases. Figure 3.15 shows a distribution of the as-constructed ride quality across various hot mix projects. The initial roughness values generally ranged from 30 to 145 inches/mile, indicating a wide range of ride quality (immediately after the construction) on hot mix pavements. The larger variability is due to the pooling of various hot mix project ride-quality values that fall under the three schedules. Figure 3.16, Figure 3.17, and Figure 3.18 shows ride quality distribution on hot mix projects constructed under Schedules 1, 2, and 3 respectively. Clearly the majority of the contractors are able to achieve smoother ride immediately after construction. Figure 3.19 shows a distribution of the awarded ride quality bonuses/penalties from the previous projects. A major portion of contractors received a ride-quality-related bonus, which is very similar to the previously mentioned pay factor schemes.

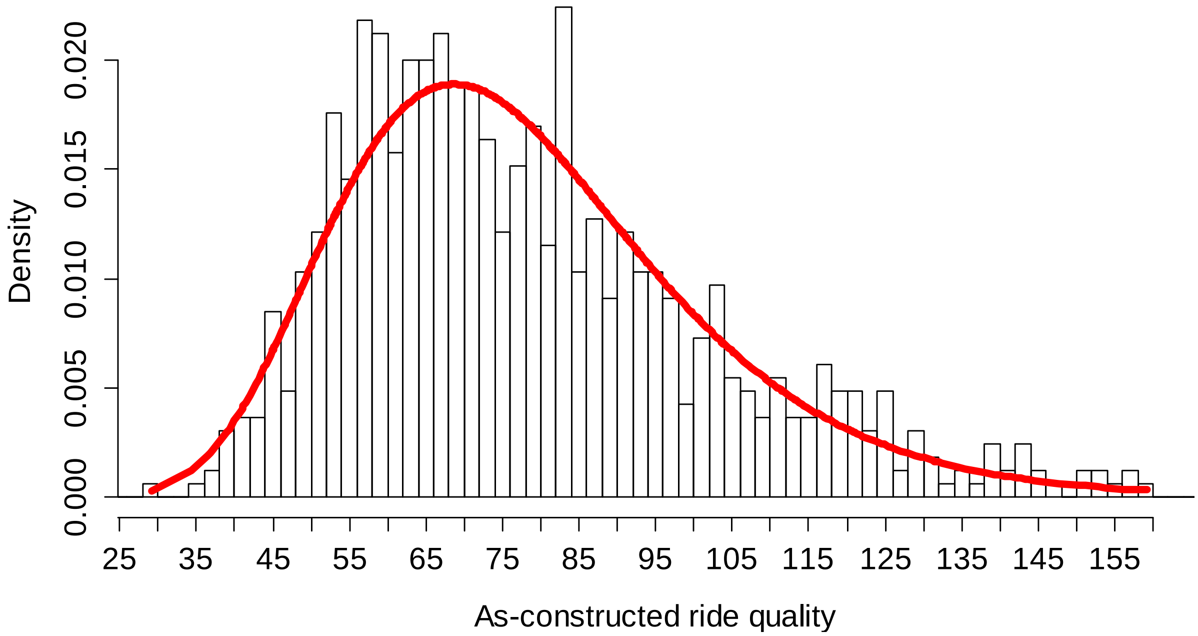

Figure 3.15: Actual and modeled distributions of as-constructed ride quality

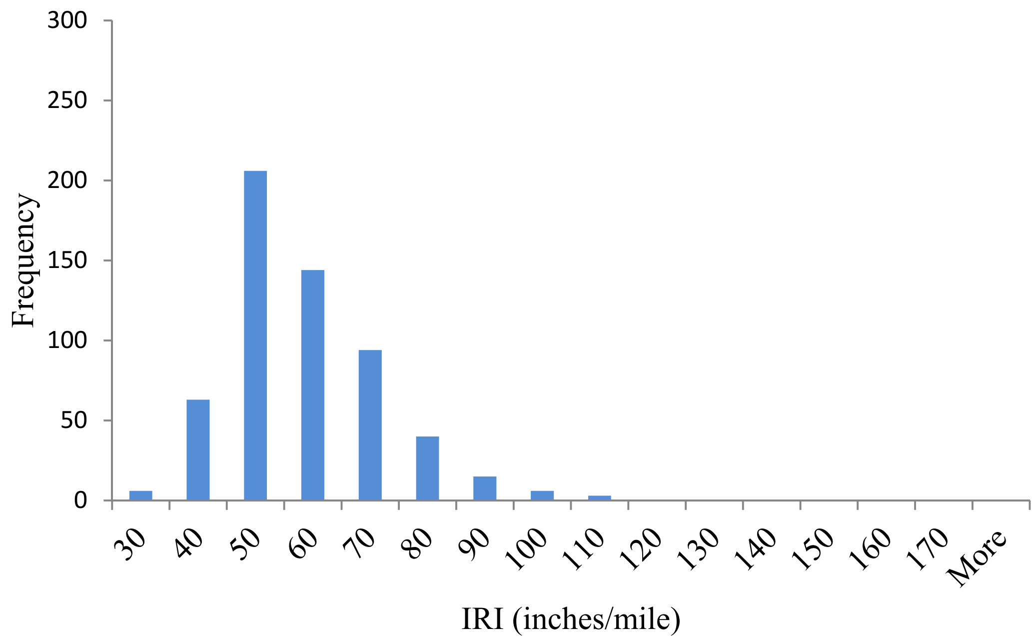

Figure 3.16: Distribution of as-constructed ride quality (HMA: Schedule 1)

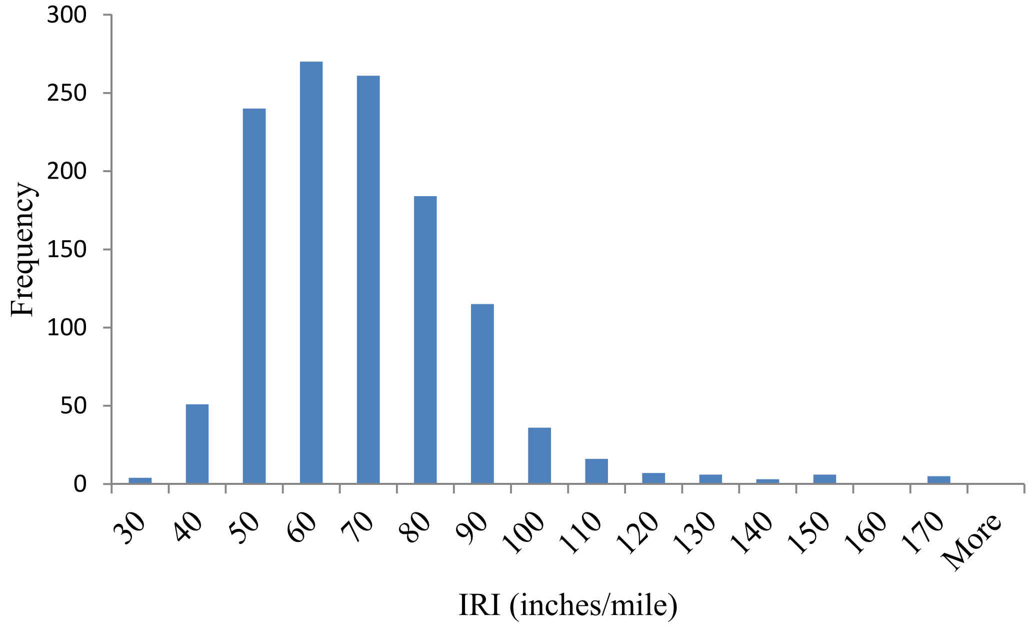

Figure 3.17: Distribution of as-constructed ride quality (HMA: Schedule 2)

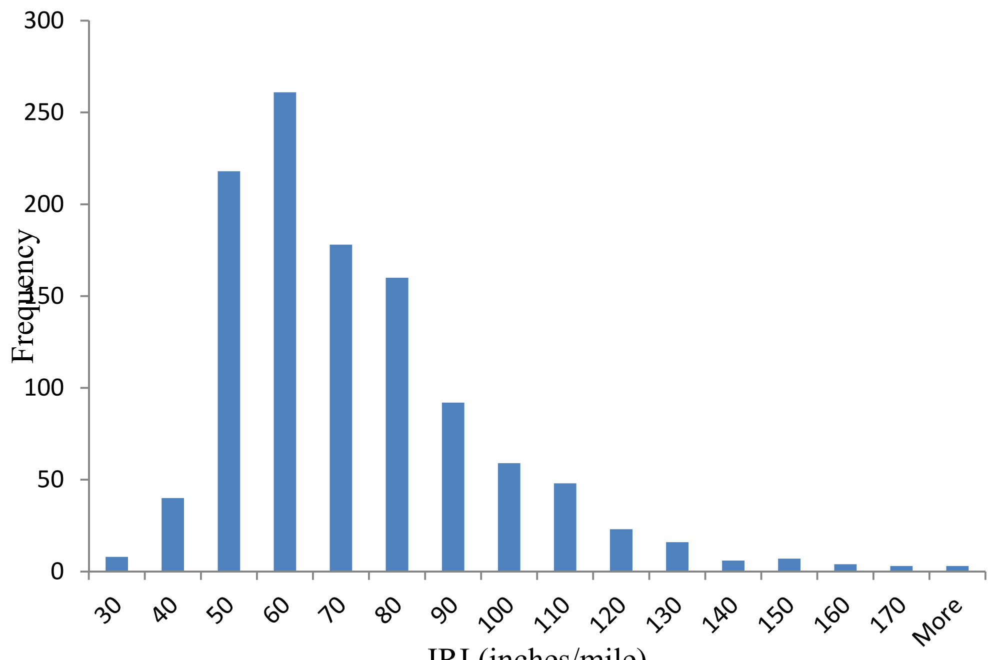

Figure 3.18: Distribution of as-constructed ride quality (HMA: Schedule 3)

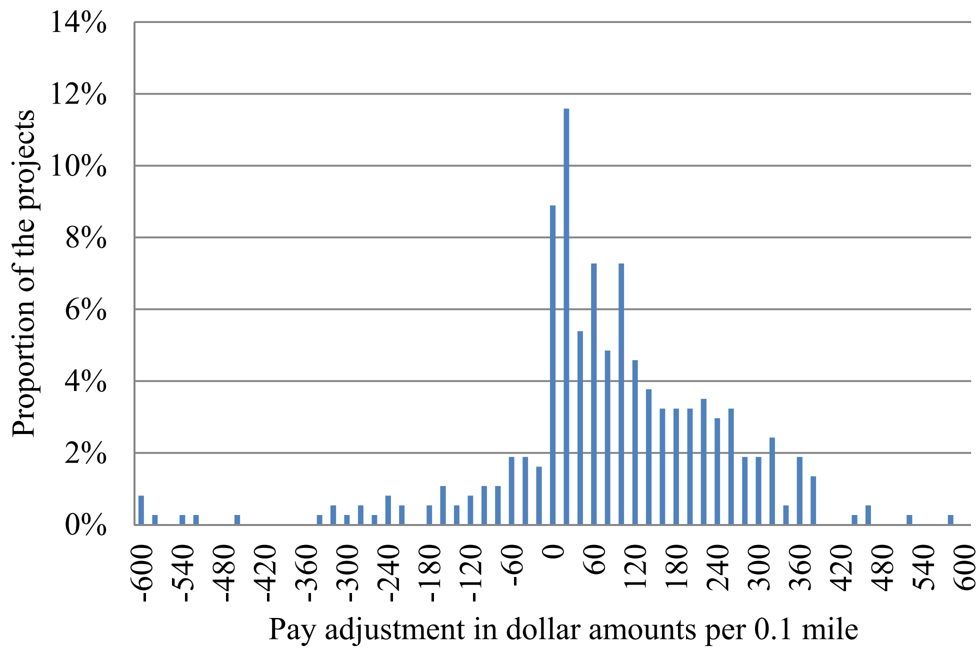

Figure 3.19: Ride quality pay factors

3.4.6 As-constructed Ride Quality (Concrete Pavements)

Despite consistent efforts throughout the project duration, the researchers were not able to gather adequate data resources and only limited data was available for the ride quality of concrete pavements. The researchers requested information from the concrete performance database maintained by Texas Tech University but this information was not forthcoming—the researchers were led to believe the database was still under development. A selected number of concrete pavements were obtained from the TxDOT SM database. A detailed description of the ride quality and PAFs is provided in Chapter 4. Figure 3.20 indicates the facility proportions of the selected 31 concrete projects identified for the analysis.

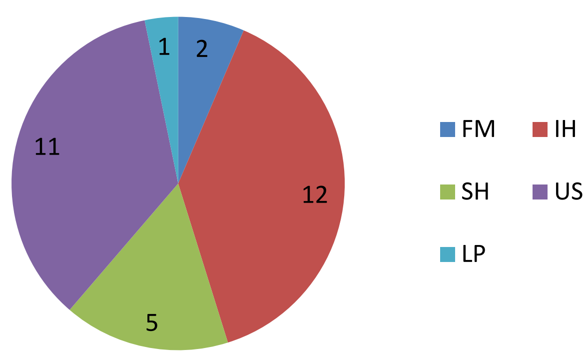

Figure 3.20: Concrete projects per facility

Chapter 4. Pay Adjustment Factor Models

As discussed in Chapter 3, a comprehensive dataset was developed as part of the study to investigate contractor tendencies and to examine the relationship between measurable construction QC parameters and the corresponding pavement performance. This chapter presents the adopted framework for developing this relationship and to quantify the sensitivity of several construction parameters on pavement performance. The chapter is broadly divided into two parts focusing on HMA-related and concrete-related pay adjustment specifications. Econometric techniques were employed for identifying the statistical association between the QC parameters (production, placement, and ride quality) and pavement performance regarding hot mix projects. Due to inadequate data resources, it was not feasible to establish an empirical relationship between the concrete ride quality and respective pavement performance. However, the researchers identified trends that are potentially useful, particularly in conjunction with adequate engineering judgment, in the development of a revised pay adjustment specification for the concrete projects. These trends are not supported by any statistical procedures due to scarcity of the concrete pavement data.

This chapter provides an overview of the methodology used for establishing the relationship between QC parameters and the corresponding pavement performance. The statistical modeling framework and estimation procedures employed by the study team are then presented, along with a comprehensive discussion on the model estimation results. A brief discussion on the strength of the statistically significant relationships that is based on variance-based sensitivity analysis and marginal effects is provided. A discussion on the concrete projects is included towards the end of the chapter. In summary, this chapter identifies the performance-related QC parameters for both hot mix and concrete projects. A revised pay adjustment specification was developed based on the performance-sensitive parameters discussed in a subsequent chapter.

4.1 Analysis Methodology

The primary intention behind rewarding or penalizing a contractor is to equally distribute the benefit or cost of a superior or inferior construction practice between TxDOT and the contractor. In general, TxDOT requires that the reward/penalty be determined during the project construction period—i.e., during (or at the end of) the contracting period. The agency obviously has no knowledge of the future performance of the as-constructed pavement at the end of the contracting period (or immediately after construction). However, the agency does have knowledge about the as-constructed quality of the pavement in terms of routinely measured quality parameters. Intuitively, the future performance of the pavement is dependent on the as-constructed quality. The agency may use the available knowledge on the QC parameters to draw conclusions about the future performance of the project, contingent on the existence of a relationship between these parameters. The researchers investigated the possibility of developing a relationship between future performance and routinely measured QC parameters. A methodology was developed to understand the existence and strength of those relationships by exploring available data and knowledge. The methodology requires gathering two types of information: 1) pavement performance information, and 2) as-constructed pavement quality information.

4.1.1 HMA Production, Placement, and Ride Quality

The study team developed a project-level deterioration index based on the rate of change of IRI over time to gauge pavement performance. The team chose IRI over other available performance indicators, such as cracking and rutting measures, due to the better reliability of the network-level IRI data. Equivalent deterioration indices that account for the progression of rutting and cracking distresses over time can easily be developed by reproducing the proposed methodology on the availability of reliable information. The project-level deterioration index is calculated by annually aggregating all the available performance data in terms of ride quality (in IRI) across the project. On the other hand, as-constructed pavement quality information, such as plant-produced mixture properties and field compaction characteristics, are typically recorded during the construction period on a daily basis. The daily measures of QC parameters are also aggregated to obtain a project-level measure of the as-constructed pavement quality. Additionally, an averaged IRI at the project level, which is typically measured after completion of pavement projects to assess the ride-quality immediately after the construction, is also employed as a QC parameter. Note that each pavement project is regarded as one observation in a project-level dataset. Subsequently, an empirical relationship between project-level performance and project-level QC parameters was developed using a Tobit (Type I) specification accounting for potential endogeneity. The sensitivity of each of the QC parameters that governs the performance was calculated using respective marginal effects. This methodology was used to identify the QC parameters that are statistically significantly associated with pavement performance as well as the strength of these respective relationships. These findings will be incorporated into the development of the revised pay adjustment specifications for production, placement, and ride quality of the HMA projects.

4.1.2 Concrete Ride Quality

The researchers attempted to identify an empirical relationship between the as-constructed ride quality of a concrete project and the corresponding field performance. It was not feasible to implement a procedure that is identical to the hot mix projects due to the unavailability of adequate datasets that would allow establishing statically significant findings. For this reason, the research team instead opted to learn from the available scarce data sources via manual inspection accompanied by minimal statistical analysis. The performance of the concrete projects was tracked using network-level ride quality information (IRI) as in the case of HMA projects; ride quality information is used due to the availability of reliable network-level ride quality data. The rate of change of the ride quality on the identified concrete projects was evaluated by inspection to identify any concrete projects deteriorating faster than is typical. The initial ride quality on these projects was also obtained using multiple data resources primarily from the SM database. A rough comparison between the initial ride quality and the respective rate of deterioration was reported. The empirical findings were combined with engineering judgment to develop recommendations for revising the concrete ride pay adjustment specification.

4.2 Analysis of HMA Projects

The researchers linked two pavement databases to form an integrated database for the study: 1) PMIS, the database of annual pavement performance, and 2) SiteManager (SM), the database of as-constructed pavement quality information. The integrated database includes project-level information such as the averaged construction QC parameters and the field performance of 824 HMA overlay projects located across Texas. Records with missing volumetric data were removed, leaving a total of 614 HMA overlay projects for analysis. This study emphasizes the need to control the quality of information and validate the SM database inputs. A detailed description of the dataset, including the various interesting characteristics and descriptive statistics, was provided in Chapter 3.

4.2.1 Correlation Analysis

A preliminary correlation study was carried out to understand any potential relationship between the field performance (rate of deterioration) and the QC measures. Figure 4.1 to Figure 4.5 show the correlation plots between various QC parameters and the rate of deterioration (in IRI). These graphs indicate that none of the QC parameters is perceivably related with field performance. It should be pointed out that similar poor correlations were found between the volumetric parameters and all other performance measures in PMIS, including rutting, cracking, and the various distress scores. Note that a pool of projects with typical quality is exhibiting relatively poor field performance. Apart from the as-constructed quality, several other factors potentially influence the field performance of the newly constructed pavements. The intricate interactions between as-constructed quality, traffic, and weather are possibly responsible for the obscure relationship shown in the correlation plots between the QC parameters and field performance. This shows that unveiling the empirical relationship between as-constructed quality and pavement performance is more complex than simple correlation analysis. Consequently, the need for sophisticated statistical techniques was identified in order to better understand the obscure datasets and to identify the underlying relationships between the as-constructed quality and the corresponding pavement performance. The following subsection describes the development of a statistical model that is most suitable for the current scenario.

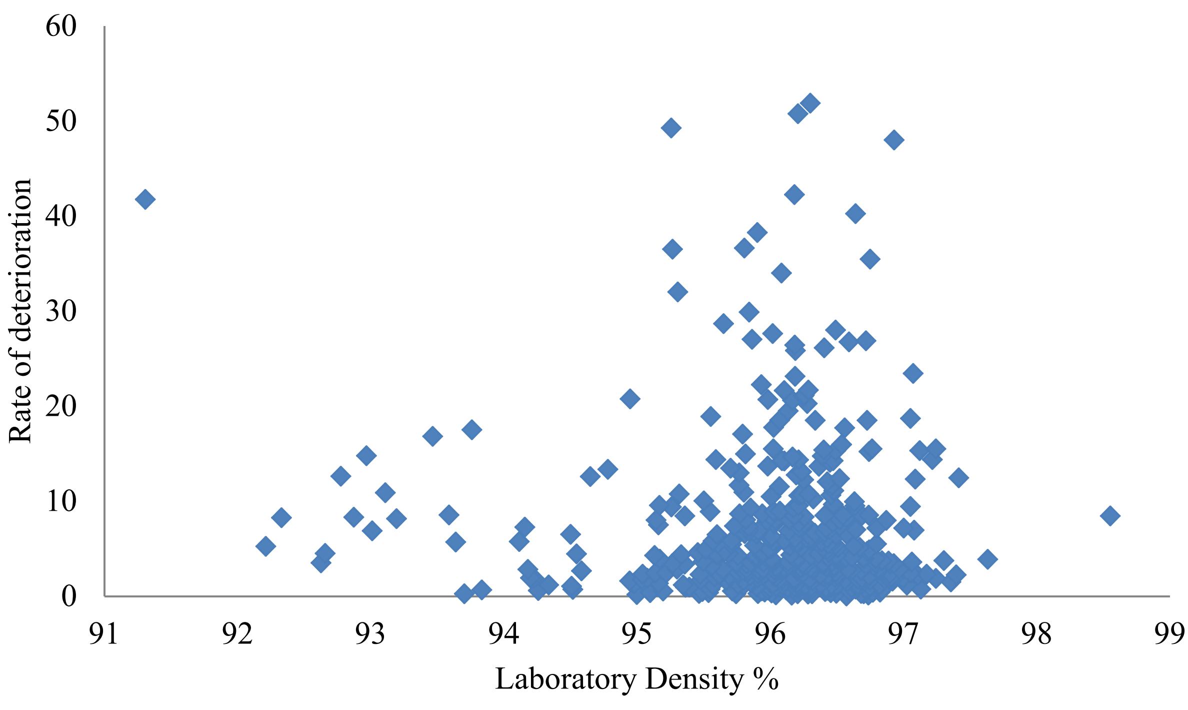

Figure 4.1: Correlation plot between laboratory density and deterioration rate

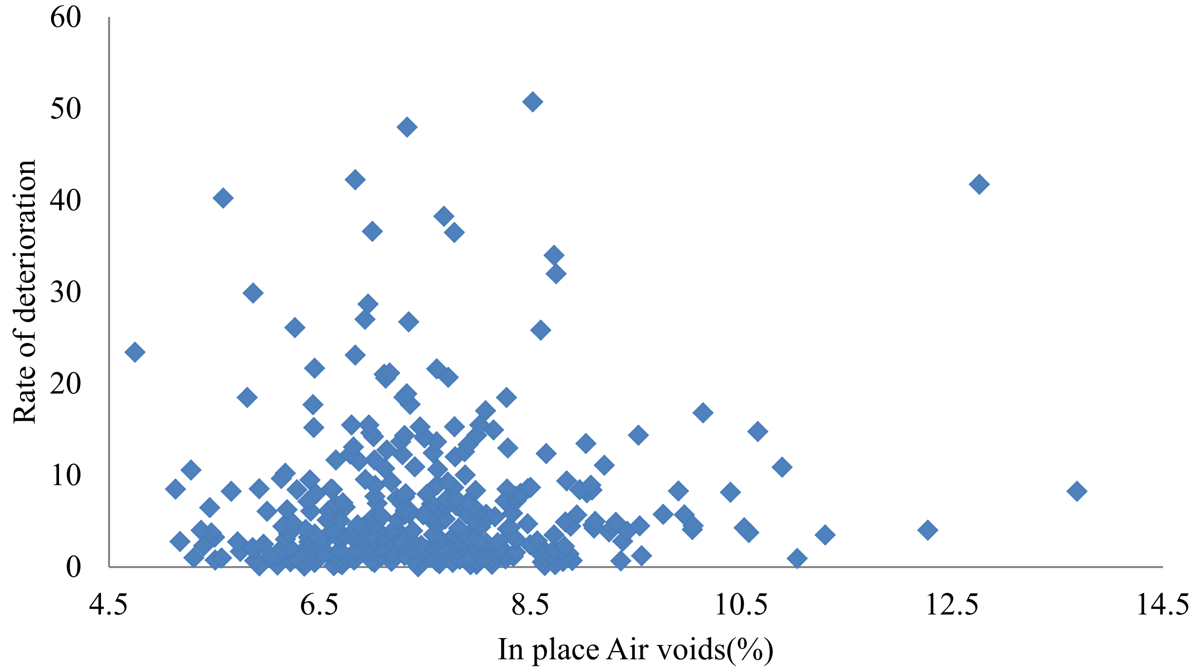

Figure 4.2: Correlation plot between in-place air voids and deterioration rate

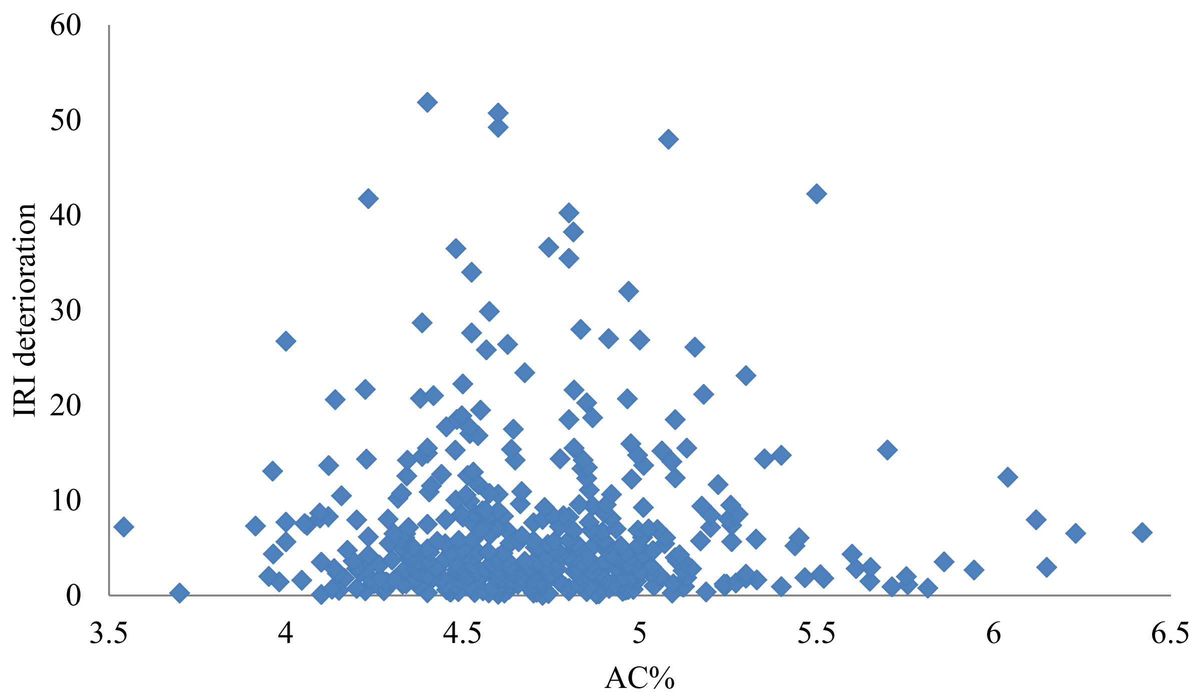

Figure 4.3: Correlation plot between asphalt content (%) vs. IRI rate of progression

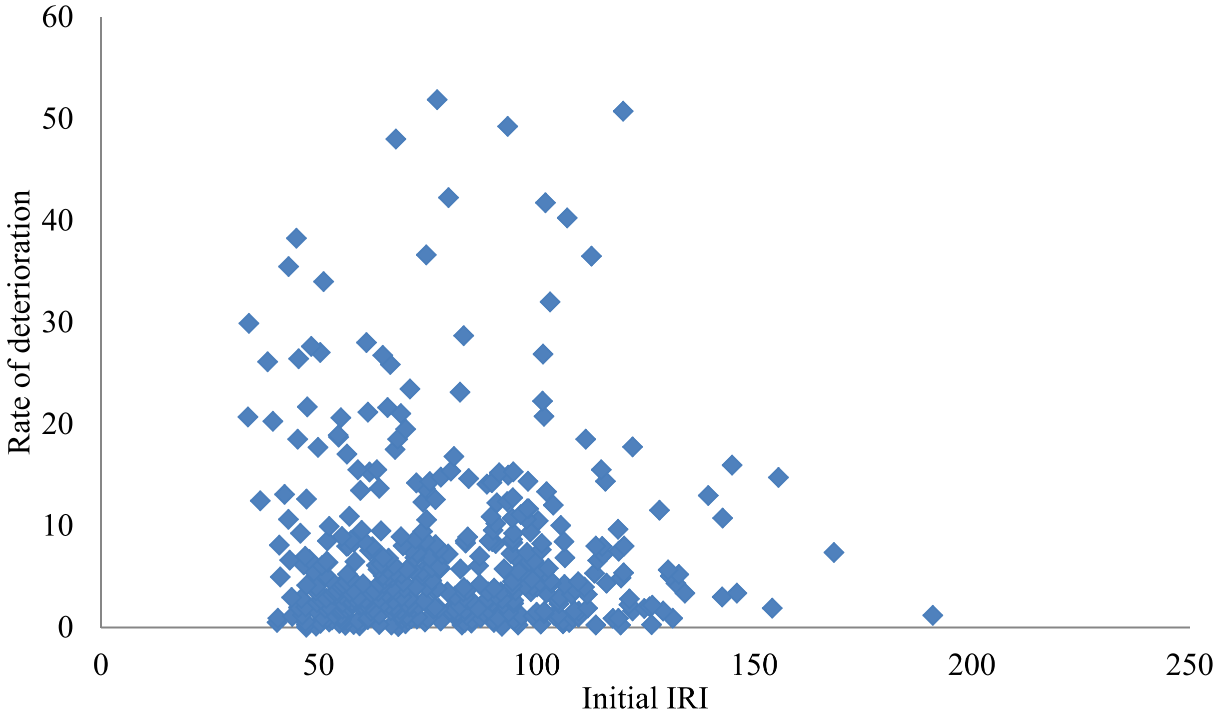

Figure 4.4: Correlation plot between initial IRI vs. IRI rate of progression

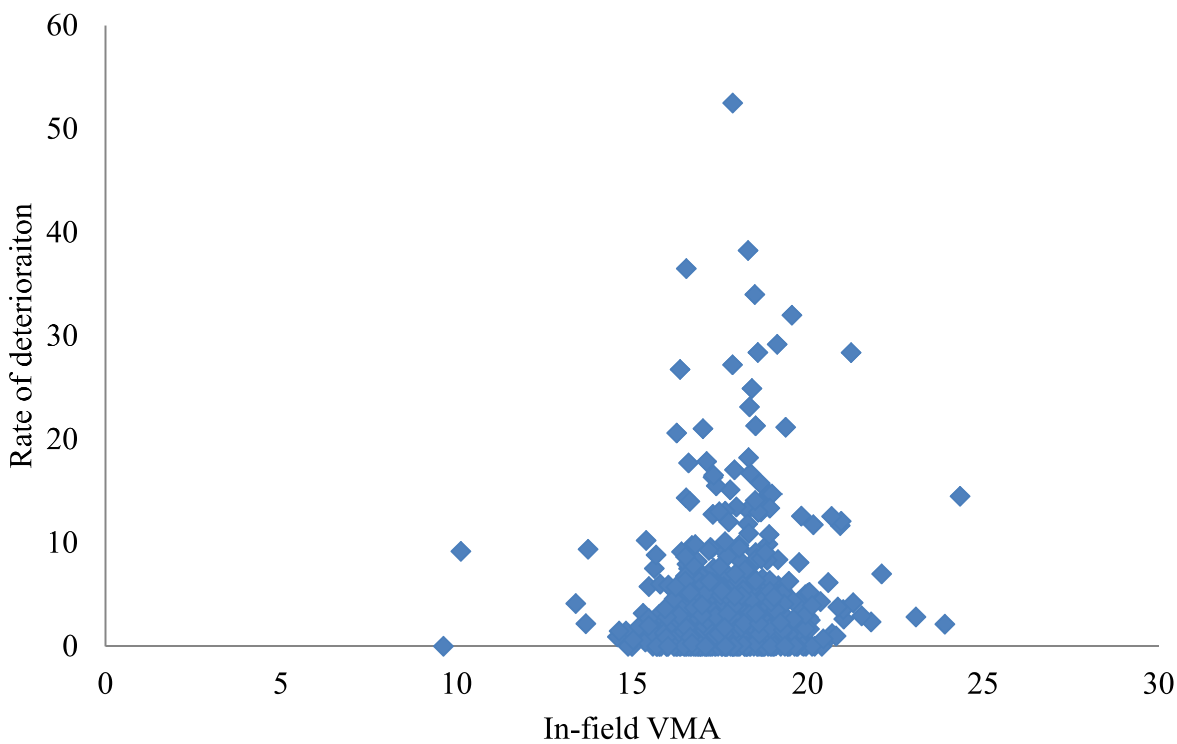

Figure 4.5: Correlation plot between in-field VMA vs. IRI rate of progression

4.2.2 Model Development

The overall goal of the statistical model development exercise is to predict the pavement field performance (i.e., deterioration rate in terms of ride quality) of a project using the project-specific data—such as volumetric properties (QC/QA), ride quality, traffic, etc.—as explanatory variables. Consequently, the empirical predictive model may be used to ascertain the underlying relationship between the QC parameters and pavement performance. It is important to understand the underlying distribution of the dependent variable (deterioration rate) as it plays a vital role in the selection of the model structure. In this study, the deterioration rate (dependent variable) takes either a zero or non-zero value, which corresponds to projects not showing or showing signs of deterioration during the analysis period. A Type I Tobit model structure was initially selected, typically used for handling dependent variables dominated by a particular response (zero in this case); these are so-called *corner solution problems* in econometrics. A standard Type I Tobit model can be written as shown in Equation 4.1–4.5.

$$y_i = \max(0, y_i^*) \qquad (4.1)$$

$$y_i^* = X_i\beta + u_i \qquad (4.2)$$

$$u_i \sim \text{Normal}(0, \sigma^2) \qquad (4.3)$$

where:

y_i: Observed deterioration rate of i^{th} project
y_i^*: Latent deterioration rate
X_i: Vector of i^{th} project attributes
β : Vector of regression coefficients
u_i: Idiosyncratic error term
σ: Standard deviation of the error term

$$E(y_i|y_i > 0) = E(y_{1i}|y_{1i} > 0) = \mu + \sigma\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \qquad (4.4)$$

$$P(y_i = 0|X) = \Phi(X_i\beta) \qquad (4.5)$$

The idiosyncratic error term associated with the dependent variable compounds the unobserved attributes that may potentially influence the deterioration rate apart from the effect of observable characteristics of the underlying pavement (pooled into the matrix X_i). The features of the pre-existing pavement (prior to construction) influence its deterioration rate as well, but they are unfortunately unobservable; thus their effect is being compounded into the idiosyncratic error term. Also, the characteristics of the existing pavement govern the initial ride quality to a large extent. Therefore, there is a chance of correlation between initial IRI and the idiosyncratic error term, leading to potential violation of the underlying assumptions of the aforementioned model (shown in Equations 4.1, 4.2, and 4.3). Additionally, the initial IRI and the deterioration rate were estimated by regressing the annual PMIS measurements of ride quality (IRI) over time. This also potentially induces a correlation between the estimated initial IRI and the unexplained portion of the respective deterioration rate (in other words, model error). Existence of correlation between the explanatory variables and the idiosyncratic error is called *endogeneity*; ignorance of the underlying endogeneity produces a bias in the estimates of the model parameters. For these reasons, the potential endogeneity associated with initial IRI was accounted for using endogenous Tobit specifications that are typically used in econometrics. The aforementioned model is reformulated into a two-stage specification using selected instrumental variables; the latent deterioration rate y_i^* is modified as shown in Equation 4.6–4.8:

$$y_i^* = y_{2i}\alpha_1 + X_{1i}\beta_1 + \varepsilon_i \quad (4.6)$$
$$y_{2i} = X_{1i}\gamma + X_{2i}\beta_2 + \delta_i \quad (4.7)$$

$$(\varepsilon_i, \delta_i) \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\tau_1^2 & \eta_1\\ \eta_1 & \tau_2^2\end{bmatrix}\right) \quad (4.8)$$

where:

X_{1i}: Vector of i^{th} project attributes
β_1: Vector of regression coefficients corresponding to X_1
γ: Vector of regression coefficients corresponding to X_1 while using as an instrument
X_{2i}: Vector of i^{th} project instruments
β_2: Vector of regression coefficients corresponding to X_2
y_{2i}: Endogenous covariate (natural logarithm of initial IRI)
ε_i: Error term in the structural model (Equation 1.4)
δ_i: Error term in the reduced form model (or first-stage model in Equation 1.5)
τ_1^2: Variance of ε_i
τ_2^2: Variance of δ_i
η_1: Covariance of ε_i and δ_i

It was important to ensure that the dataset respects the underlying statistical assumptions of the proposed model. For instance, a test of endogeneity was carried out to statistically support the necessity of the two-stage model structure; the results are shown in the subsequent subsections. Also, it is important to include the exogenous attributes that are directly affecting the deterioration rate (i.e., X_1 in Equation 4.5) as instruments in the first-stage regression along with the other potential instruments (i.e., X_2 Equation 4.5). Omission of the exogenous attributes as instruments in the first-stage model for endogenous variable (i.e., y_{2i}) generates bias in the coefficients of the exogenous attributes as well as instrumented variables (i.e., initial IRI) in the main model. The natural logarithm of the initial IRI was utilized in the proposed model; this is solely to ensure non-negativity of the IRI data.

Both magnitude and the corresponding standard errors of the aforementioned Tobit model parameters were obtained through maximum likelihood (ML) estimation. ML estimation of the model simultaneously estimates both first- and second-stage regression coefficients, unlike two-stage estimation techniques such as Newey's estimator. The implementation of the ML procedure was relatively faster, as the number of endogenous variables is not more than one; otherwise, a two-stage estimation procedure such as Newey's estimator would be a better choice in terms of convergence efficiency and computational speed. A final specification was chosen carefully based on a rigorous model development process. Model refinement was carried out through exclusion of statistically insignificant variables by following standard step-wise procedures and statistical tests (e.g., F-test). A careful and intuitive selection of instruments is crucial in the model refinement process. Practical considerations played a role in the removal of insignificant variables, rather than solely adopting a statistics based mechanical approach. Table 4.1 presents the final specification estimates of the Tobit regression with endogenous covariates.

Table 4.1: Model estimation results

Model	Variable Description	Mean	Std.er	Z	P > Z
Deterioration (Main Tobit Regression)	Initial IRI	9.00	3.13	2.87	0.00
	Indicator Variable: PG 76	1.37	0.58	2.34	0.02
	Laboratory Density	-2.04	0.64	-3.21	0.00
	Analysis Period	-0.71	0.19	-3.84	0.00
	Indicator Variable: Low Speed Facility	2.63	1.30	2.02	0.04
	Indicator Variable: IH Facility	-2.30	0.77	-2.98	0.00
	Indicator Variable: Rural area	1.49	0.64	2.33	0.02
	Constant	163.65	65.11	2.51	0.01
Initial IRI (First-stage)	Indicator Variable: PG 76	-0.004	0.022	-0.18	0.86
	Laboratory Density	0.018	0.025	0.72	0.47
	Analysis Period	-0.031	0.007	-4.57	0.00
	Indicator Variable: Low Speed Facility	0.103	0.046	2.25	0.02
	Indicator Variable: IH Facility	-0.010	0.037	-0.29	0.78
	Indicator Variable: Rural area	-0.059	0.024	-2.43	0.02
	Asphalt Content	-0.070	0.029	-2.46	0.01
	In-field VMA	0.025	0.009	2.94	0.00
	Maintenance Cost Per Mile (in million $)	1.599	0.000	3.97	0.00
	Annual Average Daily Traffic (in millions)	2.491	0.000	3.95	0.00
	Traffic Load Estimate	-0.003	0.001	-5.21	0.00
	Indicator Variable: Project Length < 2 miles	0.119	0.023	5.21	0.00
	Constant	2.947	2.384	1.24	0.22
Other Model Parameters	τ_1	6.71-			
	τ_2	0.259			
	η_1	-0.818			

Note: Number of Observations: 614
113 left censored at 0 and 501 uncensored observations
Instrumented: Initial IRI
Likelihood at Convergence: -1740.66
Wald Test of exogeneity: Chi-Square(1) = 13.5; Prob > Chi-Square = 0.00

The marginal effects corresponding to $P(y=0 \mid X)$ and $E(y|y>0,X)$ are also calculated using the Delta-method and reported in a later portion of this document

4.2.3 Results and Discussion

A brief discussion on both magnitude and sign of the estimated coefficients corresponding to the final specification shown in Table 4.1 is provided in this subsection.

First-Stage Regression

First-stage regression involves modeling of the endogenous covariate—i.e., initial IRI using selected instrumental variables, such as, X_2 and the other exogenous variables, such as X_1. The magnitude and sign of the coefficients corresponding to the instrumental variables (i.e., X_2) are of primary interest within this model, while the coefficients of exogenous variables are included only to avoid potential bias.

First, among all the QC parameters, the average asphalt content of the plant-produced mix (measured using the ignition oven) and in-place VMA (measured using field cores after the placement operation) were statistically significant instruments—i.e., they were indirectly associated with the deterioration rate through their effect on as-constructed ride-quality or initial IRI. The statistically significant negative coefficient corresponding to the average asphalt content of the plant-produced mix indicates that pavement sections constructed with plant mixes that are richer in asphalt content appear to be associated with smoother post-construction ride and vice versa. Intuitively, richer mixes are more workable and easier to compact; this is supported by the empirical evidence. The positive sign on the coefficient corresponding to the in-place VMA indicates that pavements with higher overall in-place VMA (i.e., under-compacted during placement operation) tend to be rougher immediately after the construction. It should be noted that the applicability of these findings is limited to the narrow range of the existing hot mix specifications. Extrapolation of these findings outside this range is uncertain and leads to erroneous conclusions. The reported empirical association is encouraging for the development of PRSs. In summary, a few of the QC parameters that are routinely measured are associated with the deterioration rate through their association with the IRI immediately after construction.

Second, the positive coefficient corresponding to the annual maintenance cost shows that pavement sections with higher annual maintenance cost per unit mile are associated with rougher post-construction surface. The annual maintenance cost is the cumulative sum of the various maintenance costs accumulated during a one-year period, typically reported in the PMIS database. A higher annual maintenance cost indicates a problematic pavement in need of maintenance, explaining the reported positive association.

Third, traffic volume as well as load both appear to be associated with the initial IRI. The HMA pavements constructed on facilities carrying higher traffic volumes tend to be rougher immediately after construction. On the other hand, pavements constructed on facilities carrying heavier traffic tend to be smoother immediately after construction. Pavements carrying heavier traffic are typically better designed, with a more structurally sound underlying foundation, and thus are associated with a lower initial IRI. It is important to note that the traffic volume and load being considered are averaged across the analysis period following the construction year. Also, the traffic load and volume are acting as proxies for the strength of the underlying pavement and thereby affecting the post-constructed initial IRI.

Finally, the data suggested that projects that are less than 2 miles long are typically associated with larger initial roughness values immediately after construction relative to longer projects. Intuitively, a longer project stabilizes the plant and field operations, resulting in a superior initial IRI, which is supported by the empirical finding. An interesting contractor tendency is reflected in the positive coefficient corresponding to the indicator variable related to the length of the pavement project. Contractors may tend to allocate better resources to longer projects in order to receive a larger bonus by providing a smoother initial ride quality.

Main Tobit Regression

Second-stage regression involves modeling of deterioration rate utilizing the instrumented endogenous variable—i.e., predicted initial IRI (using the first-stage model)—and the other exogenous variables (i.e., X_1) as covariates. The intuition behind the magnitude and sign of these coefficients corresponding to statistically significant covariates is described below.

Data suggested that among the measurable as-constructed QC parameters, the initial ride quality measured immediately after construction and a production-related QC variable, laboratory density, are directly associated with the field performance or rate of deterioration. *First*, the positive sign of the coefficient corresponding to the initial IRI indicates that HMA pavements with larger initial IRI tends to deteriorate faster and vice versa, everything else remaining unchanged. The empirical evidence of this relationship is encouraging and supports the current TxDOT contractual specifications, which reward the contractor providing a pavement with lower initial surface roughness. *Second,* the negative sign on the coefficient corresponding to the laboratory density indicates that plant-produced mixes with higher laboratory density are associated with a slower rate of deterioration and vice versa. It should be noted that the applicability of this finding is limited to the narrow range of the existing hot mix specifications. Extrapolation of these findings outside this range is uncertain and leads to erroneous conclusions. This is an important finding regarding the implementation of the specification. It is encouraging to see this relationship, as the current TxDOT PAF system controls the production process during pavement construction by rewarding/penalizing the contractor based on laboratory density.

The positive sign corresponding to the PG grade indicator variable indicates that stiffer binders are associated with faster deterioration rates, keeping everything else fixed.

The available performance data following the construction year is used in the analysis, which results in dissimilar analysis periods for different HMA pavement projects. The analysis period is included in the model to normalize against such dissimilarity among the individual projects. The magnitude and coefficient of the covariate corresponding to analysis period is not of any interest, although it is important to include the analysis period in the model as a control variable.

The results suggest that low-speed facilities (with speed limits of less than 45 mph) are apparently deteriorating faster than the high-speed facilities. The negative coefficient corresponding to the facility indicator indicates that Interstate Highways are deteriorating slower than the other facility types, keeping everything else fixed. Intuitively, Interstate Highways are designed to higher standards and tightly controlled for longer duration with structurally sound underlying pavement structures; thus, their deterioration is expected to be slower. The positive coefficient of the indicator variable corresponding to the rural indicator indicates that the rural HMA pavement projects are deteriorating faster than urban projects.

In summary, this subsection highlighted the empirical evidence for the existence of a relationship between measureable QC parameters and field performance. In order to qualify a QC parameter as a candidate to potentially include in PRSs, a statistically significant relationship with field performance is necessary. This significance, however, must be coupled with a significant influence to be of practical value. This is measured using a variance-based sensitivity analysis as described below.

4.3 Sensitivity Analysis

In order to build a PRS, it is critical to identify the construction QC parameters that are most sensitive to performance. Regarding HMA-related specifications, a variance-based sensitivity analysis was implemented to calculate marginal effects to identify the performance-sensitive QC parameters.

4.3.1 Variance-based Sensitivity Analysis

The goal of a variance-based sensitivity analysis is to identify the attributes (i.e., QC parameters) that explain the maximum variation in the response variable, which can be a performance measure of the respective project. Previously, the empirical relationship between the QC parameters and the respective field performance of HMA projects was established using econometric techniques. This empirical relationship or statistical model serves as the underlying mathematical model to perform the sensitivity analysis on statistically significant QC parameters.

Firstly, a performance measure is defined as the expected rate of deterioration of the ride quality measured using IRI in inches per mile based on the pool of projects that exhibited some signs of early deterioration; this is termed *performance index* in this document. In order to estimate variance-based sensitivity, knowledge of the distributional characteristics of the QC parameters is essential. The distributions of the statistically significant explanatory variables in the aforementioned empirical relationship are modeled using the most reasonable probability distributions; the distributional parameters are obtained using ML estimation. QC parameters and the project-specific features are simulated using these distributions, thereby creating 10,000 virtual HMA projects. The projects are regarded representative as they are being generated/simulated based on the real HMA project data features. The performance index is calculated for each of these 10,000 HMA projects using the estimated model parameters. As the performance index is a non-linear function of the QC parameters, a polynomial surface is created using the statistically significant attributes and the corresponding performance index as the response variable. A coefficient of variation based sensitivity index is calculated for each of the QC parameter and subsequently, QC parameters that are sensitive to the performance are identified. A revised PAF system may be developed by combining the results of the sensitivity analysis as well as the previously described empirical relationships.

Structural and Reduced Form Equations

An insight into the causal structure of the aforementioned model is essential to examine the influence of the explanatory variables or project attributes on the corresponding performance. Figure 4.6 shows the layout of the underlying causal structure of the Tobit endogenous regression model.

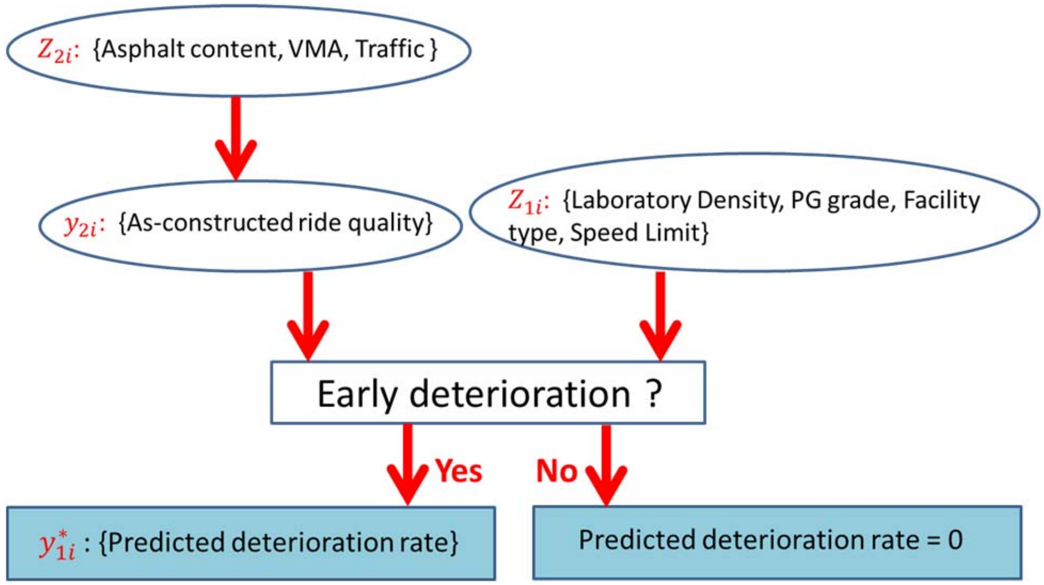

Figure 4.6: Insight into the model structure

The set of explanatory variables contained in X_{2i} act as instrumental variables or predictors of variable y_{2i} (or initial ride quality) in the second stage linear regression model (see Equation 4.7). Also, X_{1i} and y_{2i} are predictors of the deterioration rate, conditional on the event that the associated project exhibited a positive deterioration within the analysis period. In other words, the model predicted deterioration rate (in Equation 4.6) is directly dependent on the set of attributes X_{1i} and indirectly dependent on the set of attributes X_{2i}. It is to be noted that the influence of y_{2i} cannot be studied under the existing causal network of the model, because y_{2i} is being generated by the model using X_{2i}as the predictors. To reiterate, the goal of this sensitivity analysis is to study the influence of X_{1i}, X_{2i} and y_{2i} (or ride quality) on y_{1i} (deterioration rate) for any given i^{th} HMA project. Two different versions of the proposed model were constructed in order to perform the desired sensitivity analyses, while respecting the underlying casual structure of the model; a structural form and reduced form.

Structural Form

The structural model is formulated by excluding the second level regression model that involves instrumental variables (Equation 4.7); in other words, the set of instruments denoted by X_{2i} are excluded from the causal structure described earlier. Therefore, X_{1i} and y_{2i} directly influence the response variable y_{1i}(or deterioration rate); see Equations 4.9–4.11.

$$y_i = \begin{cases} y_i^* \ \ if \ y_i^* \ > 0 \\ 0 \ \ if \ y_i^* \ \leq 0 \end{cases} \quad (4.9)$$

$$y_{1i}^* \ = y_{2i}\alpha_1 + X_{1i}\delta_1 + u_i \quad (4.10)$$

$$u \sim N(0, \tau_1^2) \quad (4.11)$$

The structural model allows studying the direct influence of the ride quality (y_{2i}) and the set of explanatory variables represented by the X_{1i}. It is important to note that the estimates of the regression coefficients (i.e., α_1 and δ_1 in the structural model) are originally obtained using ML estimation of the full model described by set of equations (4.1, 4.6, and 4.7); this accounts for the potential endogeneity bias. The performance index is calculated using Equation 4.12.

$$E(y_i|y_i > 0) = E(y_{1i}|y_{1i} > 0) = \mu + \sigma\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \tag{4.12}$$

where:

$$\mu = y_{2i}\alpha_1 + Z_{1i}\delta_1$$

$$\sigma = \tau_1$$

Reduced Form

The structural model excludes the instrumental variables within the model structure and does not allow studying the influence of the instrumental variables on the deterioration rate. The instrumental variables are included in the original model (see Equations 4.1, 4.6 and 4.7); however, the direct influence of these variables on the deterioration cannot be studied. The reduced form model essentially combines the first and second stage instrumental regression models into one model, which allows studying the direct influence of the instrumental variables. Therefore, $X_{1\text{i}}$ and $X_{2\text{i}}$ directly influence the response variable $y_{1\text{i}}$ (or deterioration rate); see Equation 4.13–4.16.

$$\text{y}_\text{i} = \begin{cases} \text{y}_\text{i}^* & \text{if } \text{y}_\text{i}^* > 0 \\ 0 & \text{if } \text{y}_\text{i}^* \leq 0 \end{cases} \tag{4.13}$$

$$\text{y}_{1\text{i}}^* = (\text{Z}_{2\text{i}}\delta_2 + \text{v}_\text{i})\alpha_1 + \text{Z}_{1\text{i}}\delta_1 + \text{u}_\text{i} \tag{4.14}$$

$$\text{y}_{1\text{i}}^* = (\text{Z}_{2\text{i}}\delta_2)\alpha_1 + \text{Z}_{1\text{i}}\delta_1 + \text{w}_\text{i} \tag{4.15}$$

$$\text{w}_\text{i} \sim \text{N}(0,\ \tau_1^2 + \alpha_1^2\tau_2^2 + 2\alpha_1\eta_1) \tag{4.16}$$

The second level of the regression equation involving the instrumental variables is directly substituted into the first level equation. The error term associated with the second-level equation (i.e., δ_i) is being pooled into the error term associated with the first-level equation upon the substitution (Equation 4.16). Therefore, the variance of the error term in the reduced form model (see Equation 4.14) is inflated as shown in Equation 4.16. The performance index is calculated using Equation 4.17.

$$E(y_i|y_i > 0) = E(y_{1i}|y_{1i} > 0) = \mu + \sigma\frac{\phi\left(\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} \tag{4.17}$$

where:

$$\mu = (Z_{2i}\delta_2)\alpha_1 + Z_{1i}\delta_1$$

$$\sigma = \tau_1^2 + \alpha_1^2\tau_2^2 + 2\alpha_1\eta_1$$

4.3.2 Model Implementation

The primary goal of a variation-based sensitivity analysis is to identify the model attributes that explain the maximum variation in the response variable. The performance index is a non-linear function of various model attributes both in the structural and reduced form models outlined above (see Equations 4.12 and 4.18). A distribution of the performance indices may be obtained for the different QC parameters and project-specific attributes (traffic, facility type, etc.). It is important to ensure that the sensitivity analysis is run using realistic project attributes; for instance, the volumetric properties should be within practical limits. For this reason, the distributions of the QC parameters and other model attributes were evaluated using the actual dataset in greater detail. Empirical distributions are constructed for each of the QC parameters by fitting the most suitable probability distributions to the observed data using the classical ML estimation approach. Figures 4.7 to 4.11 indicate histograms based on the real dataset superimposed on the simulated data. Additionally, any potential correlations (based on sample correlations) among the QC parameters were evaluated and accounted for using bivariate joint distributions. For instance, asphalt content and in-field VMA were moderately correlated (see Figure 4.12), which will be accounted for within the total sensitivity index. These empirically constructed probability distributions allow simulating a large number of virtual projects (20,000) with a wide range of QC parameters; indeed the simulated projects are realistic as they are constrained by the information in the observed dataset. Performance index is calculated for each of the simulated projects by using the respective QC parameters in the corresponding equation.

A polynomial basis function is used to develop a linear relationship between the performance index and the model attributes including QC parameters. The reason for simulating 20,000 virtual projects is to build a fairly accurate linear approximation of the relationship between performance index and the model attributes. The coefficients of the polynomial basis function are estimated using multiple regression analysis and coefficients of determination are stored (R_X^2: see Equation 4.19). In order to calculate the sensitivity of the i^{th} model attribute, another polynomial basis function including all the model attributes except the i^{th} attribute is estimated using multiple regression analysis and the corresponding coefficient of determination is also stored ($R_{X_{\sim i}}^2$: see Equation 4.20). The sensitivity of the model attributes is calculated using total-sensitivity index (see Equation 4.20), a commonly used sensitivity index. The index accounts for the potential correlation among the model attributes; it was essential to account for the correlation among QC parameters for this dataset.

Regression with X :

$$Y = \beta_0 + x_1\beta_1 + x_2\beta_2 + \cdots \dots \dots x_2^2\beta_1 + \dots . \tag{4.18}$$

Regression with $X_{\sim i}$:

$$Y = \beta_0 + x_1\beta_1 + \cdots . x_{i-1}\beta_{i-1} + x_{i+1}\beta_{i+1} + \cdots \dots \dots x_2^2\ \beta_1 + \dots . \tag{4.19}$$

$$S_{T_i} = R_X^2 - R_{X_{\sim i}}^2 \tag{4.20}$$

where:

S_{T_i}: Total sensitivity index
$R^2_{X_{\sim i}}$ is coefficient of determination without including the attribute X_i.
R^2_X is coefficient of determination of the full model.

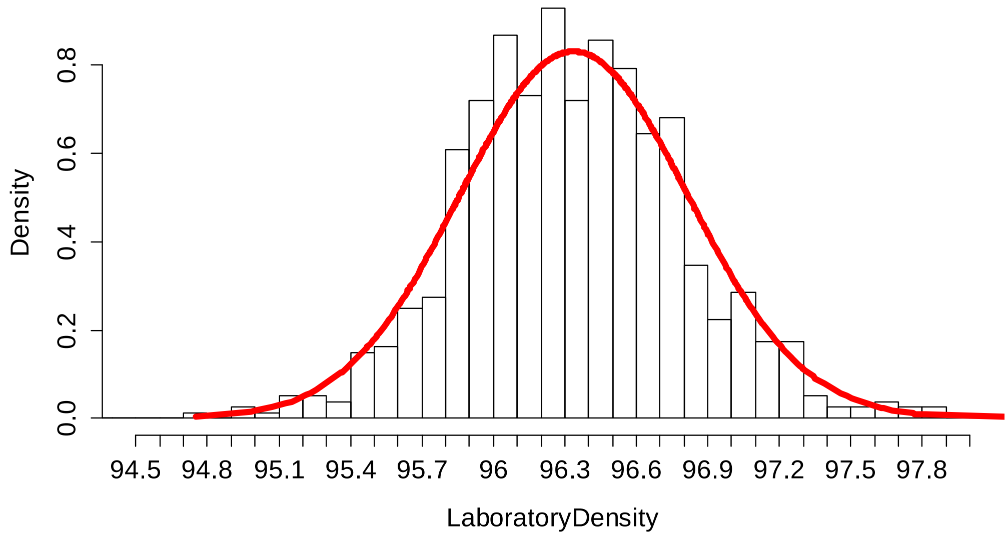

Figure 4.7: Actual and modeled (Log-normal) distributions of laboratory density

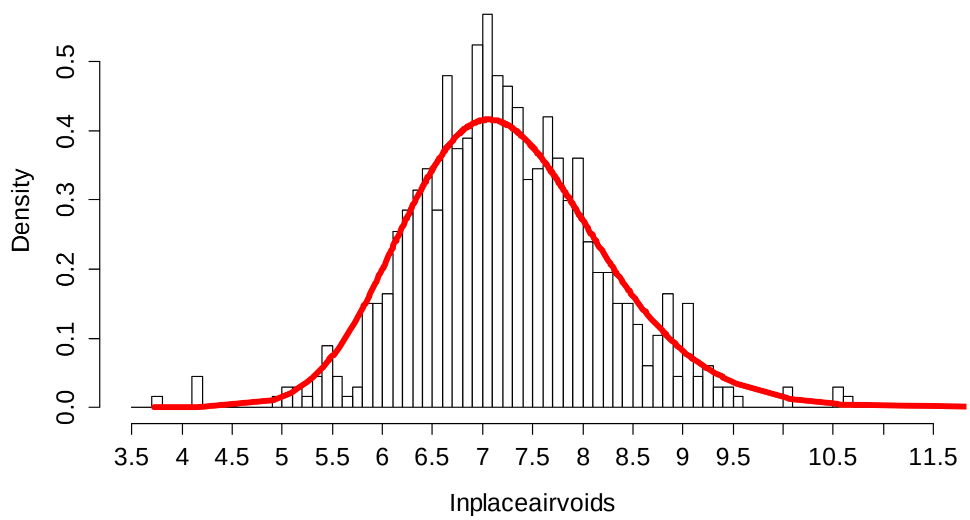

Figure 4.8: Actual and modeled (Log-normal) distributions of in-place air voids

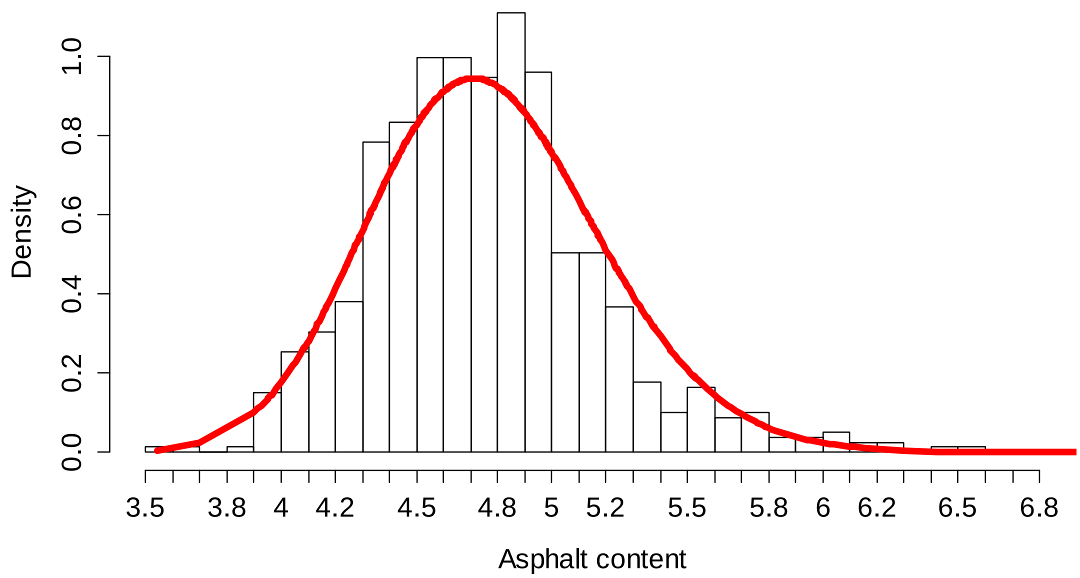

Figure 4.9: Actual and modeled (Log-normal) distributions of asphalt content

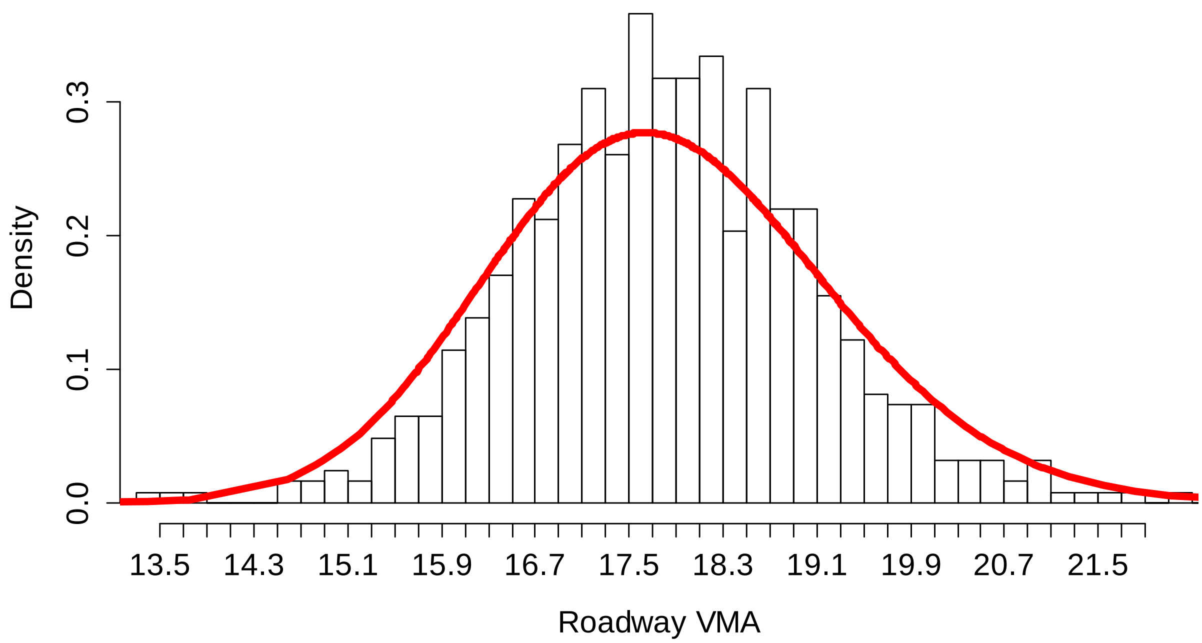

Figure 4.10: Actual and modeled (Log-normal) distributions of in-field VMA

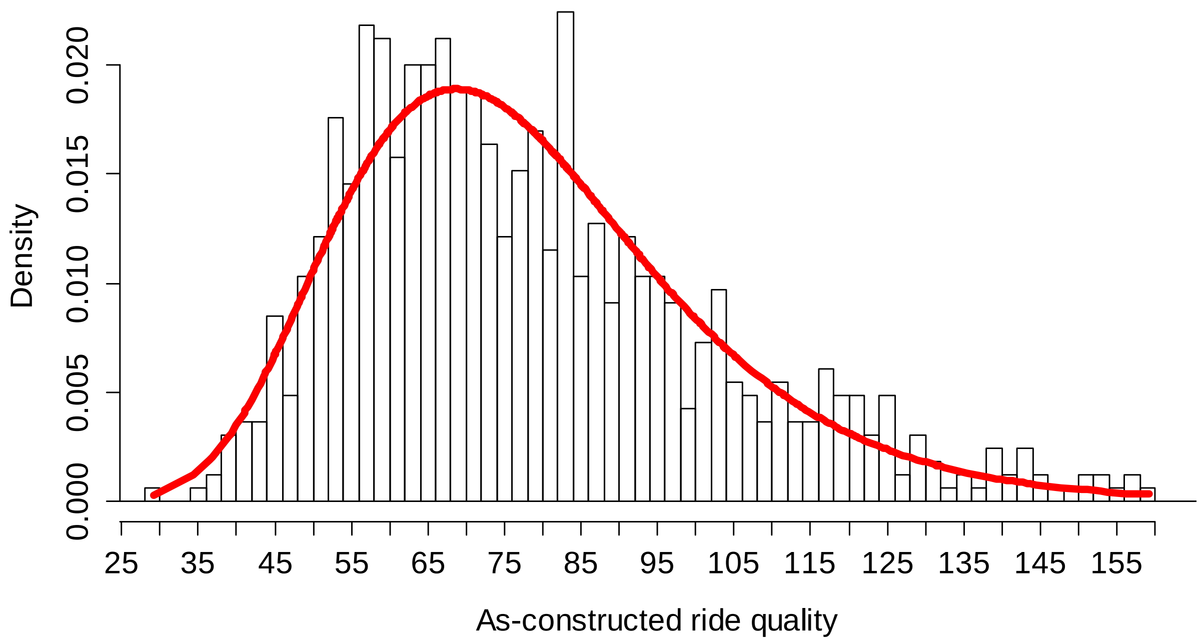

Figure 4.11: Actual and modeled (Log-normal) distributions of as-constructed ride quality

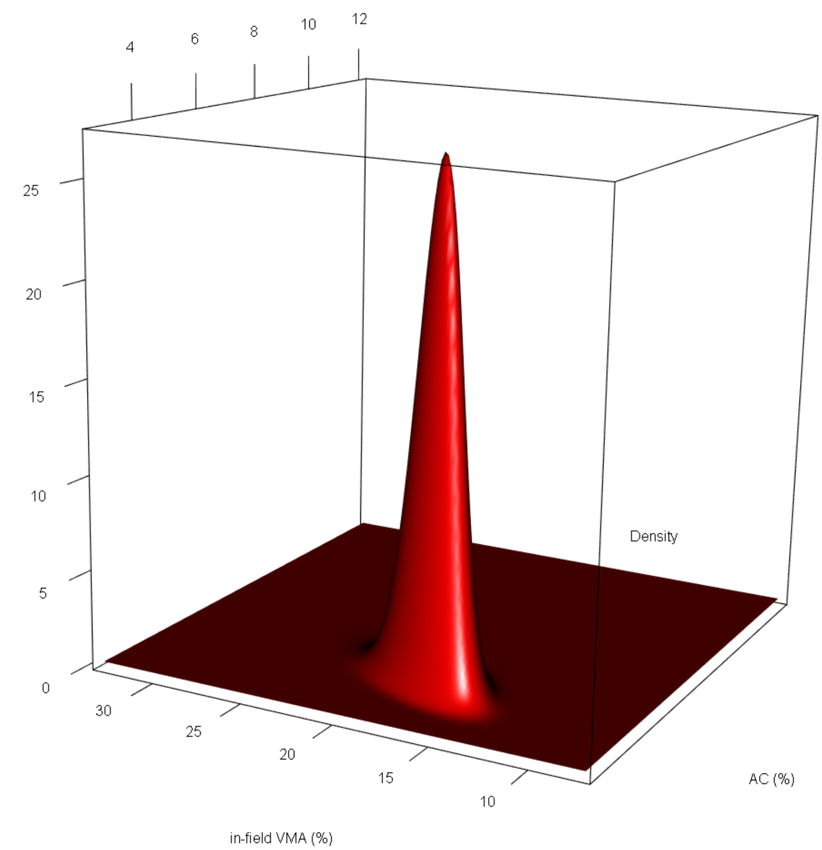

Figure 4.12: Actual and modeled (Bivariate Log-normal) distributions of asphalt content and in-field VMA

Results of Sensitivity Analysis

The sensitivity analysis was focused on identifying the sensitivity of initial ride quality, laboratory density, asphalt content and in-field VMAs with respect to performance. The total sensitivity index corresponding to initial ride quality and laboratory density is calculated based on the performance index that is evaluated using the structural model (see Equation 4.12); it is calculated based on performance index that is evaluated using the reduced form model in the case of the asphalt content and in-field VMA. The total sensitivity analysis index for both structural form and reduced form models are shown in Table 4.2. As indicated, the total sensitivity index is a measure of the proportion of the variation in the performance index explained by the corresponding model attribute. A higher total sensitivity index indicates that changing the model attributes produces a considerable change in the resulting performance of the pavement. This is very useful information in the development of PRSs.

Table 4.2: Sensitivity analysis results

Underlying model	Variable name	Total sensitivity index
Structural model	As-constructed ride quality	60%
	Laboratory density	8%
Reduced form	Asphalt content	1%
	VMA	1%
	Laboratory density	12%

The sensitivity analysis results highlight the importance of ride quality immediately after construction. The total sensitivity index indicates that the as-constructed ride quality explains about 60% of the variation in the performance index. Ride quality immediately after construction is not only associated with the performance, but improving initial ride quality produces practically significant improvements in the pavement performance. As shown in Table 4.2, laboratory density explains 8% of the variation in the performance index that is calculated based on the structural model and 12% of the variation in the performance index that is calculated based on the reduced form model. This suggests that a change in the laboratory density does produce a difference in performance of the corresponding pavement. However, the influence of production QC is lower than that of the smoothness control immediately after construction. The results also indicate that asphalt content and in-field VMA explain only 1% variation in the performance. In other words, a change in asphalt content and in-field VMA does not necessarily improve performance considerably, although they are statistically associated with pavement performance. This result is attributable to the narrow ranges of these properties within the dataset evaluated. It is important to note that the sensitivity analysis results are only valid conditional on the fact that the QC parameters (particularly volumetric properties) are within the current specification limits.

4.3.3 Marginal Effects

Marginal effect, a well-known measure of the responsiveness of a dependent variable to the changes in explanatory variables, is adopted for quantifying practical significance. It is defined as the quantifiable change in the dependent variable due to a unit change in the explanatory variable—everything else remaining unchanged. As a PRS analyst, one would be interested in quantifying both probability of a project to exhibit any signs of early deterioration—i.e., $\mathrm{P(y = 0|X)}$—as well as the extent of deterioration conditional on the fact that the project exhibited signs of deterioration during the analysis period—i.e., $\mathrm{E(y|y > 0}, X)$. The marginal effect corresponding to P(y=0 | X) depicts the change in the probability of the HMA project to deteriorate within the analysis period due to a unit change in the corresponding QC parameter. The marginal effect corresponding to E(y|y>0,X) represents the change in the level of deterioration rate, conditional on the fact that the pavement exhibited signs of deterioration, due to a change in the corresponding QC parameter. Mathematically, we are interested in changes in the conditional distributional properties of the y (conditioning on X)—i.e., $\mathrm{P(y = 0|X)}$ and $\mathrm{E(y|y > 0}, X)$—corresponding to a unit change in the respective explanatory variables. The marginal effects corresponding to P(y=0 | X) and E(y|y>0,X) are reported in Table 4.3.

Table 4.3: Marginal effects

Sensitivity type	Variable description	dy/dx	Std.er	Z	P-value
	Initial IRI	0.053	0.019	2.76	0.01
	Indicator Variable: PG 76	0.646	0.283	2.28	0.02
	Laboratory Density	-0.988	0.309	-3.20	0.00
$\mathbf{E(y|y > 0}, \boldsymbol{X})$	Analysis Period	-0.346	0.090	-3.82	0.00
	Indicator Variable: Low Speed Facility	1.201	0.646	1.86	0.06
	Indicator Variable: IH Facility	-1.088	0.374	-2.90	0.00
	Indicator Variable: Rural Area	0.647	0.301	2.15	0.03
	Initial IRI	0.005	0.002	2.96	0.00
	Indicator Variable: PG 76	0.066	0.029	2.3	0.02
	Laboratory Density	-0.101	0.032	-3.2	0.00
	Analysis Period	-0.035	0.009	-3.78	0.00
$\mathbf{P(y = 0| X)}$	Indicator Variable: Low Speed Facility	0.123	0.067	1.82	0.07
	Indicator Variable: IH Facility	-0.111	0.038	-2.93	0.00
	Indicator Variable: Rural Area	0.066	0.030	2.2	0.03
	Constant	0.005	0.002	2.96	0.00

The marginal effects analysis suggests that an increase of 10 inches/mile in the initial IRI or as-constructed ride quality increases the average deterioration rate by 0.53 inches/mile/year, while the probability of an HMA project showing any signs of deterioration increases by 0.05. The additional deterioration rate diminishes the time taken by a pavement surface to reach a threshold IRI value, thereby requiring earlier maintenance. The deterioration rate drops by 0.987 inches/mile/year with each additional 1% increase in the laboratory density, while the probability of pavement project showing any signs of deterioration decreases by 0.1. An average difference in deterioration rate between the mixes with stiffer binder (PG76) and the other mixes is 0.65 inches/mile/year, while the difference in probability of the project to exhibit any signs of deterioration is 0.066; it is to be noted that, for this dataset, stiffer binder is associated with faster deterioration in terms of roughness. The deterioration rate of the pavement project constructed on low speed routes is about 1.2 inches/mile/year higher than that of the high-speed routes, while the probability of the pavement exhibiting any signs of deterioration is about 0.12 higher. The roughness deterioration rate of the pavement project constructed on Interstate Highways is about 1.09 inches/mile/year lower than that of other facilities, while the probability of the pavement exhibiting any signs of deterioration is about 0.11 lower. The deterioration rate of the pavement project constructed in a rural area is about 0.65 inches/mile/year higher than that of other facilities, while the probability of the pavement exhibiting any signs of deterioration is about 0.07 higher. Marginal effects cannot be calculated for the instrumental variables using the full model. Therefore, the structural and reduced form equations were used to build a continuous relationship between the QC parameters and the performance index. Some of the important findings are described below.

Figure 4.13 shows a relationship between the as-constructed ride quality and the corresponding deterioration rates (performance index within the figure), keeping the other relevant continuous attributes at their respective mean values. Two different curves corresponding to different facility types and project location are shown in Figure 4.13. This bifurcation emphasizes the importance of incorporating facility type and other project-specific attributes into the pay adjustment specification. This also supports the idea of pay schedules (1, 2, and 3), currently practiced by TxDOT, to account for such project-specific attributes. It can be inferred from the figure that reducing the as-constructed roughness levels from 60 inches/mile to 30 inches/mile corresponds to an average increase of 50% in pavement life. As another example, reducing the as-constructed roughness levels from 90 inches/mile to 60 inches/mile translates into an average increase of 30% in pavement life. This indicates that the gain in pavement life reduces non-linearly with the as-constructed roughness levels, which further emphasizes the need for building smoother pavements.

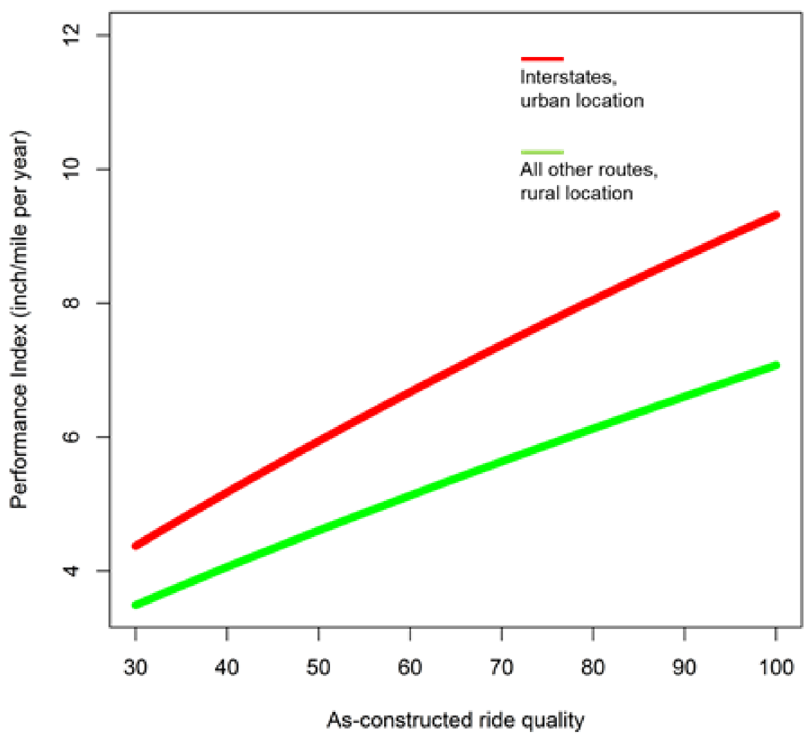

Figure 4.13: Expected deterioration (E[y|X,y>0]) vs as-constructed ride quality

Figure 4.14 shows the relationship between the laboratory density and the corresponding performance index or deterioration rate. Two different curves corresponding to the different facility types and project location are provided as before. It can be inferred from the figure that increasing the laboratory density of the plant-produced hot mix by 1% translates into an average increase of 13% in pavement life.

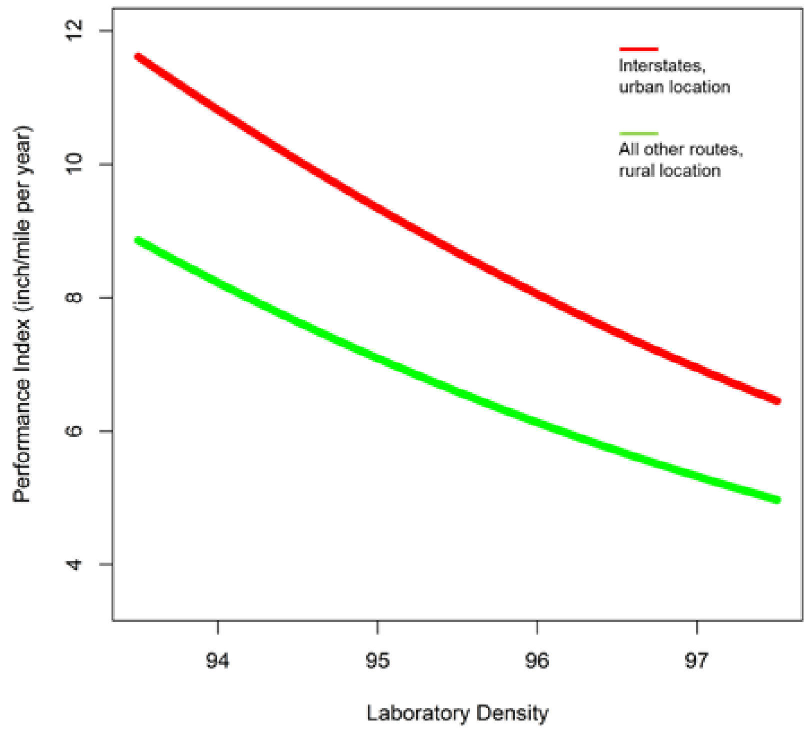

Figure 4.14: Expected deterioration (E[y|X,y>0]) vs laboratory density

Figure 4.15 shows the relationship between the asphalt content and the corresponding performance index or deterioration rate. It can be inferred from the figure that increasing the asphalt content of the plant-produced hot mix by 1% translates into an average increase of 5% in pavement life.

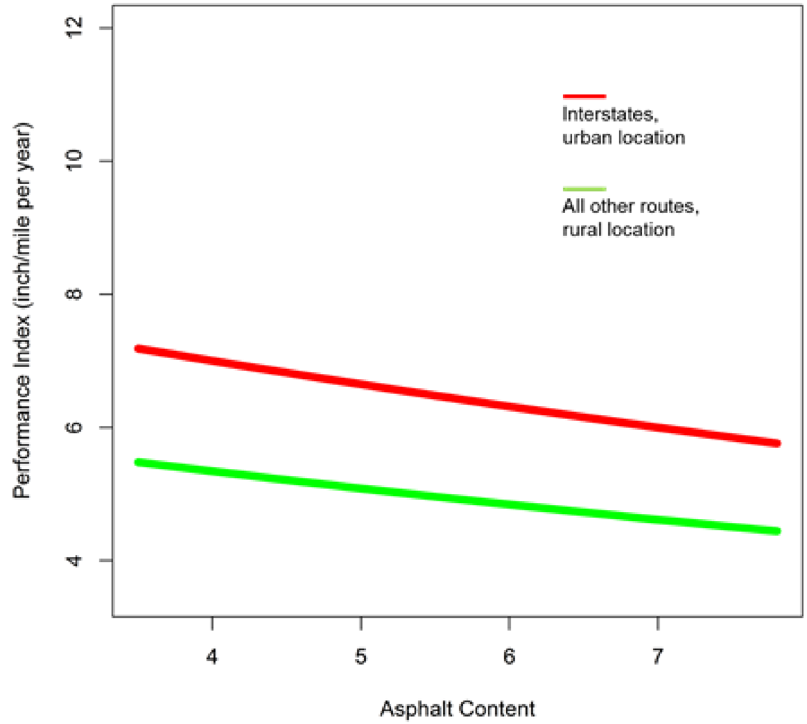

Figure 4.15: Expected deterioration (E[y|X,y>0]) vs asphalt content

Figure 4.16 shows the relationship between the in-field VMA and the corresponding performance index (or deterioration rate). It can be inferred from the figure that increasing the in-field VMA of the plant-produced hot mix by 1% translates into an average increase of 2% in pavement life.

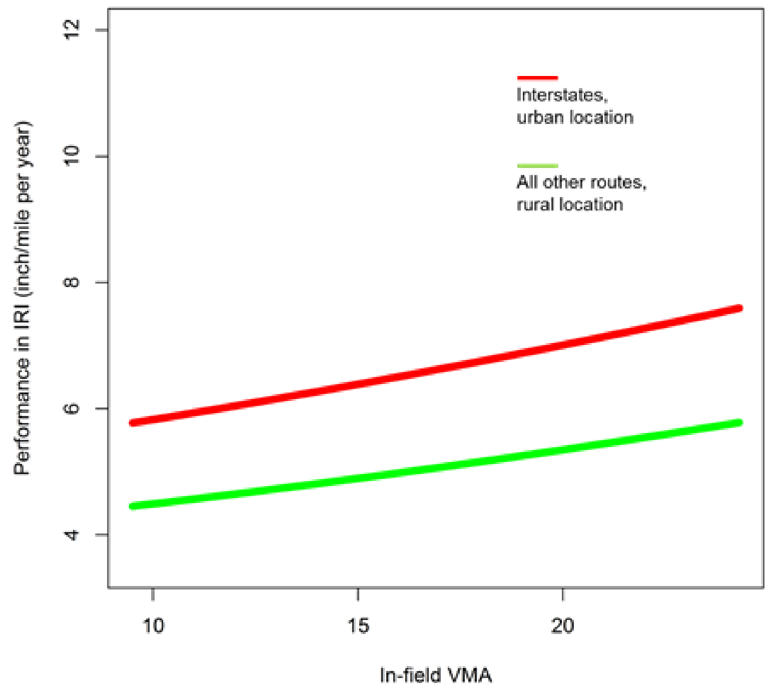

Figure 4.16: Expected deterioration (E[y|X,y>0]) vs in-field VMA

In summary, the sensitivity analyses emphasize that the pay adjustment specification of the ride quality is most important. The bonus/penalty awarded as part of Item 585 is economically justifiable as smoother roads are evidently producing long lasting pavements. Based on this finding, emphasis must be placed on the maximum penalty/bonus on the ride specification. On the other hand, production and placement QC parameters moderately influence performance. It is also evident that the influence of the production QC is marginally higher than that of placement QC. A revised pay adjustment scheme on production and placement is also necessary, although the maximum and minimum payments need not be changed. A discussion of the revised production, placement, and ride quality PAFs for HMA projects is provided in the next chapter.

4.4 Analysis of Concrete Project Performance

The concrete pavement pay adjustment specification currently relies on as-constructed ride quality. It is important to establish an empirical relationship between the as-constructed ride quality and the respective pavement performance in order to financially justify the current pay adjustment specification. To develop any empirical relationship that is adequately supported by the statistical theories, a representative sample pool of concrete projects is needed. Unfortunately (for the reasons described in the previous chapter), the concrete project data sources were not available to the research team despite efforts to gather a reasonable sample size. Information was collected for 32 concrete projects. Performance data spanning a minimum of 3 years is required to evaluate deterioration rates. This constraint reduced the sample size further down to 25 projects. Performance data were collected from the PMIS database for these 25 projects to evaluate the deterioration rate of concrete pavements as described in the previous chapter. The as-constructed ride quality information for these projects was obtained to facilitate a comparative analysis to explain the desired empirical relationships. As indicated before, a rough comparative analysis was employed rather than applying any formal statistically sound procedures, solely due to the lack of data.

The available as-constructed ride quality information was processed using simple data manipulation and descriptive statistics in order to perceive typical concrete pavement practices. Figure 4.17 shows a histogram of the as-constructed ride quality from a total of 19 concrete projects; data was missing for the remaining 6 projects. Figure 4.17 suggests that an average as-constructed ride quality of 78 inches/mile is being reported on a newly constructed concrete overlay/pavement. Empirical evidence also suggests that concrete pavements are constructed at higher initial roughness levels compared to HMA pavements. Figure 4.18 shows the majority of concrete projects are penalized, as these pavements are being constructed as rougher surfaces. Figure 4.19 reports an average penalty/bonus (per 0.1 mile) for each pay schedule; it also highlights that concrete pavements are mainly constructed under Schedule 2. These important findings reflect the concrete pavement contractor response to the existing pay adjustment specification. Arguably, contractors opt to pay a penalty rather than deliver a smoother concrete pavement or apply any corrective action to the rougher as-constructed pavement surface.

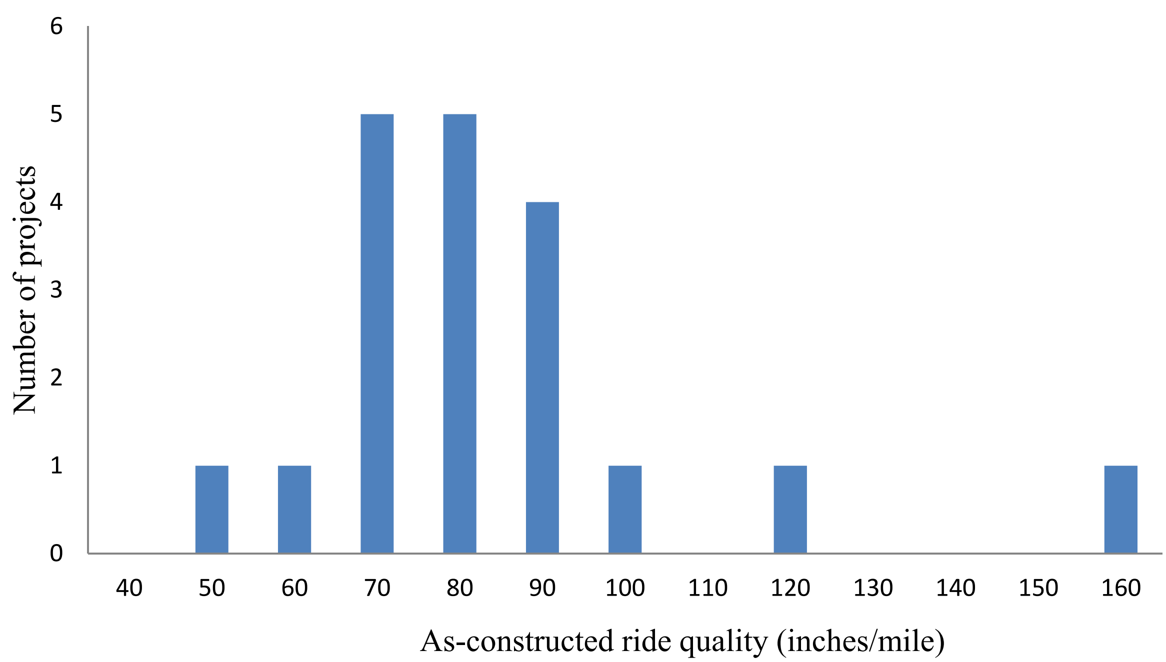

Figure 4.17: As-constructed ride quality on concrete projects

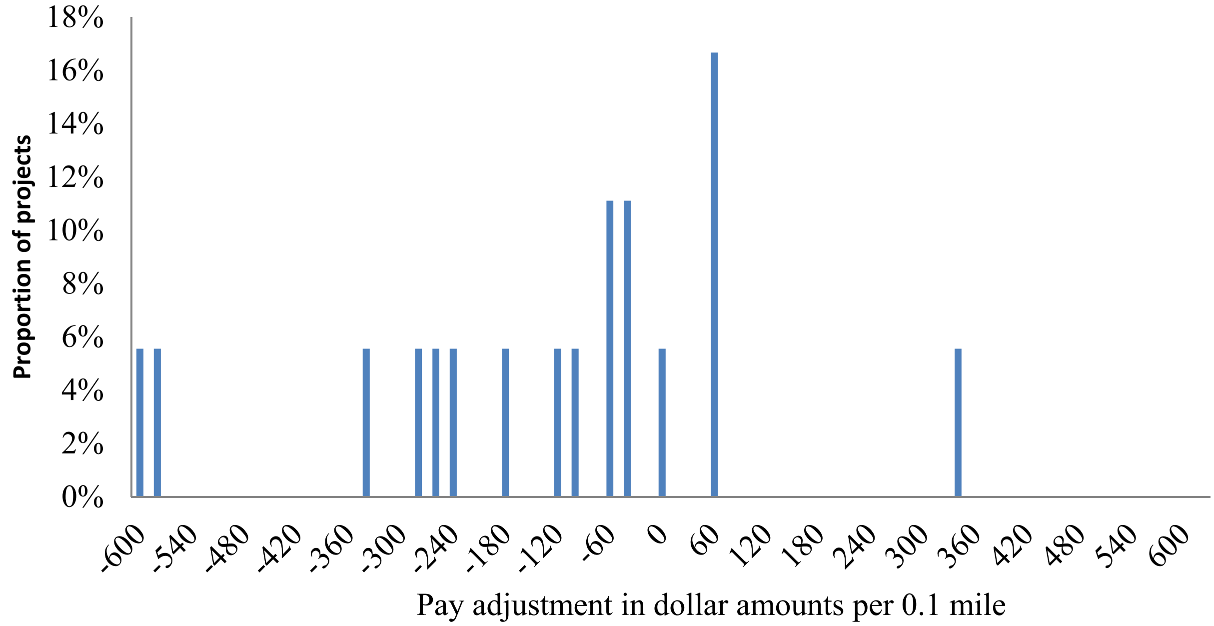

Figure 4.18: Awarded pay adjustments on concrete projects

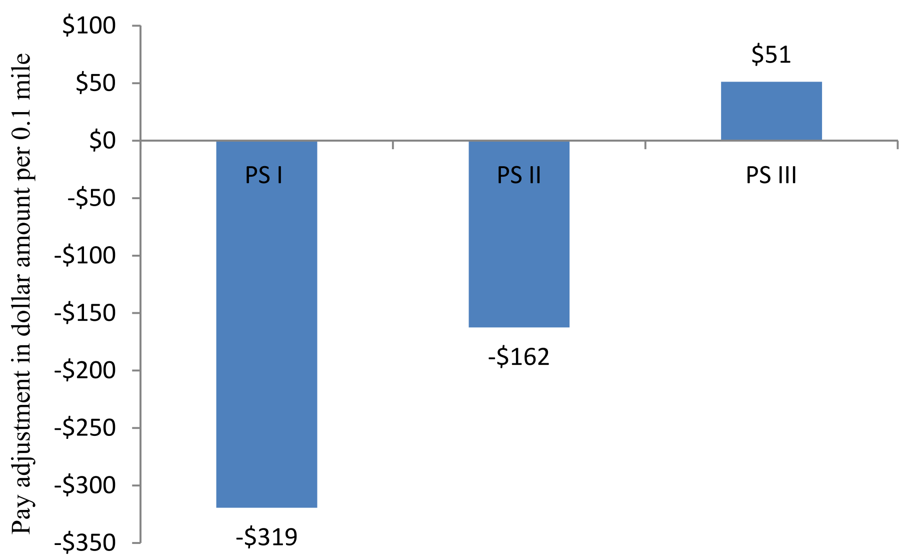

Figure 4.19: Awarded pay adjustments for concrete ride quality per 0.1 mile

The performance of the aforementioned 25 concrete projects was tracked using the PMIS ride quality and condition data. This involved producing project-level pavement performance plots indicating a time series of performance measurements spanning 2003–2012. The rate of change of the ride quality with time was evaluated by inspection with an objective to identify any concrete projects deteriorating faster than typical rates. No concrete projects were identified with practically significant deterioration rates. For instance, Figure 4.20 shows a typical deterioration curves of various concrete projects across Texas. An addendum in the form of a CD-ROM (attached to the printed version of this report) includes performance trends of all the identified concrete projects in this study. These plots suggest that the concrete projects did not exhibit any signs of deterioration in terms of ride quality within the analysis period; this is expected for structurally sound concrete pavements. For this reason it is recommended that the field performance of older concrete pavements be used to evaluate the desired empirical relationships to develop PRSs for concrete pavements.

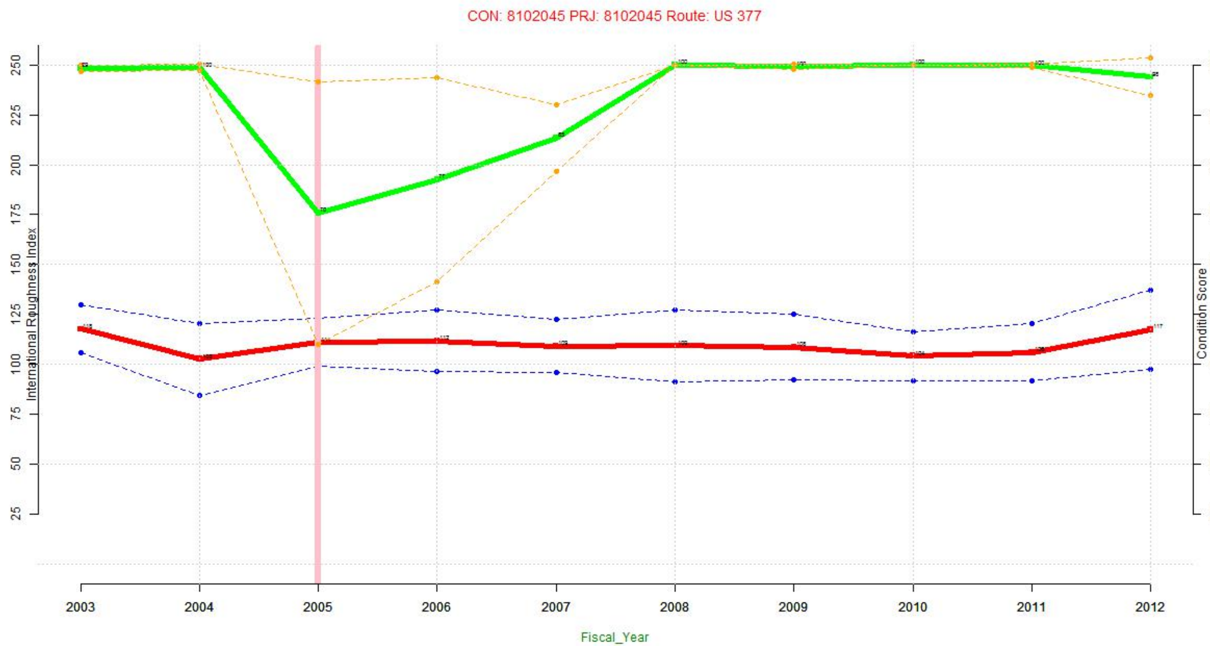

Figure 4.20: Deterioration of a concrete project (vertical line indicating the construction year)

Chapter 5. Revised Pay Adjustment System

In order to develop a sound basis for proposing a pay adjustment schedule for the ride quality of pavements as well as for the production and placement quality of HMA, it is necessary to understand the importance of potential savings that the highway agency may obtain through superior construction practices and better paving materials. Ideally, the bonus that is awarded to the contractor should be no more than the maximum savings to the agency due to the extended service life of the pavement and the penalty levied on the contractor should be no less than the increased cost of maintenance due to premature failure of the pavement. It is often difficult or impractical to perform a conventional economic analysis to quantify the actual savings to the state agency. A recent NCHRP project (10-79) also recommended creating performance-based pay adjustment schedules combined with engineering judgment rather than conducting an economic analysis. A methodology was developed to propose a revised PAF system; the method blends the key findings from the empirical relationships, sensitivity analysis, expert opinion, and engineering judgment.

This chapter is organized as follows. First, the development of the revised PAF system corresponding to HMA production, placement, and ride quality is described. The later section discusses pay adjustment recommendations with regard to the ride quality of concrete projects. Despite the unavailability of data sources for concrete projects, a revised ride pay adjust factor system for concrete ride specification is proposed, primarily guided by available empirical evidence and engineering judgment.

5.1 HMA Production PAF System

The current production PAF is based on the laboratory-molded density using the Engineer's test results. A lot refers to the quantity of HMA making up one days production, typically about 2,000 tons. The number of lots varies depending on the size of the project. The plant material is sampled from each sublot (typically 500 tons) during a normal day's production and the lab-molded density is determined for each sublot. PAFs are determined for each sublot using the absolute deviation from the target laboratory-molded density; i.e., AAD. The final production PAF for completed lots is the average of the PAFs from the four sublots sampled within that lot.

According to the current HMA specification (Item 341), the deviation of the laboratory density shall be within $\pm 1\%$ of the target laboratory density. The current PAF system provides incentive for being within the specification; a maximum bonus of 5% is awarded on lots with laboratory density that is within $\pm 0.2\%$ of the target density. As shown in Figure 5.1, the production pay factor is symmetric about the zero deviation (or exactly achieving target density) and drops linearly to a pay factor of 1.0. The penalty for exceeding the specification limits ($\pm 1\%$) increases at a higher rate than that of the incentive portion of the PAF curve. The current PAF encourages the contractor to achieve the target density closely in the production process and does not reflect the relationship of the laboratory density with pavement performance by any means.

Empirical evidence suggested that mixes with higher laboratory densities are expected to produce better performing pavements. Therefore, a revised PAF system is proposed that financially encourages contractors to produce hot mix with higher laboratory densities within the current specification thresholds ($\pm 1\%$). The sensitivity analysis indicated that controlling or modifying the laboratory density (being within the specification) does produce a change in performance, although the change is not appreciable. Based on this finding, it was opted not to change the existing incentive and penalty per unit deviation of the laboratory density from the target density. The proposed PAF system for HMA production is developed simply by right-skewing the existing pay factor system. The proposed pay factor system reduces the incentive for producing hot mix with laboratory density that is lower than the target density and increases the bonus for producing higher densities. However, it is to be noted that the penalty portion is unchanged. Data suggested that a majority of the contractors are able to achieve production densities close to the target densities. Thus, implementing the proposed pay adjustment system does not penalize the contractors receiving bonus, but rewards the contractors providing higher densities than the target. As the total bonus or penalty is not altered, but shifted towards the right hand side (larger densities than target), the financial implications due to the implementation of the revised specification on the TxDOT front are unchanged on average. The proposed revisions reflect the relationship of laboratory density with field performance and thereby encourage contractors to produce mixes with laboratory densities that are higher than the target densities (while being within the specification). It is believed that the revised specification financially motivates the contractors to slightly increase the asphalt contents in the mix to produce laboratory densities that are higher than the respective targets. In summary, the revised production pay adjustment system is arguably financially justifiable as it rewards superior performance (or higher laboratory densities) and penalizes inferior performance (or lower laboratory densities), thereby portraying the underlying philosophy of PRSs.

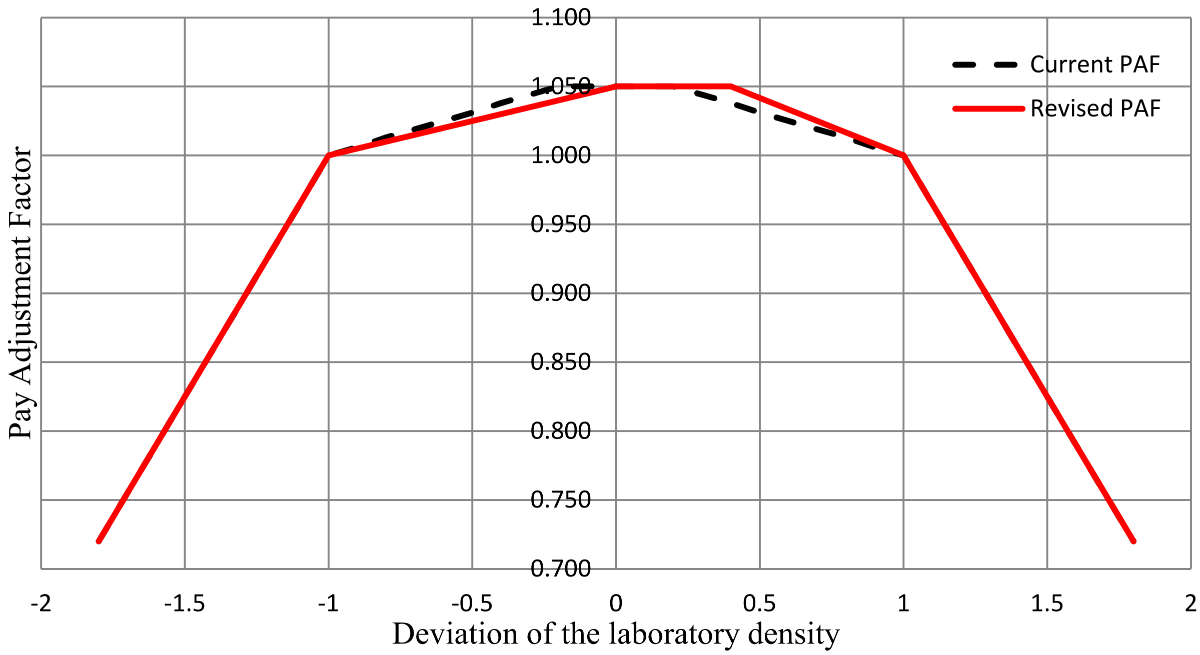

Figure 5.1: Current and revised HMA production PAF systems

5.2 HMA Placement PAF System

The placement PAF is based on in-place air voids of field-cored specimens that are obtained at randomized paving locations. The current placement PAF system uses data from two field-cored specimens within each sublot for determining the bonus and penalty corresponding to that sublot. The lot level placement PAF is the average of the placement PAFs for each of the four sublots within that lot. The system rewards moderate compaction levels (in-place air voids between 4.7% and 8.5%) and penalizes both under-compaction and excessive compaction levels. The researchers met with the core team that developed the existing placement PAF system to discuss the underlying rationale behind the current system. The development of the existing PAF system was primarily based on the premise that an intermediate placement density ensures better quality pavements, although there was no empirical evidence. The current PAF system does not reflect the relationship between placement density and the corresponding pavement performance.

Findings from the study suggested that a lower in-field VMA (of the field cores) is associated with better ride quality immediately after construction, which in turn enhances pavement performance. In-field VMA arguably acts as a proxy for in-place air voids for a given asphalt content of the hot mix. Therefore, one may argue that lower in-place air voids are associated with better pavement performance. The sensitivity analysis does provide evidence for the influence of the placement quality on the pavement performance, although quantified improvement per unit compaction effort is practically insignificant (being within the specification limits). Based on these findings, a left-skewed PAF system is proposed for HMA placement with enhanced rewards on higher compaction levels and increased penalty for under-compacted pavements (see Figure 5.2). The revised placement PAF system financially encourages contractors to increase compaction effort during the placement operation as this produces pavements with a longer life. It was opted not to reduce the bonus levels on the under-compaction side of the PAF system to avoid practical and contractual issues that may arise in moving from a bonus to a large penalty for a given placement density. The proposed PAF system is largely supported by the empirical evidence (based on 614 HMA projects) and arguably ensures extended performance for an awarded unit bonus on the placement operation. In summary, the revised placement pay adjustment system is financially justifiable as it rewards superior performance (or lower in-place air voids) and penalizes inferior performance (or higher in-place air voids), thereby portraying the underlying philosophy of PRSs.

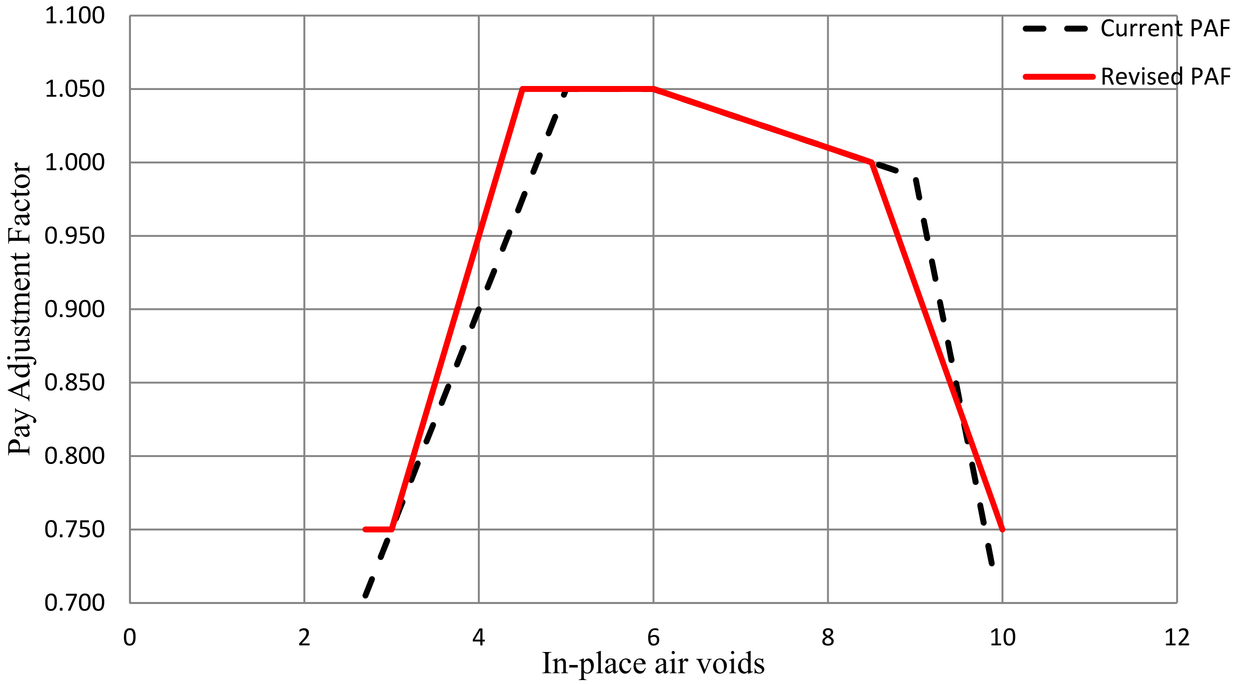

Figure 5.2: Revised and current HMA placement PAF systems

5.3 Combining Placement and Production PAF

According to the current specification, total adjustment pay (TAP) is based on the applicable PAFs for production and placement for each lot and is calculated as follows:

$$\text{TAP} = (\text{A} + \text{B})/2$$

Where:

A = Bid price × production lot quantity × average PAF for the production lot
B = Bid price × placement lot quantity × average PAF for the placement lot

It is recommended to continue using this approach, as it encourages equal importance for both production and placement operations. Currently, if the production or placement PAFs for three consecutive lots falls below 1, production is suspended until further test results conform to the satisfaction of the engineer so that the next material produced or placed will result in pay factors of at least 1.

5.4 HMA and Concrete Ride Quality Pay Adjustment System

TxDOT uses the guidelines provided as part of Item 585 specifications (TxDOT, 2004) to determine the pay adjustment schedule for ride quality requirements for HMA as well as concrete pavements. A roughness index of less than 60 inches/mile is considered as excellent while anything beyond 95 inches/mile will require some kind of corrective action according to the current guidelines. The pay adjustment schedules that are currently in practice are shown in Figure 5.3. However, a key decision involves selection of the correct pay adjustment schedule. Currently TxDOT uses a set of guidelines that are provided in Table 2.8 in Chapter 2.

In general, the requirements are more stringent for new constructions as well as scenarios that present more than one smoothness opportunity (for example, diamond grinding and placing a single lift of asphaltic concrete). The guidelines also provided some room for on-the-spot modification where the pavements have a history of high roughness index. For example, a pavement that has a roughness index in excess of 170 inches/mile before rehabilitation will allow the engineer to make adjustments to the pay adjustment schedule (marked with *). Figure 5.3 demonstrates that the current penalty schedule is most stringent for Schedule 1, followed by Schedule 2, and there is no penalty for Schedule 3.

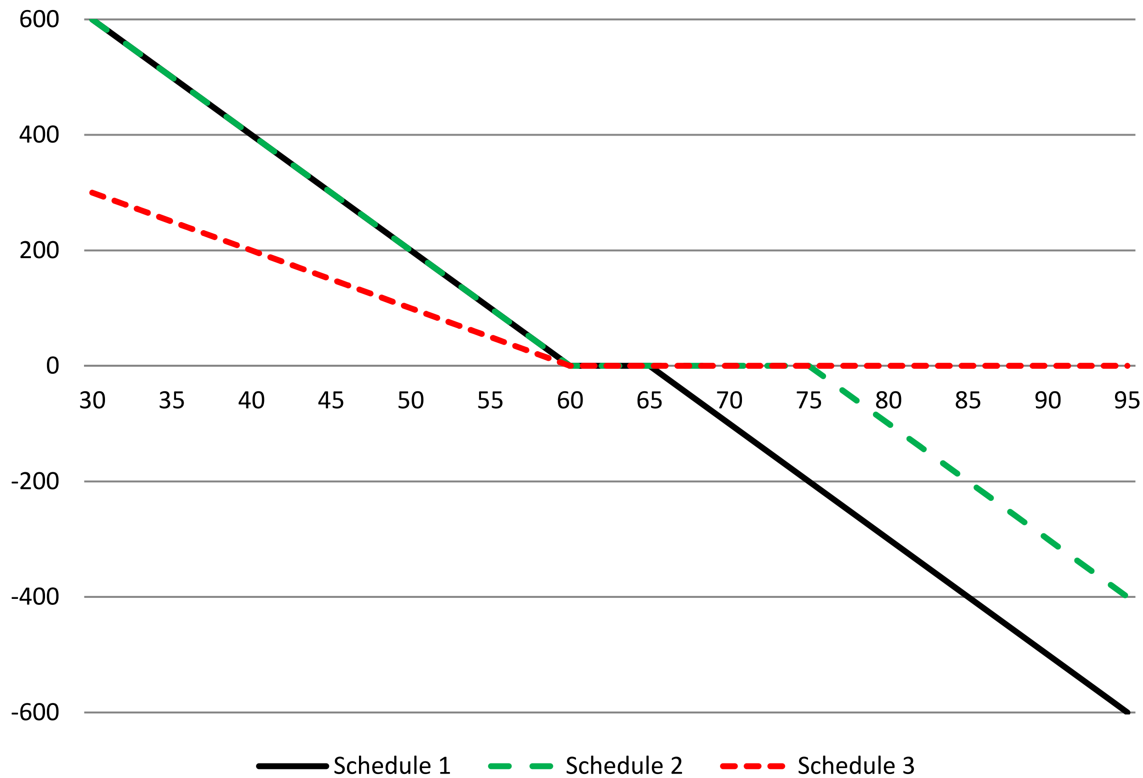

Figure 5.3: Graphical illustration of current ride specification

Empirical evidence suggested a positive association between the as-constructed ride quality and the respective pavement performance. In other words, a newly constructed pavement with a smoother surface immediately after construction evidently lasts longer. The existing ride specification (Item 585) does encourage paying bonuses for smoother surfaces and penalizes rougher surfaces. Schedule 1 of the current ride specification rewards HMA projects that are smoother than 60 inches/mile (IRI), while it penalizes projects that are rougher than 65 inches/mile. Schedule 2 of the specification is slightly less restrictive, only penalizing projects that are rougher than 75 inches/mile. Schedule 3 does not penalize any projects, while the bonus is reduced to one-half of the bonus awarded by the Schedules 1 and 2 (see Figure 5.3).

The previously reported ride quality data that is measured immediately after construction was analyzed from more than 800 HMA projects that were awarded a bonus or penalty under different schedules. Figure 5.4, Figure 5.5, and Figure 5.6 provide histograms of the ride quality (in terms of IRI in inches/mile) that was achieved on the various HMA projects under Schedules 1, 2, and 3. The data suggests that a majority of the contractors were actually able to achieve the existing thresholds for each of the schedules; therefore, the majority of the contractors are receiving a bonus on HMA projects. The vision to improve the ride quality of newly constructed roads during the development of the existing specification has been fulfilled. Empirical findings suggest that it is important to raise the standards as contractors have already *"learnt"* to achieve the existing requirements for winning incentives. The researchers propose selecting the thresholds based on the quintile analysis of the observed ride quality data (shown in Figure 5.4, Figure 5.5, and Figure 5.6). The thresholds are selected such that the upper 50% of the contractors providing smoother surfaces are rewarded, while the lower 25% of the contractors delivering rougher pavements are penalized under each schedule. The revised thresholds financially encourage the contractors to deliver smoother (relative to the current standards) pavement surfaces immediately after construction. As it is shown in Figure 5.7, Figure 5.8, and Figure 5.9, the revised ride specification rewards HMA projects smoother than 50 inches/mile (IRI), while penalizing projects that are rougher than 60 inches/mile under Schedule 1. Under Schedule 2, the bonus threshold is unchanged at the existing level of 60 inches/mile, although the revised specification penalizes projects that are rougher than 70 inches/mile. On the other hand, Schedule 3 is essentially unchanged in terms of the bonus/penalty thresholds.

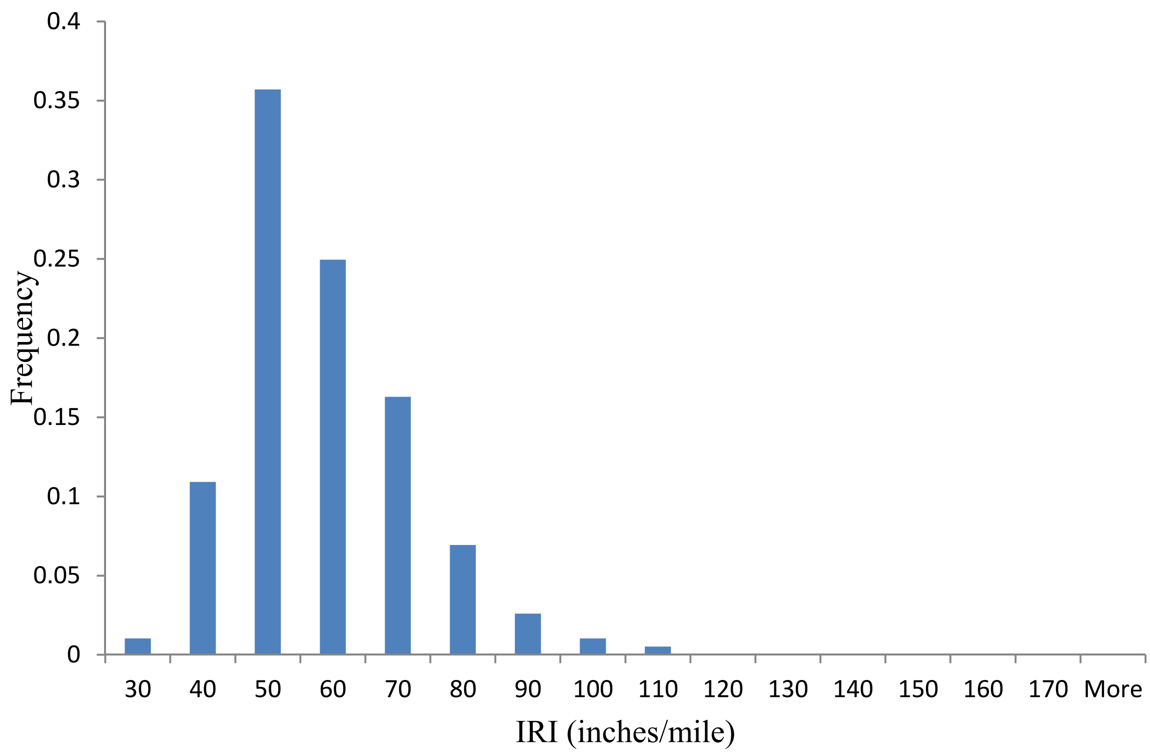

Figure 5.4: As-constructed ride histogram (Schedule 1): HMA projects

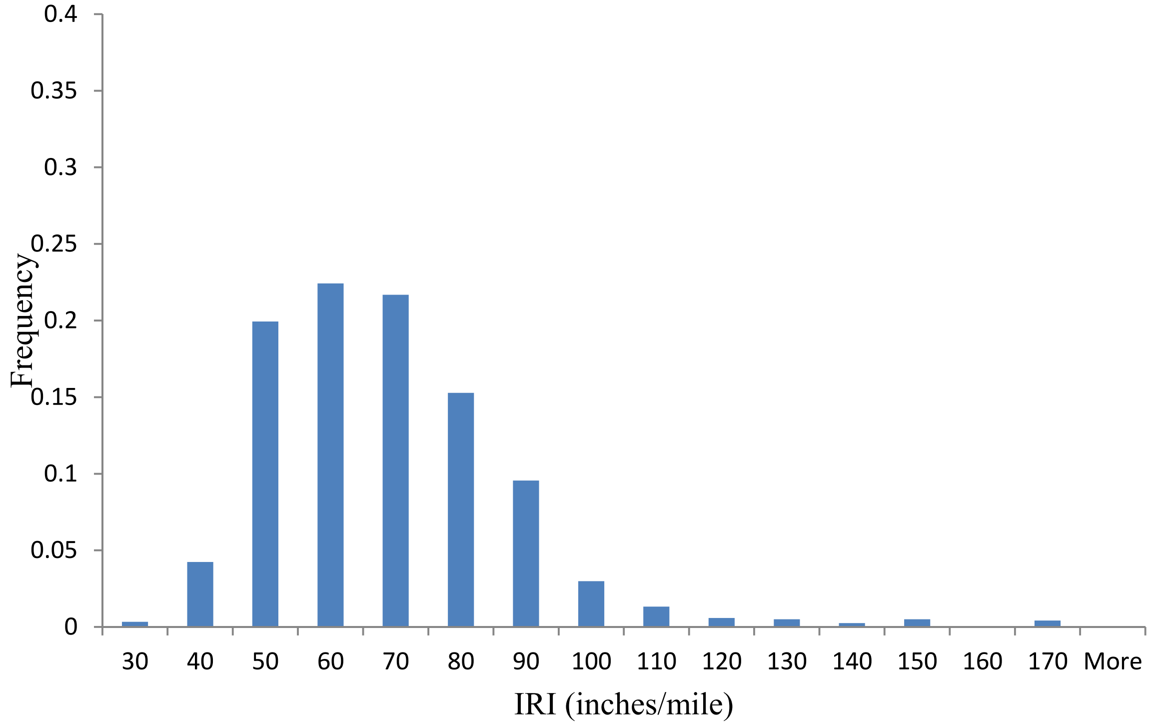

Figure 5.5 As-constructed ride histogram (Schedule 2): HMA projects

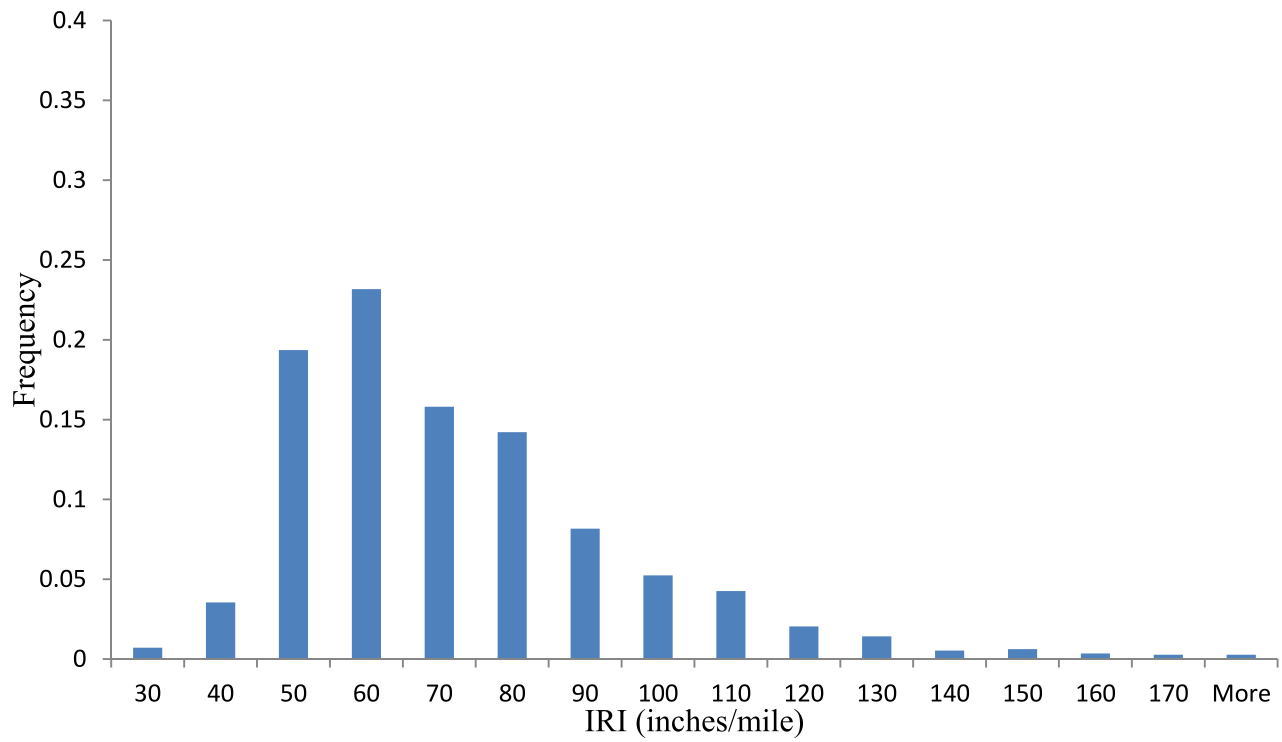

Figure 5.6 As-constructed ride histogram (Schedule 3): HMA projects

Apart from the aforementioned thresholds for no bonus/penalty, the magnitude of the bonus/penalty paid per unit increase in as-constructed ride quality (IRI in inches/mile) is equally important and should be financially justifiable. The existing specification simply assumed a linear relationship with a maximum bonus and penalty fixed at $600 under both Schedules 1 and 2 (Schedule 3 awards a maximum bonus of $300 with no penalty). The maximum bonus/penalty was set a long time ago and needs to be updated to account for potential inflation. A maximum bonus/penalty of $1,000 is proposed to accommodate any possible inflation under plausible assumptions on the inflation rate. The linearity assumption is also relaxed and uses the empirical relationship that reflects the real field performance data to evaluate the pay adjustment in dollars for a given ride quality level, while respecting the previously discussed thresholds. An expected performance index corresponding to each initial ride quality level at 1 inch/mile increments was evaluated using the empirical predictions for an average hot mix project in Texas. Figure 5.10 shows the model-predicted performance index corresponding to each initial ride quality for different project groups. The revised pay adjustment system is proportional to the respective expected life along with the aforementioned mentioned maximum bonus and penalty levels. For instance, an HMA project constructed at an initial ride level of 30 inches/mile lasts at least twice as long as a project constructed at initial ride level of 90 inches/mile. Therefore, the incentive/penalty corresponding to these initial ride levels of 30 and 90 inches/mile should be proportional to the respective expected lives. A similar procedure is applied to develop pay adjustment system for Schedules 1, 2, and 3.

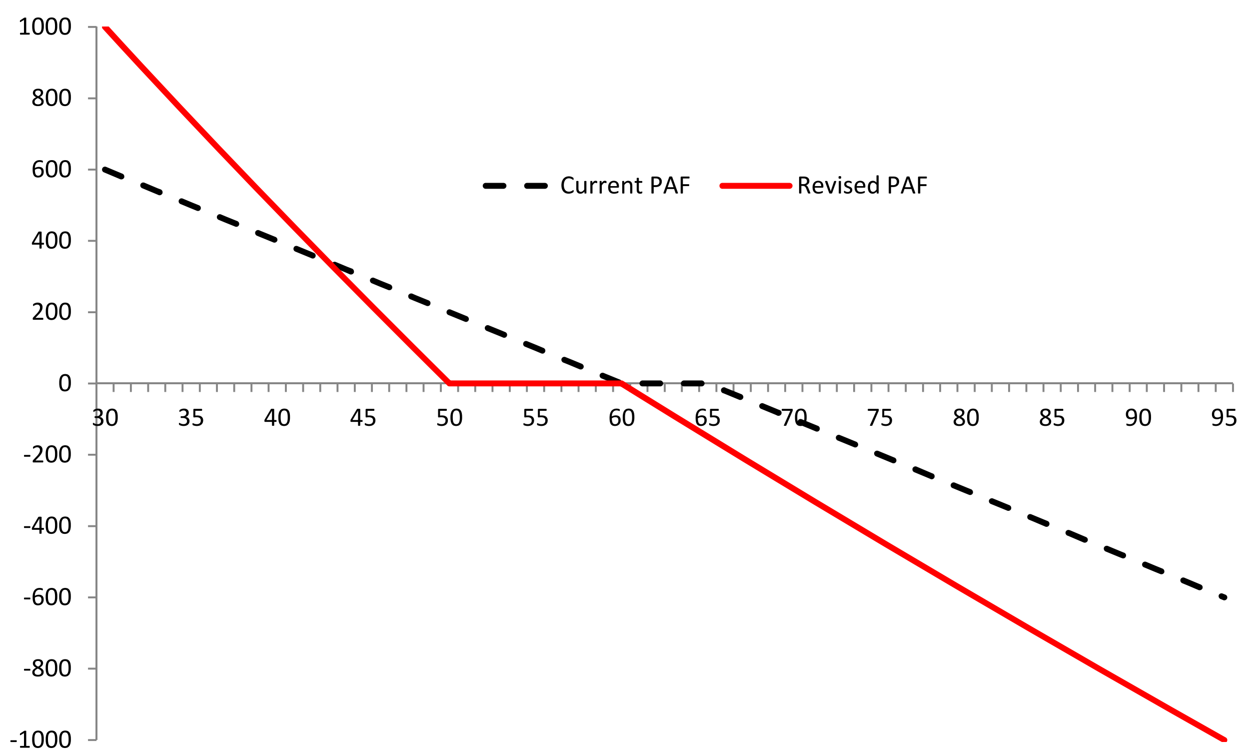

Figure 5.7 Current and revised HMA ride pay adjustment systems (Schedule 1)

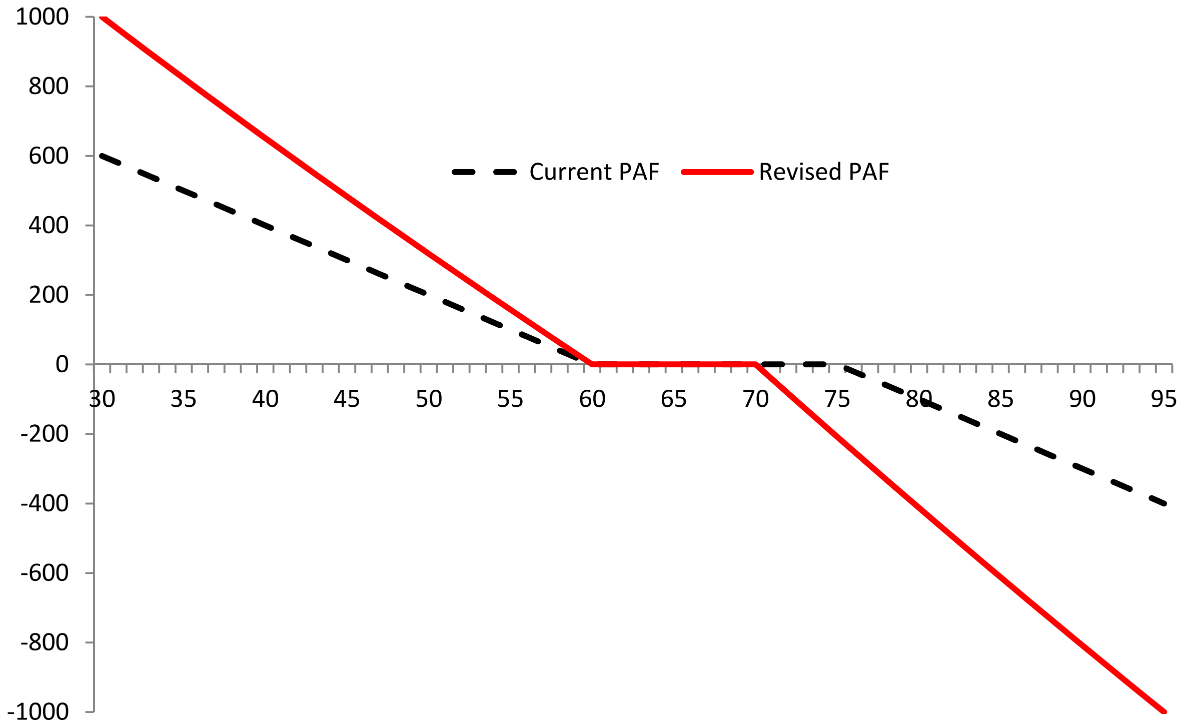

Figure 5.8 Current and revised HMA ride pay adjustment systems (Schedule 2)

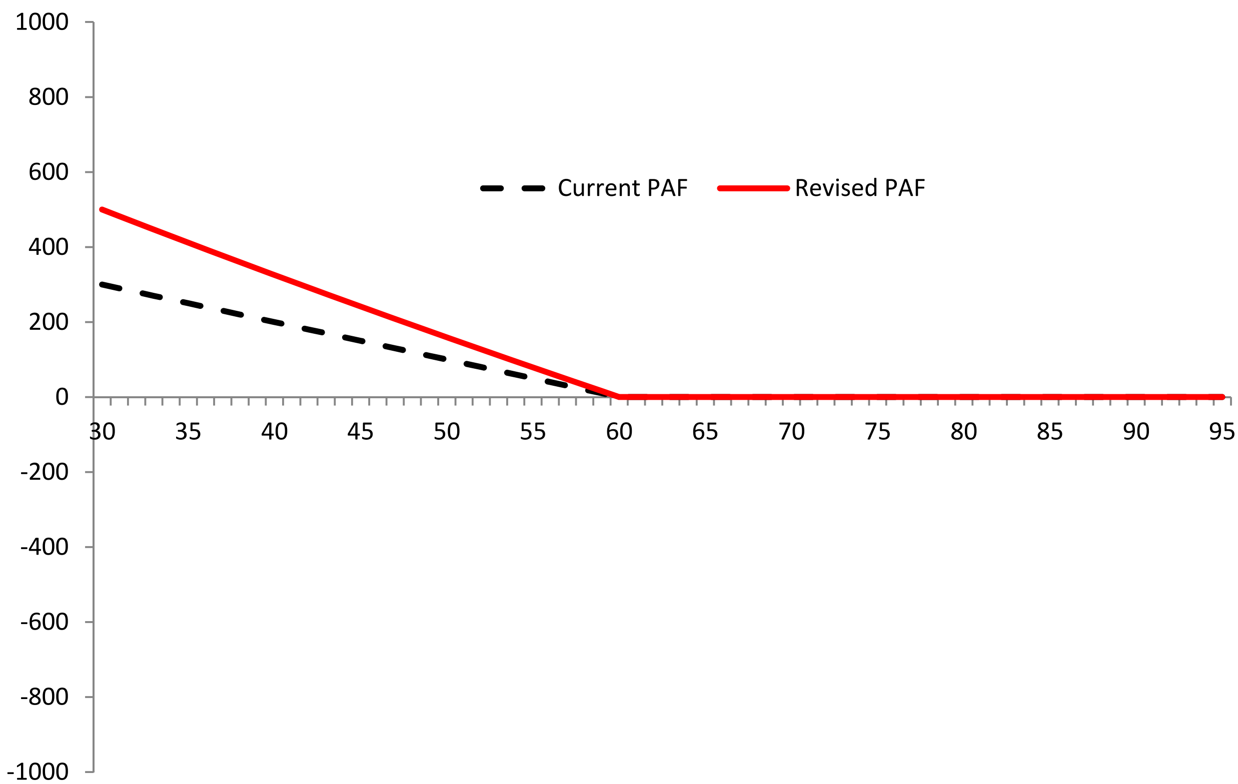

Figure 5.9 Current and revised HMA ride pay adjustment systems (Schedule 3)

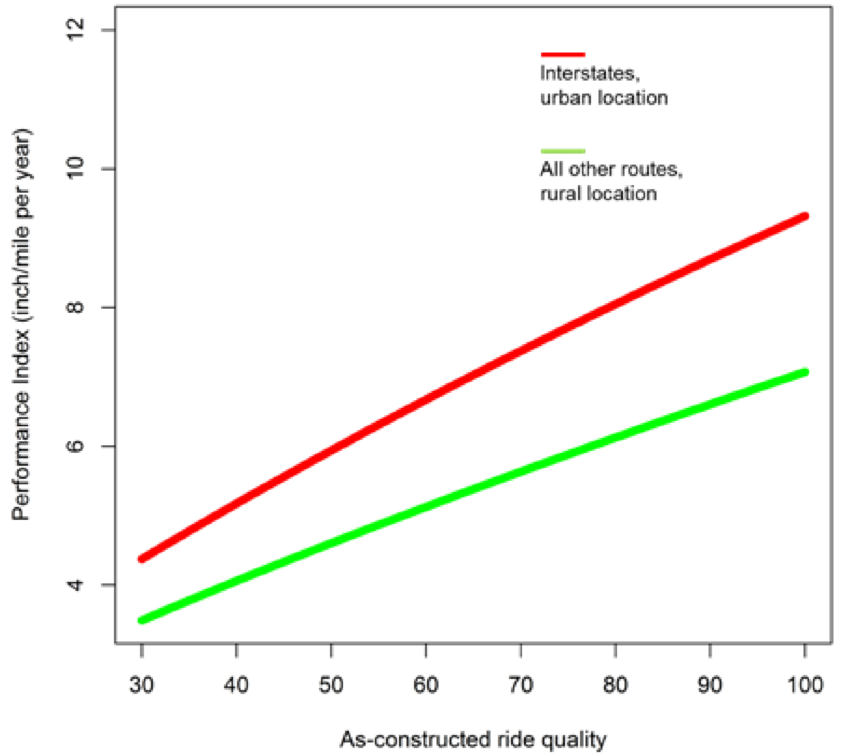

Figure 5.10: Expected performance vs as-constructed ride quality

5.5 Concrete Ride Quality Pay Adjustment System

As mentioned previously, the pay adjustment for ride quality on concrete projects is identical to that of HMA projects; however, the majority of concrete pavements are constructed under Schedules 2 and 3 as per the existing guidelines for selecting the schedule. This research project did analyze concrete pavement ride quality data and attempted to establish a relationship between the initial ride quality on concrete pavements and their field performance over time. The scarcity of the concrete project data did not provide enough statistical power to build any potential empirical relationship as for HMA projects. However, preliminary descriptive and simple statistical analysis of the performance data indicated some interesting trends and key findings. A revised pay adjustment system for ride quality on the concrete projects was developed based on these empirical findings combined with engineering judgment and the field experience of TxDOT personnel.

The data suggests that the majority of concrete project contractors are receiving penalties, as it is slightly harder to achieve smoother ride on concrete pavements, although there are opportunities to improve the smoothness of a newly constructed rigid pavement. Additionally, contractors are opting to pay the penalty rather than applying any smoothness improvement technique such as diamond grinding. Based on these findings, it can be argued that the current penalty on concrete projects is lower than the financial burden of executing a smoothness improvement action. For this reason a large penalty is proposed for levels beyond the allowable smoothness thresholds that is equivalent to constructing a hot mix overlay to improve the ride quality. This is primarily to encourage the contractors to improve the ride quality of concrete projects during the construction phase itself rather than applying any smoothness improvement actions such as diamond grinding. Indeed, it is helpful to TxDOT to save an additional opportunity to improve skid resistance in the future, as the number of grinding passes allowed is limited before exposing the underlying steel reinforcement. Figure 5.11, Figure 5.12, and Figure 5.13 show the proposed pay adjustment for concrete projects under Schedules 1, 2, and 3. Schedule 1 primarily targets newly constructed rigid pavement with room for improving the initial ride quality, including enhanced base construction before concrete layer placement. The specification is essentially similar to the HMA specification with larger penalties. A maximum penalty of $6,000 is imposed on concrete projects that are rougher than 60 inches/mile under Schedule 1, and 70 inches/mile under Schedule 2. On the other hand, Schedule 3 is similar to the existing HMA ride specification. The lack of performance data for concrete pavements is the only reason for proposing a pay adjustment system identical to that of HMA projects.

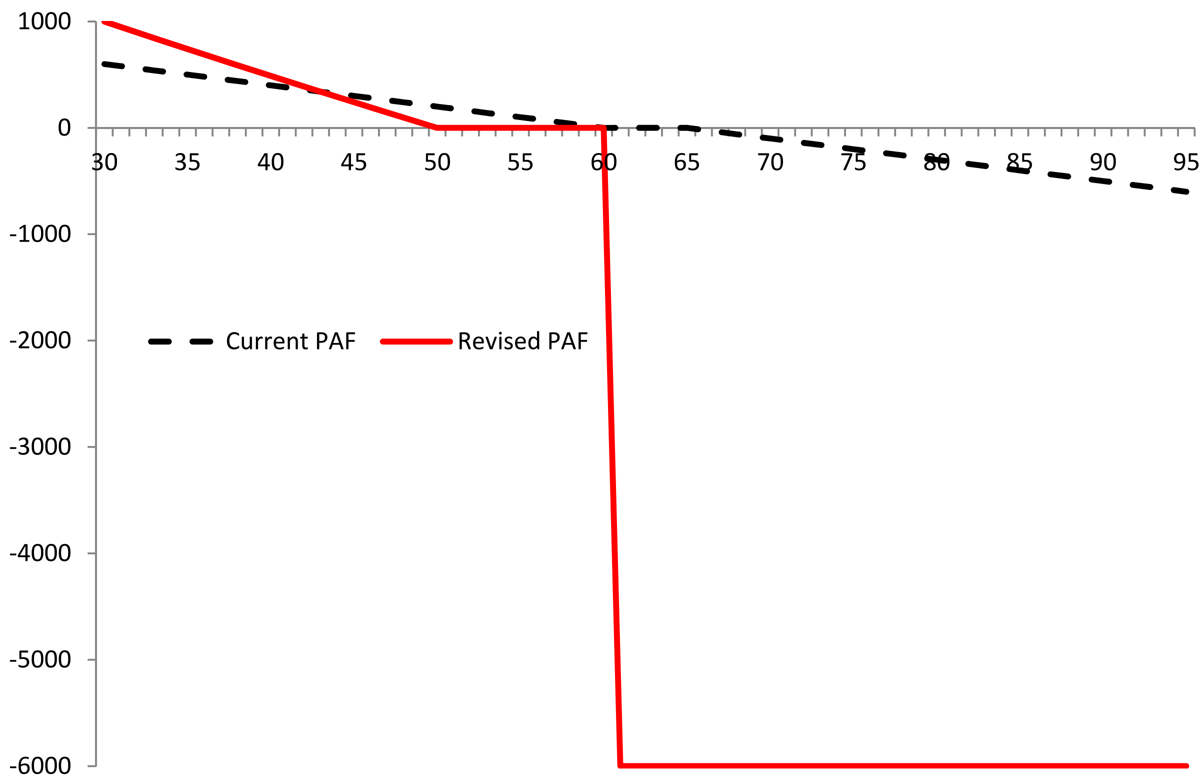

Figure 5.11: Current and revised concrete ride pay adjustment systems (Schedule 1)

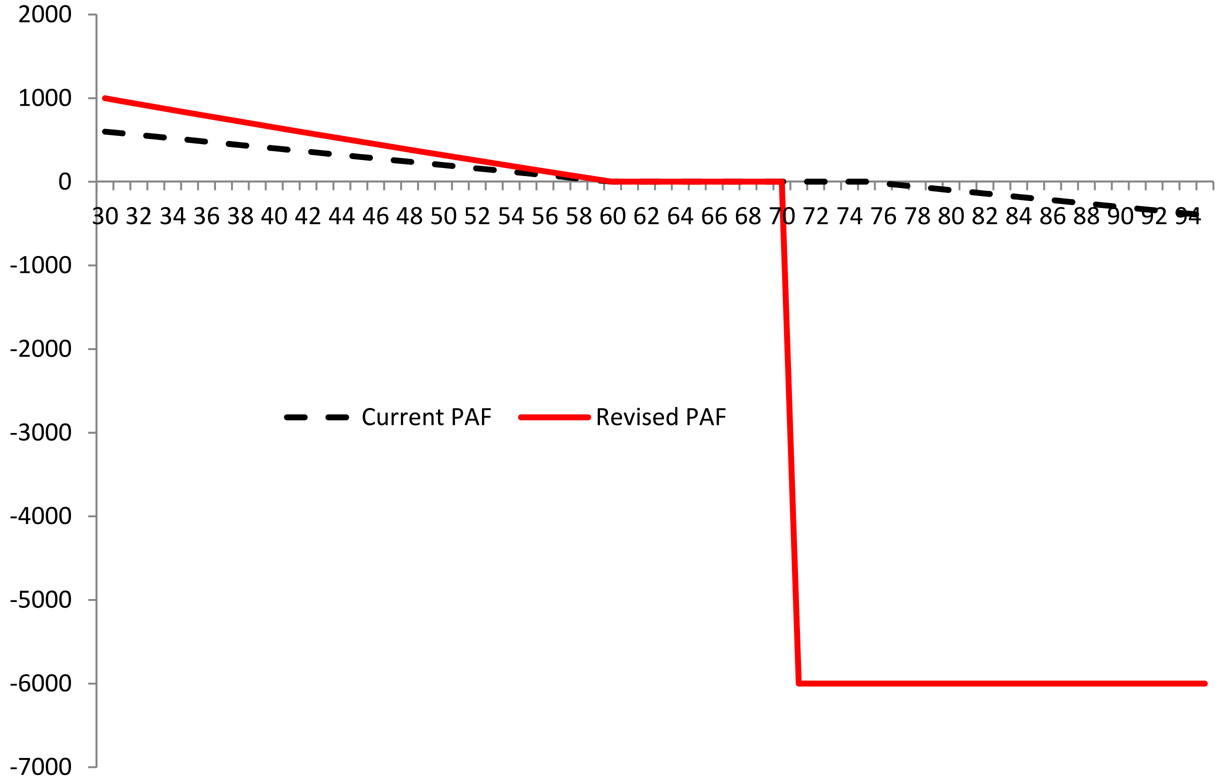

Figure 5.12: Current and revised concrete ride pay adjustment systems (Schedule 2)

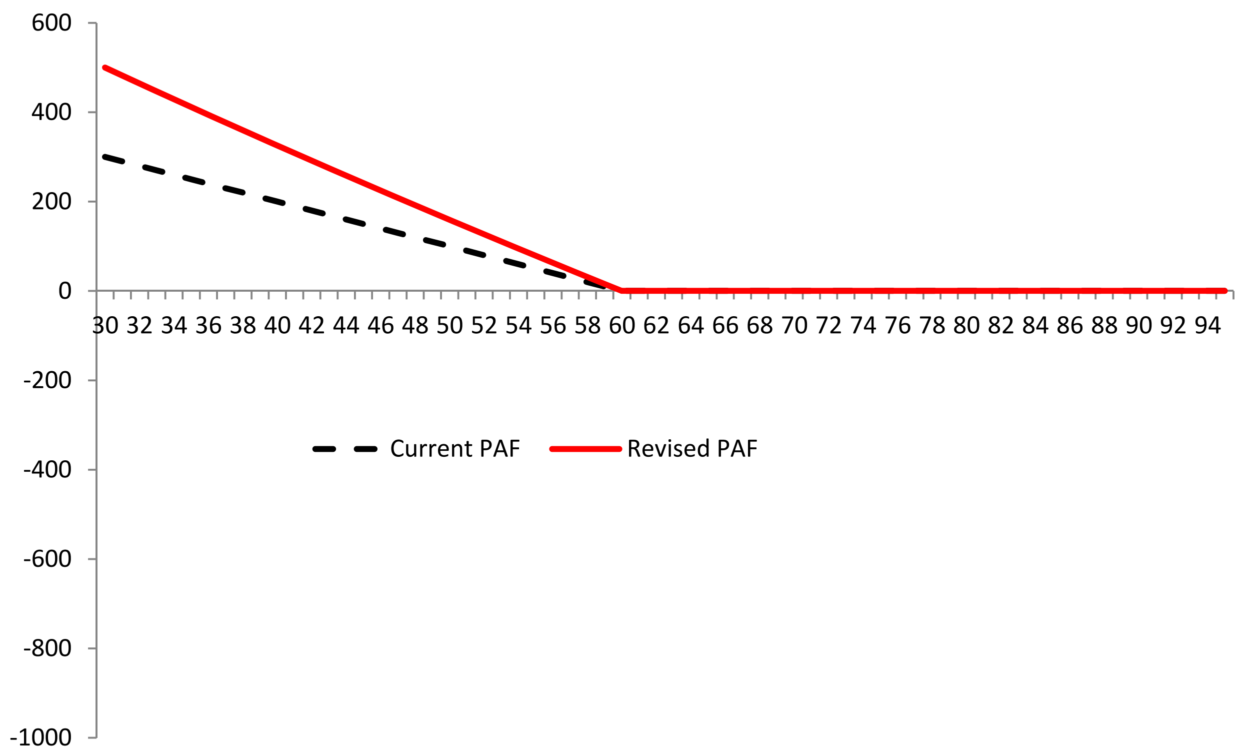

Figure 5.13: Current and revised concrete ride pay adjustment systems (Schedule 3)

The intention behind a separate concrete ride specification (with three schedules) is to cultivate a culture for improved ride quality during each stage of the construction, possibly by measuring the ride starting from the base construction and correcting for rougher areas at each stage of construction. The proposed specification encourages the concrete contractors to improve the ride quality as part of the construction process and restricts the execution of unwanted smoothness improvement actions such as diamond grinding. Under the revised ride specification, a concrete project contractor delivering a product with an inferior initial ride is bound to construct an asphalt overlay, which greatly improves the initial ride quality. It is believed that contractors will *learn the system* with time as in the case of HMA projects. Thus, the overall smoothness levels on concrete pavements will be improved across Texas in the long run and save significant user costs and vehicle operating expenses on concrete pavements.

Chapter 6. Validation of Sample Size and Sampling Methods

Two important aspects that do affect the final calculation of the PAFs are sample size and sampling techniques. Altering the sampling size or method arguably affects the test results of QC parameters such as laboratory density and in-place air voids, which in turn influence the consistency and reliability of the production and placement PAFs. Russell et al. (2001) mentioned that the number of samples within a lot can affect risk levels for both the agency and contractor. Intuitively, the determination of quality parameters will be more precise and reliable as the sample size increases. However, larger samples sizes in turn increase the cost of testing. Therefore, it is plausible to hypothesize the existence of a point of diminishing return at which further increments of sample size are redundant because the cost of testing outweighs the benefit of capturing the variability of the quality. This research attempted to identify the optimal sample size and recommended changes in terms of sample size and sampling techniques.

This chapter is organized as follows. Firstly, a brief note on the sampling methods from the literature including the practices in Texas is provided. Subsequently, a testing plan to evaluate the adequacy of the current sampling scheme is discussed along with the support of a statistical analysis. The chapter concludes by highlighting recommendations to the existing sampling schemes used by TxDOT.

6.1 Sample Methods

The choice of sampling method often significantly impacts the results of the quality tests. Elseifi (2007) discussed that even though the testing can be conducted according to specifications, the sampling method is equally important to ensure that the tested material is representative of the installed product. Various methods have been suggested to properly collect HMA samples in the past few decades; these include samples taken either at the plant from a loaded truck, or on the roadway behind the paver. However, because segregation and contamination of the collected samples can easily occur, care must be exercised to ensure success of the sampling process. For instance, the coarse aggregate coated with asphalt binder usually tends to roll down the side of the pile of an HMA mixture and accumulate next to the sides and the ends of the truck bed. As another example, the contamination of sampling mix with the tack coat material may potentially hamper the test results and should be minimized. Literature suggested that selection of sampling method is dependent on the type of QC parameter that is being measured. For instance, Kvasnak et al. (2005) reported that truck sampling is appropriate for measuring laboratory density, Rice densities, and asphalt binder content, whereas the ring and plate method is suitable for measurement of VMAs. The sampling method is also dependent on the type of hot mix sampled.

In Texas, typically truck sampling is specified at the plant and is normally executed by the contractor. Truck sampling involves removal of the upper layer of hot mix material followed by scooping adequate sampling material using a square shovel from individual trucks; each hot mix sample will be obtained from more than one truck load. Segregation of the material is a problem, as it produces inaccurate test results. For example, a segregated sample typically indicates erroneously lower asphalt binder contents during testing (Roberts et al. 1996). The truck sampling method is very quick and inexpensive as it only requires a shovel, bucket, and sampling platform. However, it is often criticized as not being representative of the hot mix sample as most of the sample is taken from the top of the mound and on the side closest to the sampling platform. In addition, the truck sampling method does not account for any additional asphalt absorption during transportation and compaction. It is recommended that the truck sampling as currently employed be continued as it is inexpensive and contractors are familiar with this procedure; however, TxDOT should ensure strict control of the track sampling protocol and should nullify the pay adjustment in case of any violations.

6.2 Sample Frequency

6.2.1 Data Collection

The fundamental idea behind the sampling data collection is to replicate the existing QA process but with a larger number of samples. UT Austin and UT Tyler research teams have independently collected hot mix samples from different plants across Texas. The UT Austin team collected the following mixes: 1) Type B mix from Austin, 2) Type D mix from Austin, 3) Type C mix from El Paso. The mix samples were collected at random intervals during a typical production day. It is to be noted that the sample asphalt mixes are obtained from truck loads for Type B and Type D mixes, whereas the Type C mix samples are collected from the belt that feeds the silos. The UT Austin team prepared 6 to 18 equivalent laboratory samples for all these mixes while TxDOT currently uses 3 laboratory samples for QA; Table 6.1 shows detailed sampling frequencies of these mixes. Laboratory density, asphalt content, and gradation were measured at the UT Austin and TxDOT asphalt laboratories.

Table 6.1: Sampling frequencies

Mix Type	Production at Sampling	Laboratory Density		Asphalt Content		Gradation	
		UT	TxDOT	UT	TxDOT	UT	TxDOT
Type B	500 ton	18	3	3	1	3	1
	1,000 ton	18	3	3	1	3	1
Type C	50 Ton	6	3	3	1	3	1
	500 ton	6	3	3	1	3	1
	1,000 ton	6	3	3	1	3	1
Type D	500 ton	18	3	3	1	3	1
	1,000 ton	18	3	3	1	3	1
	1,500 tons	18	3	3	1	3	1

In addition, UT Tyler collected Type C hot mix sample data for 2 days from the Armor Materials plant located on route FM 206 in Tyler, TX. The truck sampling technique was employed as following: 2 Lots (one full day per lot) * 4 Sublots per lot * 2 Samples per sublot * 3 specimens per sample (from 3 different trucks) resulting in a total of 48 specimens. A lot is considered one day's production, which was 1,000 tons during data collection. Each specimen was collected from the vehicle containing the 41st ton of the lot.

6.2.2 Testing Results

Test results are presented in two subsections corresponding to experiments conducted by the UT Austin and UT Tyler teams consecutively.

UT Austin

As mentioned earlier, UT Austin collected Type B, Type C, and Type D mixes from different sources and replicated the QA process but with larger sampling frequency (18 specimens). The study team measured laboratory density, asphalt content, gradation. Table 6.2 shows laboratory density and asphalt contents for different sublots for the three commonly used mix types; Table 6.3 shows the gradation of the corresponding mixes. The sublot level standard deviation is considerably lower (on both laboratory density and asphalt content), indicating the consistency of the hot mix production operation. The sublot level sampling frequency arguably needs no further increase with such small variation in the laboratory density. However, a formal statistical test procedure is necessary to support this argument. Additionally, it is important to test the differences between sublots to evaluate the possibility of a lot level QC/QA that reduces the testing costs drastically (almost to one-fourth). ANOVA analysis is a well-suited technique for identifying any possible differences between sublot populations; the results and conclusions of the statistical tests are presented below.

Table 6.4 highlights the lack of statistical difference in the majority of the QC parameters for Type C and Type D mixes; note that Type B mixes belonging to different sublots differed significantly from each other in terms of the QC parameters. In other words, QC was essential at least once in each sublot for Type B mix production. On the other hand, Type C and Type D mixes were observed to be much more consistent across a day's production process (or lot), particularly in terms of asphalt content and gradation. However, filler content (Sieve #200) was different for the Type D mixes produced in different sublots. The data suggested that laboratory density varied significantly between sublots for all type of mixes (Types B, C, and D), which suggests that sublot level measurement of the laboratory density is essential for all types of mixes. In summary, the empirical findings highlight that the laboratory density measurement should be conducted at least once per sublot; in other words, the existing testing interval (i.e., three times per sublot) is adequate but cannot be reduced to lot-level measurements. On the other hand, the frequency of asphalt content and gradation measurement can be reduced to once per each lot, particularly for Type C and Type D mixes; however, sublot level measurement of asphalt content and gradation is necessary in the case of Type B mixes.

In addition, a further analysis to determine the optimum number of samples to be tested per sublot was carried out. A total of 18 samples per sublot were available in the Type B and Type D mix categories, while 6 samples per sublot were available for Type C mix. Different sets containing up to 18 samples are obtained using random draws from the original set of 18 (or 6 in the case of Type C) samples. For example, a set of 3 samples was drawn repeatedly with replacement from the original set of 18 samples (or 6 samples in the case Type C mix). The variation in each of these sets was plotted to understand the relationship between variability and the sample size. Of course, the variability reduces with larger sample size. The research team sought to identify the optimal sample size beyond which there is no significant gain in capturing the variability. Figure 6.1 to Figure 6.8 show the plots of variability versus sample size. The analysis suggested that a sample size of three per sublot is adequate for measuring laboratory density (same as the existing testing frequency). In other words, the current specification seems adequate and shouldn't be reduced further.

Table 6.2: Laboratory density and AC measurements

Mix Type	Production at Sampling	Laboratory Density		Asphalt Content	
		Mean	St.Dev	Mean	St.Dev
Type B	500 ton	96.1	0.3	5.72	0.16
	1,000 ton	95.9	0.4	5.04	0.09
Type C	50 Ton	96.9	0.1	5.08	0.14
	500 ton	97.4	0.2	5.21	0.04
	1,000 ton	97.0	0.2	5.18	0.01
Type D	500 ton	96.8	0.5	5.77	0.12
	1,000 ton	97.3	0.5	5.92	0.05
	1,500 tons	97.7	0.5	5.75	0.22

Table 6.3: Gradation measurements

Mix Type		Type B		Type C			Type D		
Production at Sampling		500 ton	1,000 ton	50 Ton	500 ton	1,000 ton	500 ton	1,000 ton	1,500 tons
Sieve # 1	Mean	100%	99%	-NA-	-NA-	-NA-	-NA-	-NA-	-NA-
	St.Dev	0%	2%	-NA-	-NA-	-NA-	-NA-	-NA-	-NA-
Sieve # 3/4	Mean	97%	92%	100%	100%	100%	97%	98%	96%
	St.Dev	97%	2%	0%	0%	0%	1%	1%	1%
Sieve # 3/8	Mean	78%	69%	84%	82%	79%	89%	92%	89%
	St.Dev	1%	5%	2%	3%	4%	1%	1%	1%
Sieve # 4	Mean	56%	46%	57%	58%	56%	58%	64%	62%
	St.Dev	1%	3%	2%	3%	3%	1%	2%	2%
Sieve # 8	Mean	41%	33%	40%	40%	40%	37%	42%	41%
	St.Dev	0%	2%	2%	2%	2%	0%	1%	1%
Sieve # 30	Mean	30%	24%	24%	25%	24%	25%	29%	28%
	St.Dev	0%	1%	1%	1%	1%	0%	1%	1%
Sieve # 50	Mean	25%	21%	17%	17%	16%	21%	24%	23%
	St.Dev	0%	1%	0%	0%	1%	0%	1%	1%
Sieve # 200	Mean	7%	6%	4%	4%	4%	6%	6%	6%
	St.Dev	0%	1%	0%	0%	0%	0%	0%	0%

Table 6.4: ANOVA analysis results sublot basis

QC Parameter	P-Value			Minimum recommended testing frequency		
	TY-B	TY-C	TY-D	TY-B	TY-C	TY-D
Laboratory Density	0.04	0.00	0	1 per sublot	1 per sublot	1 per sublot
AC%	0.00	0.40	0.35	1 per sublot	1 per lot	1 per lot
Sieve #1 (%passing)	0.37	-NA-	-NA-	1 per lot	-NA-	-NA-
Sieve #3/4 (%passing)	0.02	1.00	0.22	1 per sublot	1 per lot	1 per lot
Sieve #3/8(%passing)	0.05	0.40	0.05	1 per lot	1 per lot	1 per lot
Sieve #4 (%passing)	0.00	0.86	0.32	1 per sublot	1 per lot	1 per lot
Sieve #8 (%passing)	0.00	0.99	0.22	1 per sublot	1 per lot	1 per lot
Sieve #30 (%passing)	0.00	0.70	0.23	1 per sublot	1 per lot	1 per lot
Sieve #50 (%passing)	0.00	0.80	0.13	1 per sublot	1 per lot	1 per lot
Sieve #200 (%passing)	0.22	0.23	0.03	1 per lot	1 per lot	1 per sublot

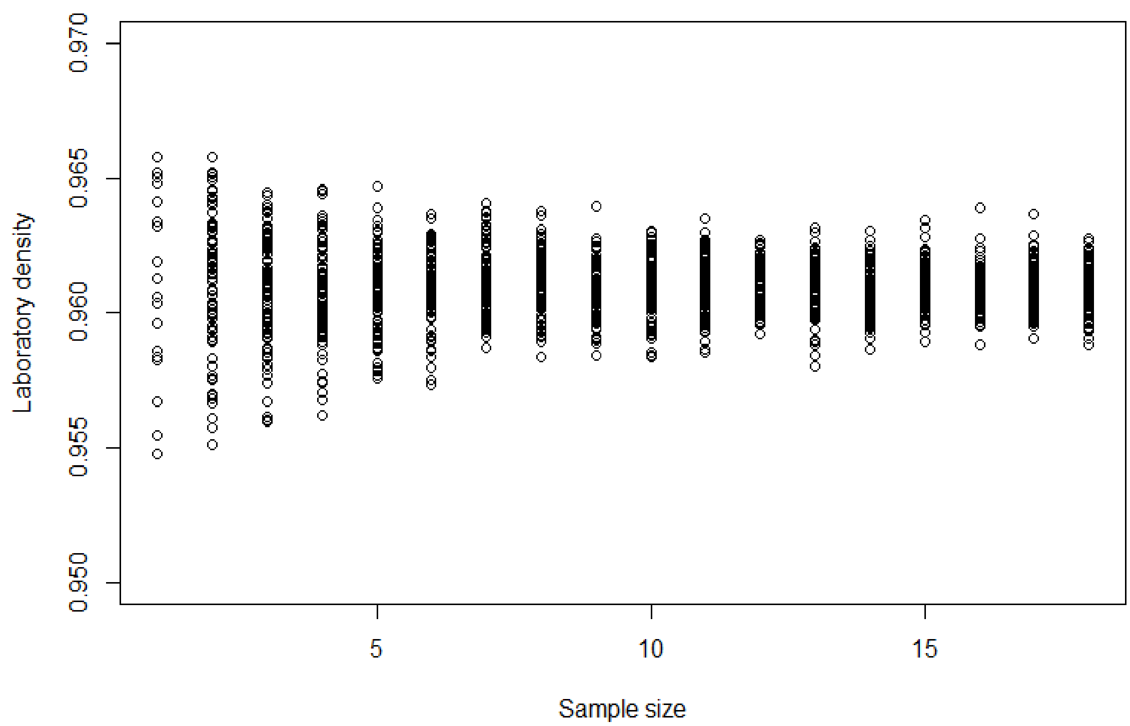

Figure 6.1: Laboratory density – sampling frequency – Type B mix; Sublot 1

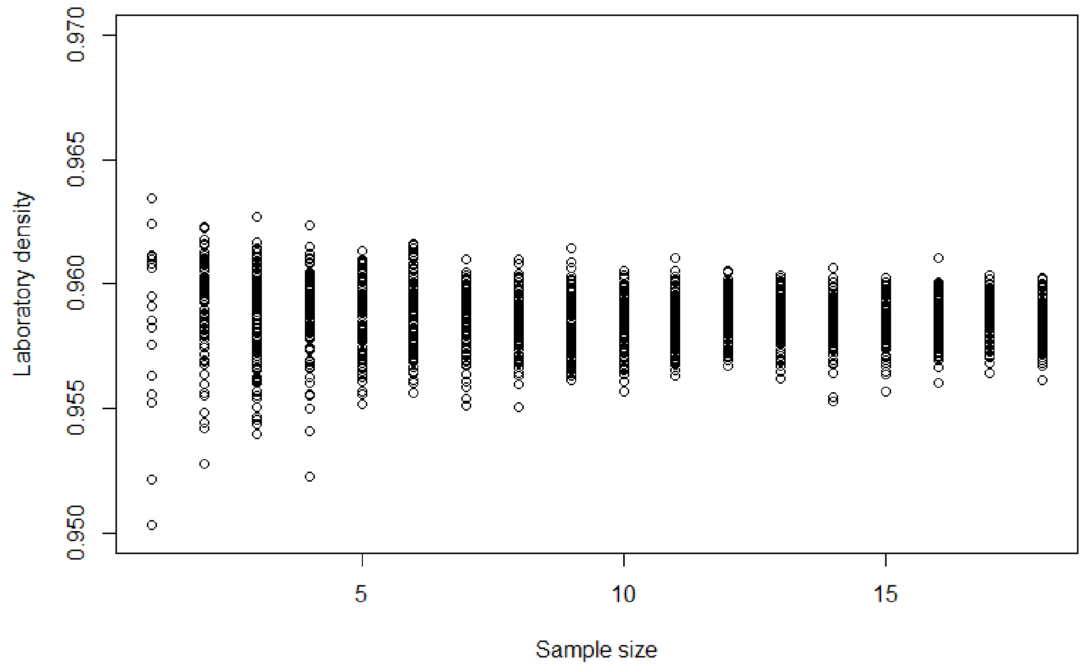

Figure 6.2: Laboratory density – sampling frequency – Type B mix; Sublot 2

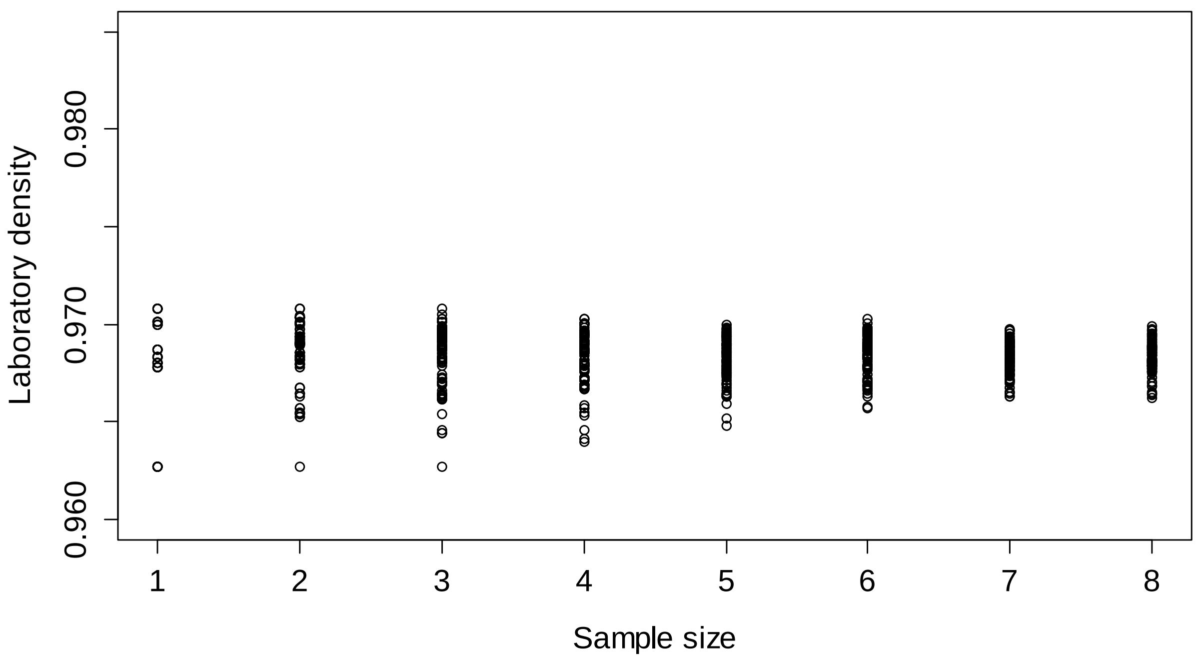

Figure 6.3: Laboratory density – sampling frequency – Type C mix; Sublot 1

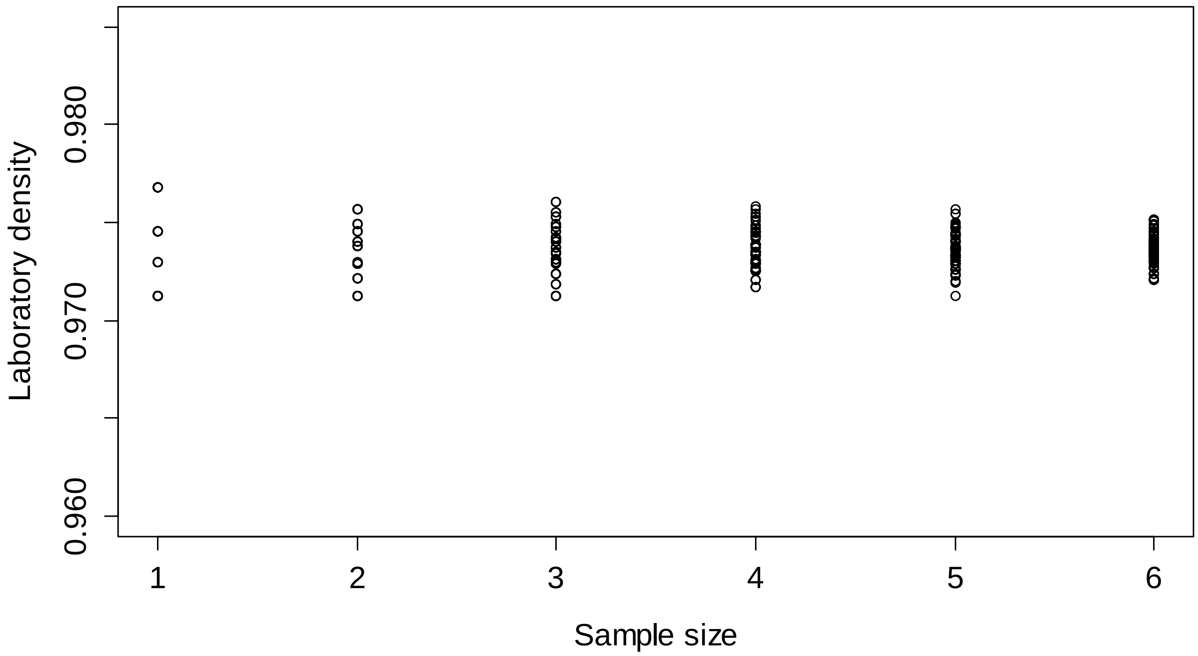

Figure 6.4: Laboratory density – sampling frequency – Type C mix; Sublot 2

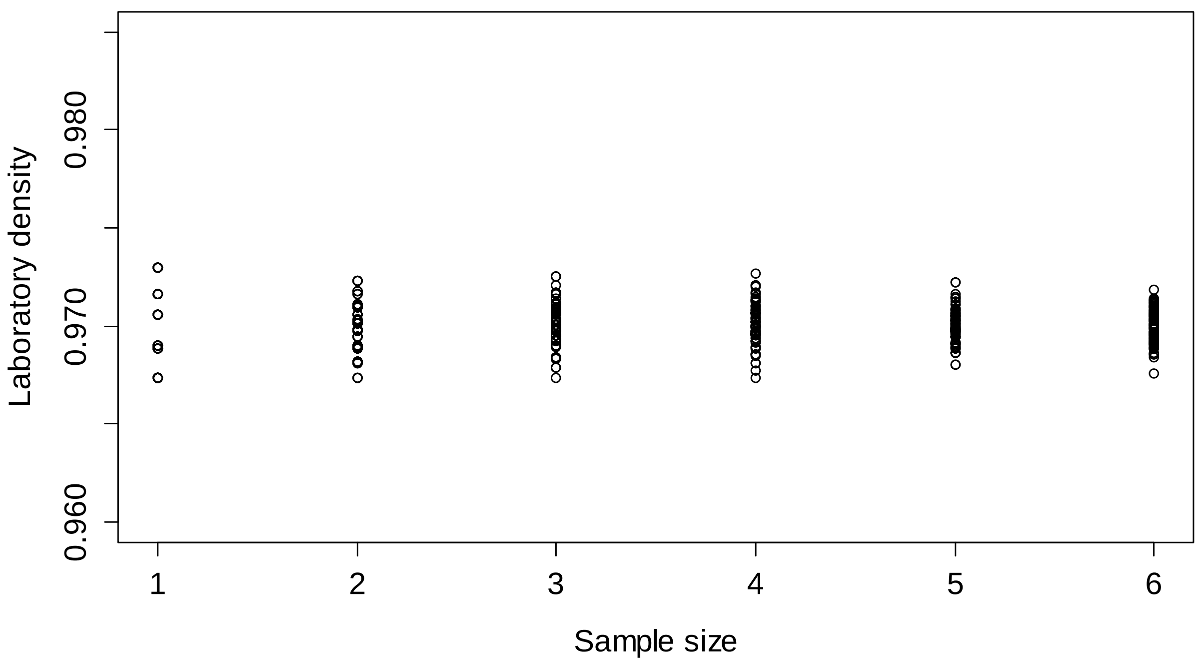

Figure 6.5: Laboratory density – sampling frequency – Type C mix; Sublot 3

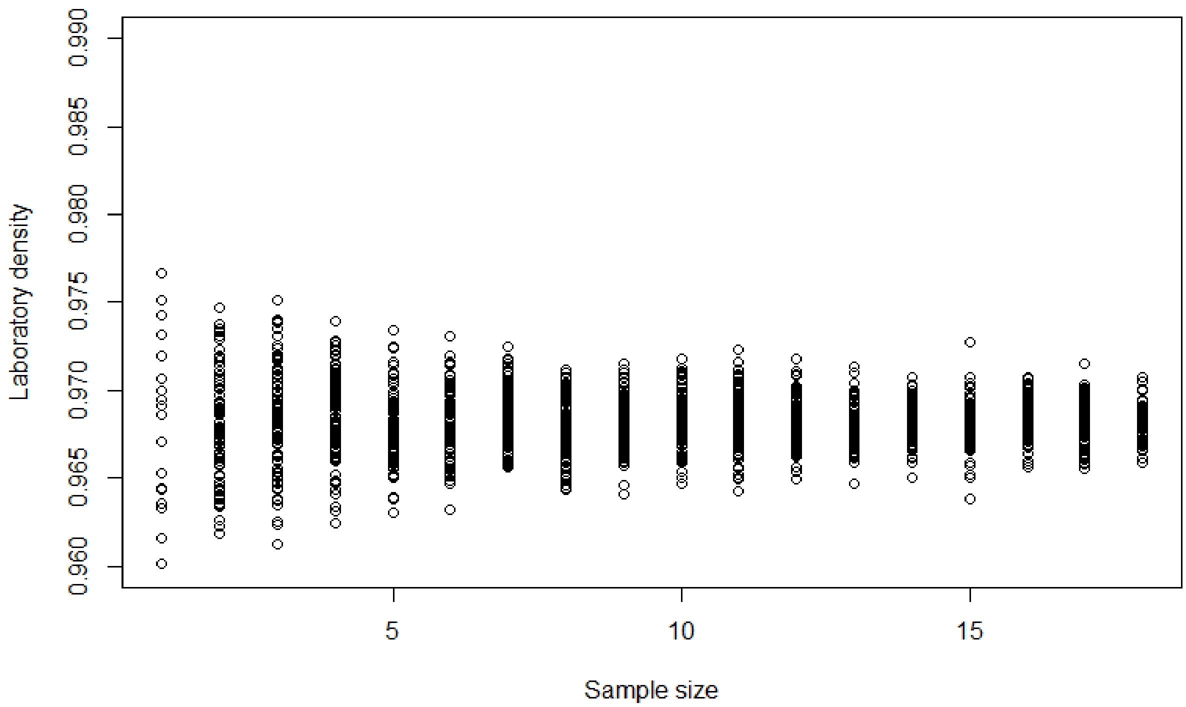

Figure 6.6: Laboratory density – sampling frequency – Type D mix; Sublot 1

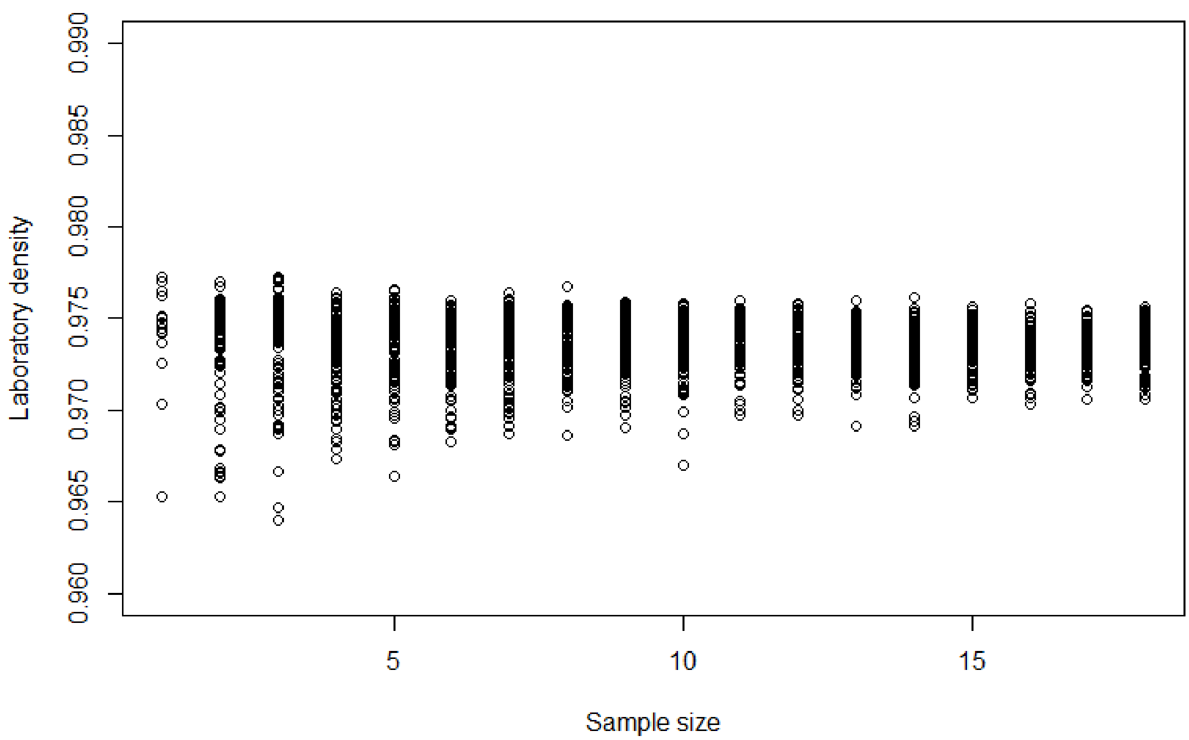

Figure 6.7: Laboratory density – sampling frequency – Type D mix; Sublot 2

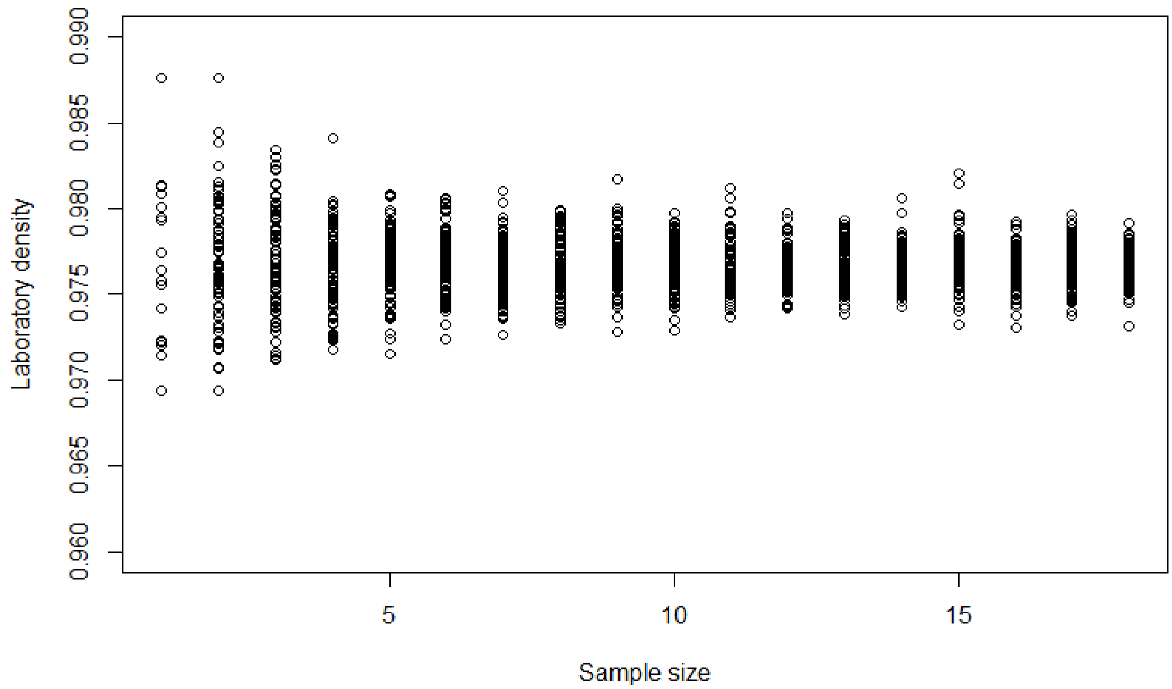

Figure 6.8: Laboratory density – sampling frequency – Type D mix; Sublot 3

UT Tyler

As mentioned earlier, UT Tyler collected Type C mix from one hot mix plant and replicated the QA process but with a larger sampling frequency (48 specimens). The study team measured laboratory density, asphalt content, and gradation. Figure 6.9 to Figure 6.12 show results of the hot mix volumetric properties from the 48 specimens. An average asphalt content of 4.56% with a standard deviation of 0.38% is reported. ANOVA was employed to determine if differences in means existed among the various subsets of the data. These subsets included comparing individual lots and sublots within a project.

First, the ANOVA analysis was carried out to evaluate any statistically significant differences between the mean QC parameters measured in two different lots. Table 6.5 shows the ANOVA analysis results corresponding to asphalt content, laboratory density, and gradation properties. The results suggest that measurements of asphalt content and laboratory density statistically differed across the two different lots, whereas the gradation measurements are not statically different across different lots. Next, the ANOVA analysis was repeated to evaluate any statistically significant differences between the mean QC parameters measured in different sublots.

Table 6.6 shows the ANOVA analysis results corresponding to asphalt content, laboratory density, and gradation properties. The results suggest that measurements of gradation, asphalt content, and laboratory density statistically did not differ across the four different sublots.

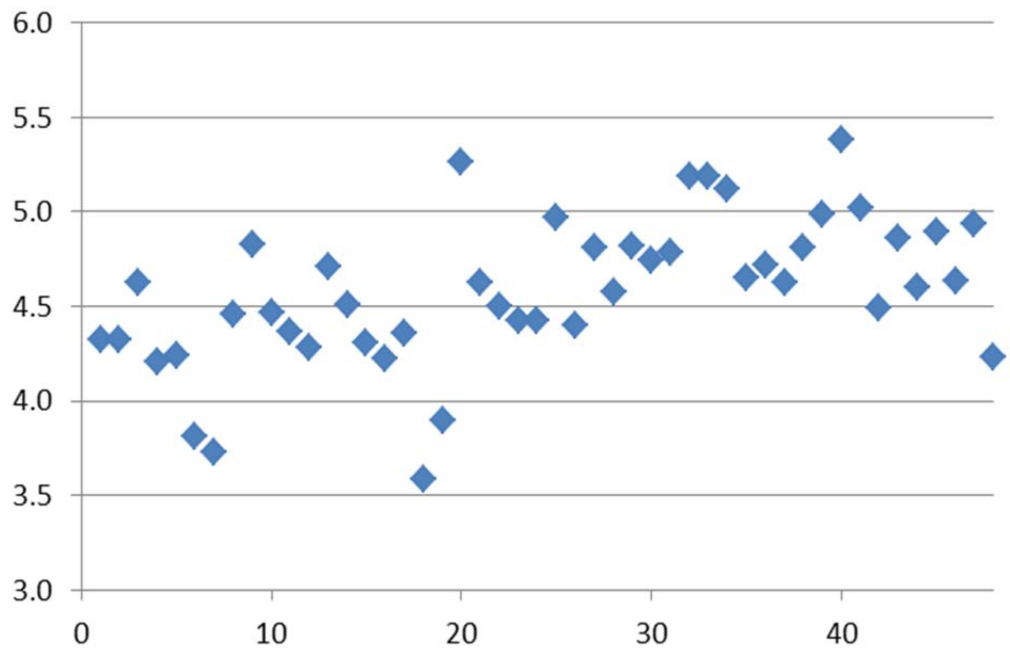

Figure 6.9: Asphalt contents (X-Axis: sample index, Y-Axis: AC%)

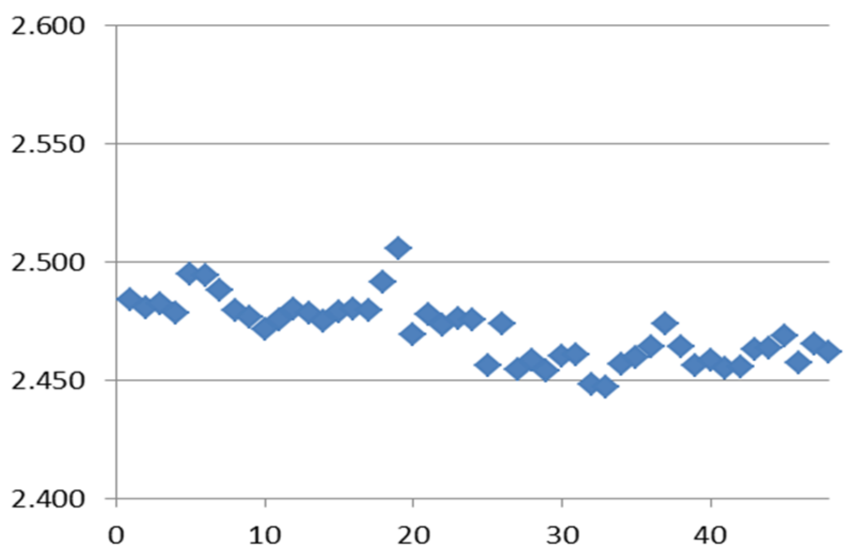

Figure 6.10: Maximum theoretical specific gravity (X-Axis: sample index, Y-Axis: Rice density)

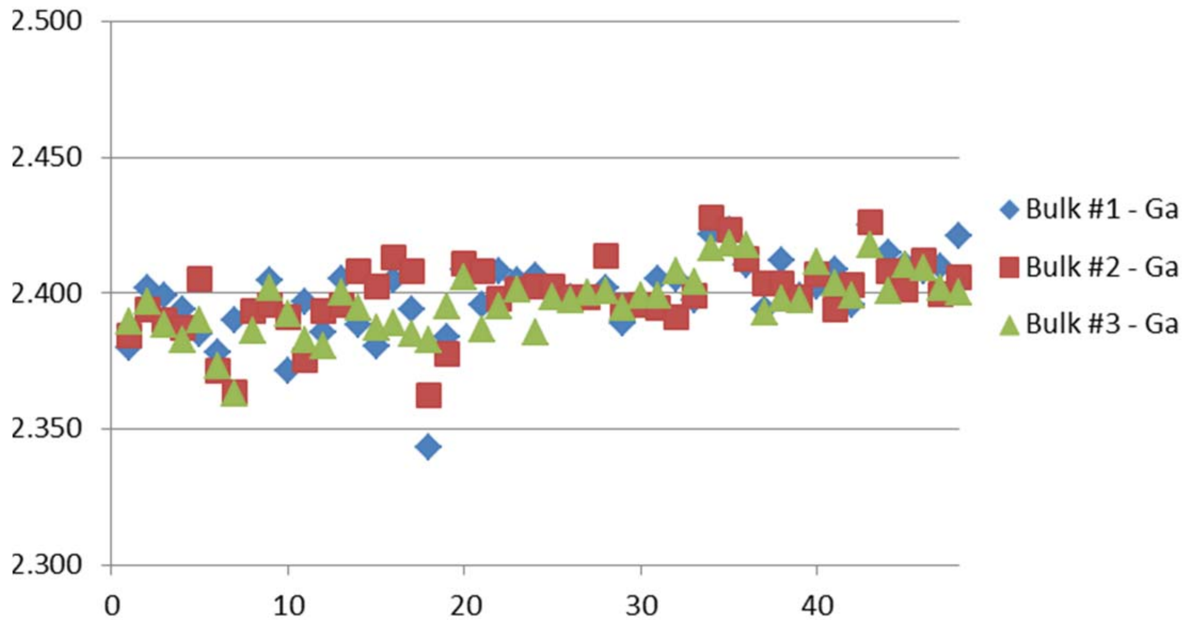

Figure 6.11: Bulk specific gravity (X-Axis: sample index, Y-Axis: Bulk density)

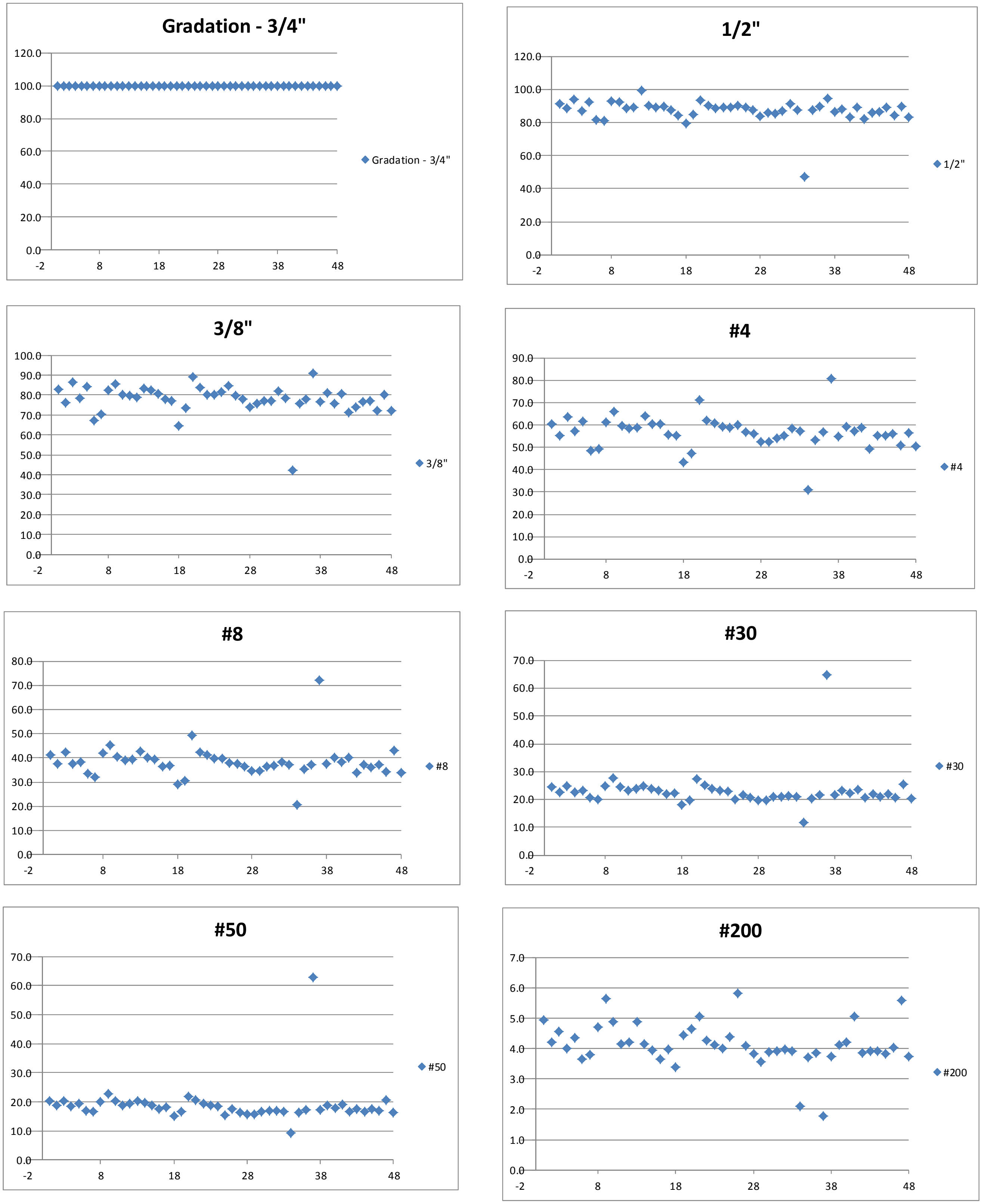

Figure 6.12: Gradation of HMA (X-Axis: sample index, Y-Axis: % passing from respective sieve)

Table 6.5: SAS ANOVA analysis result lot basis

Class Level Information					
Class	**Levels**	**Values**			
Lot	2	1, 2			
Number of Observations Read	48				
Number of Observations Used	48				
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Aggregate Gradation (% Passing)					
Sieve-3/4''	1	0.00000004	0.00000004	0	0.9596
Sieve-1/2''	1	128.278723	128.278723	2.72	0.1061
Sieve-3/8''	1	116.803155	116.803155	2.21	0.1435
Sieve- #4	1	104.204901	104.204901	2.07	0.1574
Sieve-#8	1	16.744494	16.744494	0.37	0.5447
Sieve-#30	1	3.555411	3.555411	0.08	0.7802
Sieve-#50	1	1.362167	1.362167	0.03	0.8662
Sieve-#200	1	1.64778205	1.64778205	3.37	0.0727
Asphalt Content					
AC(%)	1	2.49593216	2.49593216	24.92	<.0001
Density					
MaxTheoGravity	1	0.00543583	0.00543583	96.07	<.0001
Bulk1	1	0.00250527	0.00250527	15.96	0.0002
Density1	1	24.19615016	24.19615016	57.66	<.0001
Bulk2	1	0.00165127	0.00165127	10.66	0.0021
Density2	1	20.58825886	20.58825886	48.79	<.0001
Bulk3	1	0.00271651	0.00271651	36.44	<.0001
Density3	1	25.02213833	25.02213833	96.22	<.0001
Average Density(%)	1	23.22790234	23.22790234	76.22	<.0001

Table 6.6: SAS ANOVA analysis result sublot basis

Class Level Information					
Class	**Levels**	**Values**			
Sublot	4	1, 2, 3, 4			
Number of Observations Read	48				
Number of Observations Used	48				
Variable	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Aggregate Gradation (% Passing)					
Sieve-3/4''	3	0.00000741	0.00000247	0.15	0.9321
Sieve-1/2''	3	29.287697	9.762566	0.19	0.9032
Sieve-3/8''	3	64.867683	21.622561	0.38	0.7651
Sieve- #4	3	50.182974	16.727658	0.31	0.818
Sieve-#8	3	95.371212	31.790404	0.7	0.5554
Sieve-#30	3	134.830528	44.943509	1.02	0.3943
Sieve-#50	3	143.396165	47.798722	1.03	0.3884
Sieve-#200	3	1.28205537	0.42735179	0.82	0.4878
Asphalt Content					
AC(%)	3	0.16743209	0.0558107	0.35	0.7864
Density					
MaxTheoGravity	3	0.00018508	0.00006169	0.35	0.7924
Bulk1	3	0.00161681	0.00053894	2.92	0.0443
Density1 (%)	3	3.21121343	1.07040448	1.17	0.3323
Bulk2	3	0.00063943	0.00021314	1.15	0.3385
Density2 (%)	3	1.18899974	0.39633325	0.45	0.719
Bulk3	3	0.0004378	0.00014593	1.12	0.3493
Density3 (%)	3	1.27348135	0.42449378	0.52	0.6687
Average Density (%)	3	1.63550415	0.54516805	0.67	0.5728

6.3 Recommendations

First, the researchers recommend using the quick and inexpensive truck sampling method to collect hot mix for QC/QA testing. Although the QC/QA test results will be more precise and reliable as the sample size increases, the current sampling frequency employed by TxDOT is deemed to be adequate. A simulation exercise using the sampling data from 18 specimens (for each mix type tested) confirmed that a sampling frequency of three per sublot is sufficient to capture the variability and cannot be reduced further. ANOVA analysis of the sampling data confirmed that QC/QA measurements significantly differ between different both lots and sublots. The collection of QC/QA data in each sublot is recommended (at least 500 tons of production) with a sampling frequency of three specimens per sublot. Note that these recommendations are applicable to Type B, C, and D mixes.

Chapter 7. Conclusions, Recommendations, and Implementation

Currently, TxDOT uses a pay adjustment factor (PAF) system for the production and placement of hot mix asphalt (HMA) and the ride quality of HMA and concrete pavements; this system has been in existence for almost a decade. A review of data collected as part of this study suggests that TxDOT is consistently rewarding hot mix contractors, which indicates that the contractors have mastered the system to gain financial incentives over time. This trend reflects the appreciable success of the existing pay adjustment specification in consistently achieving better quality pavements over time. Continuing the use of the pay adjustment specification that consistently rewards all the contractors is not financially justifiable. This research project identifies the need to change the existing pay adjustment specification and proposes incorporation of pavement performance into the pay adjustment specifications. A performance-related specification (PRS) is arguably financially justifiable as it is primarily based on the relationship between measurable parameters in a construction project and expected performance. It also ensures that bonuses awarded to contractors don't exceed the benefits to the highway agency, and that the penalties levied on contractors don't fall short of the potential losses incurred by the agency. A rational performance-based pay adjustment system improves the contractual relationship by balancing the risks between TxDOT and contractors to benefit both parties in the long run.

A comprehensive review of the literature on development of performance-related PAFs is provided and key findings were highlighted in the beginning of the report. The literature review suggested that a PRS should employ only the QC parameters that are associated with and sensitive to the performance for pay adjustments. This report is broadly divided into two parts in terms of content: 1) research on hot mix pay adjustment specifications (production, placement, and ride quality), and 2) research on concrete pay adjustment specifications (ride quality).

Regarding the hot mix pay adjustment system, a comprehensive database of hot mix projects was established by integrating the SM and PMIS databases. An empirical relationship between the QC parameters that are commonly measured during hot mix projects and the respective field performance is developed using the aforementioned dataset and appropriate econometric techniques. The data analysis suggested that as-constructed ride quality, laboratory density, asphalt content, and in-place voids in the mineral aggregates (VMA) (calculated using in-place air voids) are associated with field performance, while controlling for all the other potential factors affecting the field performance, such as traffic, facility type, etc. In order to qualify a QC parameter as a candidate to potentially include in a PRS, a statistically significant relationship with field performance is necessary, but not sufficient. A change in the QC parameter should translate into a practically appreciable pavement performance. Consequently, performance-sensitive QC parameters that appreciably affect performance were identified using a variance-based sensitivity analysis. It is important to note that the applicability of the empirical findings of this research is restricted to the existing specification limits on each of the QC parameters, which are generally narrow ranges.

Analysis of the data suggested that lower as-constructed roughness and higher laboratory densities are associated with better field performance. Also, it was empirically evident that higher asphalt contents, lower in-field VMA, or equivalently, lower in-place air voids are associated with better pavement performance. Based on the sensitivity analysis, reducing the as-constructed roughness levels from 60 inches/mile to 30 inches/mile translates to an average increase of 50% in pavement life. As another example, reducing the as-constructed roughness levels from 90 inches/mile to 60 inches/mile translates into an average increase of 30% in pavement life. This result indicates that the gain in the pavement life reduces non-linearly with the as-constructed roughness levels, which further emphasizes the need for building smoother pavements. The analysis also suggested that increasing the laboratory density of the plant-produced hot mix by 1% (being within the specification limits) translates into an average increase of 13% in pavement life. Sensitivity analysis also revealed that changes in asphalt content and in-field VMA or equivalently in-place air voids do not necessarily produce a practically appreciable improvement in performance, despite the statistically significant association of these variables with performance. It is believed that placement-related QC parameters such as in-place air voids are affecting as-constructed ride quality and indirectly influencing the field performance through their effect on the as-constructed ride quality.

Regarding the concrete pay adjustment system, adequate reliable data resources within the project duration could not be identified; however, a relatively smaller database of 25 projects was populated with initial ride quality data and field performance data spanning at least 3 years. These sources were evaluated by manual inspection accompanied with minimal statistical analysis and engineering judgment.

7.1 Revised Pay Adjustment System

A revised PAF system was developed by synthesizing the key findings from the empirical relationships, sensitivity analysis, expert opinion, and engineering judgment. The revised pay adjustment system (for hot mix) is largely supported by the empirical evidence (based on 614 HMA projects) and arguably ensures extended performance for an awarded unit bonus. The revised pay adjustment tables corresponding to production, placement, and ride quality of HMA and ride quality of concrete are provided in Appendix A. Recommendations are described below.

7.1.1 Revised Production Pay Factors: HMA

The current pay adjustment system for hot mix production is based on the deviation of the laboratory density from the target density. A revised PAF system is developed by right-skewing the existing pay adjustment system to reflect the positive association of the laboratory density with field performance. The new system reduced the incentive for producing hot mix with laboratory density that is lower than the target density and increased the bonus for producing higher densities. However, it is to be noted that the penalty portion is unchanged. It is anticipated that the contractors will achieve higher laboratory densities by using slightly higher asphalt contents.

7.1.2 Revised Placement Pay Factors: HMA

The current placement PAF is based on in-place air voids of field-cored specimens that are obtained at randomized paving locations. A left-skewed PAF system is proposed with enhanced rewards on higher compaction levels and increased penalties on under-compacted pavements relative to the existing placement pay adjustment specification. The revised placement PAF system financially encourages the contractors to increase compaction efforts during the placement operation to produce pavements with longer life. The bonus levels on the under-compaction side of the PAF system are not reduced, so as to avoid practical and contractual issues that may arise in moving from a bonus to a large penalty for a given placement density.

The research team recommends giving equal weight to both production and placement pay factors to calculate the combined pay factor that is applied to the unit rate of the asphalt mix. In other words, the unit bid rate of the hot mix will be multiplied by a factor that is calculated by averaging production and placement PAFs, as is currently specified.

7.1.3 Revised Ride Quality Pay Adjustment: HMA

The current ride quality pay adjustment provides a fixed dollar amount (bonus/penalty) for achieving a given as-constructed ride quality that is measured in terms of IRI inches/mile per 0.1 mile length of the project. The amount of bonus/penalty changes linearly with as-constructed ride quality with a maximum possible bonus of $600 at 30 inches/mile and penalty of $600 at 95 inches/mile. The pay adjustment system is divided into three schedules, which are employed depending on the ease of achieving the desired post-construction ride quality in a given project. Schedule 1 of the current ride specification rewards HMA projects that are smoother than 60 inches/mile (IRI), while it penalizes projects that are rougher than 65 inches/mile. Schedule 2 of the specification is slightly less restrictive, only penalizing projects that are rougher than 75 inches/mile. Schedule 3 does not penalize any projects, while the bonus is reduced to one-half of the bonus awarded by the Schedules 1 and 2. Contractors have mastered the existing pay adjustment system over time and are consistently winning incentives on hot mix projects. The current system does not account for field performance.

A revised ride specification is recommended to incorporate the relationship between the as-constructed ride quality and field performance. ***First***, the no bonus/penalty zones on the three schedules have been made more stringent to force the contractors to deliver even smoother pavements, as they are able to achieve the existing requirements. The revised ride specification rewards HMA projects smoother than 50 inches/mile (IRI), while penalizing projects that are rougher than 60 inches/mile under Schedule 1. Under Schedule 2, the bonus threshold is unchanged at the existing level of 60 inches/mile, although the revised specification penalizes projects that are rougher than 70 inches/mile. On the other hand, Schedule 3 is essentially unchanged in terms of the no bonus/penalty zone. ***Second***, a maximum bonus/penalty of $1,000 is proposed to accommodate any possible inflation under realistic assumptions on the inflation rate. ***Third,*** the linearity assumption of bonus/penalty is relaxed; proposed instead is the empirical relationship that reflects the true field performance data in evaluating the pay adjustment in dollars for a given as-contructed ride quality level, while respecting the previously mentioned thresholds. Thus, the revised pay adjustment system is proportional to expected life along with the aforementioned mentioned maximum bonus and penalty levels. The pay adjustment system for revised Schedules 1 and 2 are very similar, except for the location of the no bonus/penalty zones. The revised Schedule 3 bonus is equal to half of the bonus under Schedule 1 and 2 and does not penalize inferior ride quality.

7.1.4 Revised Ride Quality Pay Adjustment: Concrete

Due to the unavailability of sufficient data sources for concrete projects, the revised ride pay adjust factor system for concrete ride specification proposed here is primarily guided by intuition and engineering judgment. The current ride pay adjustment for concrete projects is identical to that of the hot mix specification (Item 585). Typically, concrete pavement projects follow mainly Schedule 2 and in few cases Schedule 3 for ride quality pay adjustments. It was noticed that a majority of the concrete contractors are opting to pay a penalty rather than deliver a smoother surface or at least apply a corrective action such as diamond grinding. Based on empirical findings, a revised pay adjustment system for concrete projects is proposed that is similar to that of the revised hot mix ride specification except for the penalty zones. The revised concrete ride specification comprises three different schedules, just as for the hot mix ride specification. The bonus zones and the no bonus/penalty zones of the concrete ride specification are identical to that of the hot mix ride specification. However, the penalty for delivering a rougher pavement is increased by imposing a $6,000 penalty per 0.1-mile length of the project for delivering a pavement that is rougher than the no bonus/penalty zones. This penalty reflects the typical cost of an asphalt overlay placed to correct unacceptable ride quality. The revised concrete ride specification encourages contractors to explore other opportunities for improving ride quality at the various stages of the concrete pavement construction. Additionally, the revised specification restricts the contractors from grinding rougher pavements, which reduces layer thickness. Smoother as-constructed ride quality reduces the accumulation of vehicle operating and user costs over time.

In summary, the revised production pay adjustment system is arguably financially justifiable as it rewards superior performance and penalizes inferior performance, thereby supporting the underlying philosophy of PRSs.

7.2 Recommendations on Sampling Frequency and Methods

This report also documents the recommendations on sampling frequency and sampling methods for QC/QA in hot mix projects, particularly for Type B, C, and D mixes. It is recommended that TxDOT continue to use the quick and inexpensive truck sampling method to collect hot mix for QC/QA testing. In addition TxDOT should continue to collect QC/QA data in each sublot with a sampling frequency of three specimens per sublot, per current specification requirements.

7.3 Implementation: Guidelines for a Validation Experiment

It is important to validate the proposed pay adjustment schedule for ride quality for rigid and flexible pavements and HMA production and placement against field observations before implementing the revised system. The validation of the proposed performance-based PAFs can only be based on currently available performance data, which is 2 years within the scope of this research project. Two years are not nearly sufficient to make any accurate judgment in terms of pavement performance unless early failures occur. For this reason, the research team designed an experiment to establish the medium- and long-term validation of the proposed PAF. Since it is not possible to ask contractors to run a mix through the plant that will result in them paying a penalty, to stop rolling before achieving the desired density, or to deliver a pavement with poor riding quality, a different approach to a typical experimental design is proposed for validation.

Information on actual projects that were constructed during the 2 years of the research project was collected; this information includes laboratory density, in-place air void content, and as-constructed ride quality. The performance of the identified pavement sections will have to be monitored for a period of no less than 5 years. A total of 33 pavement sections were identified for monitoring to evaluate the revised PAF for HMA production, placement, and as-constructed ride quality. Concrete pavement sections will not deteriorate as fast as HMA pavements. In most cases, concrete pavements retain their ride quality for longer periods. For this reason the research team opted not to monitor concrete pavements based on the empirical finding from the available performance dataset on older concrete pavements. At the end of this analysis period, the data from the 33 hot mix projects should be analyzed and necessary adjustments to the revised hot mix PAFs should be made and implemented. A brief description of the selected project locations, volumetric characteristics, distributions of the awarded pay adjustments, and the guidelines for monitoring the selected pavement sections are provided below.

7.3.1 Identified HMA Projects

Thirty-three HMA projects that were constructed during 2011–2012 were identified across Texas; the SiteManager (SM) database was used for project identification. Table 7.1 and Table 7.2 show the number of projects and pavement sections in each district as well as the facility-wide distribution.

The revised pay adjustment system requires laboratory density for production pay factor calculation, in-place air voids for placement pay factor calculation, and initial ride quality data for ride pay adjustment calculation. These properties were collected from the SM database for each of the identified projects. Table 7.3 includes descriptive statistics of the identified HMA sections. The ride quality immediately after construction varied between 42 inches/mile to 110 inches/mile with a mean value of 68 inches/mile. The ride pay adjustment per 0.1-mile length of an HMA project varied between a penalty of $3,545 and a bonus of $18,410 with a mean pay adjustment of $1,592 (bonus). The dollar amounts corresponding to the ride pay adjustments include the bump penalty. The laboratory density varied between 95.9 and 97.2% with a mean value of 96.8%. The placement density varied between 4.8% and 7.5% with a mean value of 6.5%. The mean production and placement pay factors are also provided in Table 7.3.

In addition to the descriptive statistics, the distributions of the volumetric properties and awarded pay adjustments are provided in Figure 7.1 through Figure 7.5. These distributions highlight the adequately diverse range of the identified validation sections. It is hypothesized that the pool of projects with higher laboratory densities, lower in-place air voids, and lower as-constructed ride quality will perform better over time compared to the complementary pool of projects. A validation study that tracks the performance of these identified HMA sections over time will provide potential empirical evidence for evaluating the researcher team’s hypothesis.

Table 7.1: District distribution of the validation projects

District	HMA # sections
Amarillo	0
Atlanta	1
Austin	2
Beaumont	0
Brownwood	1
Bryan	3
Corpus Christi	1
Dallas	5
El Paso	1
Fort Worth	4
Houston	0
Laredo	1
Lubbock	1
Paris	2
Pharr	5
San Angelo	1
Tyler	0
Waco	1
Wichita Falls	4

Table 7.2: Facility distribution of the projects

Facility	HMA # sections	Concrete #Sections
FM	6	2
IH	4	6
LP	2	1
RM	1	0
SH	5	1
SL	1	0
SP	1	0
US	13	0

Table 7.3: Statistical summary of the validation projects

Feature	Min	Max	Mean	Std.Dev
Length (miles)	0.3	10.2	3.9	2.7
Ride Quality (inches/mile)	42	110	68	15
Laboratory density (%)	95.9	97.2	96.8	0.4
In-place air voids (%)	4.8	7.5	6.4	0.7
Plant VMA	13.2	16.6	14.7	0.7
AC %	4.5	5.5	5	0.3
Production PAF	1.01	1.05	1.031	1.01
Placement PAF	1	1.03	1.019	1
Ride pay (total $ amount) (includes bumps)	-1160000	1990000	154270	595597
Ride pay ($ average per 0.1 mile) (includes bumps)	-3545	18410	1592	4409

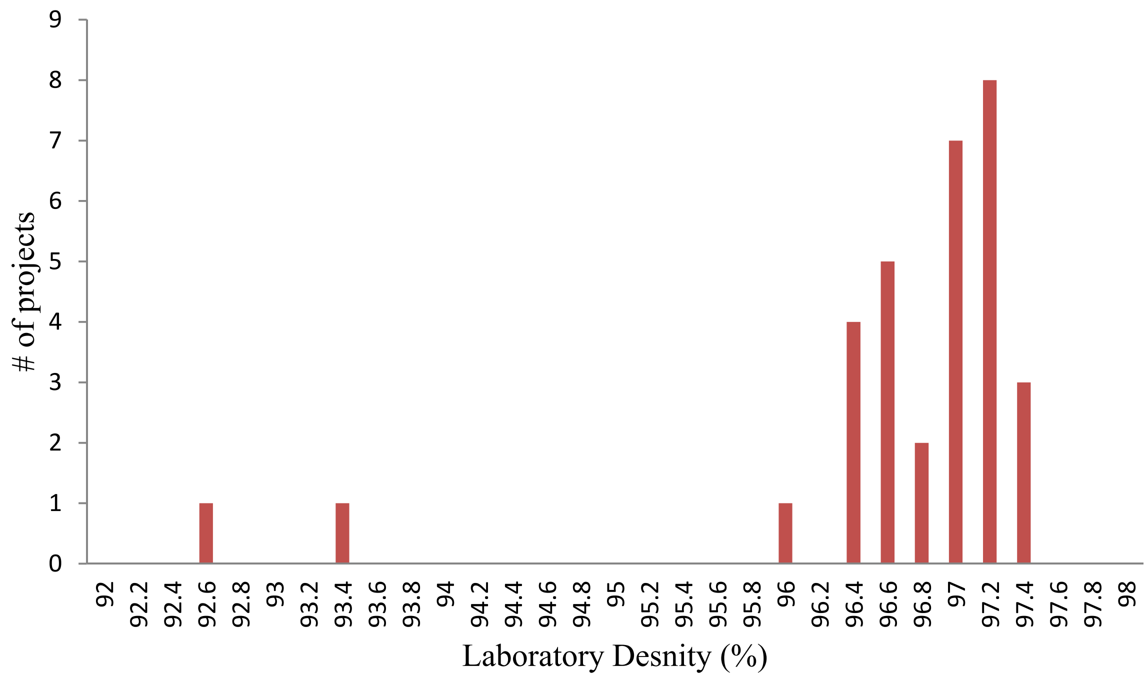

Figure 7.1: Distribution of the laboratory densities

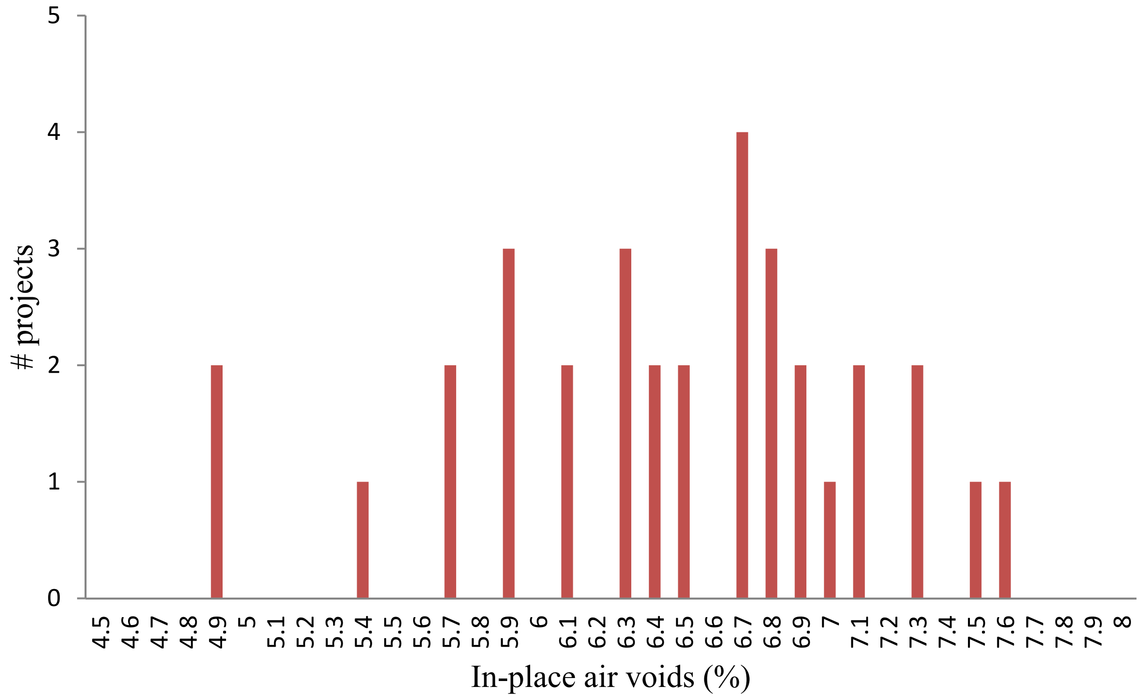

Figure 7.2: Distribution of the in-place air voids

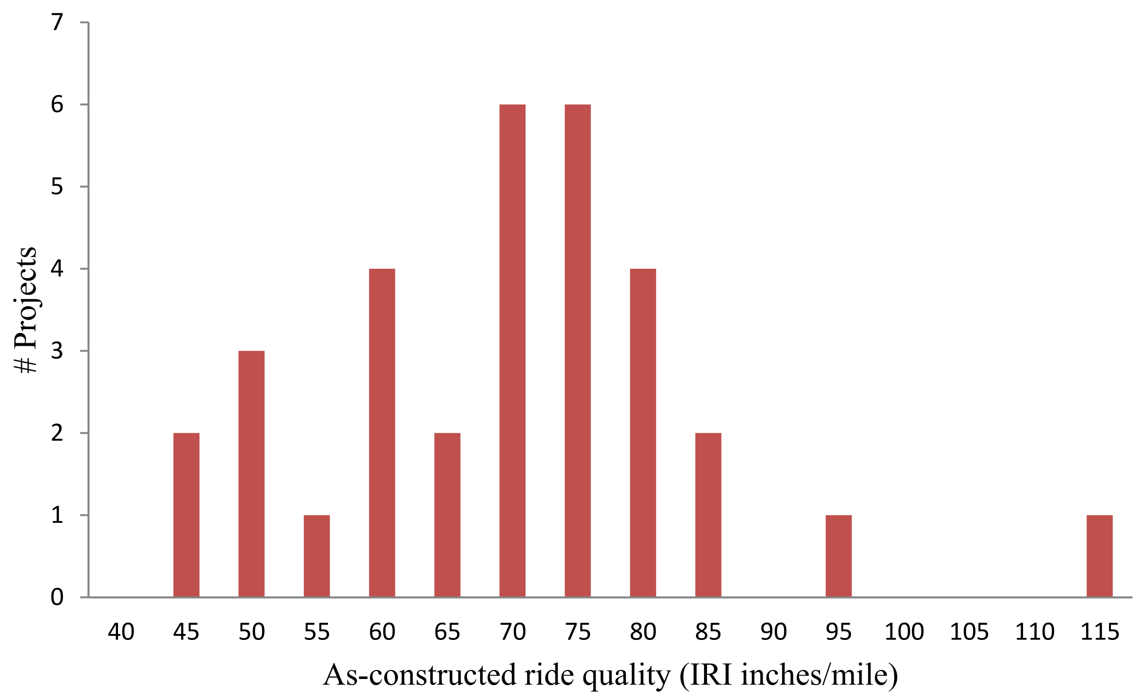

Figure 7.3: Distribution of the as-constructed ride quality

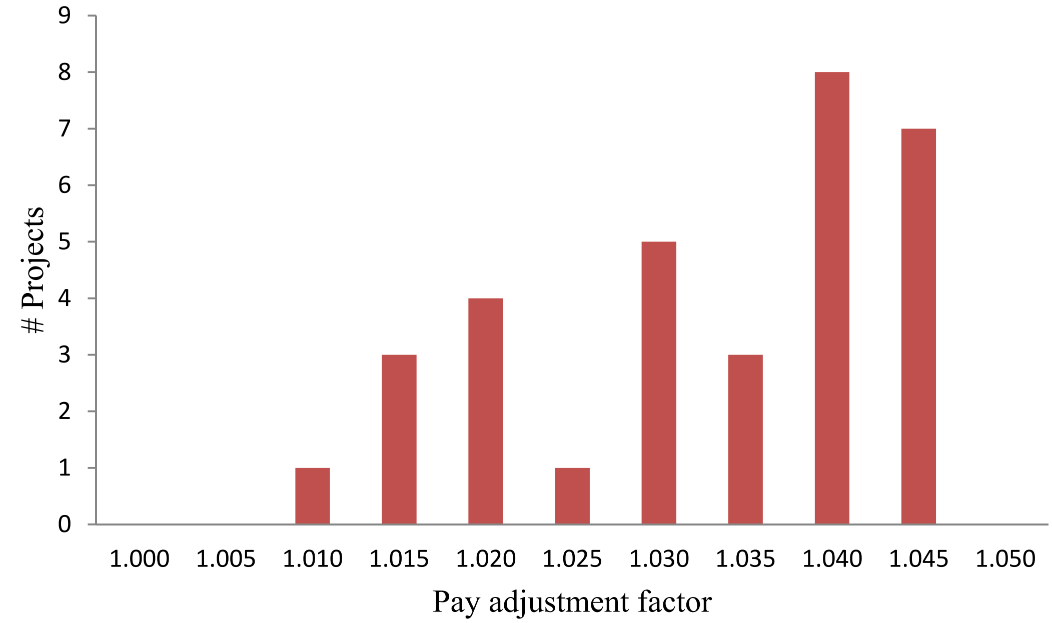

Figure 7.4: Distribution of awarded production PAF

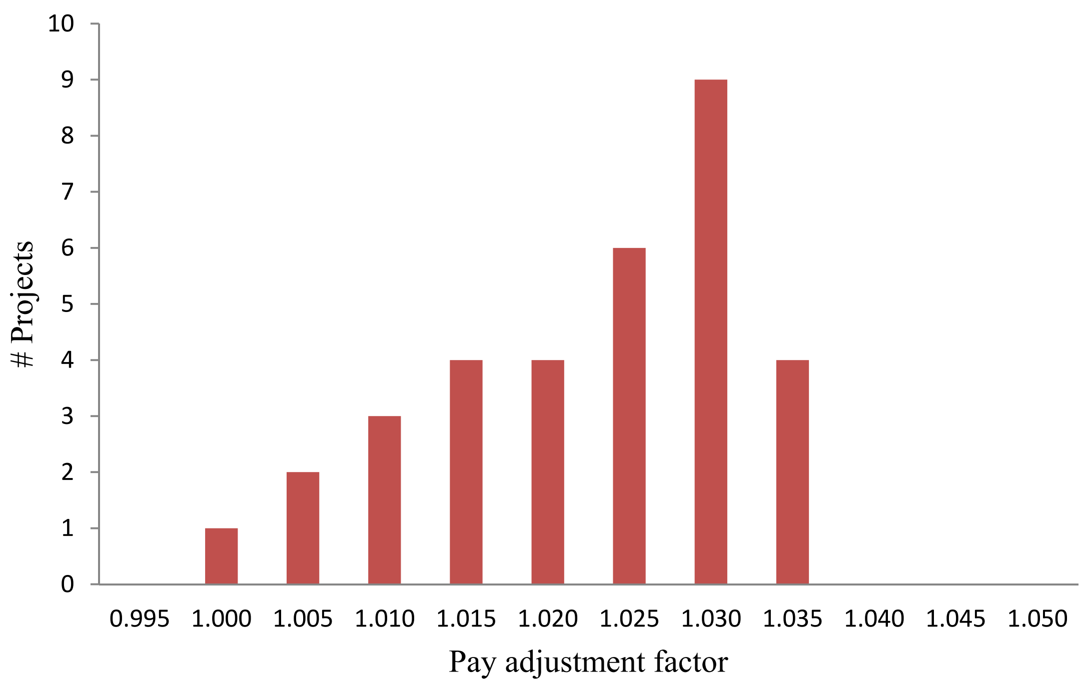

Figure 7.5: Distribution of awarded placement PAF

References

AASHTO. AASHTO Implementation Manual for Quality Assurance. American Association of State Highway and Transportation Officials, Washington, DC (1996).

AASHTO T-168-03. Sampling Bituminous Paving Mixtures. American Association of State Highway and Transportation Officials, February 2010.

Anderson, D.A., Luhr, D.R., and Antle, & C.E. Framework for Development of Performance-related Specifications for Hot mix Asphalt Concrete, NCHRP Report 332, Transportation Research Board, National Research Council, Washington, D.C, December 1990, 118 pp.

Arambula, E., Martin, A.E., Park, E.S., & Spielgelman, C. Assessment of the TxDOT Binder Quality Assurance Program. The Texas Transportation Institute, Texas Department of Transportation, Research and Technology Implementation Office, College Station, TX, December 2003.

Arizona State University, Furgo Consultants Inc. A Performance-Related Specification for Hot-Mixed Asphalt. NCHRP Report 704, Transportation Research Board, National Research Council, Washington, D.C., 2011.

Banerjee, A, Smit, A., & Prozzi, J.A. Influence of Operational Tolerances on HMA Performance. Construction and Building Materials 27, no. 1, pg. 15–23, February 2012.

Burati, J. & Weed, R. Accuracy and Precision of Typical Quality Measures. In Transportation Research Record: Journal of the Transportation Research Board, No. 1946, TRB, National Research Council, Washington, D.C., pg. 82–91, 2006a.

Burati, J. & Weed, R. Estimating Percent within Limits for Skewed Populations. In Transportation Research Record: Journal of the Transportation Research Board, No. 1946, TRB, National Research Council, Washington, D.C., pg. 71–81, 2006b.

Burati, J.L., Weed, R.M., Hughes, C.S. and Hill, H.S. Optimal Procedures for Quality Assurance Specifications. Report FHWA-RD-02-095, FHWA, U.S. Department of Transportation, 2003.

Burati, J.L., Hill, H.S., Hughes, C.S., & Weed, R.M. Evaluation of Procedures for Quality Assurance Specifications. Department of Civil Engineering, Clemson University, Federal Highway Administration, Office of Research, Development, and Technology, No. FHWA-HRT-04-046, October 2004.

Buttlar, G.W & Harrell.M. Development of End-Result and Performance-Related Specification for Asphalt Pavement Construction in Illinois. Transportation Conference Proceedings, Transportation Research Board, Washington, D.C, 1998.

Chamberlin P. W. Performance-Related Specifications for Highway Construction and Rehabilitation, NCHRP Synthesis 212, Transportation Research Board, Washington, D.C 1995.

Choi, Jae-Ho & Bahia, H.U. Life Cycle Cost Analysis-Embedded Monte Carlo Approach for Modeling Pay Adjustment at State Departments of transportation. Transportation Research Record, Journal of Transportation Research Board, No.1900, pg.86–93, Transportation Research board, Washington, D.C, 2004.

Chou, S. & Pellinen, T. K. Assessment of Construction Smoothness Specification Pay Factor Limits Using Artificial Neural Network Modeling. Journal of Transportation Engineering, Vol. 131, No. 7, July 1, 2005.

Deacon, J. A., Monismith, C. L., Harvey, J. T. & Popescu, L. Pay Factors for Asphalt Concrete Construction: Effect of Construction Quality on Agency Costs. California Department of Transportation, Berkeley, CA. February, 2001.

Deacon, J.A., Monismith, C., & Harvey, J. WesTrack: Performance Models for Permanent Deformation and Fatigue. Nichols Consulting Engineers, Reno, NV, June 2000.

Deacon, J.A., Monismith, C.L., & Harvey, J.T. Pay Factors for Asphalt-Concrete Construction: Effect of Construction Quality on Agency Costs. California Department of Transportation. Berkeley, CA, 1997.

El-Basyouny, M. & Myung, G.J. Probabilistic Performance-Related Specifications Methodology Based on Mechanistic-Empirical Pavement Design Guide. Transportation Research Record, Journal of Transportation Research Board, No.2151, pg. 93–102, Transportation Research board, Washington, D.C, 2010.

Elseifi, M. A. Evaluation of Hot-Mix Asphalt Sampling Techniques. Report No. FHWAICT-07-010, Illinois Center for Transportation, Urbana-Champaign, IL, 2007.

Epps, A.J., Hand, A., Seeds, S., Schulz, Todd., Alavi, S., Ashmore, C., Monismith, C.L., Deacon, J.A., Harvey, J.T., & Leahy, R. Recommended Performance Related Specification for Hot Mix Asphalt construction: results of the Westrack project, NCHRP Report 455, Transportation Research board, Washington, D.C, 2002.

Epps, A., Spiegelman, C., Park, E-S, Arambula, E., Ahmed, T., & Apanasovich, T. Initial Assessment of TxDOT Binder Quality Assurance Program. Texas Department of Transportation. Texas A&M, TX. December, 2001.

Federal Highway Administration (FHWA). Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance. Report to Congress, Washington, D.C, 1998, http://www.fhwa.dot.gov////policy/1999cpr/index.htm, accessed March 9, 2011.

Gillespie, T.D. Everything You Always Wanted to know about the IRI, But Were Afraid to Ask. Presented at the Road Profile Users Group Meeting, Lincoln, Nebraska, September 22–24, 1992.

Hand, A. J., Martin, A.E., Sebaaly, P, E., & Weitzel,D. Evaluating Field Performance: Case Study including Hot Mix Asphalt Performance-Related Specifications. Journal of Transportation Engineering, ASCE, April 2004.

Hughes, S.C. State Construction Quality Assurance Programs. NCHRP Synthesis 346, Transportation Research board, Washington, D.C, 2005.

Humphrey, D.N., Sandford, T.C., & Tweedie, J.J. Full-Scale Field Trials of Tire Shreds as Lightweight Retaining Wall Backfill under At-Rest Conditions. Transportation Research Record 1619, pg. 64–71., Transportation Research Board, National Research Council, Washington, D.C. 1998.

Irick, P.E. A conceptual Framework for the development of Performance-Related Materials and construction Specifications. Transportation Research Record 1126, pg. 1–27, Transportation Research Board, National Research Council, Washington, D.C., 1988,

Irick, P.E., Seeds, S.B., Myers, M.G., & Moody, E.D. Development of Performance-Related Specifications for Portland cement Concrete Construction. FHWA-RD-89-211, ARE Inc., Austin, Texas, May 1990.

Karimi, S. Increasing Durability of Hot Mix Asphalt Pavements Designed With the Superpave System. Master's Thesis, The University of Maryland, 2009.

Killingsworth, B.M. NCHRP project 9-15, Quality Characteristics and Test Methods for Use in Performance-Related Specifications of Hot-Mix Asphalt Pavements. Transportation Research Board, Washington, D.C., August 2004.

Kim, Y. R. Impact of Price Reductions on the Long-Term Pavement Performance of HMA Mixes in North Carolina. NC Department of Transportation Research and Analysis Group, Aug 2005.

Kim, Y. R., Lee, S. J., Seo, Y. & El-Haggan, O. A Mechanistic Approach to Determine Price Reduction Factors for Density-Deficient Asphalt Pavements. Proceedings of the 2006 Airfield and Highway Pavement Specialty Conference, Atlanta, GA, April 30–May 3, 2006.

Ksaibati, K., Mahmood, S. Utilizing the Long-Term Pavement Performance Database in Evaluating the Effectiveness of Pavement Smoothness. Research study conducted at The University of Wyoming, March, 2002.

Kvasnak, A. N. & Williams, R. C. Analysis of Variability in Pay Factor Systems and Sampling Methods. Journal of Testing and Evaluation, Volume 33, No. 4, July 2005.

Mahoney, J. P. & Muench, S. T. Quantification and Evaluation of WSDOT's Hot Mix Asphalt Concrete Statistical Acceptance Specification. Department of Civil and Environmental Engineering, University of Washington. 2001.

Mahoney, J.P. & Backus, A.W. QA Specification Practices. Washington State Transportation Center (TRAC), Seattle, WA, 2000.

Majizadeh, K. & Ilves, G.J. Correlation of Quality Control criteria and Performance of PCC pavements. FHWA/RD-83/014, Federal Highway Administration, Washington, D.C, March 1984.

Martin. A, E. TxDOT Binder Quality Assurance Program: Assessment and Improvement. Texas Transportation institute, Texas A&M University system, Project Summary Report 0-4047-S, 2003.

Michigan Department of Transportation, Michigan Test Method 324-99. Sampling Bituminous Paving Mixtures behind Lay Down Machine. Lansing, MI, 1999.

Mitchell T.M. Performance-Related Specification for Highway Construction and rehabilitation. Project No. 6G, FCP Annual progress Report, September 30, 1981.

Monismith C.L., Popescu, L., Harvey, J.T., & Hoover T.P. Performance Based Pay Factors for Asphalt Concrete. Proceedings of the 8th Conference on Asphalt Pavements for Southern Africa (CAPSA 04), Sun City, South Africa, September 2004.

Monismith, C.L., Harvey, J. T., Popescu, L., & T.P. Hoover Performance-Based Pay Factors for Asphalt Concrete Construction: Berkeley, CA, 2000.

Muench & Mahoney. WSDOT Acceptance Plan. 2000, http://classes.engr.oregonstate.edu/cce/winter2012/ce492/state_information/08_qa_specifications/wsdot_acceptance_plan.htm, Accessed on February 15th, 2013.

National Cooperative Highway Research Program (NCHRP). Field Procedures and Equipment to Implement SHRP Asphalt Specifications. Transportation Research Board, 9-7 Final Report, National Research Council, Washington, DC, 1996.

Palise, F., Strizki, B. and Weed, R. M. Upgrading the Asphaltic Concrete Specification of the New Jersey Department of Transportation. Transportation Research Record 1616, pg. 1–5, TRB, National Research Council, Washington, D.C., 1998.

Frazier, P. Implementation of a Statistically Based Quality Control/Quality Assurance (QC/QA) Procedure for HMA Construction. AL. No. 2019-14, Auburn University Highway Research Center, The Alabama Department of Transportation, Montgomery, April 1995.

Patel, A. Development of Fatigue Based Pay Adjustment Factors in Performance-Related Specifications for Full-Depth Asphalt Concrete Pavements. Doctorate's Thesis, The University of Illinois. October 1995.

Popescu, L. & Monismith, C. L. Performance-Based Pay Factors for Asphalt Concrete Construction: Comparison with a Currently Used Experience-Based Approach. California Department of Transportation, Division of Construction, Report # UCPRC-RR-2006-16, November, 2006.

Research and Development Program for Highway Construction Engineering Management. Report No. FHWA RD-90-34, Transportation Research Board, National Research Council, Washington, D.C., February 1990.

Roberts, F. L., Kandhal, P.S., Brown, E. R., Lee, D., & Kennedy, T. W. Hot Mix Asphalt Materials, Mixture Design, and Construction. 2nd Edition, NAPA Education Foundation, Lanham, Maryland, 1996.

Russell, J.S., Hanna, A., Nordheim, E., Russell, J., & Schmitt, R. Testing and Inspection Levels for Hot-Mix Asphaltic Concrete Overlays. National Cooperative Highway Research Program. Washington, D.C. 2001.

Seeds, Weed, S. B, Basavaraju, R. M., & Epps, R.J. Development of Performance-Related Specifications for Hot-Mix Asphalt Pavement Through Wes Track. In Transportation Research Record 1575, pg. 85–91, 1997.

Sha, S.C. Correlation of Construction Quality Criteria with performance of Asphaltic Concrete Pavement. FHWA/LA-87/208, Louisiana Transportation research Center, Baton Rouge, August 1987.

Shook, J.F, Diaz, M.A., Stroup-Gardiner, M., & Seeds, S.B. Performance related Specifications for Asphalt Concrete – Phase II. FHWA-RD-91-070, ARE Inc., Austin, Texas, January 1992.

Smit, A. F. & Prozzi, J. A. Developing a Sustainable Flexible Pavement Database in Texas. For Texas Department of Transportation Research and Technology Implementation Office. September 1, 2005 – Aug 31, 2008.

Swanlund, M. Enhancing Pavement Smoothness. Public Roads, pg. 20–22, October 2000.

Taute, A., Pretorius, D., Marais, & Grobler, J. Hot Mix Asphalt Quality control. Gaborone, Botswana, 2007.

Tong, Y. Towards Performance Based Pay Adjustment Factors. Master's Thesis, The University of Texas at Austin, August, 2009.

Turner, P. & West R. C. Evaluation of the Effect of Sampling Location on Laboratory Measured Quality Control Parameters. Transportation Research Record 1946, pg. 163–170: Journal of the Transportation Research Board, TRB, Washington, D.C., 2006.

TxDOT. Ride Quality for Pavement Structures, Item 585. Texas Department of Transportation – Construction Division, Flexible Pavements Branch, November 2004.

TxDOT. Standard Specifications for Construction and Maintenance of Highways, Streets, and Bridges. Texas Department of Transportation. 2004.

Vazquez, C, G., Aguiar-Moya, J, P., Smit, A, F., Prozzi, J. A. Laboratory Evaluation of Influence of Operational Tolerance (Acceptance Criterion) on Performance of Hot-Mix Asphalt Concrete. Texas Department of Transportation Research and Technology Implementation Office, June 2010.

Von Quintus, H.L., Rauhut, J.B., Kennedy, T.W. and Jordahl, P.R. Cost effectiveness of sampling and testing programs. FHWA/RD-85/030, Brent Rauhut Engineering Inc., Austin, Texas, January 1985.

Weed, R.M. A Rational Method for Relating As-Built Quality to Pavement Performance and Value. Transportation Research Record 1632, pg. 32–39, TRB, National Research Council, Washington, D.C., 1998.

Weed, R. M. Derivation of Equation for Cost of Premature Pavement Failure. Transportation Research Record: journal of the Transportation Research Board, No. 1761 TRB, National Research Council, Washington, D.C., 2001b.

Weed, R.M. Development of Composite Quality Measures. Transportation Research Record: Journal of the Transportation Research Board, No. 1712, pg. 103–108, TRB, National Research Council, Washington, D.C., 2000a.

Weed R.M. Method to Establish Pay Schedules for Rigid Pavement. Transportation Research Record 885, pg.18–24, Transportation Research board, National Research Council, Washington, D.C., 1982.

Weed, R. M. Mathematical Modeling of Pavement Smoothness. Transportation Research Record: Journal of the Transportation Research Board, No. 02-2223, pg. 159–163, TRB, National Research Council, Washington, D.C., 2002.

Weed, R. M. Mathematical Modeling Procedures for Performance-Related Specifications. Transportation Research Record: Journal of the Transportation Research Board, No. 1946 TRB, National Research Council, Washington, D.C., 2006.

Weed, R.M. Method to Model Performance Relationships and Pay Schedules. Transportation Research Record: Journal of the Transportation Research Board, No. 1712, pg.117–124, TRB, National Research Council, Washington, D.C., 2000b.

Weed, R. M. Multi characteristic Performance-Related Specification for Hot-Mix Asphalt Pavement. In Transportation Research Record: Journal of the Transportation Research Board, No. 1861, pg. 53–59, TRB, National Research Council, Washington, D.C., 2003.

Weed, M. R. New Jersey's Superpave Specification. In Transportation Research Record: Journal of the Transportation Research Board, No. 1761, pg.10–14, TRB, National Research Council, Washington, D.C. 2001a.

Weed, R. M. Practical Framework for Performance-Related Specifications. In Transportation Research Record 1654, TRB, National Research Council, Washington, D.C., pg. 81–87, 1999.

Weed, R. M. Quality Assurance Software for the Personal Computer. FHWA, Demonstration Projects Division, U.S. Department of Transportation, Washington, D.C. Transportation Research Record 1544, pg. 116–124, 1996.

Weed, R.M. Statistical Specification Development. 2nd ed., FHWA/NJ-88-017, New Jersey Department of Transportation, Trenton, March 24, 1989.

Welborn, J.Y. State of the Art in Asphalt Pavement Specifications. FHWA/RD-84/075, Sheladia Associates, Inc., Riverdale, Maryland, July 1984.

Willoughby, K. A, Mahoney, J.P., Pierce, L.M., Uhlmeyer, J.S, Anderson, K.W., Read, S.A., Muench, S.T., Thompson, T.R., & Moore, R. Construction-Related Asphalt Concrete Pavement Temperature Differentials and the Corresponding Density Differentials. Report No. WA-RD 476.1, Washington State Department of Transportation. July 2001.

Yu, J. Pavement Service Life Estimation and Condition Prediction. Doctorate's Thesis, The University of Toledo, December 2005.

Appendix A: HMA Pay Adjustments

Table A.1: Production pay adjustment

Deviation from Lab Density	Existing PAF	Revised PAF
-1.8	0.720	0.720
-1.7	0.755	0.755
-1.6	0.790	0.790
-1.5	0.825	0.825
-1.4	0.860	0.860
-1.3	0.895	0.895
-1.2	0.930	0.930
-1.1	0.965	0.965
-1	1.000	1.000
-0.9	1.006	1.005
-0.8	1.013	1.010
-0.7	1.019	1.015
-0.6	1.025	1.020
-0.5	1.031	1.025
-0.4	1.038	1.030
-0.3	1.044	1.035
-0.2	1.050	1.040
-0.1	1.050	1.045
0	1.050	1.050
0.1	1.050	1.050
0.2	1.050	1.050
0.3	1.044	1.050
0.4	1.038	1.050
0.5	1.031	1.042
0.6	1.025	1.033
0.7	1.019	1.025
0.8	1.013	1.017
0.9	1.006	1.008
1	1.000	1.000
1.1	0.965	0.965

Deviation from Lab Density	Existing PAF	Revised PAF
1.2	0.930	0.930
1.3	0.895	0.895
1.4	0.860	0.860
1.5	0.825	0.825
1.6	0.790	0.790
1.7	0.755	0.755
1.8	0.720	0.720

Table A.2: Placement pay adjustment

In-Place Voids	Existing PAF	Revised PAF
2.7	0.705	0.750
2.8	0.720	0.750
2.9	0.735	0.750
3	0.750	0.750
3.1	0.765	0.770
3.2	0.780	0.790
3.3	0.795	0.810
3.4	0.810	0.830
3.5	0.825	0.850
3.6	0.840	0.870
3.7	0.855	0.890
3.8	0.870	0.910
3.9	0.885	0.930
4	0.900	0.950
4.1	0.915	0.970
4.2	0.930	0.990
4.3	0.945	1.010
4.4	0.960	1.030
4.5	0.975	1.050
4.6	0.990	1.050
4.7	1.005	1.050
4.8	1.020	1.050
4.9	1.035	1.050
5	1.050	1.050
5.1	1.050	1.050
5.2	1.050	1.050
5.3	1.050	1.050
5.4	1.050	1.050
5.5	1.050	1.050
5.6	1.050	1.050
5.7	1.050	1.050
5.8	1.050	1.050

In-Place Voids	Existing PAF	Revised PAF
5.9	1.050	1.050
6	1.050	1.050
6.1	1.048	1.048
6.2	1.046	1.046
6.3	1.044	1.044
6.4	1.042	1.042
6.5	1.040	1.040
6.6	1.038	1.038
6.7	1.036	1.036
6.8	1.034	1.034
6.9	1.032	1.032
7	1.030	1.030
7.1	1.028	1.028
7.2	1.026	1.026
7.3	1.024	1.024
7.4	1.022	1.022
7.5	1.020	1.020
7.6	1.018	1.018
7.7	1.016	1.016
7.8	1.014	1.014
7.9	1.012	1.012
8	1.010	1.010
8.1	1.008	1.008
8.2	1.006	1.006
8.3	1.004	1.004
8.4	1.002	1.002
8.5	1.000	1.000
8.6	0.998	0.983
8.7	0.996	0.967
8.8	0.994	0.950
8.9	0.992	0.933
9	0.990	0.917
9.1	0.960	0.900

In-Place Voids	Existing PAF	Revised PAF
9.2	0.930	0.883
9.3	0.900	0.867
9.4	0.870	0.850
9.5	0.840	0.833
9.6	0.810	0.817
9.7	0.780	0.800
9.8	0.750	0.783
9.9	0.720	0.767
10		0.750

Table A.3: Ride quality pay adjustment

Avg IRI for each 0.1 mile of traffic lane	Schedule 1	Schedule 2	Schedule 3
<30	1000	1000	500
30	1000	1000	500
31	947	964	482
32	895	929	464
33	843	893	447
34	792	858	429
35	740	823	412
36	689	788	394
37	638	754	377
38	588	719	360
39	538	685	343
40	488	651	325
41	438	617	309
42	389	583	292
43	339	550	275
44	290	516	258
45	241	483	241
46	193	450	225
47	144	417	208
48	96	384	192
49	48	351	176
50	0	318	159
51	0	286	143
52	0	254	127
53	0	222	111
54	0	189	95
55	0	158	79
56	0	126	63
57	0	94	47
58	0	63	31
59	0	31	16
60	0	0	0

Avg IRI for each 0.1 mile of traffic lane	Schedule 1	Schedule 2	Schedule 3
61	-30	0	0
62	-60	0	0
63	-90	0	0
64	-119	0	0
65	-149	0	0
66	-179	0	0
67	-208	0	0
68	-237	0	0
69	-267	0	0
70	-296	0	0
71	-325	-42	0
72	-354	-84	0
73	-383	-126	0
74	-412	-167	0
75	-441	-208	0
76	-469	-249	0
77	-498	-290	0
78	-527	-331	0
79	-555	-372	0
80	-583	-412	0
81	-612	-453	0
82	-640	-493	0
83	-668	-533	0
84	-696	-573	0
85	-724	-612	0
86	-752	-652	0
87	-780	-691	0
88	-808	-730	0
89	-835	-769	0
90	-863	-808	0
91	-891	-847	0
92	-918	-885	0
93	-945	-924	0

Avg IRI for each 0.1 mile of traffic lane	**Schedule 1**	**Schedule 2**	**Schedule 3**
94	-973	-962	0
95	-1000	-1000	0
>95	Remove	Remove	Remove